Web Appendix to

Filtering without recursion and some of its uses in financial economics

Simon Donker van Heel
Econometric Institute, Erasmus University Rotterdam, and Tinbergen Institute

Neil Shephard
Department of Economics and Department of Statistics, Harvard University

2 September 2026

https://simondonkervanheel.com/filtering-web-appendix/

About this appendix

This Web Appendix accompanies "Filtering without recursion and some of its uses in financial economics". It gives the full per-asset results behind the paper's appendix "Replicating results over 54 other assets". There is one chapter per asset, 54 stocks and one Exchange Traded Fund (ETF), `SPY`, over the 22 trading days 17 September to 16 October 2024, of which the last 12 are the days we analyse; the first 10 feed each asset's rolling diurnal profile and the opening warm-up of its noise scale. The single-day exhibits use 16 October. Every asset runs through the same pipeline, the five filters and their tuning selections, with the same common values, a 3 trade half-life for the price filter and a 4 minute half-life for the noise scale, and with the Mandelbrot parameter a selected per asset.

Every chapter shows the same exhibits. These are the selected tunings and what running the common values costs in forecast loss, the noise scale by time of day, the volatility signature, the spot volatility paths,[1] the window around the largest Huber residual, the outlier direction filter and its filtered conditional probability, the tables for 16 October, the tail-robustness panel, and the diagnostics. Chapters are ordered by the average trading rate over the days we analyse.

[1] The spot volatility paths in this version are computed without the finite-sample correction $c_{t,s}$ of the paper. An updated version of this appendix will follow shortly.

Assets

1 NVDA, 70 trades per second **5**

2 TSLA, 44 trades per second **14**

3 AAPL, 24 trades per second **23**

4 the ETF SPY, 17 trades per second **32**

5 AMD, 17 trades per second **41**

6 AMZN, 16 trades per second **50**

7 MSFT, 15 trades per second **59**

8 AVGO, 10 trades per second **68**

9 MU, 9.2 trades per second **77**

10 INTC, 8.9 trades per second **86**

11 META, 8.7 trades per second **95**

12 GOOG, 8.2 trades per second **104**

13 COIN, 5.8 trades per second **113**

14 MRVL, 5.4 trades per second **122**

15 RIVN, 4.6 trades per second **131**

16 HOOD, 4.5 trades per second **140**

17 CELH, 4.4 trades per second **149**

18 QCOM, 4.3 trades per second **158**

19 PYPL, 4.3 trades per second **167**

20 CSCO, 4.0 trades per second **176**

21 PEP, 4.0 trades per second **185**

22 AMAT, 3.9 trades per second **194**

23 SBUX, 3.7 trades per second **203**

24 NFLX, 3.7 trades per second **212**

25 CRWD, 3.4 trades per second **221**

26 ADBE, 3.1 trades per second **230**

27 COST, 3.0 trades per second **239**

28 DKNG, 2.8 trades per second **248**

29 TXN, 2.7 trades per second **257**

30 PANW, 2.4 trades per second **266**

31 ABNB, 2.3 trades per second **275**

32 GILD, 2.2 trades per second **284**

33 DDOG, 2.1 trades per second **293**

34 ETSY, 1.9 trades per second **302**

35 LCID, 1.9 trades per second **311**

36 TTD, 1.6 trades per second **320**

37 ROKU, 1.5 trades per second **329**

38 TEAM, 1.5 trades per second **338**

39 INTU, 1.4 trades per second **347**

40 AMGN, 1.4 trades per second **356**

41 ISRG, 1.3 trades per second **365**

42 VRTX, 1.1 trades per second **374**

43 REGN, 0.91 trades per second **383**

44 CROX, 0.87 trades per second **392**

45 MELI, 0.83 trades per second **401**

46 CALM, 0.71 trades per second **410**

47 ORLY, 0.60 trades per second **419**

48 WING, 0.57 trades per second **428**

49 BKNG, 0.56 trades per second **437**

50 DUOL, 0.54 trades per second **446**

51 MIDD, 0.46 trades per second **455**

52 SAIA, 0.39 trades per second **464**

53 FIZZ, 0.22 trades per second **473**

54 LANC, 0.19 trades per second **482**

55 IPAR, 0.19 trades per second **491**

NVDA 1 to 16 October 2024

trades on 16 October 2024	1,521,339
average trading rate	69.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	16,314 (1.07%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.28%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.08%
direction half-life[3]	13 s [9, 21]
Mandelbrot a minimizing QLIKE	536 trades
a within 0.0005 of the minimum	329 to 873
a run in the volatility filter	563 trades
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	21% $\to$ 250%
Huber $\hat{\sigma}$, before $\to$ peak	18% $\to$ 19%
median spot volatility, Huber / linear input[5]	31.7 / 38.7 %
hyperbolic against HExp, correlation of logs	0.9987
HExp QLIKE advantage, leave-one-day-out[6]	+0.00024
15-minute signature level	20 to 57 %
Huber signature, 1 min over 1 s, 12 day average	1.03

NVDA over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.115	0.090	0.106	0.442	0.537	0.431	0.493	2.995
2	0.114	0.103	0.125	0.362	0.467	0.389	0.487	2.159
5	0.136	0.123	0.150	0.281	0.418	0.366	0.460	1.405
10	0.151	0.139	0.161	0.241	0.382	0.348	0.418	1.018
25	0.162	0.153	0.166	0.207	0.333	0.315	0.351	0.682
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.24	−0.18	−0.02	−0.48	1.73	1.64	1.74	1.77
2	0.01	0.01	−0.04	−0.01	1.60	1.48	1.91	1.80
5	−0.02	−0.02	−0.03	0.00	1.51	1.36	1.94	1.83
10	−0.01	−0.01	−0.01	0.00	1.41	1.31	1.79	1.89
25	−0.01	−0.01	0.00	0.00	1.55	1.48	1.67	2.10

Table 1: Trade-lag differences, `NVDA`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.50	2.34	2.58	3.54	38.8	35.3	62.6	245.2
1 s	1.95	1.91	1.98	2.20	33.5	32.7	40.5	77.9
5 s	1.95	1.94	1.96	2.03	34.1	33.9	35.5	44.5
10 s	1.98	1.97	1.98	2.02	34.1	34.0	34.9	46.6
1 min	1.93	1.93	1.92	1.94	33.3	33.3	33.3	35.7
15 min	2.06	2.06	2.06	2.06	37.1	37.1	37.1	37.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.26	−0.24	−0.40	−0.81	1.83	1.81	2.05	2.31
1 s	0.05	0.07	0.00	−0.37	3.32	3.38	3.81	2.64
5 s	0.08	0.09	0.06	−0.03	7.57	7.63	6.62	4.87
10 s	0.04	0.04	0.03	−0.12	8.42	8.44	8.16	6.19
1 min	0.06	0.06	0.06	−0.02				

Table 2: Clock-lag differences, `NVDA`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.30	0.38	0.41
s.d. of $\hat{\epsilon}_t$ in cents	2.18	2.16	1.84
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	1.74
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.28	-0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.17	-0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.06	-0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	0.02	-0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	-0.26,0.26	-0.28,0.28
95% C.I. for $\theta_t - X_t$	—	-0.46,0.46	-0.54,0.54
99% C.I. for $\theta_t - X_t$	—	-0.72,0.75	-1.23,1.15
s.e. of $\theta_t - X_t$ in cents	—	0.22	∞(0.56)

Table 3: Residual summaries, `NVDA`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. does not exist in theory, the sample value is in brackets.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.22	1.25	1.06 to 1.48	1.10 to 1.40
Residual divided by δ_t	1.21	1.22	1.09 to 1.47	1.10 to 1.38
Repricing windows removed	1.25	1.22	1.12 to 1.49	1.09 to 1.37
Calm windows only	1.35	1.46		
Every 10th trade	1.19	1.24	1.04 to 1.49	1.10 to 1.40
Six-trade half-life	1.27	1.29	1.09 to 1.59	1.13 to 1.50

Table 4: Hill estimator of the tail index of the Huber filter residual, `NVDA`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 10.1 to 11.6 cents over the days and removes 7% to 19% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 5.0 cents on 16 October 2024, 25% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

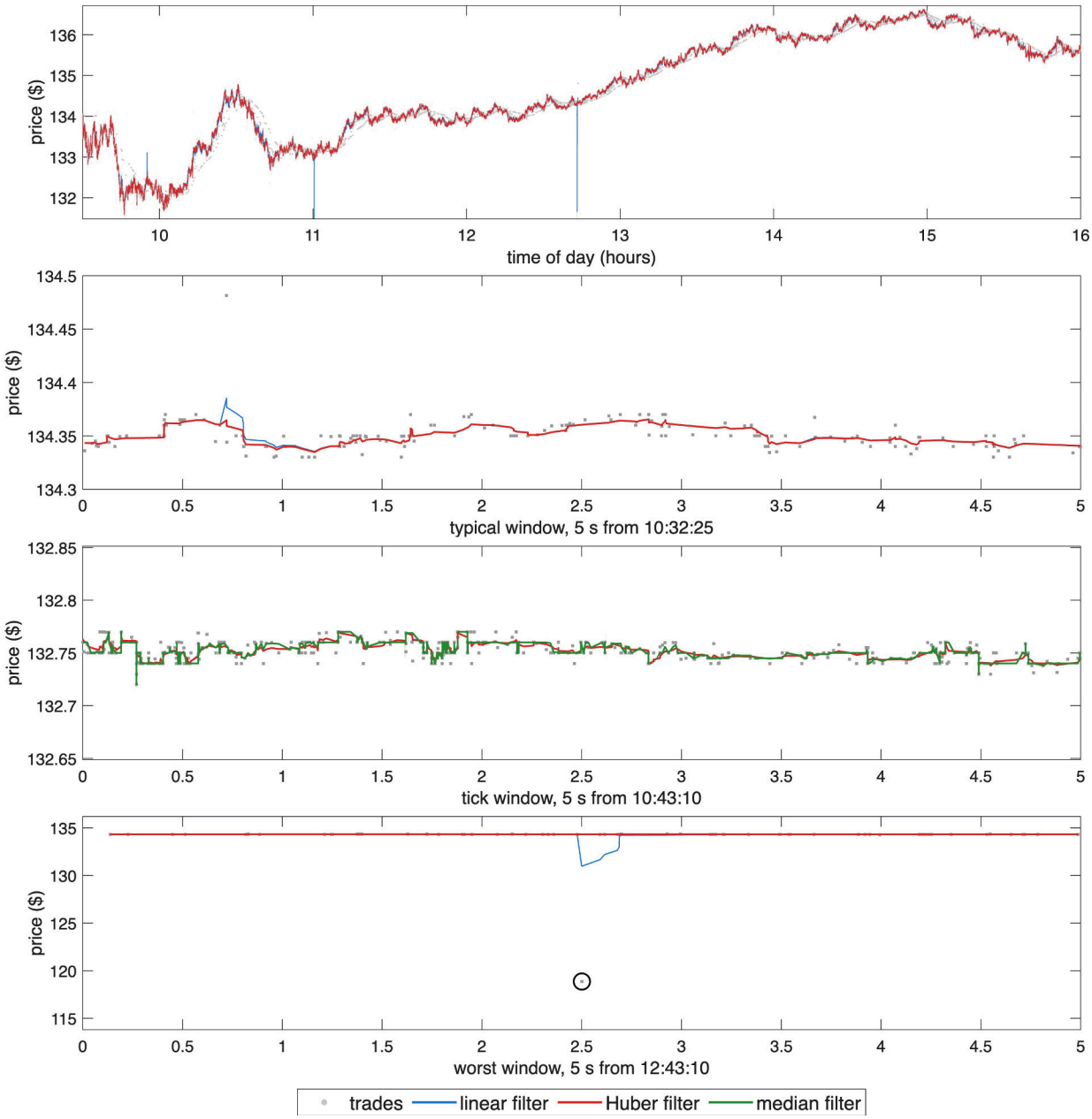

Figure 1: `NVDA` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:32:25. 3rd: 5-second window at 10:43:10 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

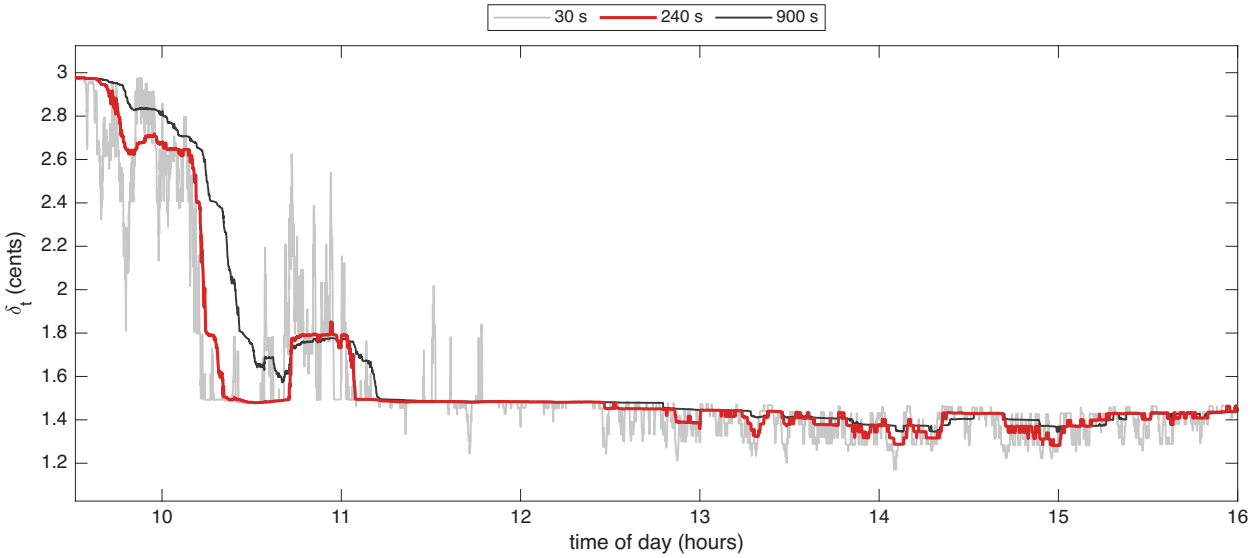

Figure 2: The Huber threshold δ_t in cents for `NVDA` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

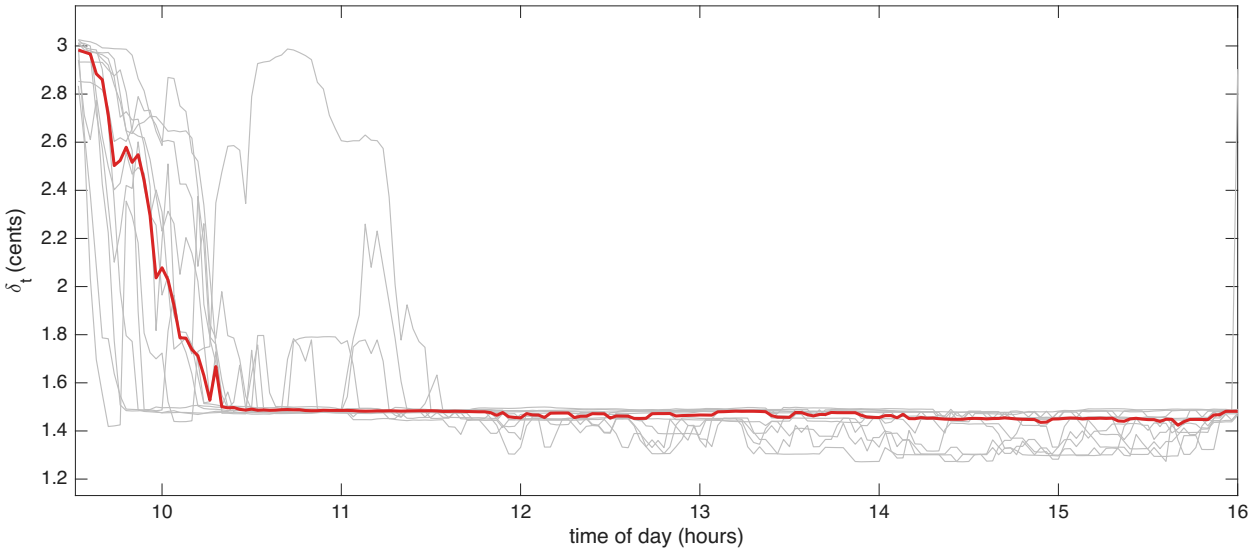

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `NVDA` trading days in Oct 2024 (grey lines) and their median (red line).

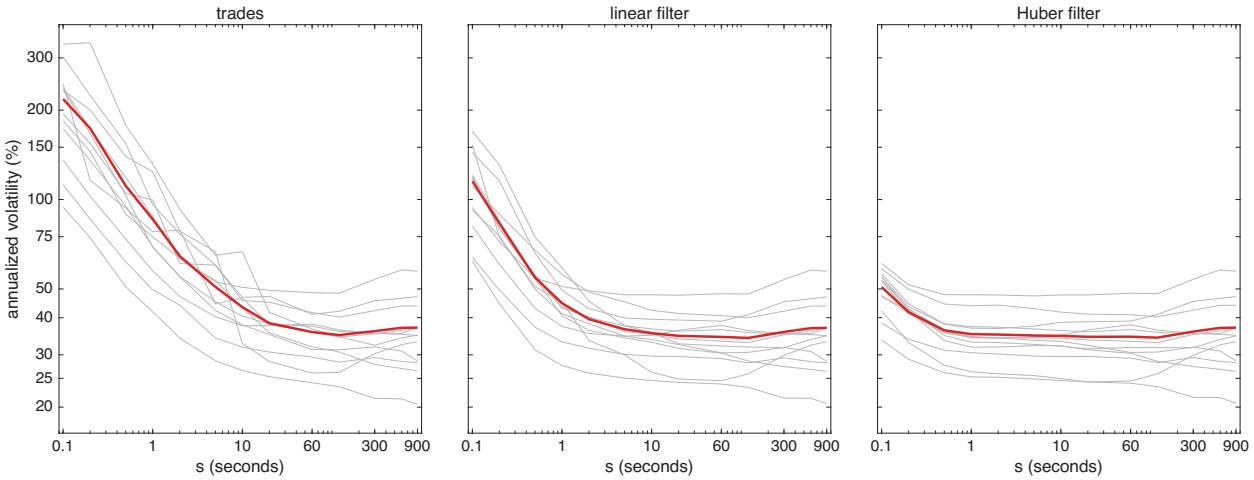

Figure 4: Volatility signature curves for `NVDA` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

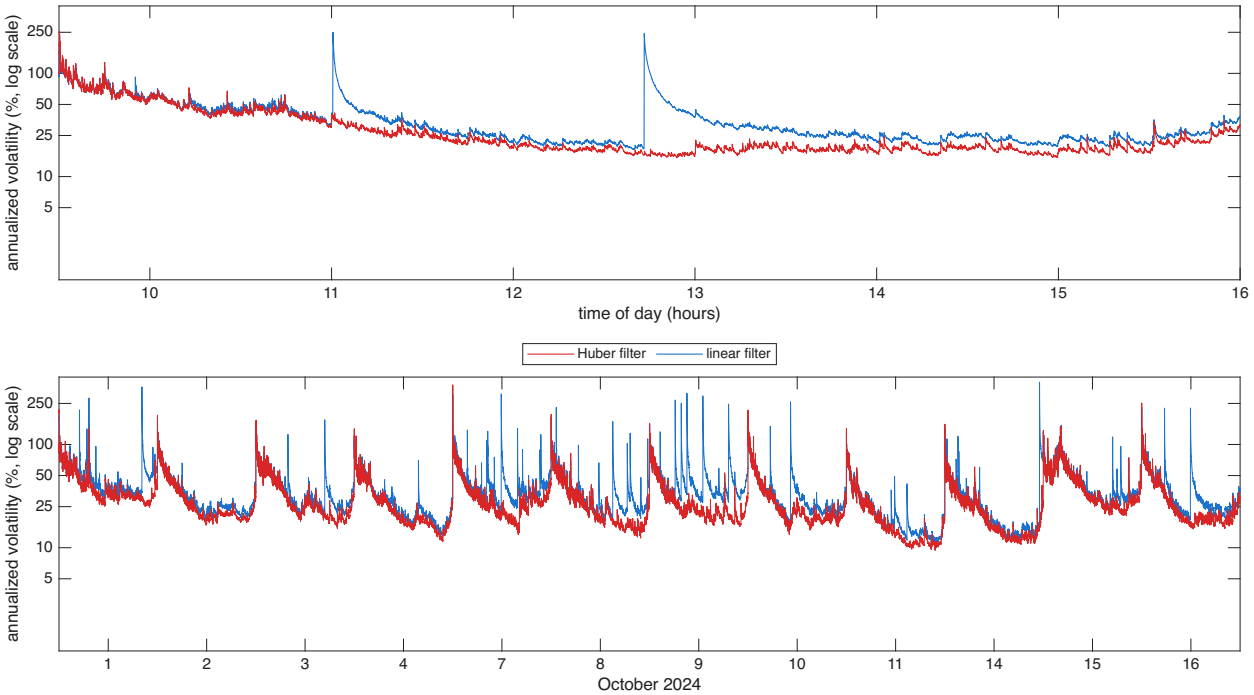

Figure 5: Filtered spot volatility for `NVDA` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 563$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

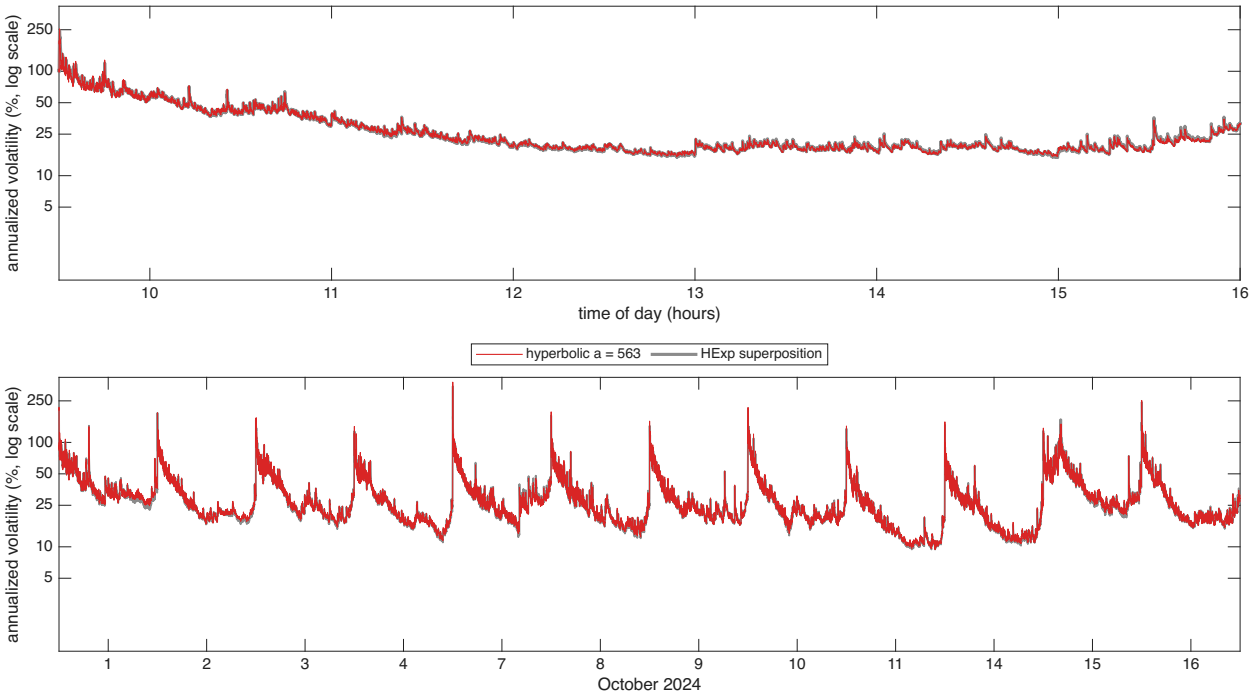

Figure 6: Hyperbolic against superposition weights for `NVDA` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 563$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

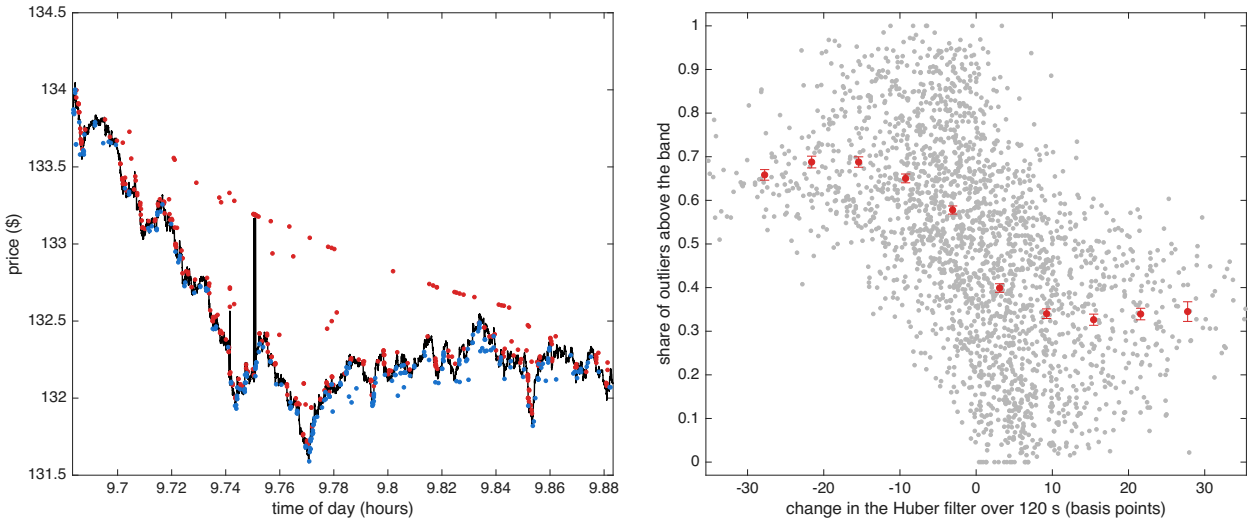

Figure 7: Outliers & robust price for `NVDA` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

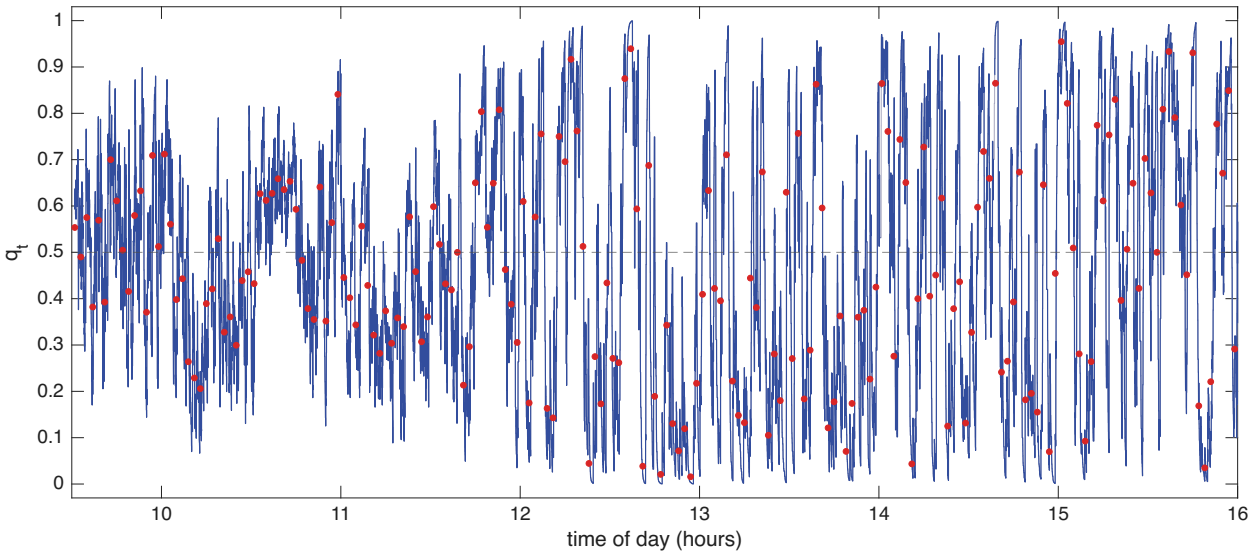

Figure 8: The directional conditional probability for NVDA, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 13 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

TSLA 1 to 16 October 2024

trades on 16 October 2024	693,626
average trading rate	44.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	16,860 (2.43%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+2.23%
scale half-life: run / minimizer[2]	240 / 45 s
cost of running 240 s	+0.18%
direction half-life[3]	14 s [11, 21]
Mandelbrot a minimizing QLIKE	359 trades
a within 0.0005 of the minimum	210 to 614
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	72% $\to$ 92%
Huber $\hat{\sigma}$, before $\to$ peak	71% $\to$ 93%
median spot volatility, Huber / linear input[5]	35.6 / 38.3 %
hyperbolic against HExp, correlation of logs	0.9985
HExp QLIKE advantage, leave-one-day-out[6]	-0.00069
15-minute signature level	27 to 54 %
Huber signature, 1 min over 1 s, 12 day average	1.09

TSLA over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:55; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.355	0.264	0.276	1.225	0.948	0.561	0.564	3.311
2	0.340	0.295	0.317	1.000	0.841	0.569	0.593	2.457
5	0.392	0.342	0.370	0.770	0.780	0.601	0.623	1.670
10	0.415	0.372	0.394	0.645	0.721	0.607	0.622	1.261
25	0.415	0.386	0.397	0.529	0.633	0.575	0.581	0.903
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.03	0.10	−0.45	2.76	2.43	2.13	2.16
2	0.05	0.07	0.04	−0.01	2.75	2.31	2.33	2.19
5	−0.01	0.00	−0.01	0.00	2.48	2.23	2.49	2.24
10	−0.01	−0.01	−0.01	0.00	2.28	2.19	2.50	2.30
25	0.00	−0.01	−0.01	0.00	2.43	2.43	2.87	2.45

Table 1: Trade-lag differences, `TSLA`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	5.69	5.18	5.35	8.29	36.6	29.5	32.7	108.3
1 s	3.64	3.51	3.58	4.29	33.1	31.3	32.4	50.9
5 s	3.55	3.51	3.52	3.73	33.8	33.4	33.5	38.0
10 s	3.43	3.40	3.41	3.52	33.8	33.5	33.5	35.7
1 min	3.27	3.26	3.26	3.28	34.4	34.3	34.2	34.5
15 min	2.50	2.50	2.49	2.50	26.8	26.8	26.7	26.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.10	−0.07	−0.12	−0.40	3.34	3.02	2.70	2.23
1 s	0.17	0.22	0.21	−0.03	3.96	4.37	4.31	3.47
5 s	0.13	0.14	0.15	0.06	9.66	9.75	9.40	9.35
10 s	0.15	0.16	0.16	0.13	8.23	8.06	8.83	7.07
1 min	−0.08	−0.08	−0.08	−0.08				

Table 2: Clock-lag differences, `TSLA`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.82	1.02	1.06
s.d. of $\hat{\epsilon}_t$ in cents	2.41	2.34	2.17
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.13
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.18	0.21	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.05	0.12	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.02	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	0.00	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.76,0.68	−0.80,0.73
95% C.I. for $\theta_t - X_t$	—	−1.35,1.25	−1.49,1.40
99% C.I. for $\theta_t - X_t$	—	−2.02,2.00	−2.47,2.46
s.e. of $\theta_t - X_t$ in cents	—	0.61	0.80

Table 3: Residual summaries, `TSLA`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.13, is above two, while the Hill index of the Huber residual itself on this day is 1.77, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.77	1.78	1.48 to 1.77	1.46 to 1.84
Residual divided by δ_t	1.62	1.74	1.43 to 1.62	1.38 to 1.74
Repricing windows removed	1.77	1.82	1.46 to 1.95	1.41 to 1.98
Calm windows only	1.67	1.85		
Every 10th trade	1.83	1.87	1.43 to 1.84	1.48 to 1.90
Six-trade half-life	1.86	1.84	1.55 to 1.88	1.53 to 1.90

Table 4: Hill estimator of the tail index of the Huber filter residual, `TSLA`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.1 to 22.5 cents over the days and removes 8% to 20% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 10.1 cents on 16 October 2024, 37% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

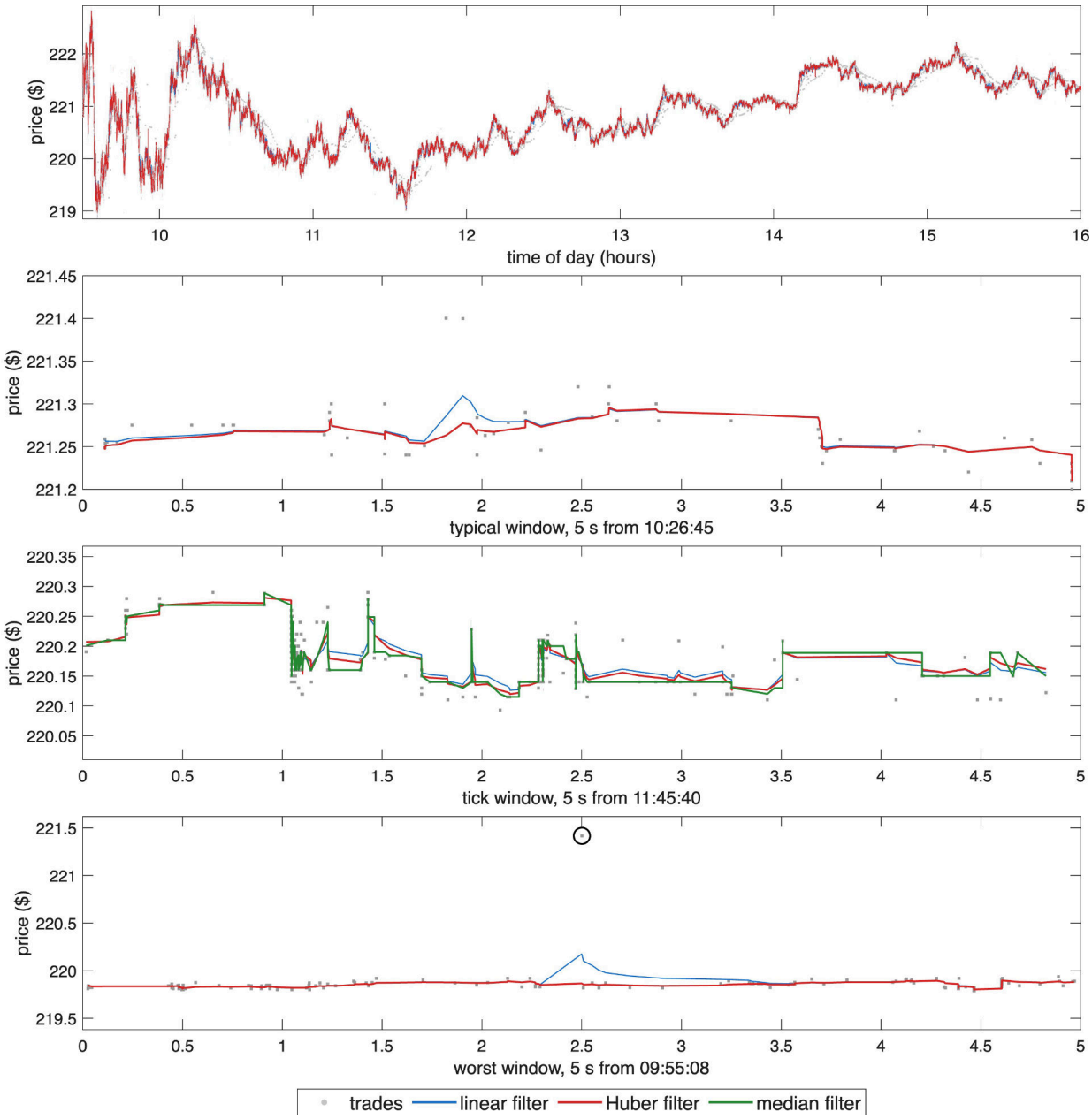

Figure 1: `TSLA` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:26:45. 3rd: 5-second window at 11:45:40 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

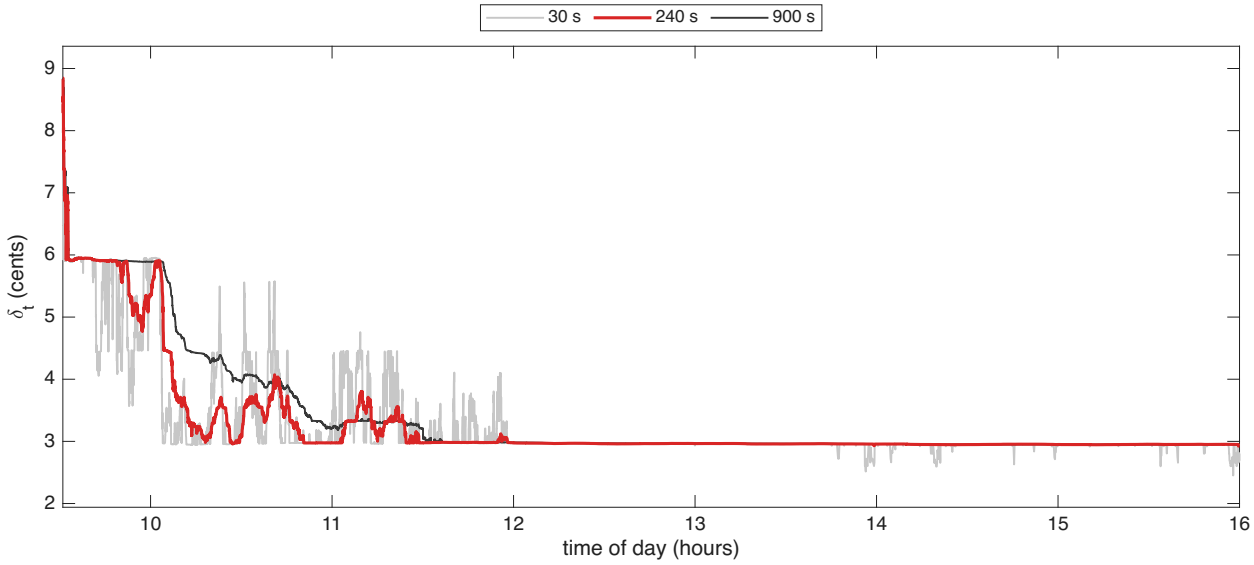

Figure 2: The Huber threshold δ_t in cents for `TSLA` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

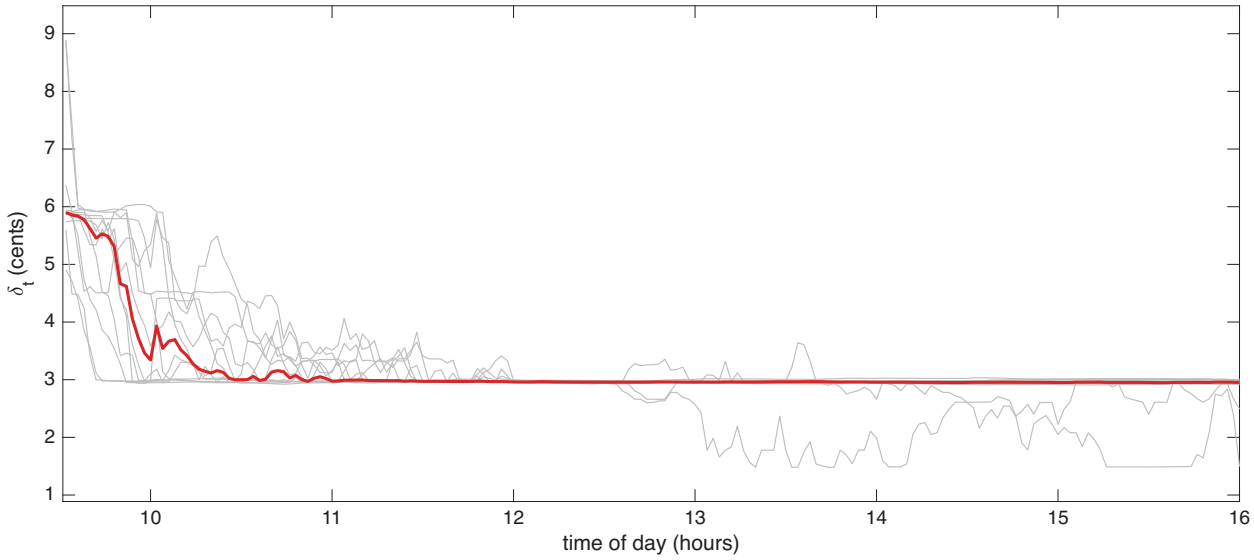

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `TSLA` trading days in Oct 2024 (grey lines) and their median (red line).

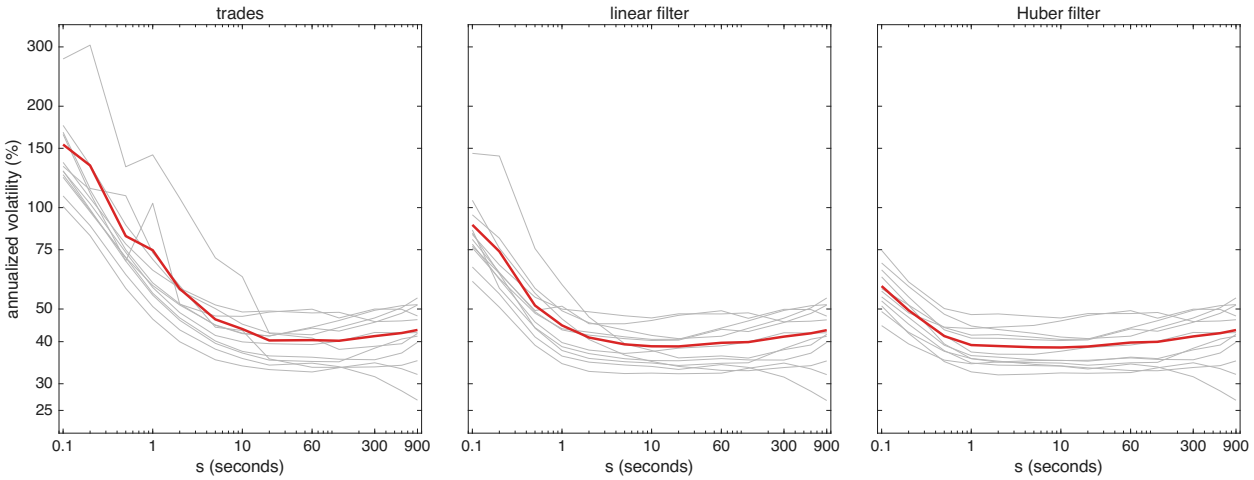

Figure 4: Volatility signature curves for `TSLA` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

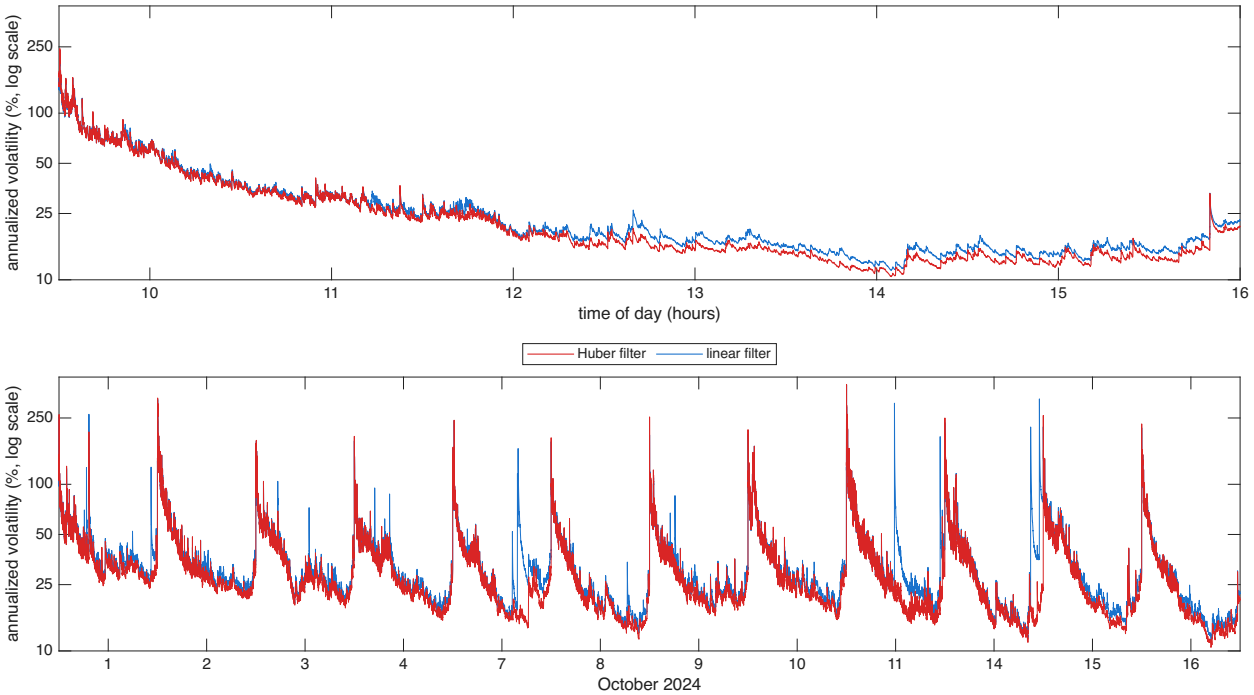

Figure 5: Filtered spot volatility for `TSLA` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 359$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

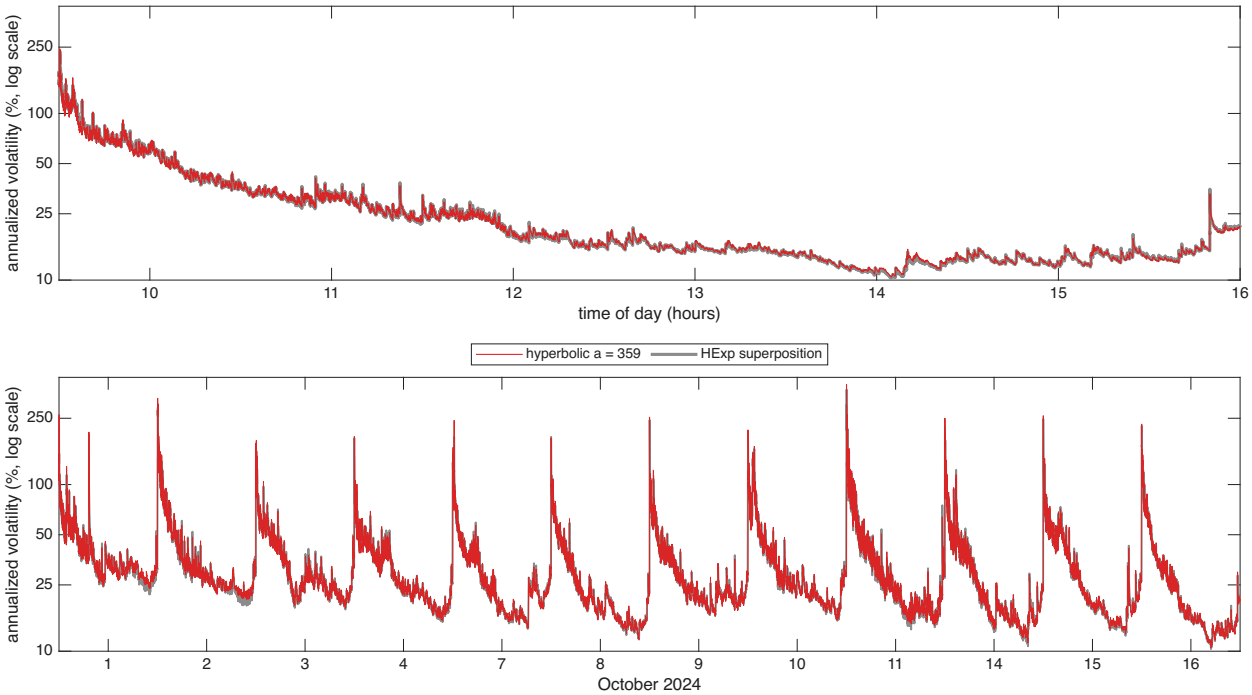

Figure 6: Hyperbolic against superposition weights for `TSLA` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 359$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

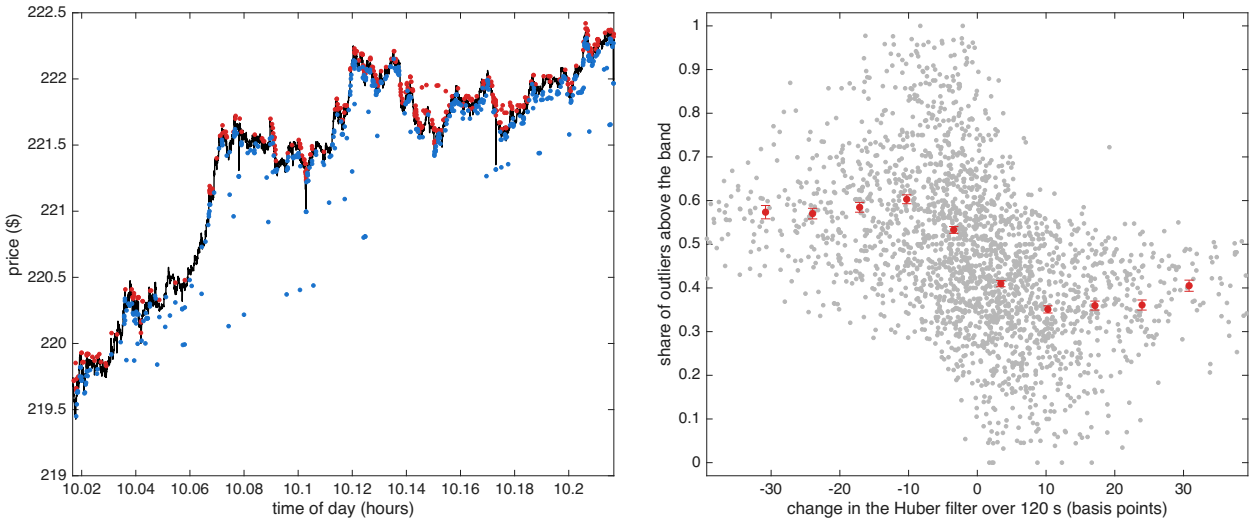

Figure 7: Outliers & robust price for `TSLA` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

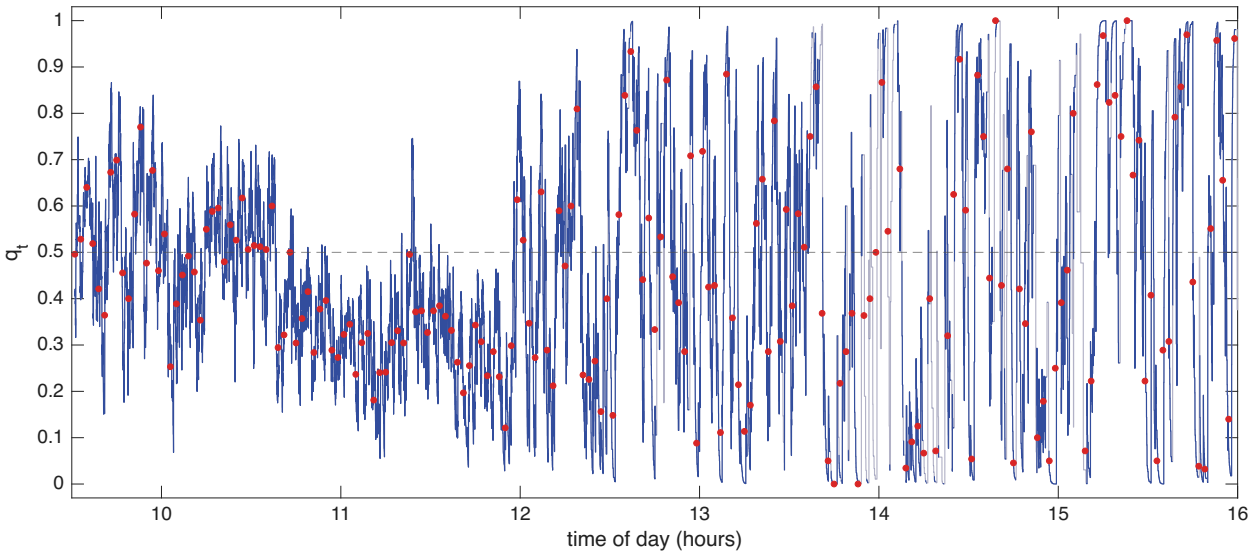

Figure 8: The directional conditional probability for `TSLA`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 14 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

On 11 October 2024, 32 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 22 per cent on any other day. Every number that day contributes to carries that burst.
The largest Huber residual on 16 October 2024 occurs at 09:35, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:55.

Diagnostics.

AAPL 1 to 16 October 2024

trades on 16 October 2024	534,087
average trading rate	23.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	6,972 (1.31%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+3.21%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.02%
direction half-life[3]	20 s [15, 27]
Mandelbrot a minimizing QLIKE	143 trades
a within 0.0005 of the minimum	92 to 234
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	12% $\to$ 54%
Huber $\hat{\sigma}$, before $\to$ peak	7% $\to$ 12%
median spot volatility, Huber / linear input[5]	13.7 / 16.6 %
hyperbolic against HExp, correlation of logs	0.9991
HExp QLIKE advantage, leave-one-day-out[6]	-0.00482
15-minute signature level	11 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.06

AAPL over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:56; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.128	0.105	0.121	0.490	0.434	0.262	0.347	2.110
2	0.131	0.123	0.146	0.409	0.383	0.261	0.352	1.539
5	0.161	0.152	0.177	0.326	0.356	0.279	0.350	1.012
10	0.184	0.174	0.195	0.287	0.339	0.291	0.339	0.750
25	0.204	0.195	0.207	0.256	0.324	0.301	0.323	0.531
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.22	0.00	0.03	−0.47	1.95	1.91	2.24	2.70
2	0.06	0.10	−0.01	−0.01	2.14	1.81	2.60	2.77
5	0.00	0.02	−0.01	0.00	2.15	1.90	2.65	2.81
10	−0.01	0.00	−0.01	0.00	2.22	2.10	2.56	2.92
25	0.01	0.01	0.00	0.00	2.51	2.47	2.67	3.09

Table 1: Trade-lag differences, `AAPL`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.34	2.18	2.33	3.54	12.9	10.5	16.2	68.9
1 s	1.75	1.71	1.76	2.05	14.8	14.2	15.8	29.8
5 s	1.61	1.59	1.60	1.72	15.5	15.3	15.5	19.4
10 s	1.59	1.58	1.58	1.66	15.4	15.3	15.5	20.5
1 min	1.50	1.49	1.49	1.51	14.2	14.2	14.0	14.8
15 min	1.40	1.40	1.40	1.41	12.8	12.8	12.8	12.9
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.00	0.03	0.02	−0.37	2.91	2.81	2.66	3.65
1 s	−0.06	−0.05	−0.06	−0.25	4.02	3.98	3.99	4.27
5 s	−0.04	−0.03	−0.03	−0.12	6.08	6.17	6.02	5.89
10 s	−0.05	−0.05	−0.05	−0.12	5.45	5.35	5.06	4.73
1 min	−0.08	−0.08	−0.08	−0.09				

Table 2: Clock-lag differences, `AAPL`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.35	0.44	0.47
s.d. of $\hat{\epsilon}_t$ in cents	1.54	1.50	1.33
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.24
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.24	0.30	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.20	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.08	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.03	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	−0.34,0.33	−0.37,0.36
95% C.I. for $\theta_t - X_t$	—	−0.62,0.61	−0.71,0.70
99% C.I. for $\theta_t - X_t$	—	−0.99,1.01	−1.38,1.39
s.e. of $\theta_t - X_t$ in cents	—	0.29	0.45

Table 3: Residual summaries, `AAPL`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.24, is above two, while the Hill index of the Huber residual itself on this day is 1.62, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.62	1.69	1.28 to 1.74	1.27 to 1.69
Residual divided by δ_t	1.69	1.71	1.33 to 1.69	1.38 to 1.84
Repricing windows removed	1.56	1.66	1.32 to 1.63	1.32 to 1.66
Calm windows only	1.53	1.65		
Every 10th trade	1.51	1.73	1.38 to 1.69	1.28 to 1.73
Six-trade half-life	1.74	1.79	1.33 to 1.89	1.32 to 1.79

Table 4: Hill estimator of the tail index of the Huber filter residual, `AAPL`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 17.3 to 20.1 cents over the days and removes 2% to 4% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 8.7 cents on 16 October 2024, 74% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

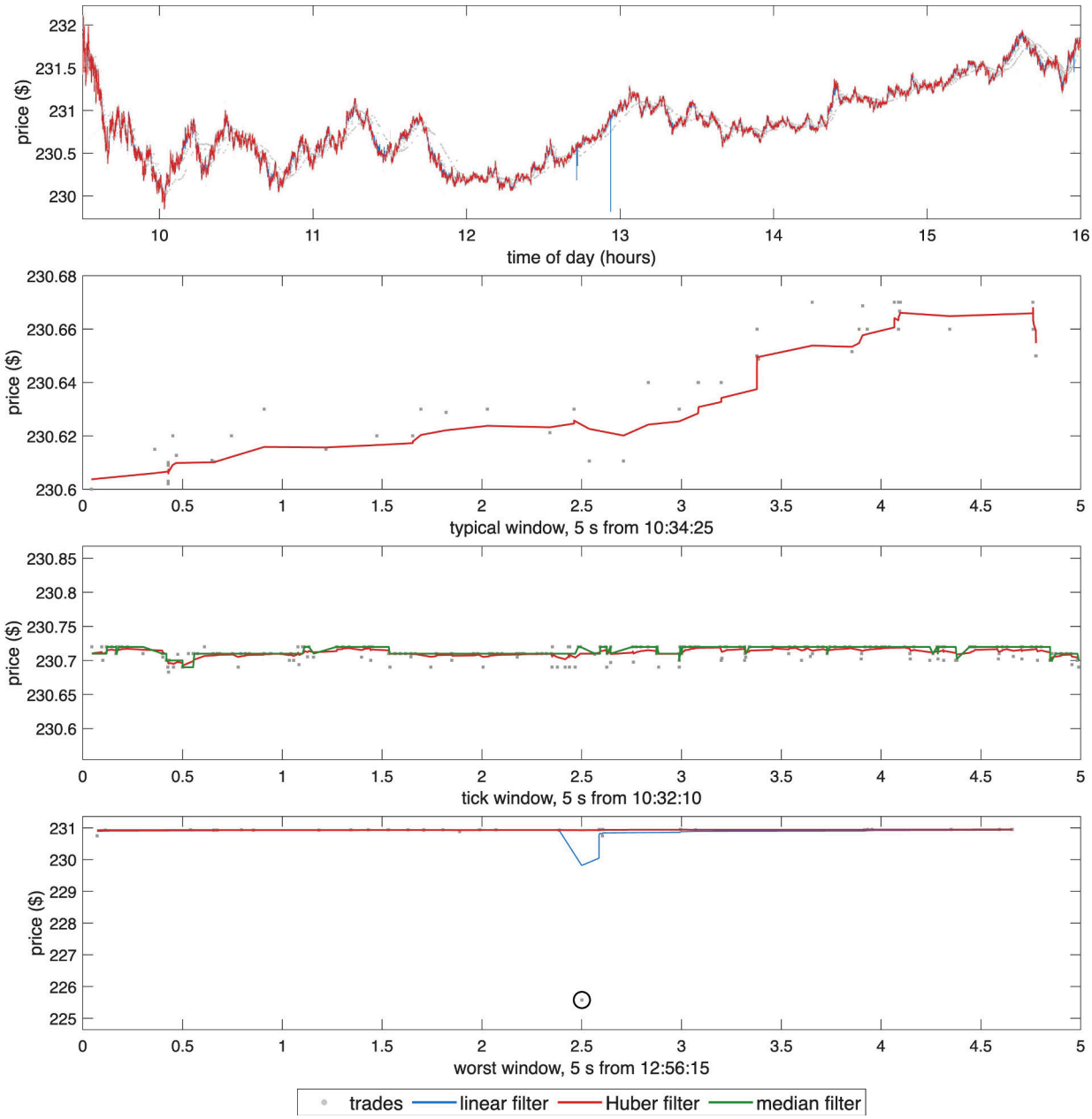

Figure 1: `AAPL` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:34:25. 3rd: 5-second window at 10:32:10 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

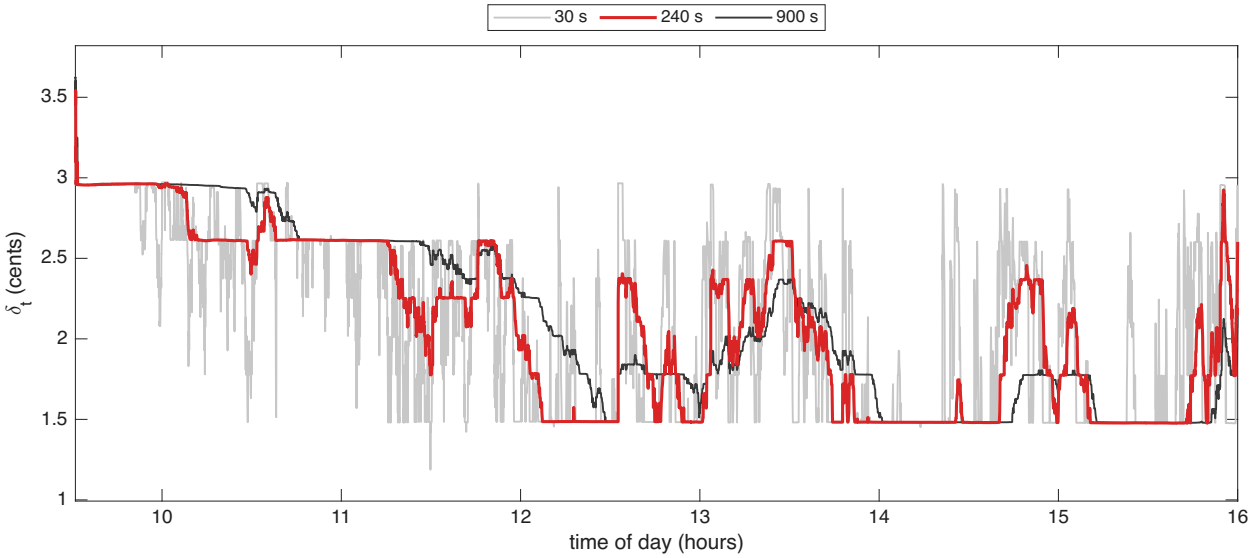

Figure 2: The Huber threshold δ_t in cents for `AAPL` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

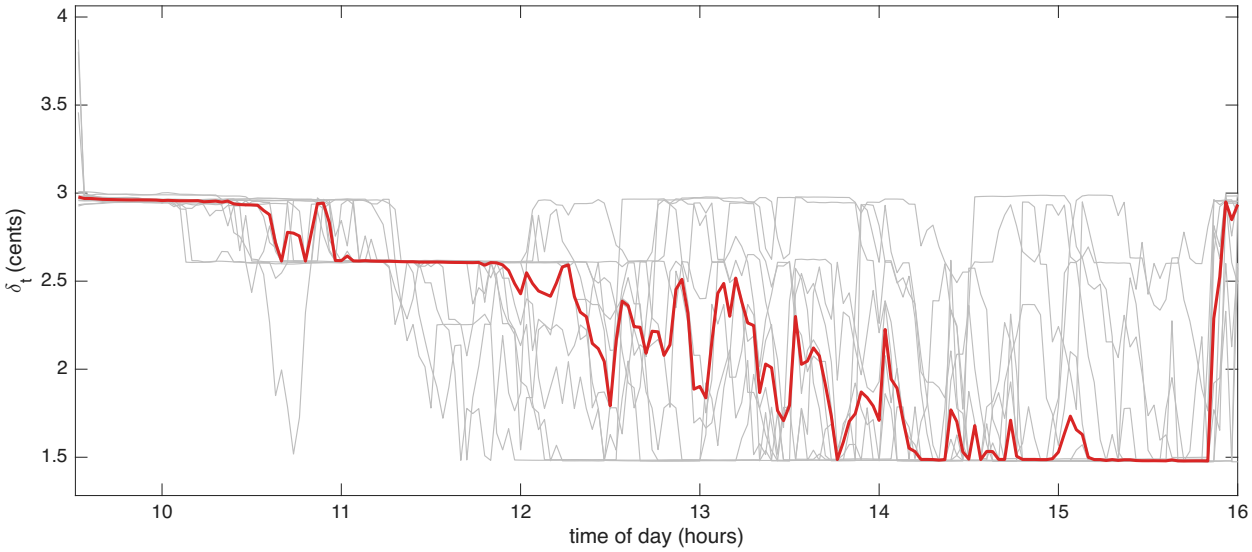

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AAPL` trading days in Oct 2024 (grey lines) and their median (red line).

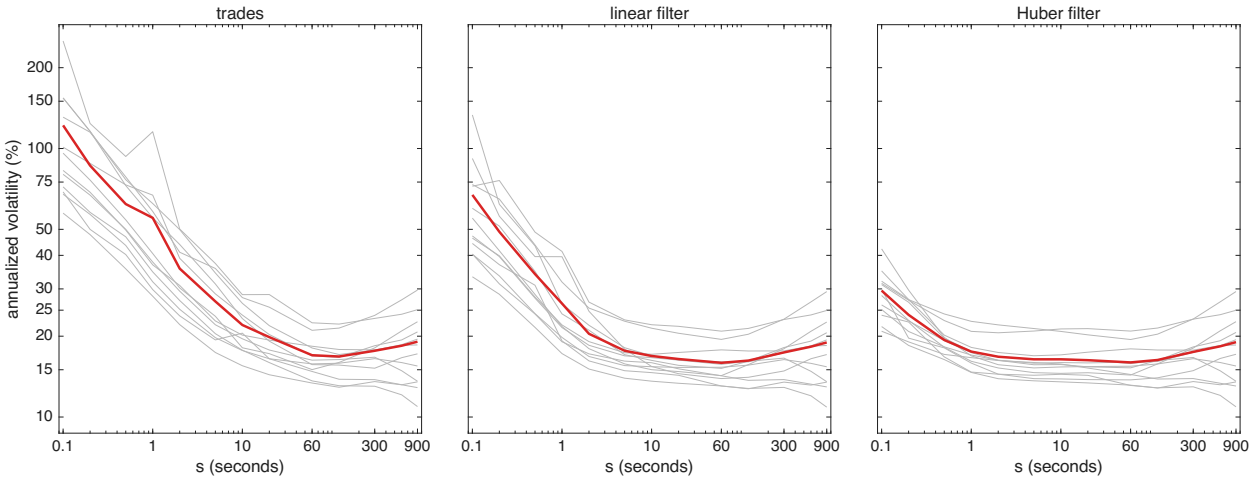

Figure 4: Volatility signature curves for `AAPL` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

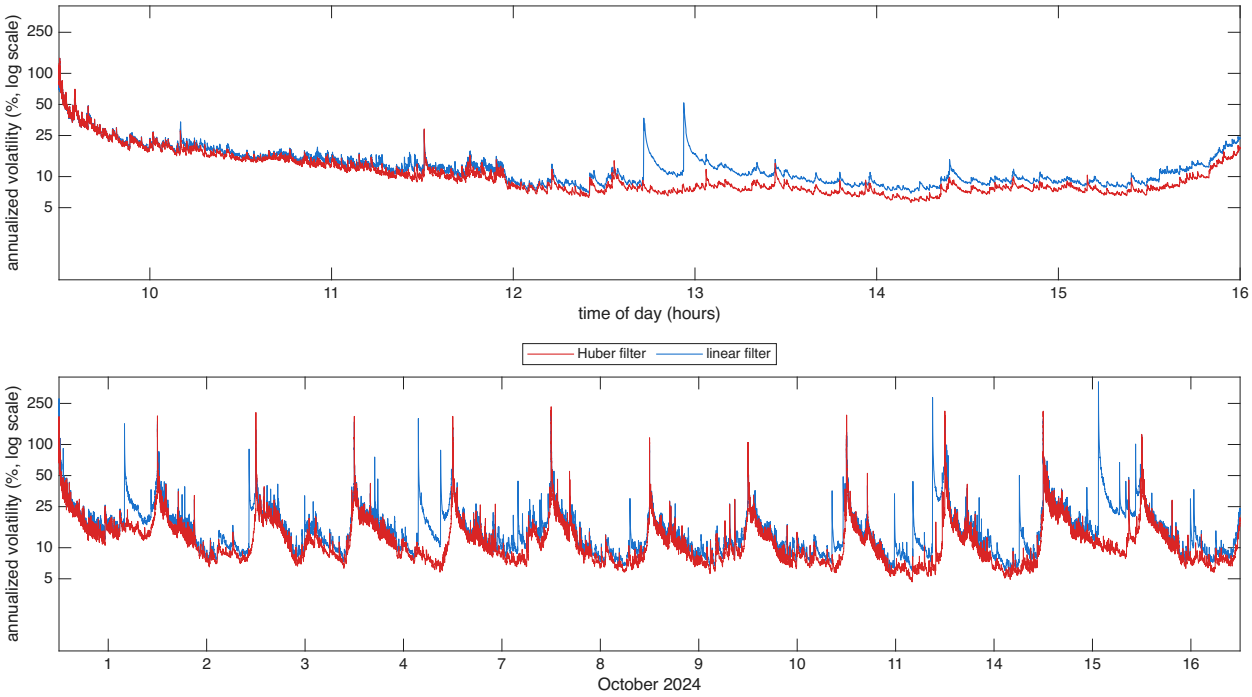

Figure 5: Filtered spot volatility for `AAPL` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 143$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

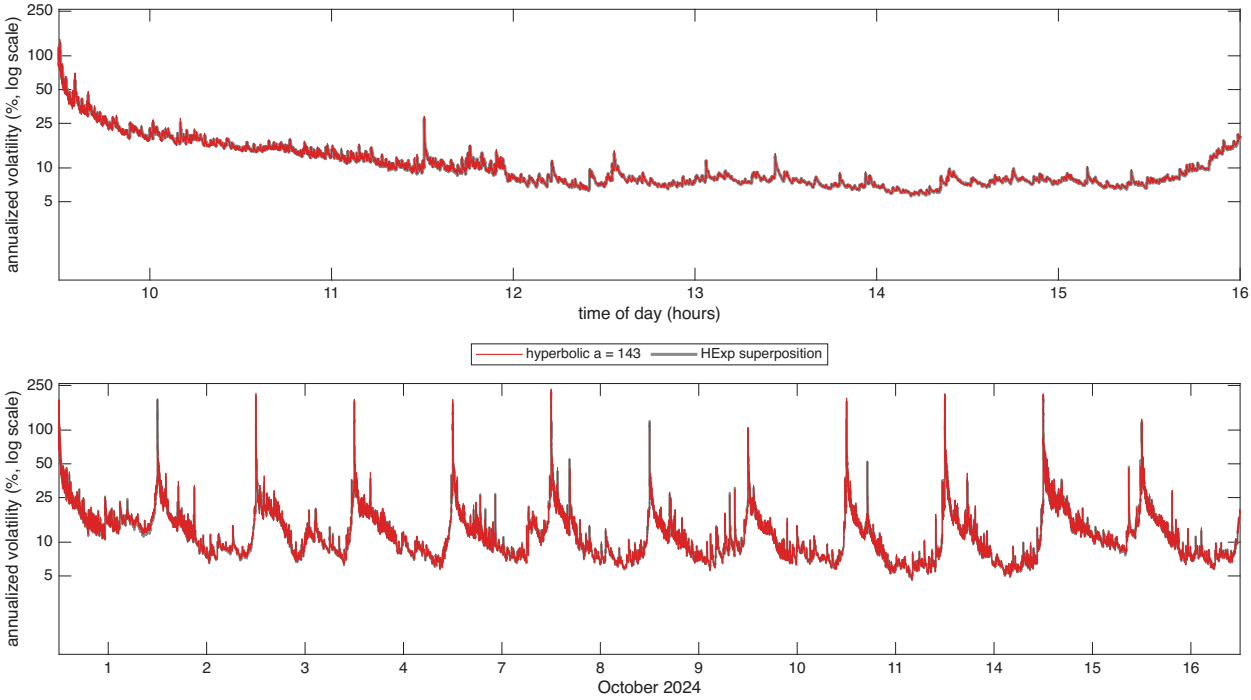

Figure 6: Hyperbolic against superposition weights for `AAPL` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 143$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

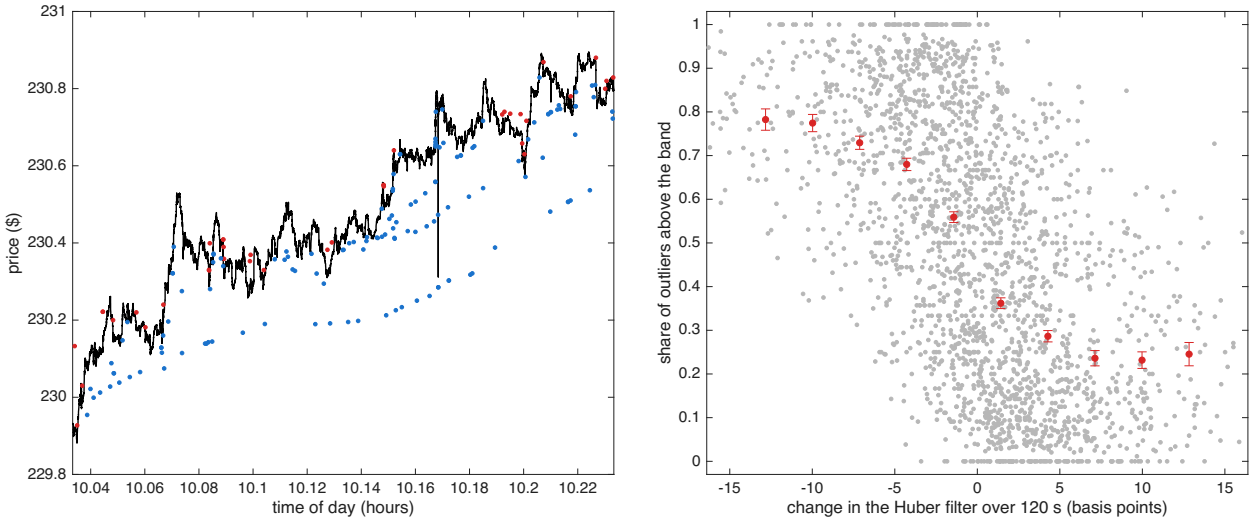

Figure 7: Outliers & robust price for `AAPL` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

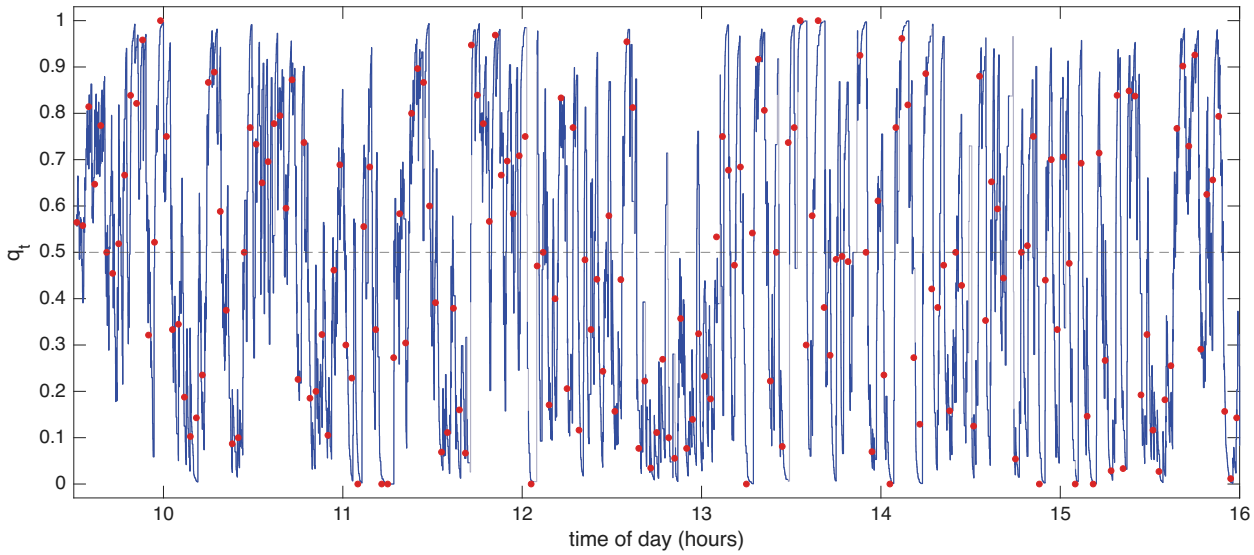

Figure 8: The directional conditional probability for `AAPL`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 20 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

Selecting a on single days gives values from 35 to 6528 trades, against 143 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 239 trades for this stock.

Diagnostics.

SPY 1 to 16 October 2024

trades on 16 October 2024	312,181
average trading rate	17.1 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,872 (0.92%)
price half-life: run / minimizer[1]	3.0 / 1.1 trades
cost of running 3 trades	+12.94%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.01%
direction half-life[3]	37 s [28, 53]
Mandelbrot a minimizing QLIKE	72 trades
a within 0.0005 of the minimum	49 to 107
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	6% $\to$ 875%
Huber $\hat{\sigma}$, before $\to$ peak	6% $\to$ 13%
median spot volatility, Huber / linear input[5]	7.5 / 9.9 %
hyperbolic against HExp, correlation of logs	0.9927
HExp QLIKE advantage, leave-one-day-out[6]	+0.00213
15-minute signature level	6 to 17 %
Huber signature, 1 min over 1 s, 12 day average	1.18

SPY over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:23; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.160	0.123	0.142	0.529	0.508	0.312	2.374	14.778
2	0.170	0.148	0.174	0.457	0.462	0.320	2.254	10.457
5	0.211	0.188	0.219	0.379	0.414	0.319	1.952	6.622
10	0.241	0.221	0.248	0.344	0.394	0.335	1.600	4.690
25	0.266	0.254	0.270	0.316	0.378	0.352	1.100	2.979
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.17	0.05	−0.10	−0.50	2.01	1.71	1.62	1.76
2	−0.02	−0.05	−0.08	0.00	1.89	1.59	1.61	1.81
5	0.01	0.04	−0.04	0.00	2.05	1.85	1.57	1.83
10	0.00	0.00	−0.01	0.00	2.50	2.45	1.68	1.87
25	0.00	0.00	0.00	0.00	3.59	3.94	2.09	1.97

Table 1: Trade-lag differences, `SPY`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.77	1.58	1.71	2.73	4.7	3.5	44.3	168.8
1 s	1.37	1.30	1.38	1.72	5.0	4.6	24.9	107.0
5 s	1.31	1.29	1.32	1.41	5.5	5.4	9.9	26.9
10 s	1.28	1.27	1.32	1.50	5.6	5.5	8.6	30.1
1 min	1.27	1.27	1.27	1.30	5.7	5.7	6.1	10.7
15 min	1.30	1.30	1.30	1.30	6.1	6.1	6.2	6.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.09	−0.05	−0.27	−0.08	4.11	4.00	3.08	2.02
1 s	0.01	0.06	−0.46	−0.66	6.81	7.04	2.73	1.93
5 s	0.02	0.04	−0.18	−0.01	9.23	9.70	4.16	2.58
10 s	0.05	0.06	−0.77	−0.95	8.23	8.10	3.43	2.23
1 min	0.05	0.05	−0.07	−0.71				

Table 2: Clock-lag differences, `SPY`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.40	0.51	0.55
s.d. of $\hat{\epsilon}_t$ in cents	10.46	10.45	8.73
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	1.62
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.31	0.42	−0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.14	0.29	−0.08
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.06	0.14	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.04	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.42,0.42	−0.43,0.44
95% C.I. for $\theta_t - X_t$	—	−0.75,0.72	−0.81,0.79
99% C.I. for $\theta_t - X_t$	—	−1.13,1.07	−1.36,1.39
s.e. of $\theta_t - X_t$ in cents	—	0.35	∞(2.24)

Table 3: Residual summaries, `SPY`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. does not exist in theory, the sample value is in brackets.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.62	1.39	1.13 to 1.69	1.26 to 1.63
Residual divided by δ_t	1.51	1.39	1.13 to 1.64	1.27 to 1.63
Repricing windows removed	1.61	1.39	1.11 to 1.68	1.24 to 1.69
Calm windows only	1.59	1.33		
Every 10th trade	1.64	1.55	1.16 to 1.67	1.31 to 1.86
Six-trade half-life	1.77	1.53	1.22 to 1.92	1.36 to 1.88

Table 4: Hill estimator of the tail index of the Huber filter residual, `SPY`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 17.9 to 20.2 cents over the days and removes 0% to 5% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 9.0 cents on 16 October 2024, 75% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

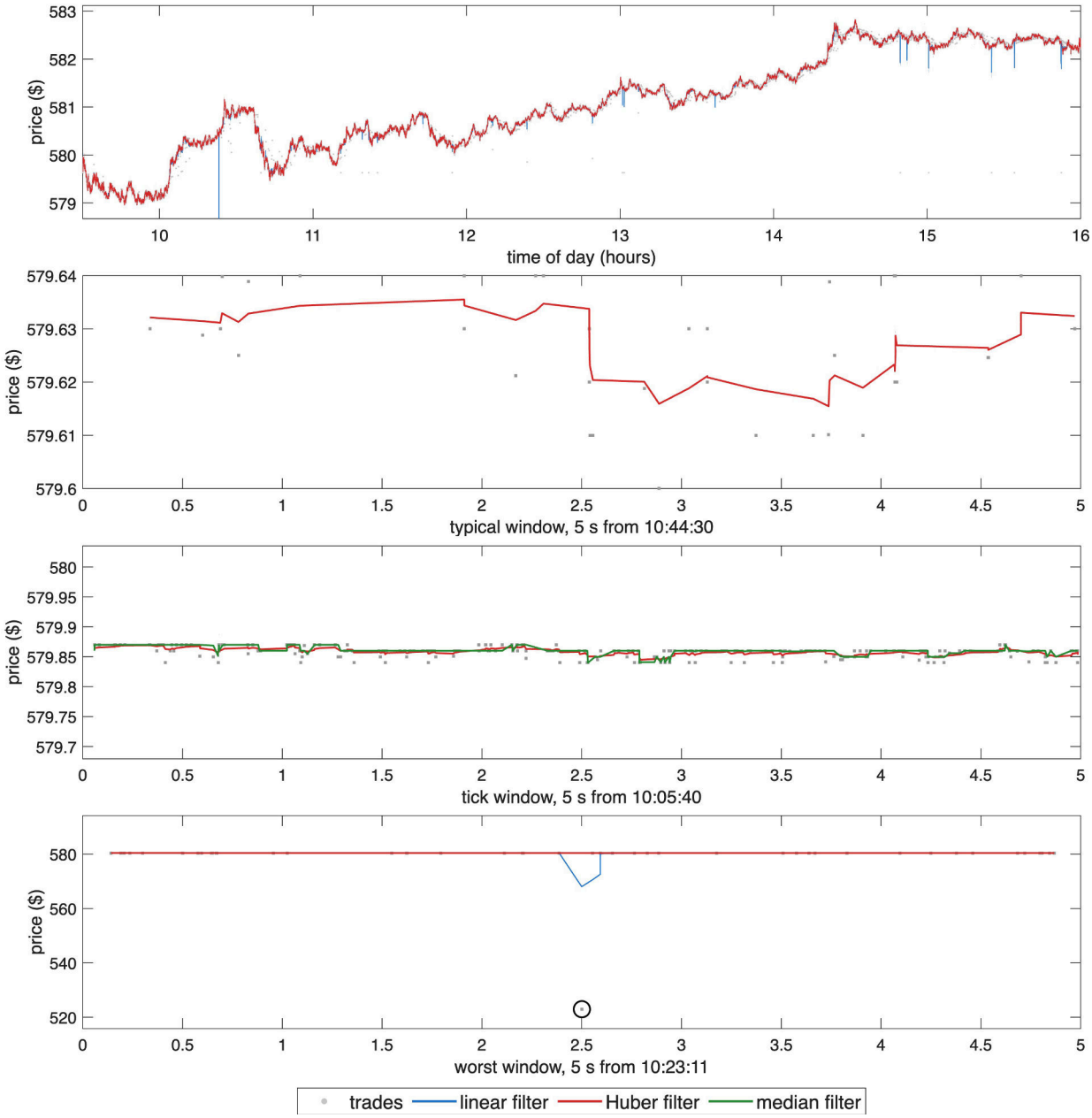

Figure 1: SPY trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:44:30. 3rd: 5-second window at 10:05:40 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

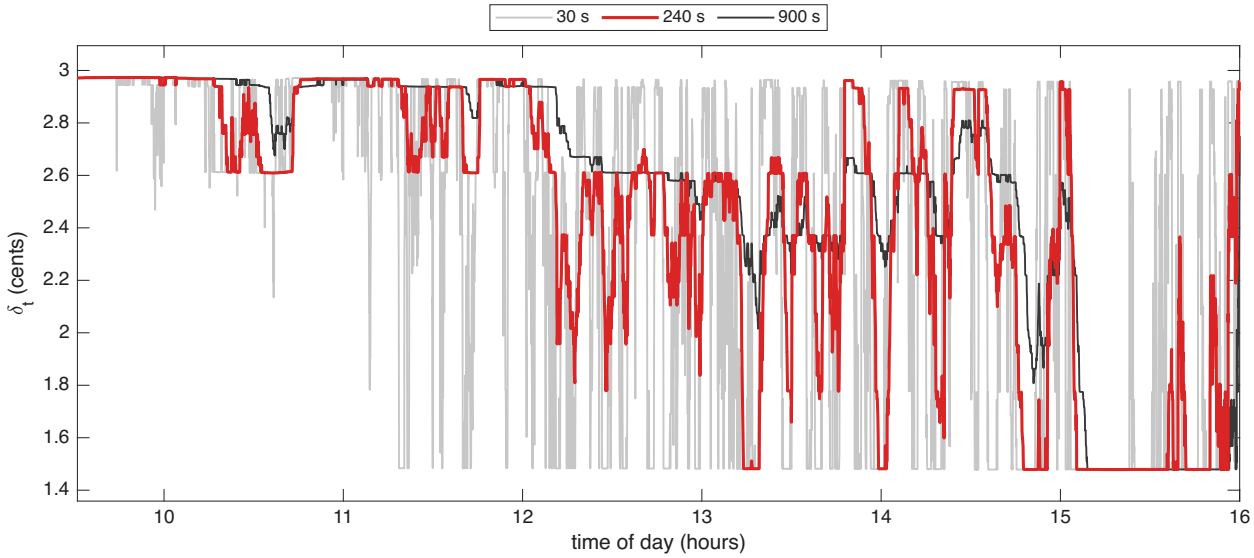

Figure 2: The Huber threshold δ_t in cents for `SPY` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

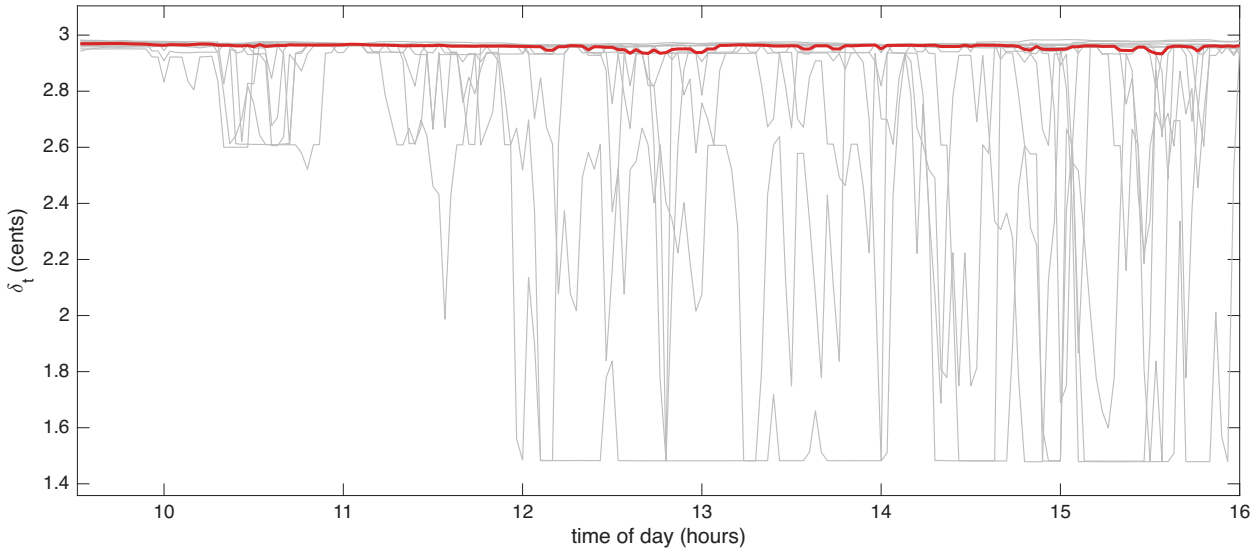

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `SPY` trading days in Oct 2024 (grey lines) and their median (red line).

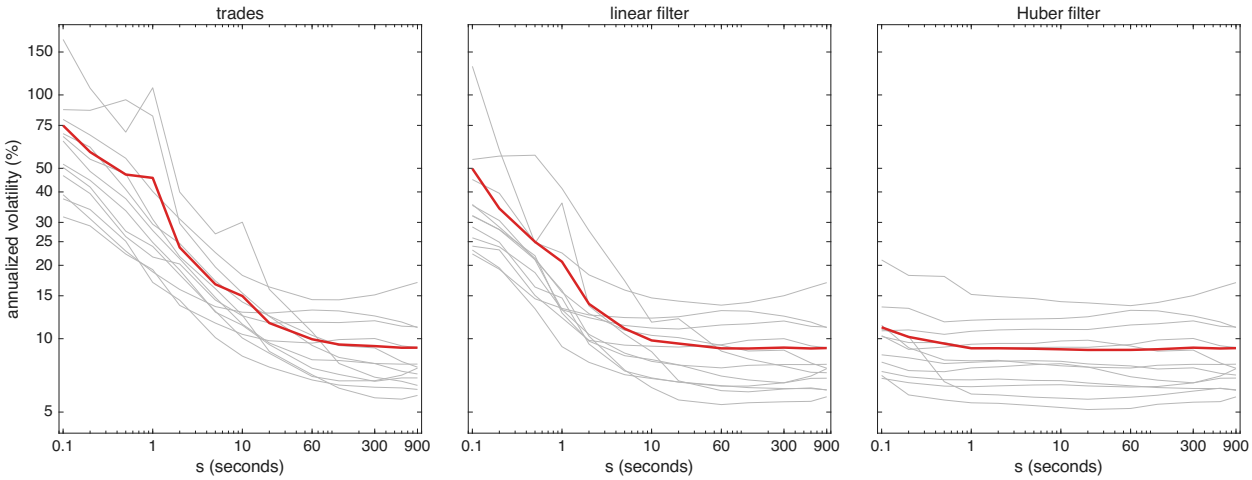

Figure 4: Volatility signature curves for `SPY` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

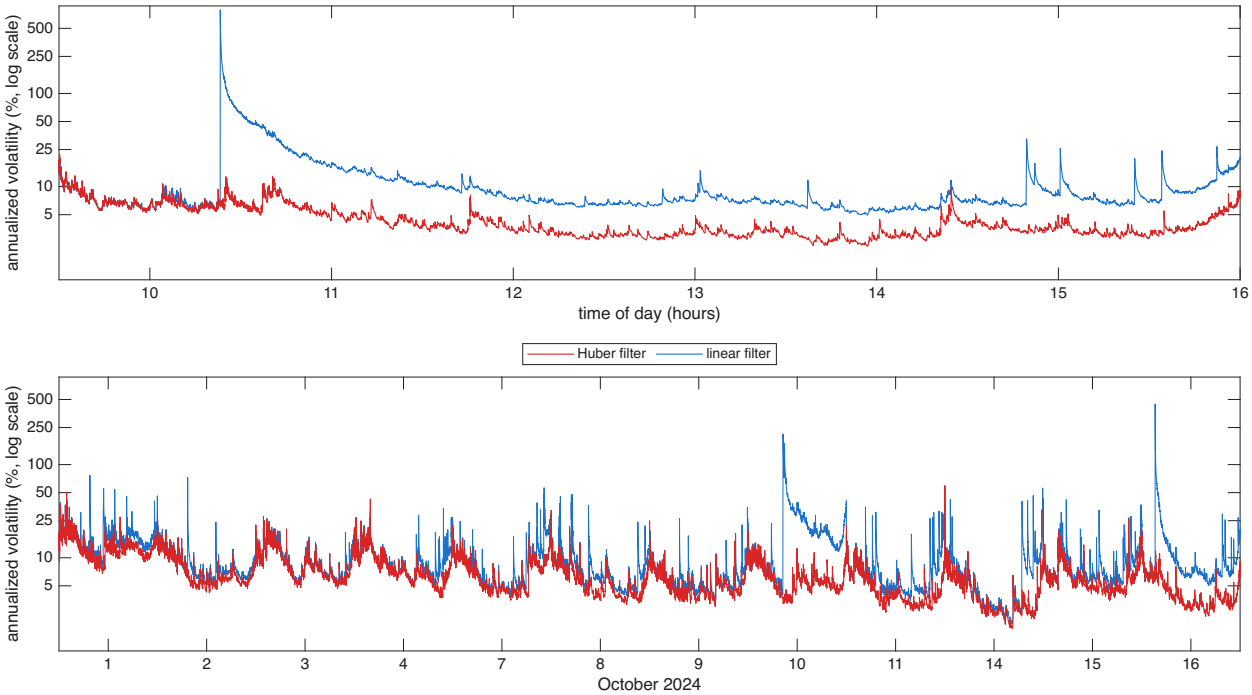

Figure 5: Filtered spot volatility for `SPY` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 72$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

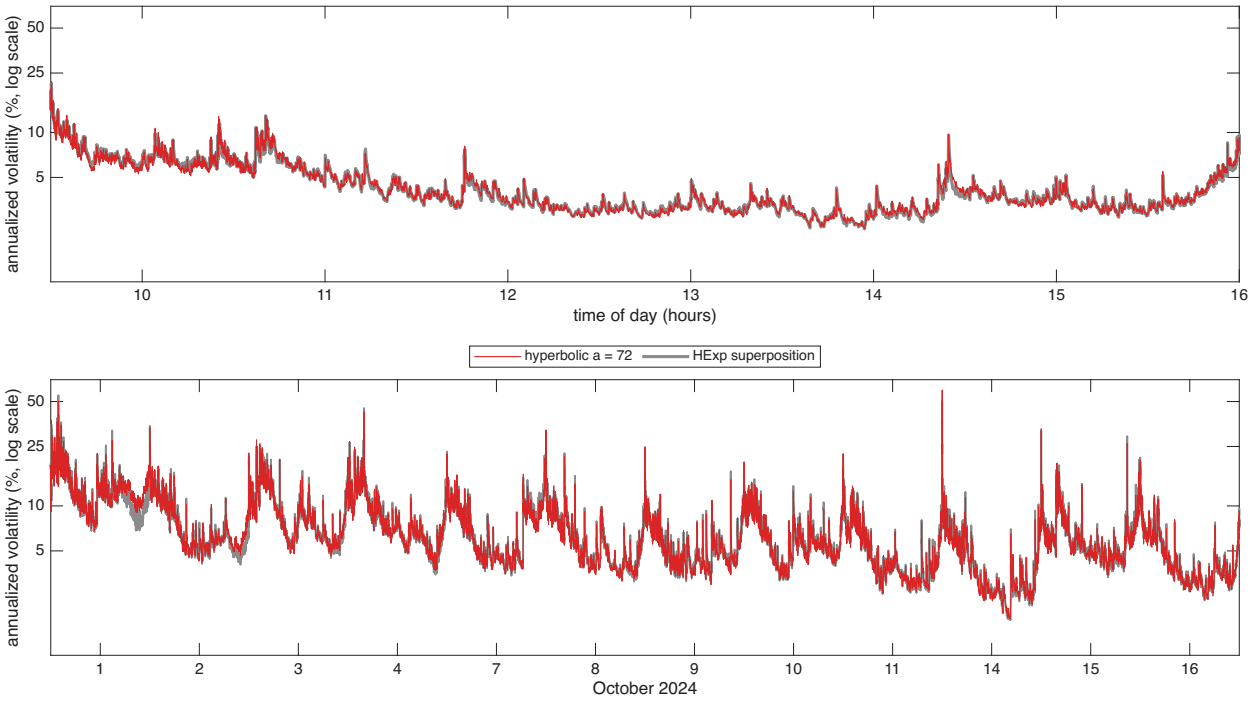

Figure 6: Hyperbolic against superposition weights for `SPY` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 72$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

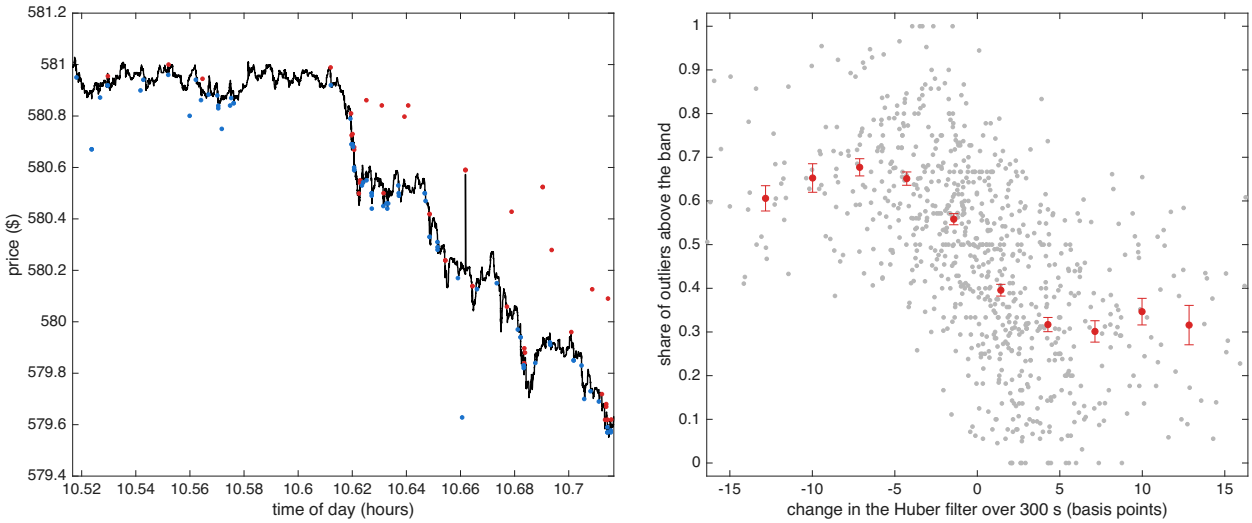

Figure 7: Outliers & robust price for `SPY` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

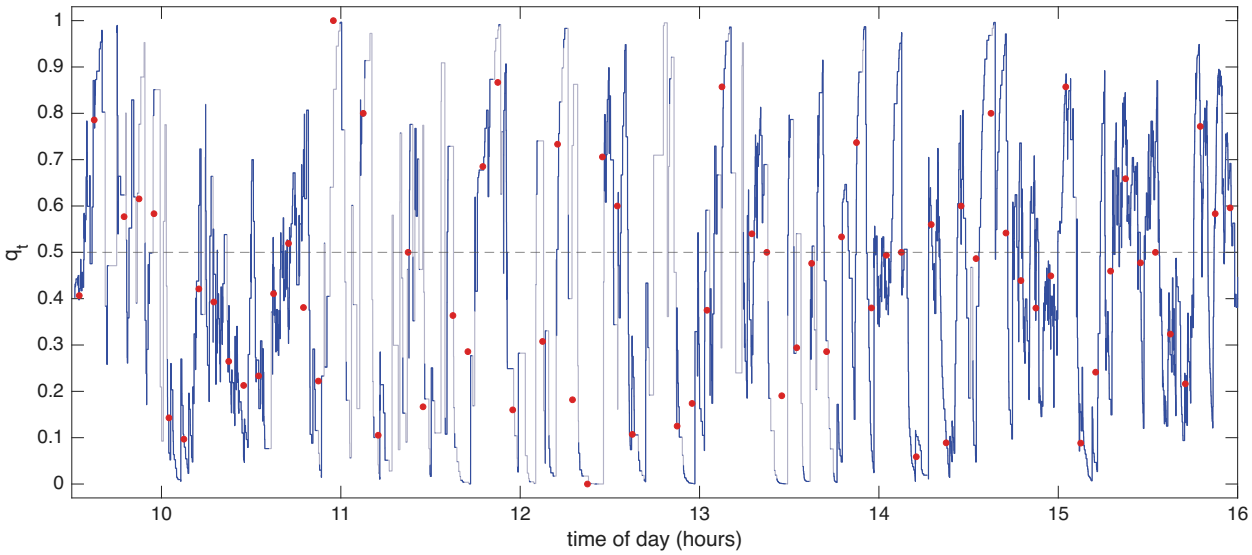

Figure 8: The directional conditional probability for SPY, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 37 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The common half-life of 3 trades raises the k-trade-ahead Huber forecast loss by 12.9% over the minimizing half-life of 1.1 trades.
At the horizons k = 2, 3, 5, 10, 30 and 100 the half-life minimizing the k-trade-ahead Huber forecast loss is 1.1, 1.1, 1.1, 1.1, 1.1 and 1.1 trades, within one grid point of the end of the 1 to 5 trade grid searched.
Selecting a on single days gives values from 13 to 562 trades, against 72 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 171 trades for this stock.
1 return(s) share a timestamp with the opening auction and are left out of the volatility sums.

Diagnostics.

AMD 1 to 16 October 2024

trades on 16 October 2024	310,663
average trading rate	17.0 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	4,981 (1.60%)
price half-life: run / minimizer[1]	3.0 / 1.6 trades
cost of running 3 trades	+4.95%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	29 s [22, 39]
Mandelbrot a minimizing QLIKE	106 trades
a within 0.0005 of the minimum	68 to 164
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	20% $\to$ 81%
Huber $\hat{\sigma}$, before $\to$ peak	16% $\to$ 16%
median spot volatility, Huber / linear input[5]	27.2 / 31.9 %
hyperbolic against HExp, correlation of logs	0.9934
HExp QLIKE advantage, leave-one-day-out[6]	+0.00443
15-minute signature level	25 to 54 %
Huber signature, 1 min over 1 s, 12 day average	1.13

AMD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.245	0.189	0.216	0.906	0.818	0.609	0.503	2.978
2	0.245	0.218	0.256	0.752	0.672	0.507	0.525	2.210
5	0.293	0.262	0.306	0.587	0.628	0.523	0.542	1.482
10	0.326	0.299	0.334	0.504	0.585	0.518	0.541	1.113
25	0.363	0.346	0.364	0.448	0.558	0.525	0.533	0.812
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.33	−0.31	0.09	−0.45	2.09	1.85	2.09	2.25
2	0.16	0.27	0.02	−0.02	2.07	1.80	2.37	2.29
5	−0.03	−0.04	−0.01	−0.01	1.85	1.68	2.34	2.31
10	−0.01	−0.01	0.00	0.00	1.77	1.75	2.34	2.34
25	0.00	−0.01	0.00	0.00	2.24	2.26	2.84	2.46

Table 1: Trade-lag differences, AMD, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.02	2.79	2.98	4.97	26.7	22.1	26.7	122.4
1 s	2.23	2.15	2.28	2.94	28.3	27.0	30.2	65.8
5 s	2.24	2.22	2.25	2.45	30.9	30.6	31.3	41.5
10 s	2.21	2.20	2.21	2.36	31.1	30.9	31.1	36.5
1 min	2.10	2.09	2.09	2.13	29.6	29.5	29.3	30.3
15 min	2.32	2.32	2.32	2.33	32.7	32.7	32.6	32.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.08	−0.01	0.00	−0.32	3.20	3.39	3.80	2.94
1 s	0.02	0.05	−0.02	−0.44	3.58	3.62	4.43	2.33
5 s	0.04	0.05	0.05	−0.08	5.51	5.89	7.47	4.07
10 s	0.08	0.09	0.07	−0.05	9.72	10.02	10.87	3.69
1 min	0.01	0.01	0.01	0.00				

Table 2: Clock-lag differences, AMD, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.63	0.78	0.83
s.d. of $\hat{\epsilon}_t$ in cents	2.21	2.16	1.94
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.08
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.30	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.20	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.10	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.05	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.57,0.55	−0.62,0.60
95% C.I. for $\theta_t - X_t$	—	−0.99,0.97	−1.16,1.15
99% C.I. for $\theta_t - X_t$	—	−1.51,1.56	−2.17,2.12
s.e. of $\theta_t - X_t$ in cents	—	0.47	0.70

Table 3: Residual summaries, `AMD`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.08, is above two, while the Hill index of the Huber residual itself on this day is 1.48, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.48	1.56	1.29 to 1.82	1.37 to 1.87
Residual divided by δ_t	1.51	1.53	1.29 to 1.92	1.34 to 1.86
Repricing windows removed	1.52	1.49	1.29 to 1.91	1.35 to 1.85
Calm windows only	1.67	1.54		
Every 10th trade	1.51	1.56	1.24 to 1.78	1.34 to 1.83
Six-trade half-life	1.54	1.68	1.31 to 1.94	1.45 to 1.98

Table 4: Hill estimator of the tail index of the Huber filter residual, `AMD`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 20.0 to 20.2 cents over the days and removes 1% to 5% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.0 cents on 16 October 2024, 65% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

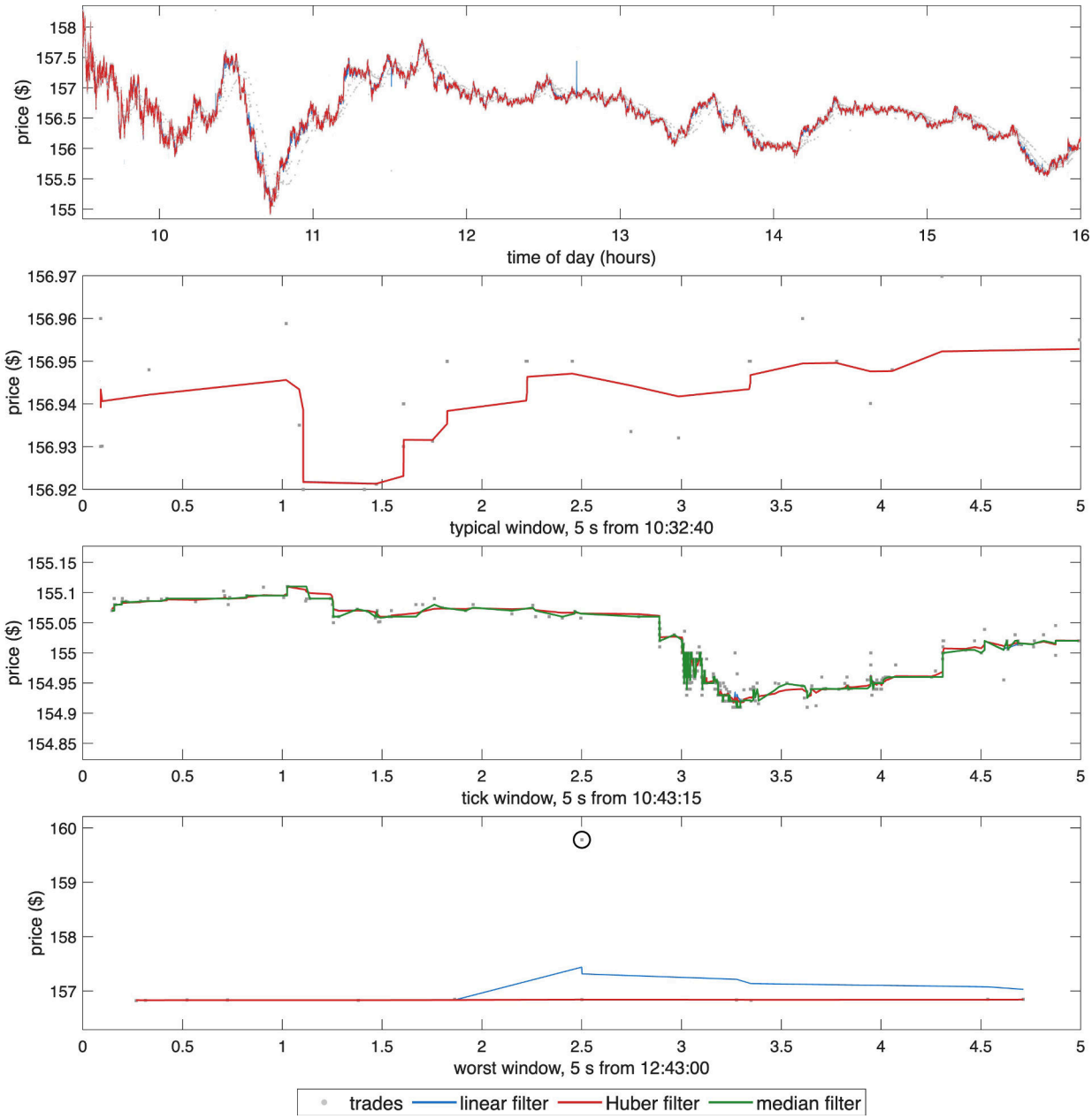

Figure 1: AMD trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:32:40. 3rd: 5-second window at 10:43:15 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

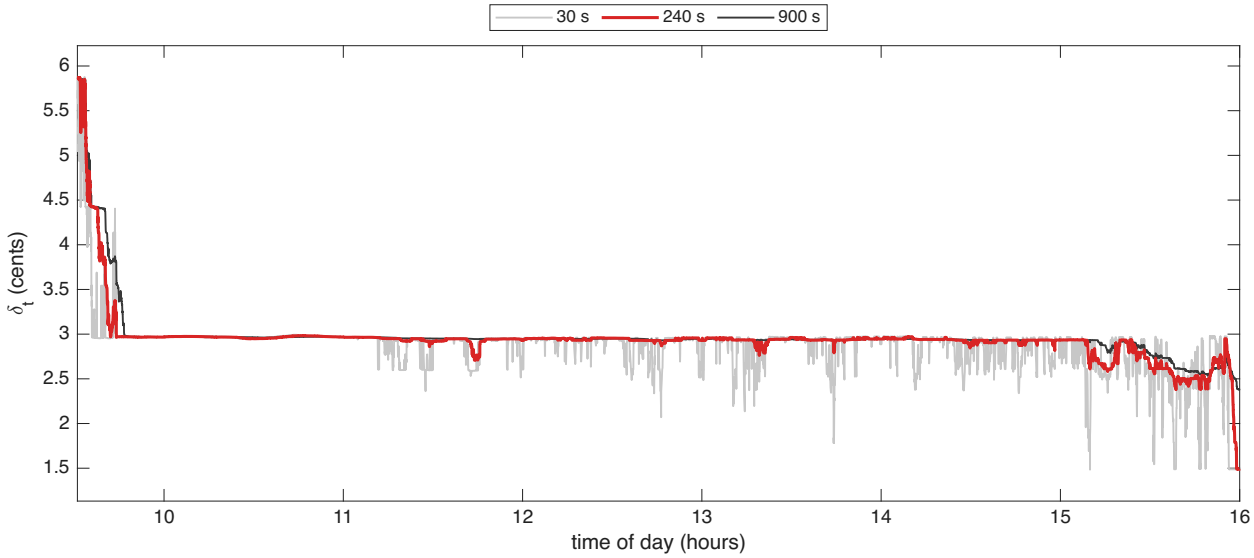

Figure 2: The Huber threshold δ_t in cents for `AMD` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

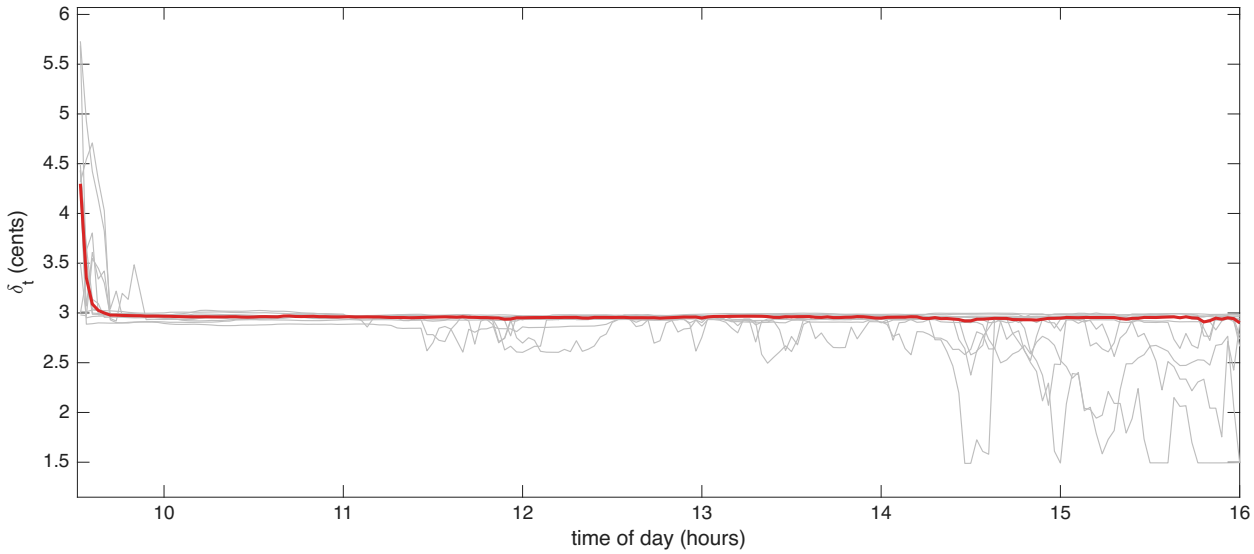

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AMD` trading days in Oct 2024 (grey lines) and their median (red line).

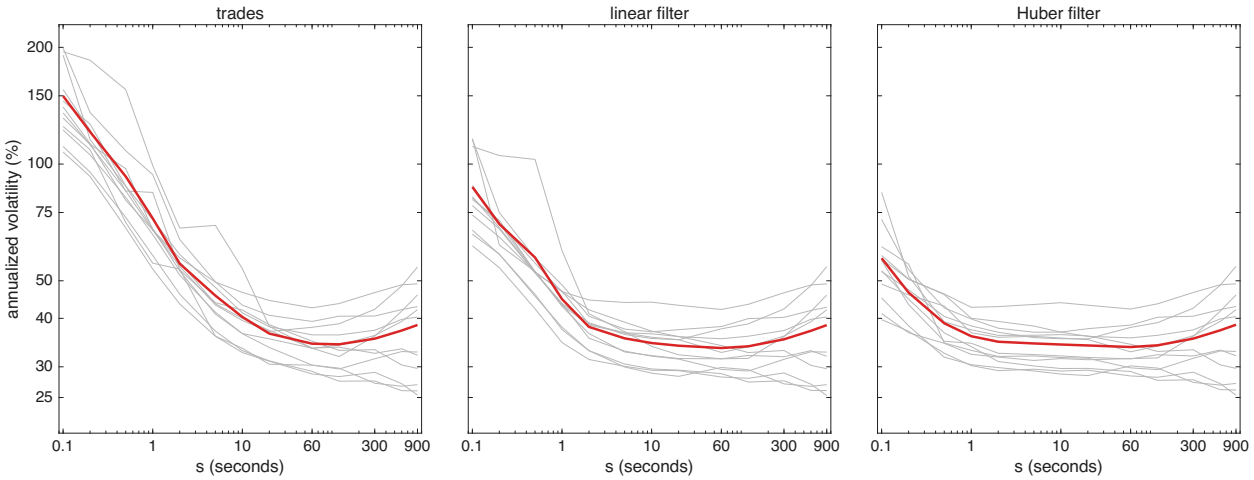

Figure 4: Volatility signature curves for AMD over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

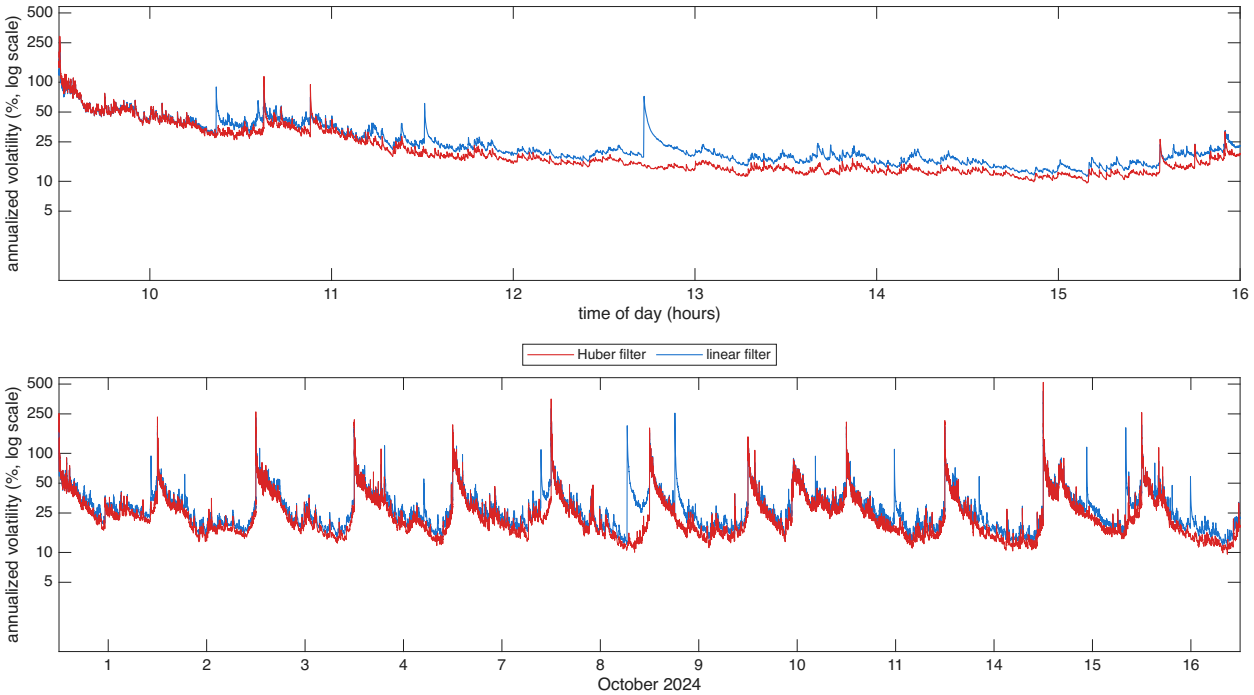

Figure 5: Filtered spot volatility for AMD over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 106$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

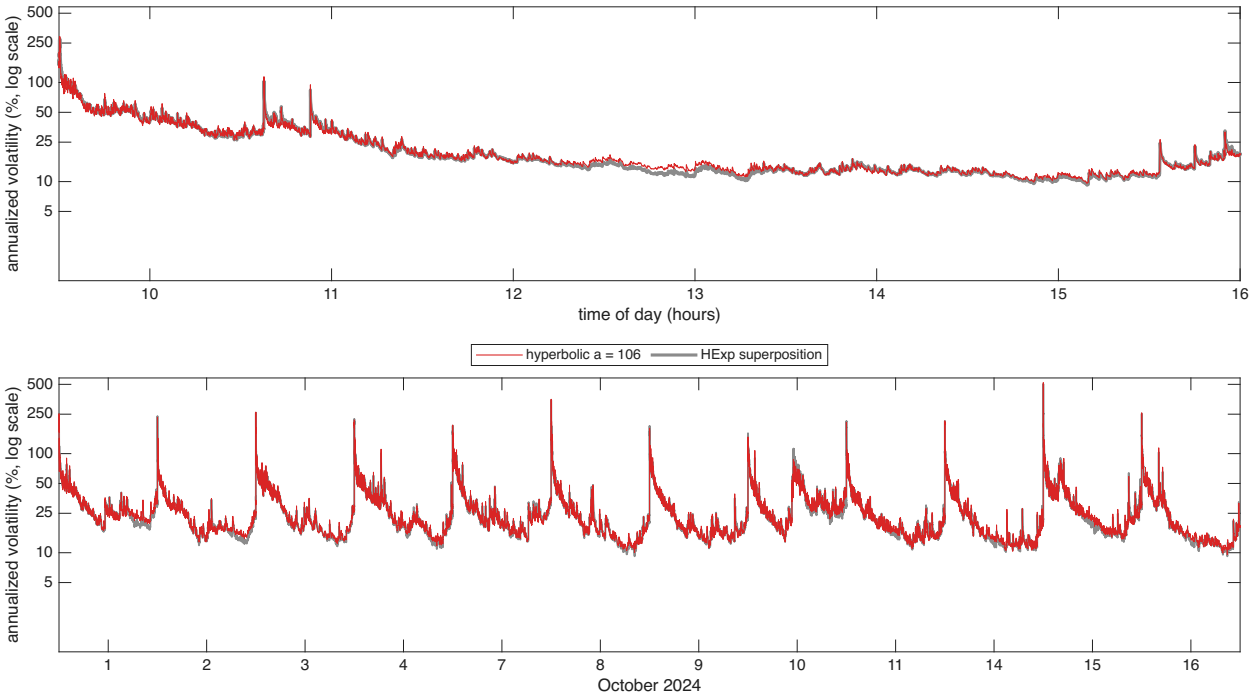

Figure 6: Hyperbolic against superposition weights for `AMD` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 106$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

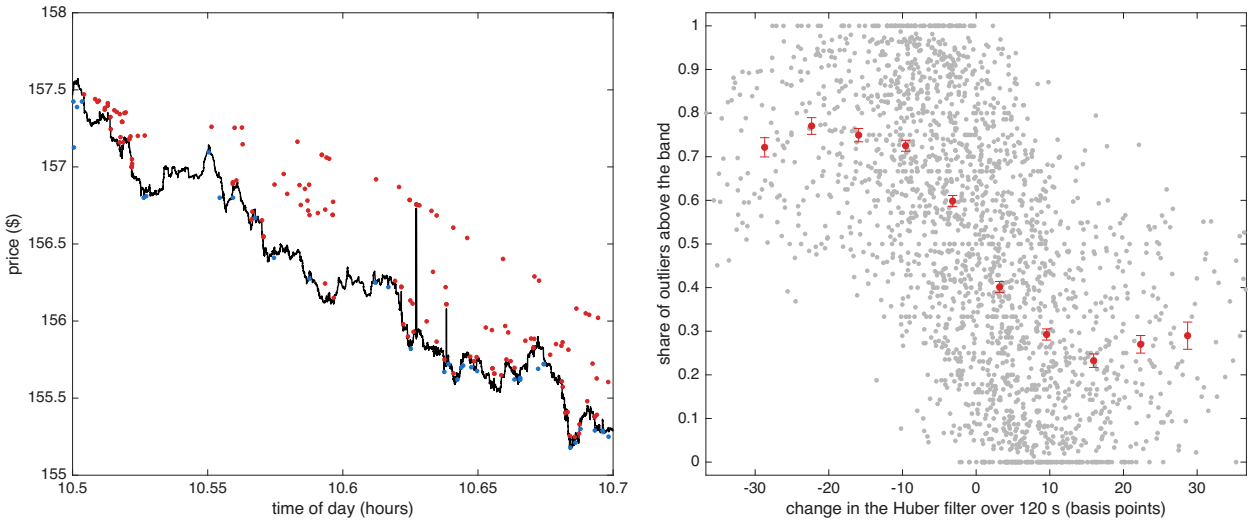

Figure 7: Outliers & robust price for `AMD` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

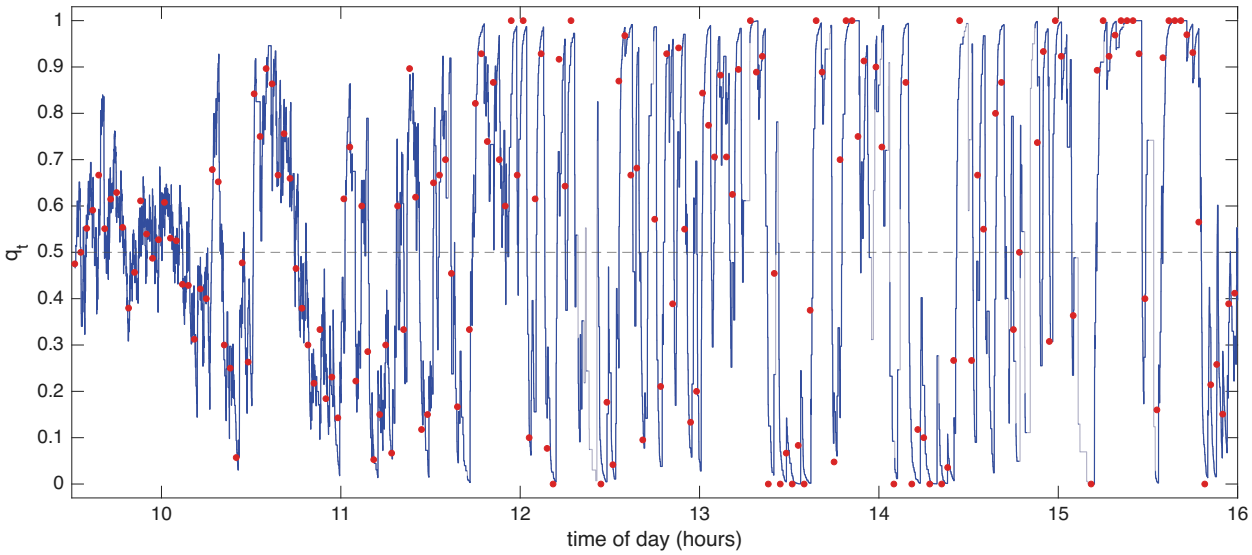

Figure 8: The directional conditional probability for AMD, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 29 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

AMZN 1 to 16 October 2024

trades on 16 October 2024	316,474
average trading rate	15.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	6,318 (2.00%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.79%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.02%
direction half-life[3]	23 s [18, 30]
Mandelbrot a minimizing QLIKE	135 trades
a within 0.0005 of the minimum	87 to 220
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	14% $\to$ 37%
Huber $\hat{\sigma}$, before $\to$ peak	13% $\to$ 16%
median spot volatility, Huber / linear input[5]	15.1 / 18.9 %
hyperbolic against HExp, correlation of logs	0.9965
HExp QLIKE advantage, leave-one-day-out[6]	-0.00139
15-minute signature level	15 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.11

AMZN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:31; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.157	0.126	0.150	0.632	0.618	0.443	0.373	2.204
2	0.161	0.146	0.179	0.523	0.566	0.440	0.389	1.627
5	0.194	0.174	0.213	0.410	0.501	0.417	0.403	1.100
10	0.213	0.195	0.226	0.346	0.461	0.406	0.397	0.826
25	0.230	0.219	0.234	0.297	0.402	0.375	0.371	0.588
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.16	−0.01	0.09	−0.45	1.53	1.38	2.32	2.58
2	−0.04	−0.06	0.03	−0.01	1.50	1.28	2.57	2.65
5	−0.01	0.00	−0.01	0.00	1.45	1.32	2.37	2.79
10	−0.03	−0.04	−0.02	0.00	1.54	1.46	2.26	2.94
25	0.01	0.01	−0.01	0.00	2.47	2.25	2.41	3.27

Table 1: Trade-lag differences, `AMZN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.73	1.56	1.68	3.08	16.4	13.6	15.9	74.9
1 s	1.18	1.13	1.21	1.63	15.2	14.2	16.3	35.9
5 s	1.12	1.10	1.12	1.28	15.6	15.3	15.7	21.5
10 s	1.12	1.11	1.12	1.22	15.7	15.6	15.7	18.9
1 min	1.09	1.08	1.08	1.10	15.7	15.6	15.4	16.0
15 min	1.15	1.15	1.15	1.15	16.7	16.7	16.7	16.7
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.34	−0.34	−0.28	−0.40	3.20	2.90	2.99	3.24
1 s	−0.24	−0.22	−0.18	−0.43	3.01	3.10	4.39	3.68
5 s	0.06	0.08	0.06	−0.14	6.15	6.23	7.26	4.91
10 s	−0.02	−0.02	−0.02	−0.13	6.16	6.20	6.30	5.08
1 min	0.07	0.07	0.08	0.05				

Table 2: Clock-lag differences, `AMZN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.44	0.53	0.58
s.d. of $\hat{\epsilon}_t$ in cents	1.66	1.60	1.43
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.32
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.30	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.19	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.08	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.03	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.36,0.34	−0.41,0.39
95% C.I. for $\theta_t - X_t$	—	−0.67,0.66	−0.83,0.80
99% C.I. for $\theta_t - X_t$	—	−1.08,1.09	−1.65,1.68
s.e. of $\theta_t - X_t$ in cents	—	0.32	0.51

Table 3: Residual summaries, `AMZN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.32, is above two, while the Hill index of the Huber residual itself on this day is 1.55, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.55	1.76	1.30 to 2.21	1.39 to 1.94
Residual divided by δ_t	1.96	2.12	1.58 to 2.23	1.63 to 2.12
Repricing windows removed	1.56	1.76	1.44 to 1.91	1.52 to 1.89
Calm windows only	1.65	2.03		
Every 10th trade	1.46	2.04	1.33 to 1.98	1.27 to 2.04
Six-trade half-life	1.62	1.80	1.41 to 2.46	1.42 to 2.03

Table 4: Hill estimator of the tail index of the Huber filter residual, `AMZN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 11.5 to 20.1 cents over the days and removes 2% to 6% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 6.0 cents on 16 October 2024, 65% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

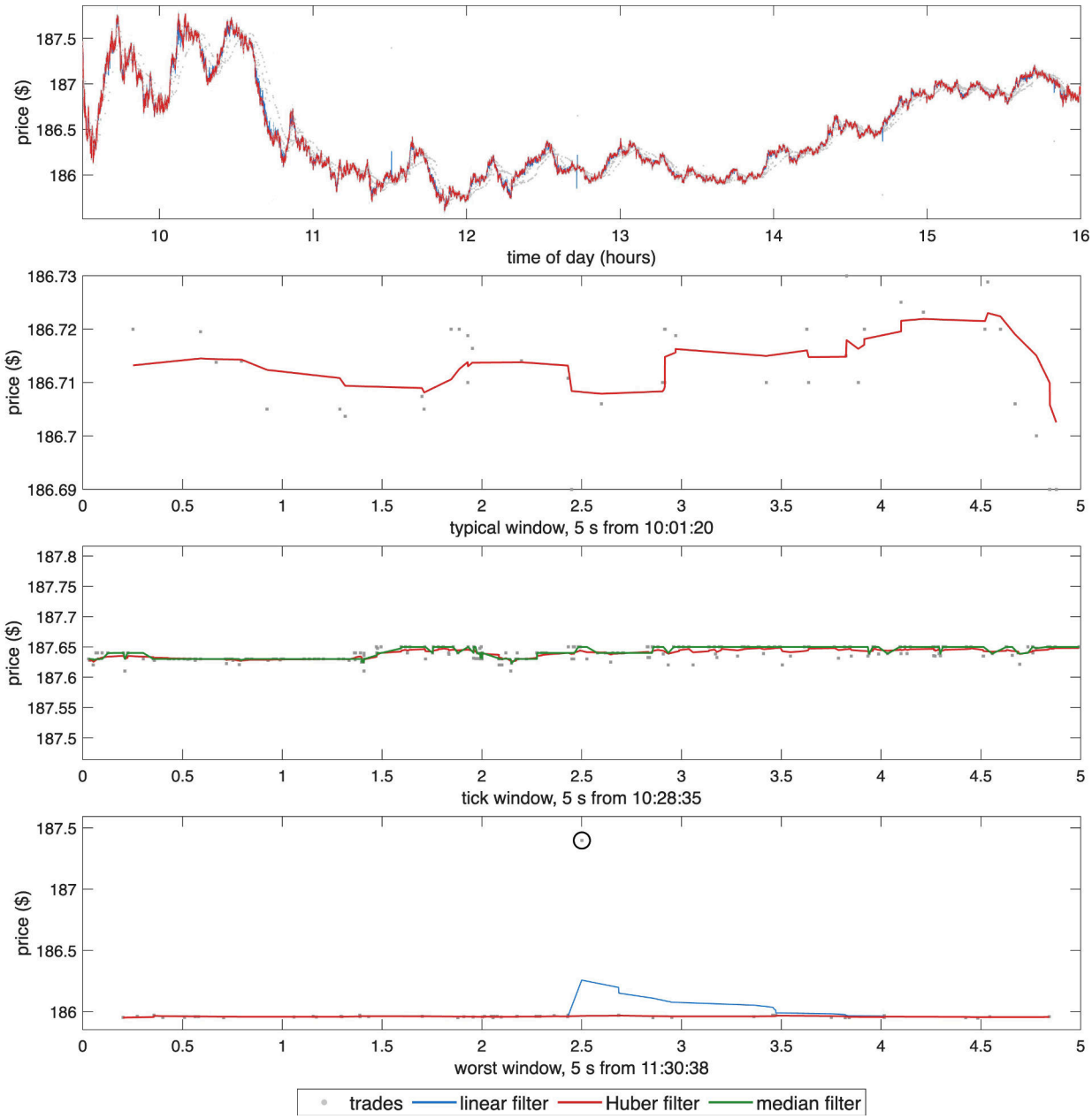

Figure 1: `AMZN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:01:20. 3rd: 5-second window at 10:28:35 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

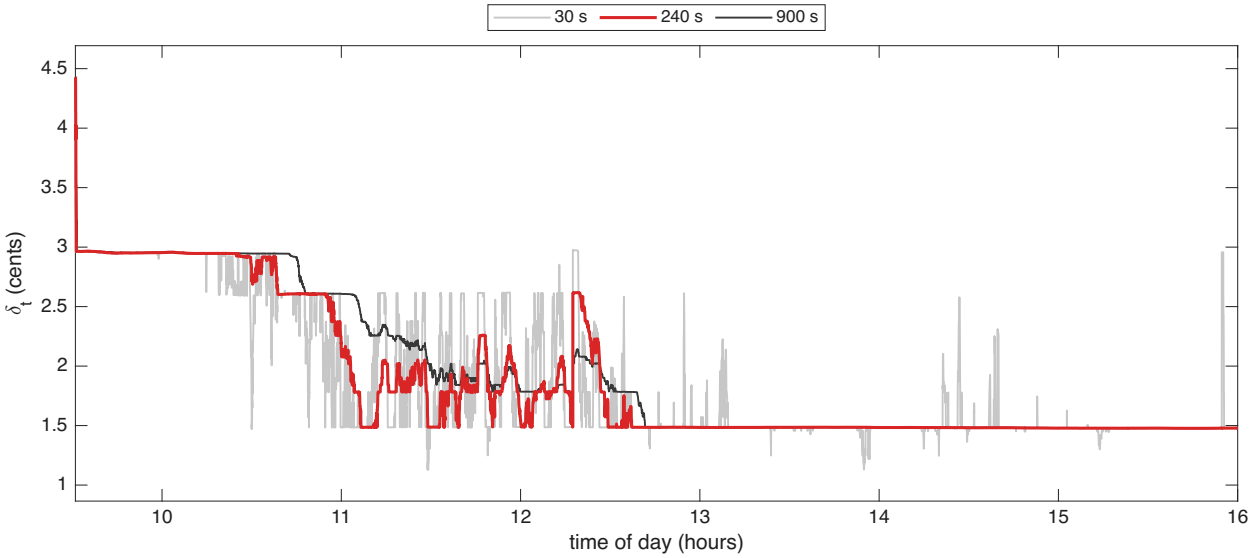

Figure 2: The Huber threshold δ_t in cents for `AMZN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

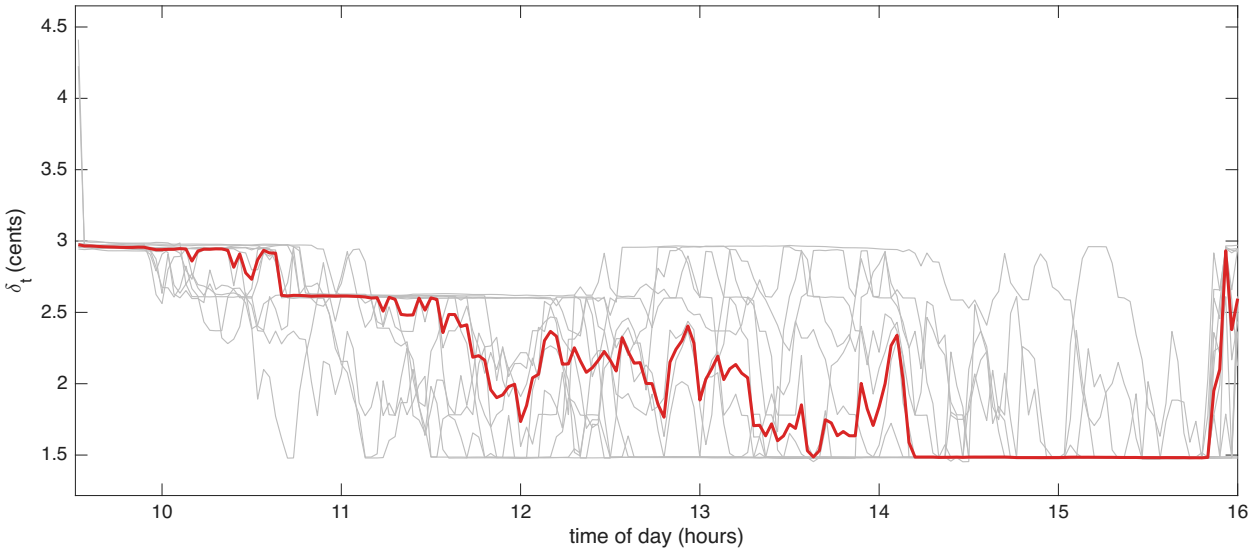

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AMZN` trading days in Oct 2024 (grey lines) and their median (red line).

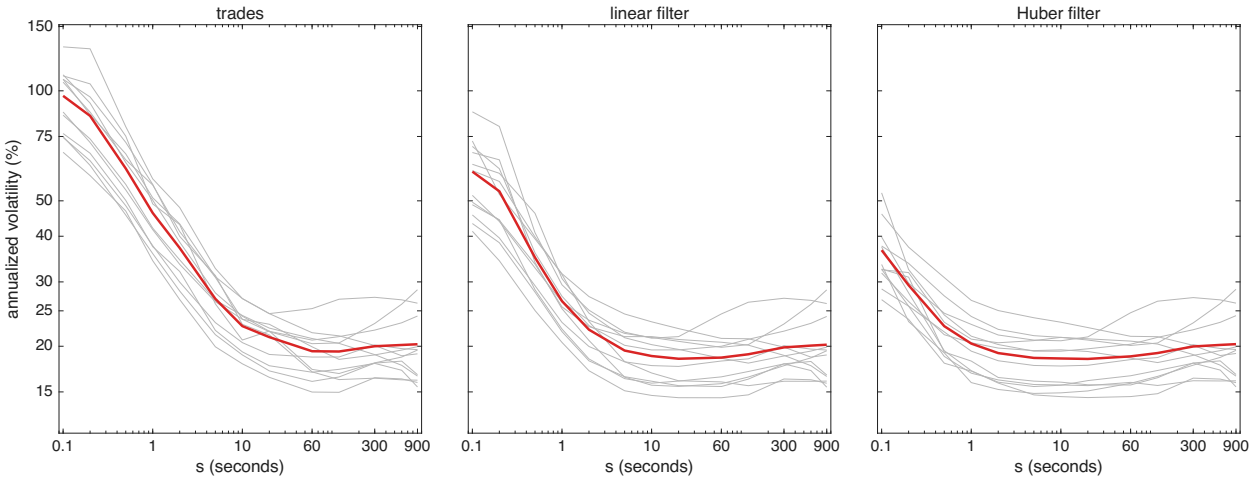

Figure 4: Volatility signature curves for AMZN over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

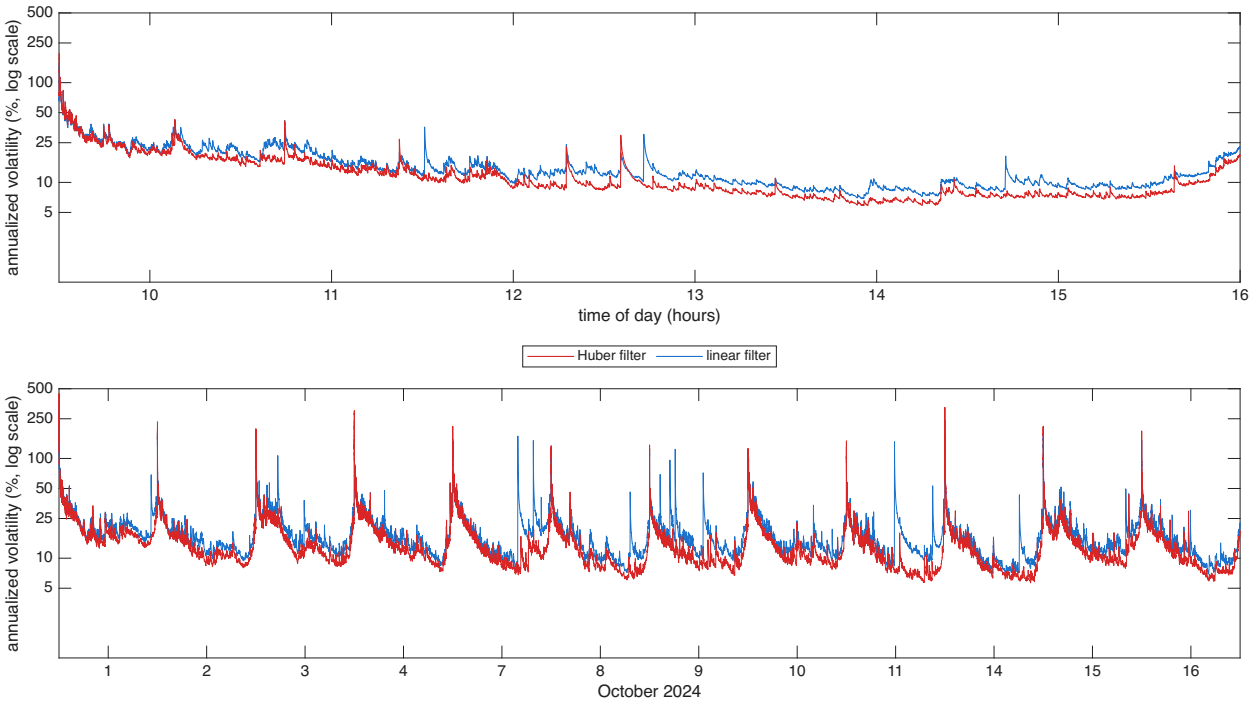

Figure 5: Filtered spot volatility for AMZN over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 135$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

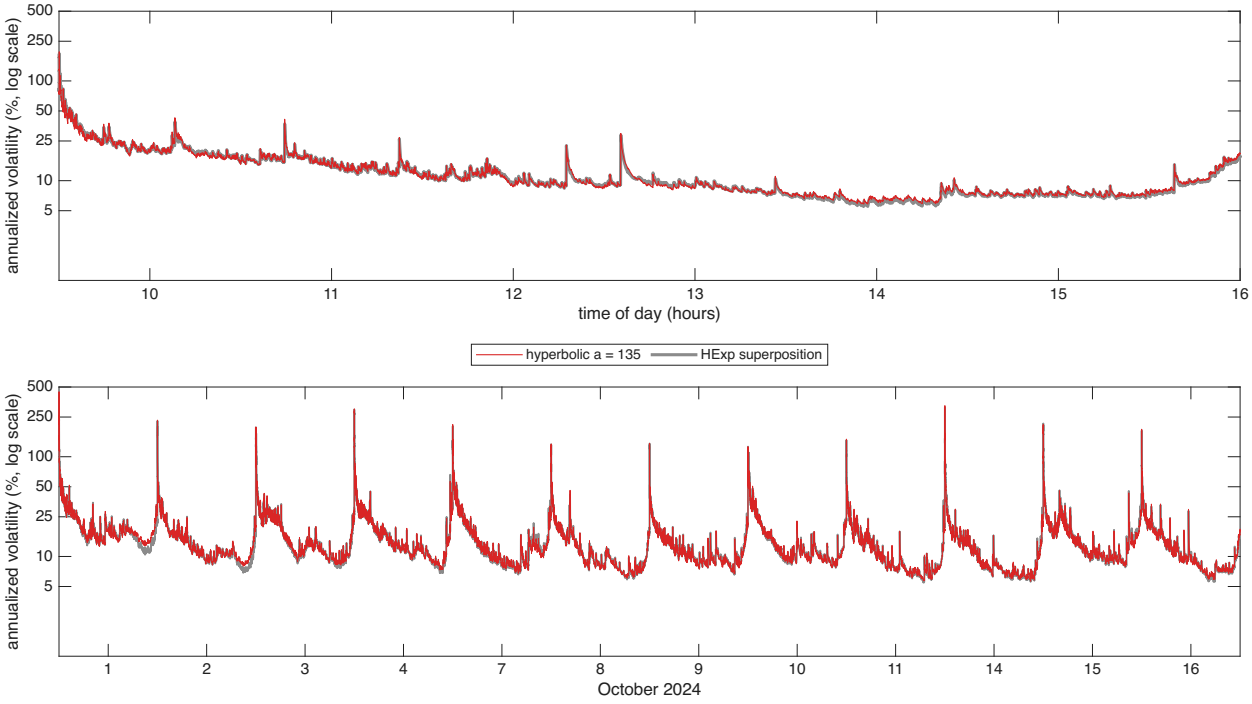

Figure 6: Hyperbolic against superposition weights for `AMZN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 135$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 12, 50, 200, 800 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

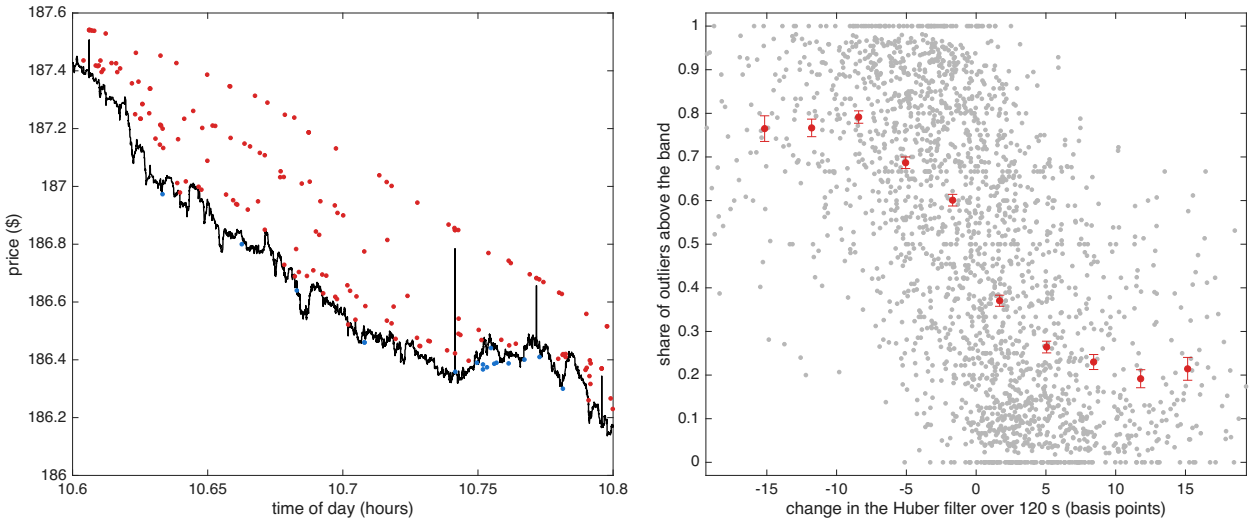

Figure 7: Outliers & robust price for `AMZN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

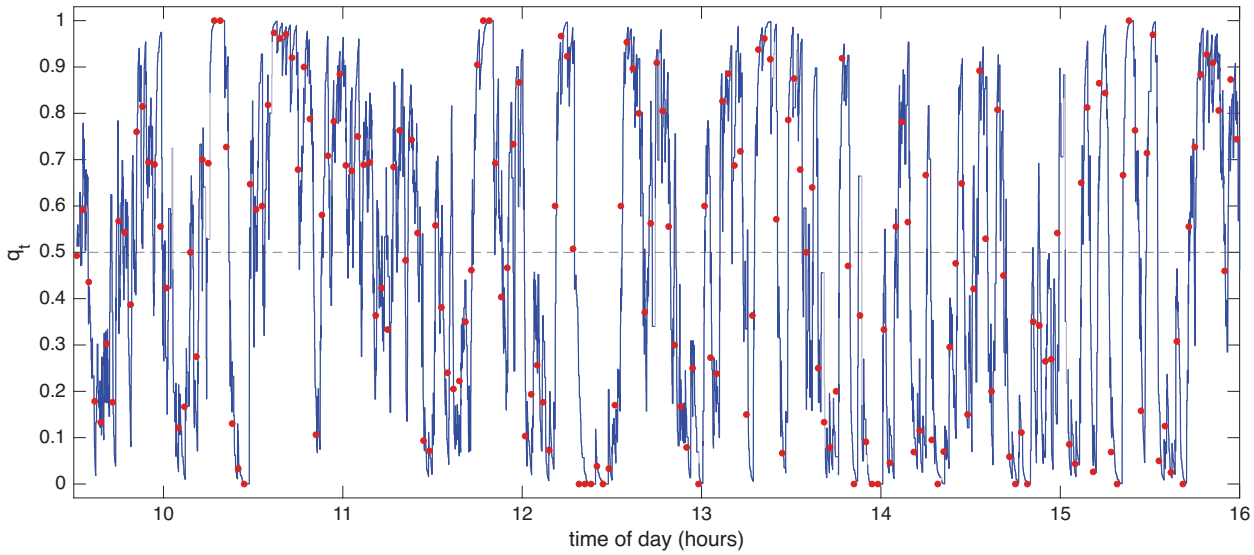

Figure 8: The directional conditional probability for `AMZN`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 23 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

Selecting a on single days gives values from 31 to 1632 trades, against 135 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 156 trades for this stock.

Diagnostics.

MSFT 1 to 16 October 2024

trades on 16 October 2024	348,397
average trading rate	14.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	7,495 (2.15%)
price half-life: run / minimizer[1]	3.0 / 2.4 trades
cost of running 3 trades	+1.96%
scale half-life: run / minimizer[2]	240 / 90 s
cost of running 240 s	+0.09%
direction half-life[3]	22 s [17, 32]
Mandelbrot a minimizing QLIKE	96 trades
a within 0.0005 of the minimum	65 to 149
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	10% $\to$ 160%
Huber $\hat{\sigma}$, before $\to$ peak	9% $\to$ 27%
median spot volatility, Huber / linear input[5]	13.3 / 16.4 %
hyperbolic against HExp, correlation of logs	0.9977
HExp QLIKE advantage, leave-one-day-out[6]	-0.00048
15-minute signature level	12 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.12

MSFT over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:14; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.414	0.321	0.349	1.514	1.452	1.013	0.973	5.952
2	0.402	0.360	0.405	1.246	1.256	0.948	0.987	4.337
5	0.466	0.419	0.470	0.965	1.151	0.949	0.983	2.875
10	0.492	0.445	0.487	0.801	1.054	0.922	0.937	2.124
25	0.484	0.450	0.471	0.633	0.906	0.837	0.831	1.453
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.25	−0.12	0.03	−0.47	1.80	1.39	2.16	2.44
2	0.04	0.05	0.00	−0.01	1.69	1.36	2.31	2.68
5	−0.02	−0.02	−0.03	0.00	1.43	1.28	2.23	2.93
10	−0.03	−0.05	−0.02	−0.01	2.04	1.88	2.19	3.16
25	−0.01	−0.02	0.00	0.00	3.61	3.41	2.45	3.66

Table 1: Trade-lag differences, `MSFT`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	5.42	4.91	5.04	8.61	18.3	14.6	18.9	98.9
1 s	3.41	3.26	3.32	4.23	15.4	14.1	18.1	48.6
5 s	3.02	2.96	2.98	3.31	15.5	15.1	16.8	21.8
10 s	3.07	3.04	3.04	3.24	15.8	15.6	16.3	21.1
1 min	3.12	3.12	3.11	3.16	16.5	16.4	16.2	17.7
15 min	3.17	3.17	3.17	3.18	16.6	16.5	16.3	16.6
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.29	−0.29	−0.26	−0.33	12.70	9.94	5.38	5.45
1 s	−0.40	−0.39	−0.35	−0.41	9.06	8.96	5.18	4.63
5 s	0.09	0.12	0.09	−0.09	6.69	6.91	4.83	4.41
10 s	0.14	0.15	0.13	0.01	10.47	10.91	7.82	6.66
1 min	−0.11	−0.11	−0.10	−0.12				

Table 2: Clock-lag differences, `MSFT`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.03	1.26	1.34
s.d. of $\hat{\epsilon}_t$ in cents	4.36	4.26	3.75
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.16
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.24	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.13	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.01	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.86,0.77	−0.93,0.85
95% C.I. for $\theta_t - X_t$	—	−1.53,1.39	−1.74,1.59
99% C.I. for $\theta_t - X_t$	—	−2.30,2.25	−3.48,3.13
s.e. of $\theta_t - X_t$ in cents	—	0.70	1.24

Table 3: Residual summaries, `MSFT`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.16, is above two, while the Hill index of the Huber residual itself on this day is 1.30, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.30	1.50	1.15 to 1.84	1.31 to 2.02
Residual divided by δ_t	1.58	1.76	1.44 to 1.99	1.68 to 2.21
Repricing windows removed	1.58	1.72	1.44 to 1.95	1.62 to 2.08
Calm windows only	1.66	1.72		
Every 10th trade	1.29	1.57	1.18 to 1.96	1.24 to 2.23
Six-trade half-life	1.33	1.53	1.18 to 1.92	1.31 to 2.09

Table 4: Hill estimator of the tail index of the Huber filter residual, `MSFT`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 21.8 to 38.0 cents over the days and removes 1% to 9% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 15.2 cents on 16 October 2024, 71% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

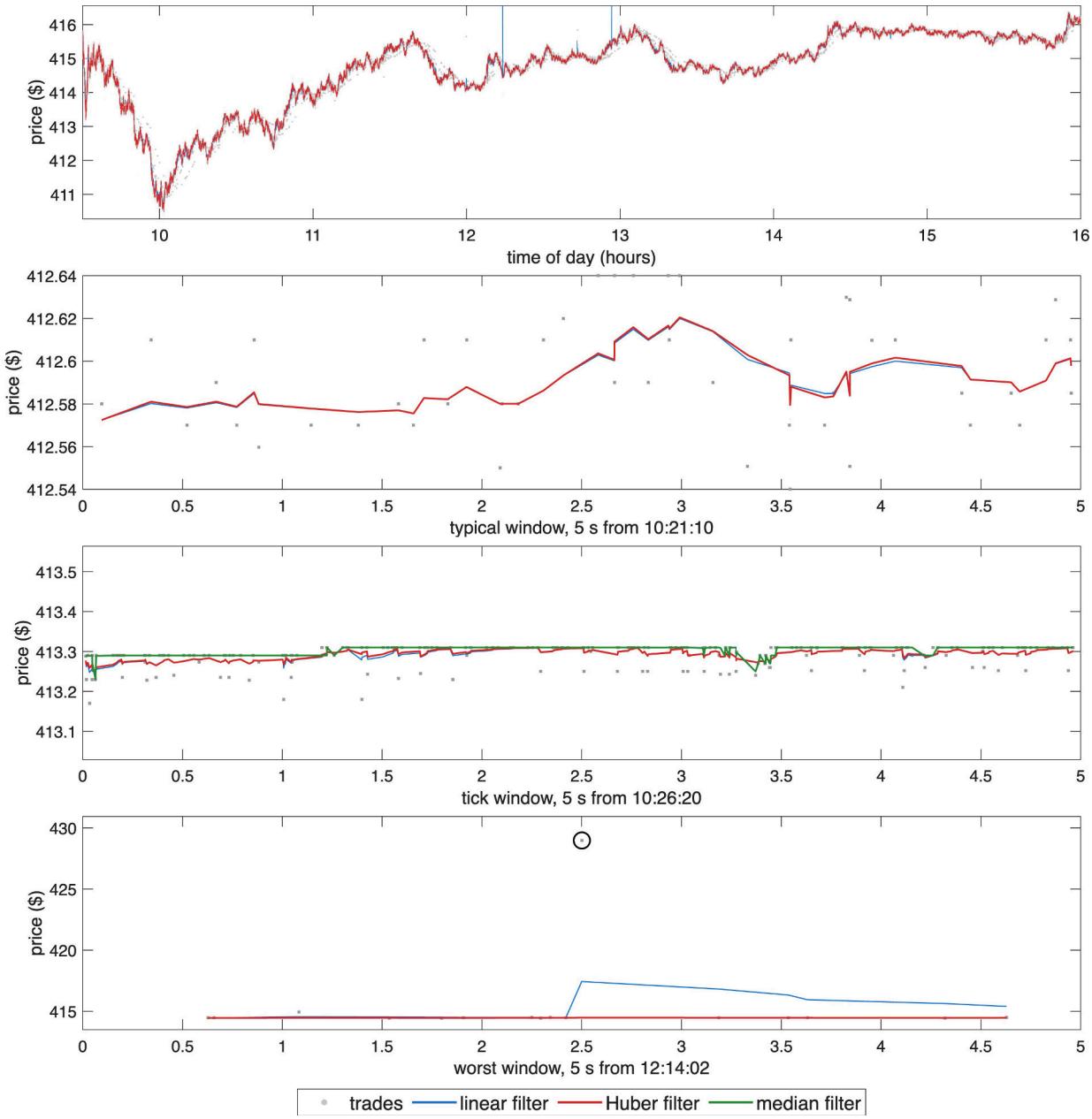

Figure 1: `MSFT` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:21:10. 3rd: 5-second window at 10:26:20 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

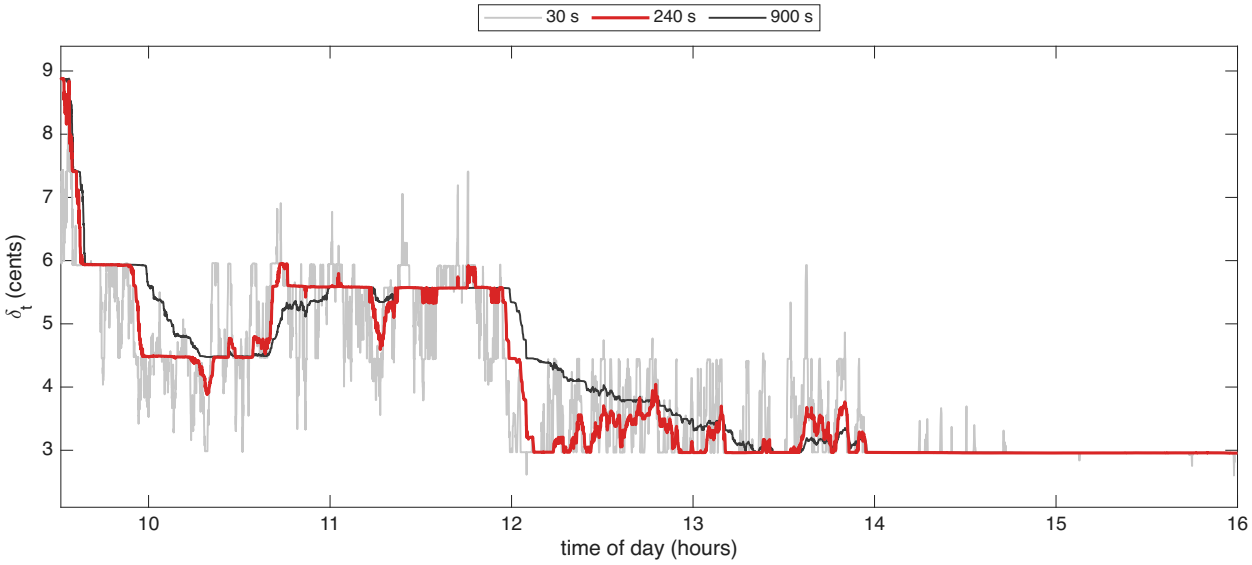

Figure 2: The Huber threshold δ_t in cents for `MSFT` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

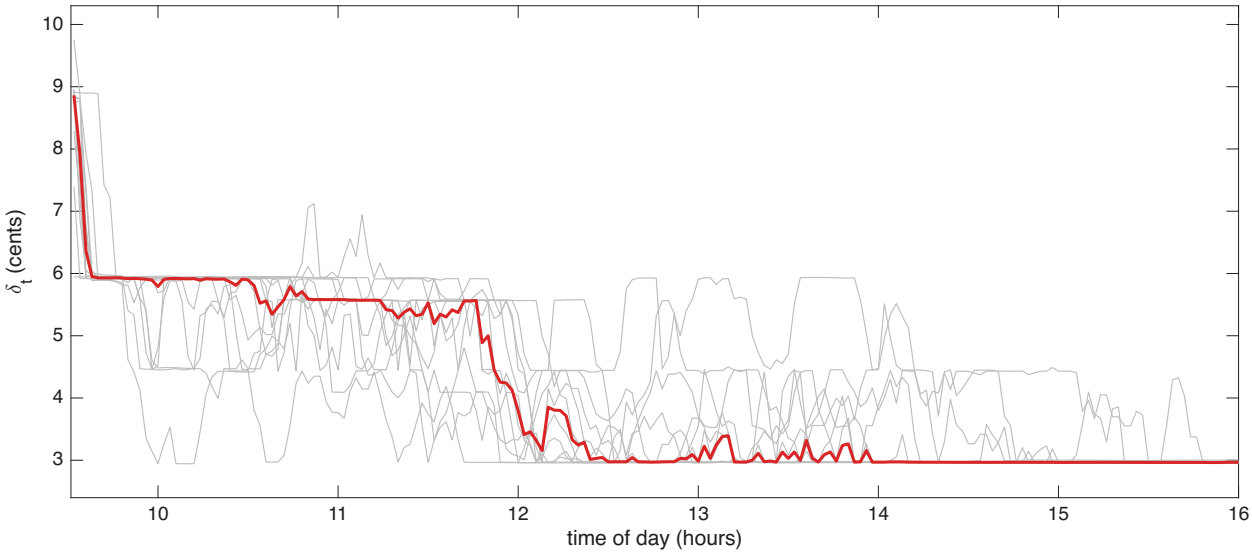

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `MSFT` trading days in Oct 2024 (grey lines) and their median (red line).

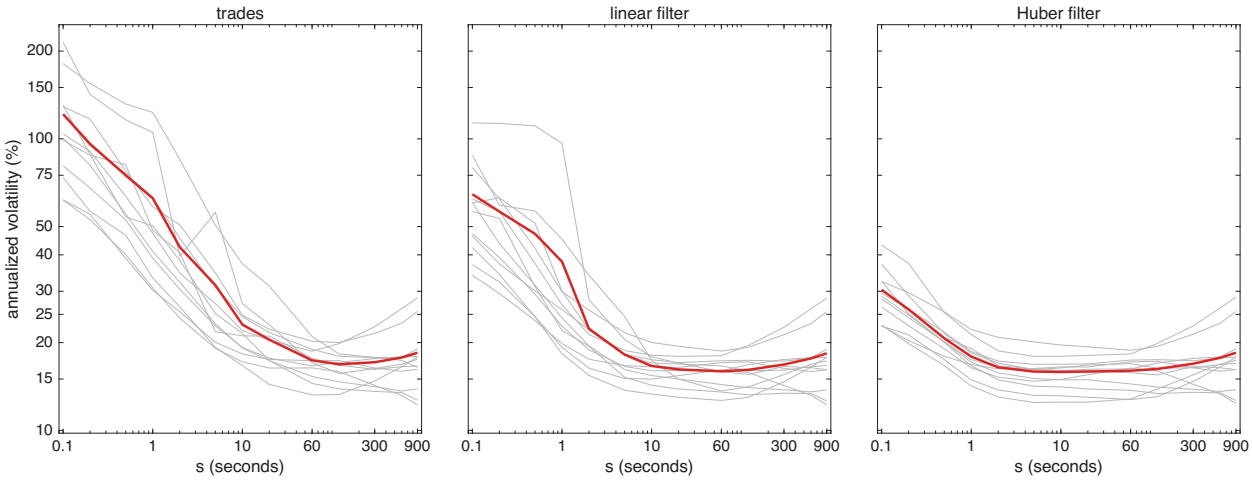

Figure 4: Volatility signature curves for `MSFT` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

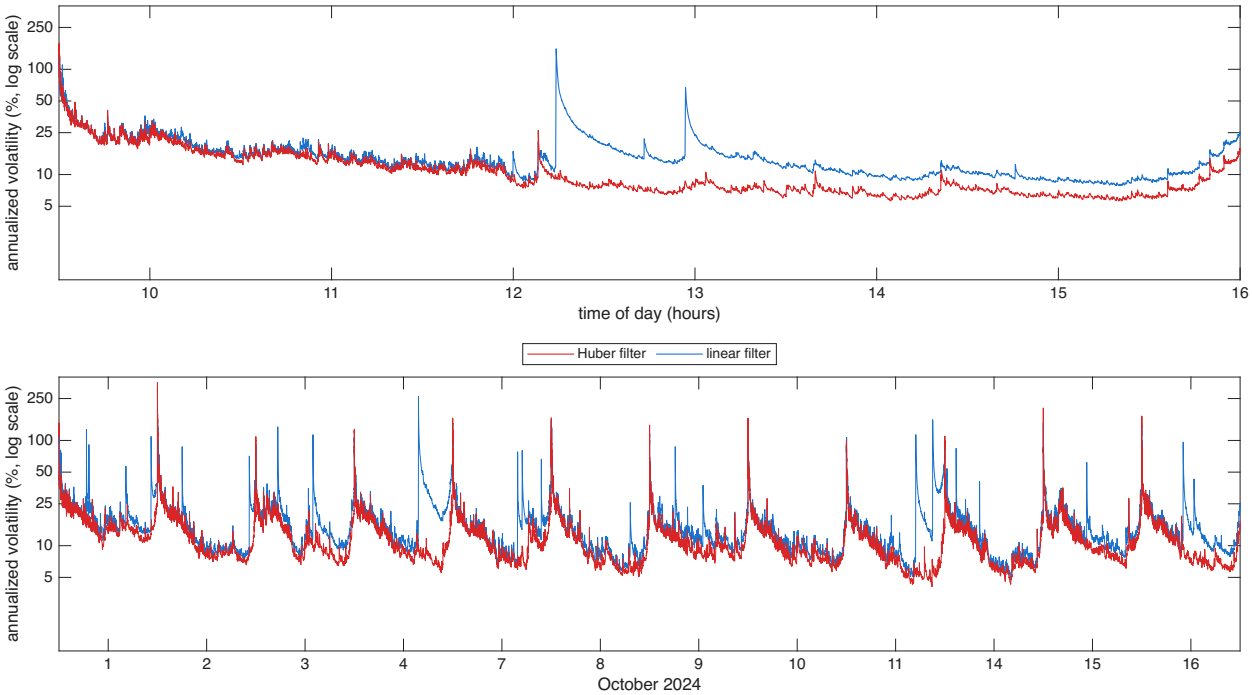

Figure 5: Filtered spot volatility for `MSFT` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 96$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

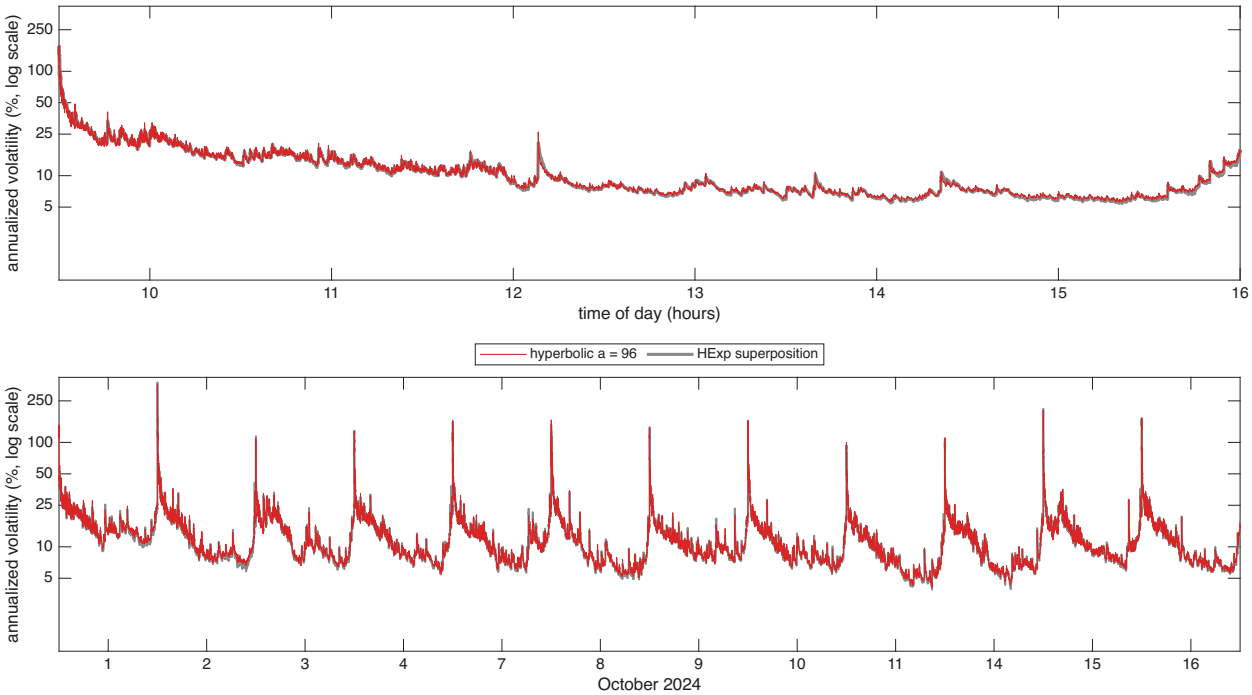

Figure 6: Hyperbolic against superposition weights for `MSFT` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 96$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 30, 150, 750, 3750 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

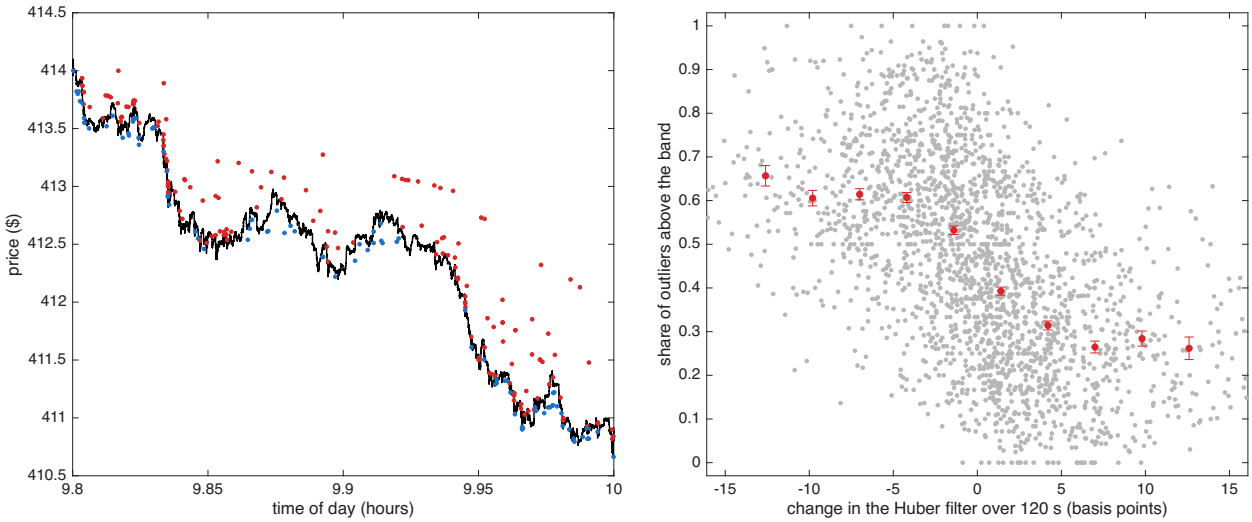

Figure 7: Outliers & robust price for `MSFT` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

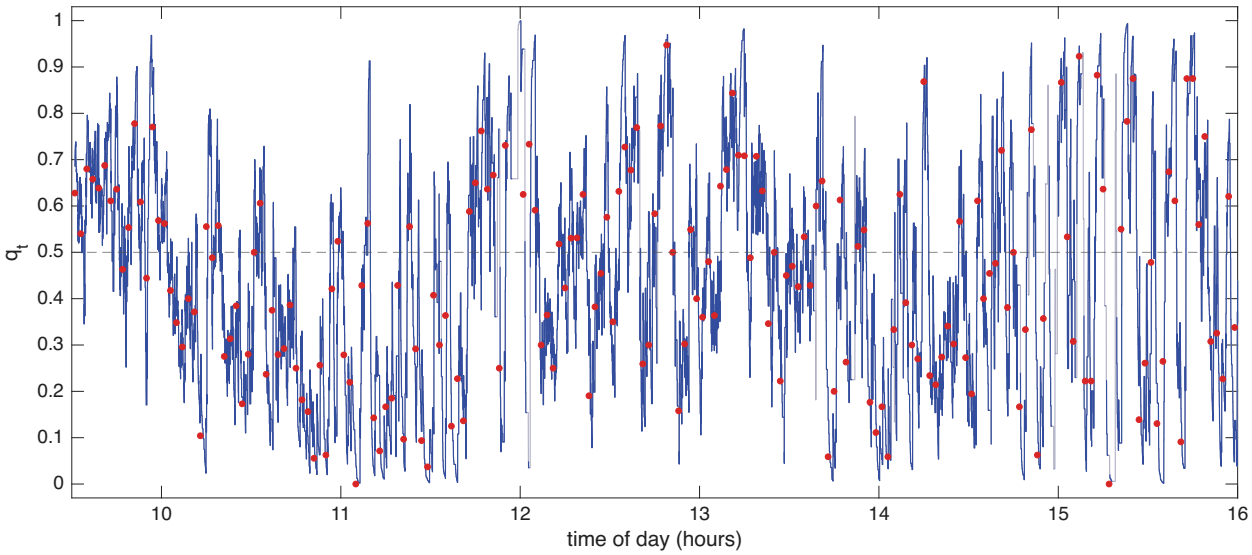

Figure 8: The directional conditional probability for `MSFT`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 22 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

AVGO 1 to 16 October 2024

trades on 16 October 2024	185,769
average trading rate	10.1 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	5,595 (3.01%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.34%
scale half-life: run / minimizer[2]	240 / 90 s
cost of running 240 s	+0.09%
direction half-life[3]	26 s [21, 34]
Mandelbrot a minimizing QLIKE	190 trades
a within 0.0005 of the minimum	111 to 310
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	15% $\to$ 143%
Huber $\hat{\sigma}$, before $\to$ peak	12% $\to$ 18%
median spot volatility, Huber / linear input[5]	25.4 / 31.6 %
hyperbolic against HExp, correlation of logs	0.9980
HExp QLIKE advantage, leave-one-day-out[6]	-0.00853
15-minute signature level	24 to 44 %
Huber signature, 1 min over 1 s, 12 day average	1.16

AVGO over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:56; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.373	0.298	0.335	1.452	1.261	0.930	0.900	5.383
2	0.374	0.340	0.394	1.208	1.158	0.916	0.921	3.973
5	0.440	0.396	0.459	0.924	1.015	0.860	0.916	2.626
10	0.482	0.441	0.490	0.778	0.916	0.819	0.868	1.932
25	0.529	0.500	0.519	0.669	0.843	0.793	0.803	1.355
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.16	−0.03	0.05	−0.46	1.80	1.53	2.11	2.17
2	−0.08	−0.09	−0.01	−0.02	1.70	1.46	2.05	2.12
5	−0.01	0.00	−0.02	0.00	1.62	1.47	1.93	2.12
10	0.00	−0.01	−0.01	0.00	1.59	1.52	1.94	2.13
25	0.00	−0.02	−0.01	0.00	1.71	1.70	2.06	2.35

Table 1: Trade-lag differences, `AVGO`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.72	3.39	3.57	6.49	30.0	24.7	27.6	158.7
1 s	2.50	2.35	2.48	3.58	26.3	24.2	31.1	78.9
5 s	2.25	2.20	2.23	2.66	28.7	28.1	31.6	66.5
10 s	2.22	2.18	2.19	2.46	29.7	29.2	29.9	40.7
1 min	2.12	2.11	2.09	2.17	29.4	29.2	28.7	31.2
15 min	2.55	2.55	2.54	2.56	36.8	36.7	36.5	36.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.33	−0.33	−0.24	−0.30	1.90	1.90	2.37	3.57
1 s	−0.07	−0.06	−0.12	−0.46	2.10	2.11	2.54	2.48
5 s	0.00	0.01	−0.01	−0.24	3.62	3.56	3.17	4.77
10 s	0.10	0.11	0.09	−0.11	7.37	8.03	6.84	6.26
1 min	−0.06	−0.05	−0.05	−0.11				

Table 2: Clock-lag differences, `AVGO`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.00	1.21	1.29
s.d. of $\hat{\epsilon}_t$ in cents	3.96	3.89	3.43
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.11
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.26	0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.15	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.06	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.03	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.82,0.78	−0.92,0.87
95% C.I. for $\theta_t - X_t$	—	−1.43,1.36	−1.72,1.60
99% C.I. for $\theta_t - X_t$	—	−2.20,2.10	−3.22,2.95
s.e. of $\theta_t - X_t$ in cents	—	0.68	1.15

Table 3: Residual summaries, `AVGO`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.11, is above two, while the Hill index of the Huber residual itself on this day is 1.66, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.66	1.86	1.66 to 2.51	1.71 to 2.23
Residual divided by δ_t	2.10	2.20	1.67 to 2.56	1.67 to 2.44
Repricing windows removed	1.91	1.96	1.61 to 2.56	1.72 to 2.18
Calm windows only	2.33	1.89		
Every 10th trade	1.80	1.93	1.64 to 2.34	1.44 to 2.24
Six-trade half-life	1.62	1.93	1.62 to 2.64	1.77 to 2.32

Table 4: Hill estimator of the tail index of the Huber filter residual, `AVGO`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 20.2 to 25.3 cents over the days and removes 2% to 6% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.1 cents on 16 October 2024, 64% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

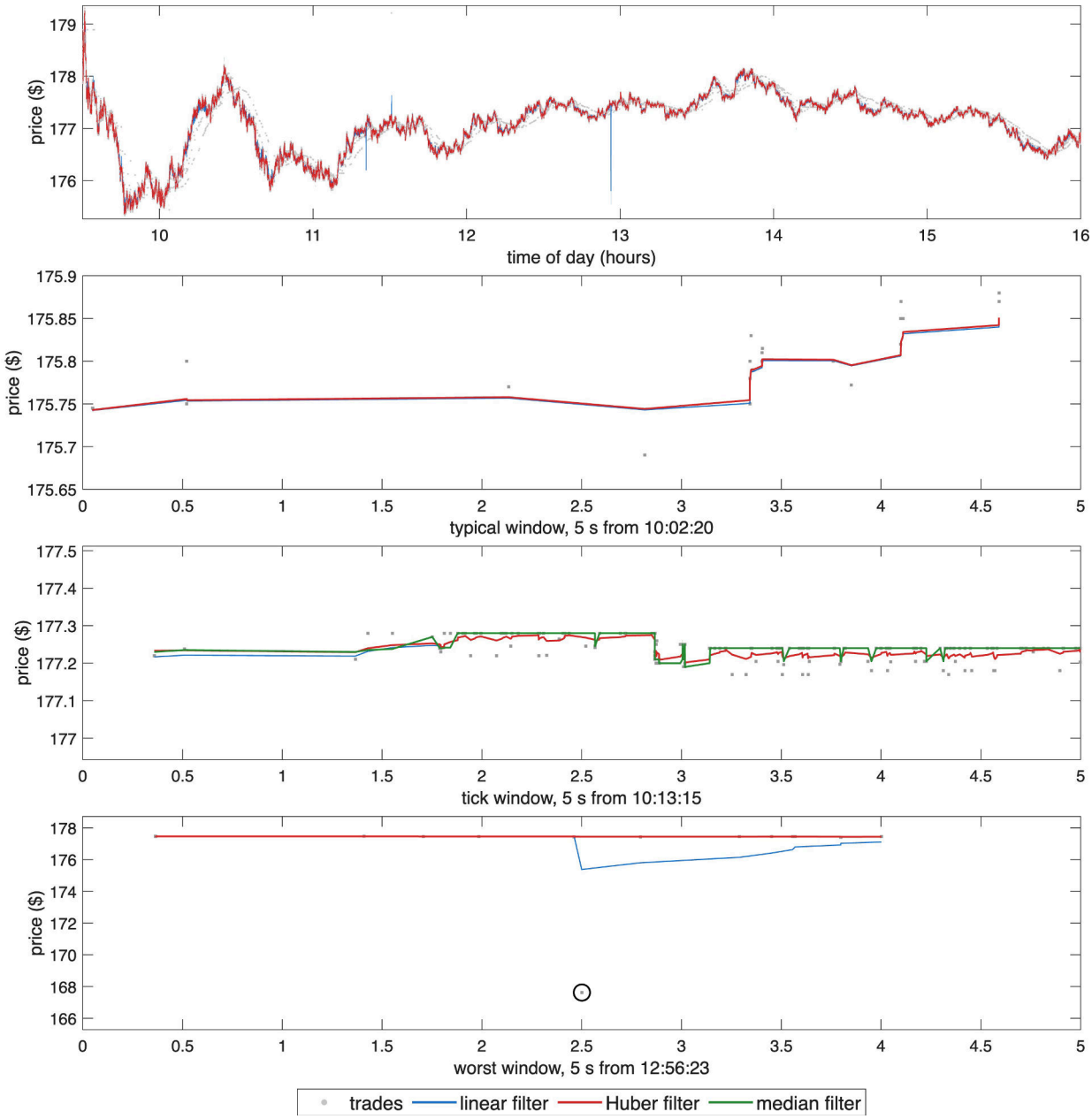

Figure 1: `AVGO` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:02:20. 3rd: 5-second window at 10:13:15 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

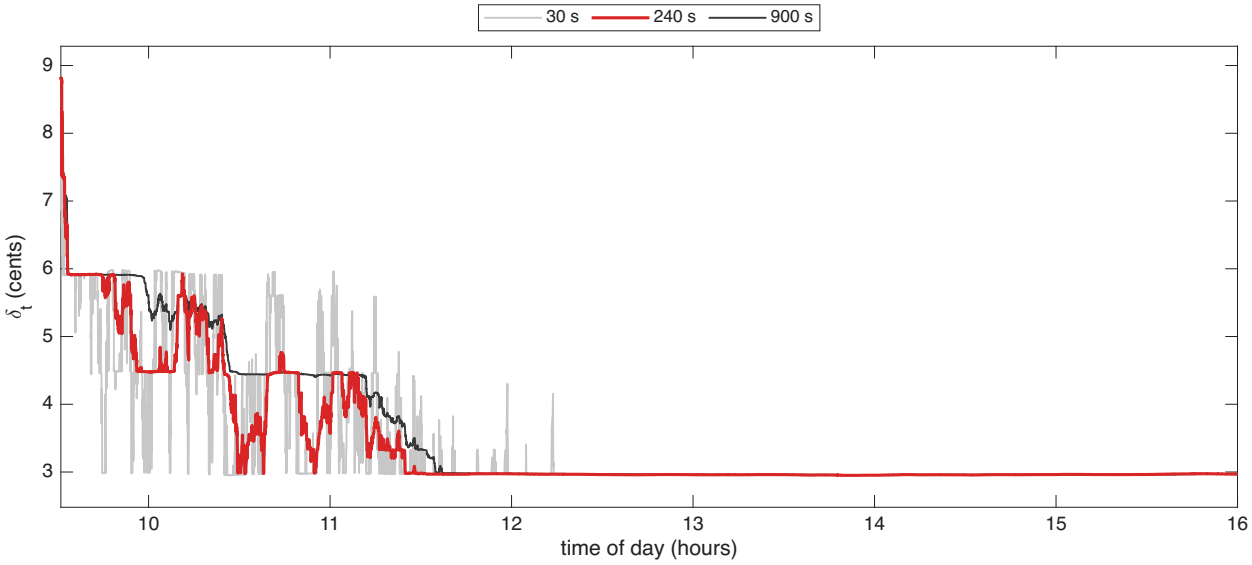

Figure 2: The Huber threshold δ_t in cents for `AVGO` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

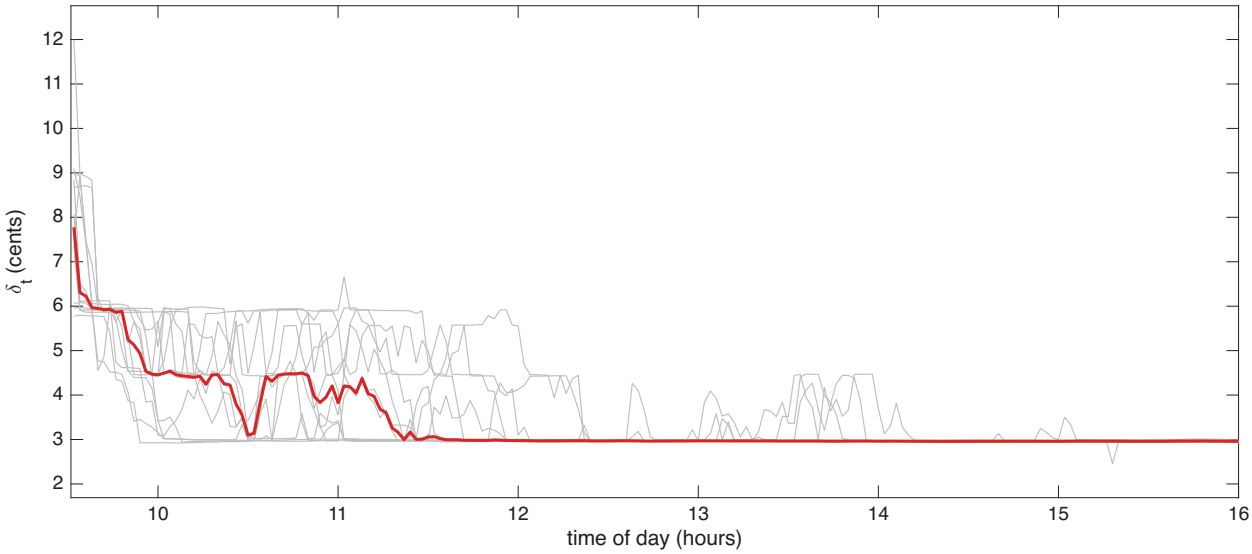

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AVGO` trading days in Oct 2024 (grey lines) and their median (red line).

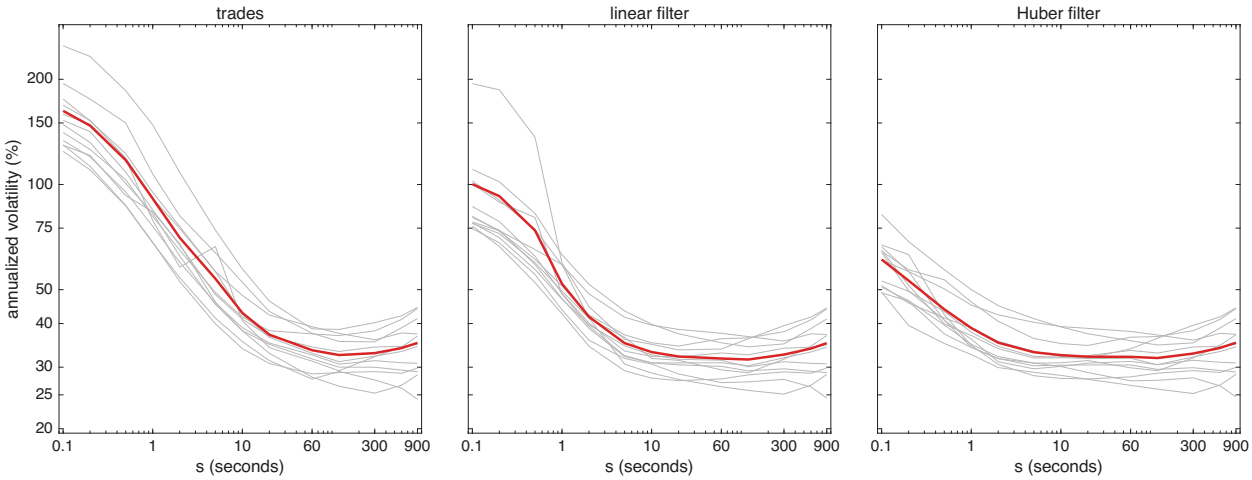

Figure 4: Volatility signature curves for `AVGO` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

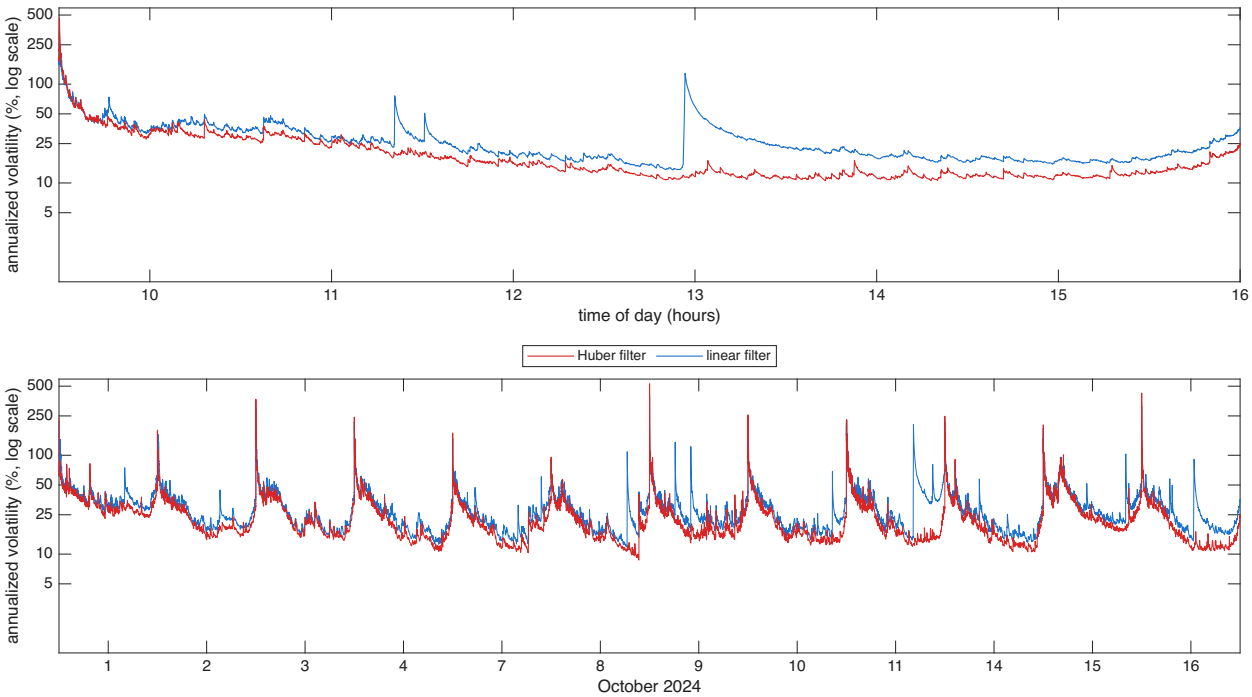

Figure 5: Filtered spot volatility for `AVGO` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 190$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

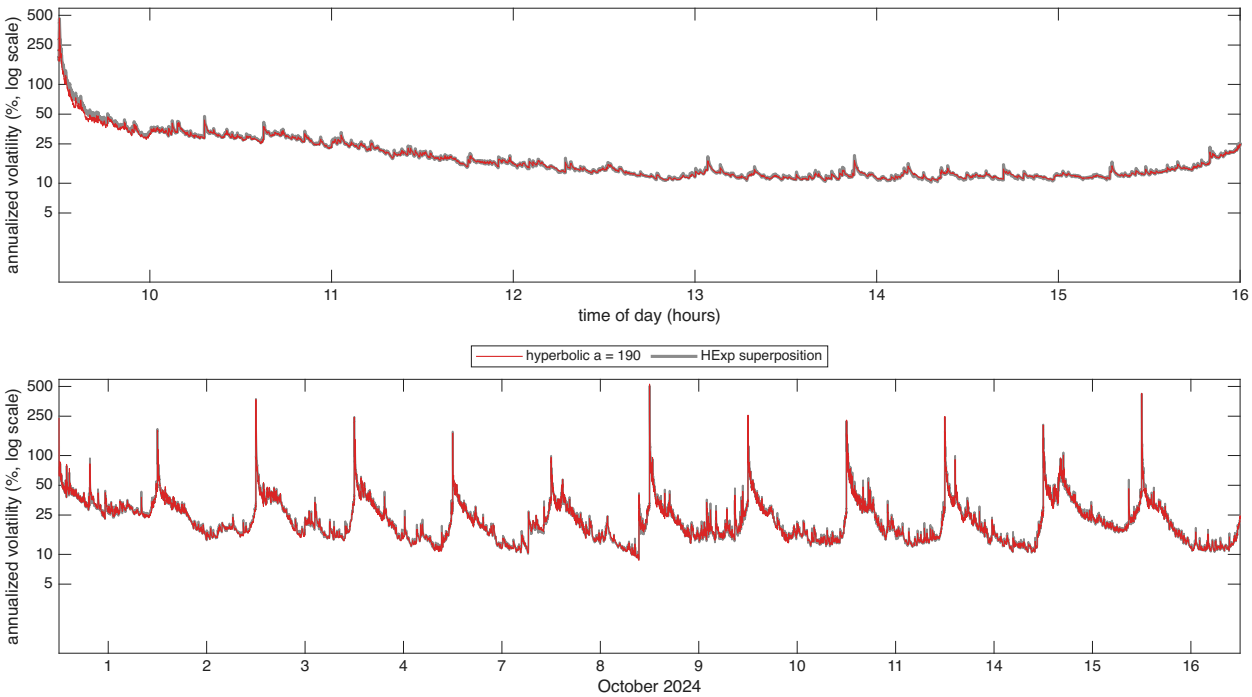

Figure 6: Hyperbolic against superposition weights for `AVGO` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 190$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

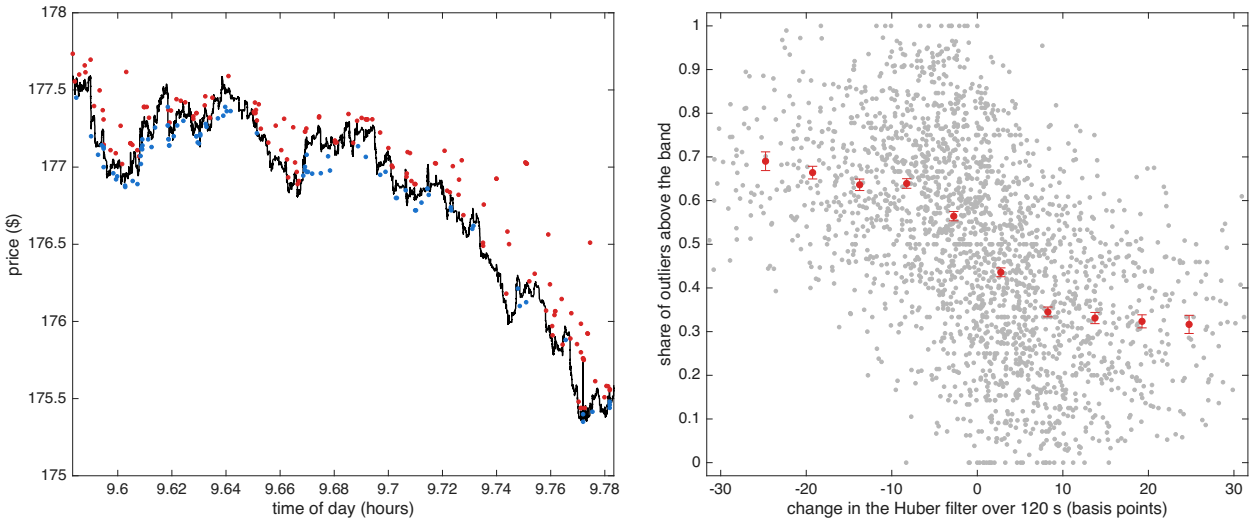

Figure 7: Outliers & robust price for `AVGO` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

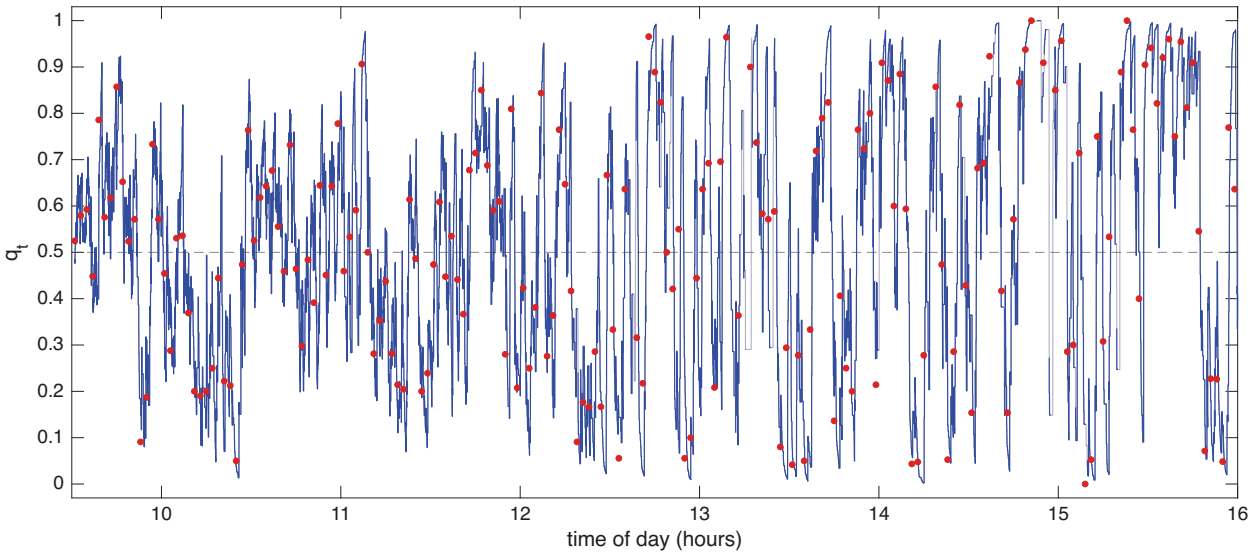

Figure 8: The directional conditional probability for AVGO, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 26 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

Diagnostics.

MU 1 to 16 October 2024

trades on 16 October 2024	215,642
average trading rate	9.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	3,464 (1.61%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+4.27%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.00%
direction half-life[3]	36 s [28, 47]
Mandelbrot a minimizing QLIKE	96 trades
a within 0.0005 of the minimum	62 to 149
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	17% $\to$ 224%
Huber $\hat{\sigma}$, before $\to$ peak	14% $\to$ 14%
median spot volatility, Huber / linear input[5]	23.2 / 29.3 %
hyperbolic against HExp, correlation of logs	0.9989
HExp QLIKE advantage, leave-one-day-out[6]	-0.00222
15-minute signature level	21 to 57 %
Huber signature, 1 min over 1 s, 12 day average	1.24

MU over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.186	0.149	0.171	0.679	0.531	0.317	0.549	3.344
2	0.195	0.177	0.207	0.576	0.483	0.331	0.547	2.414
5	0.240	0.221	0.255	0.464	0.459	0.366	0.530	1.574
10	0.275	0.257	0.284	0.411	0.447	0.391	0.506	1.157
25	0.312	0.300	0.314	0.378	0.455	0.430	0.480	0.813
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.17	0.09	−0.01	−0.48	2.41	2.13	2.11	2.31
2	0.04	0.12	−0.02	−0.01	2.13	2.07	2.36	2.37
5	0.00	0.03	−0.02	0.00	2.25	2.21	2.62	2.40
10	0.01	0.03	−0.01	0.00	2.81	2.81	2.85	2.44
25	0.00	0.00	0.00	0.00	4.11	4.39	3.98	2.56

Table 1: Trade-lag differences, MU, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.13	1.98	2.09	3.41	22.4	17.2	35.5	173.8
1 s	1.56	1.50	1.59	2.09	27.1	25.4	32.6	83.5
5 s	1.50	1.48	1.51	1.68	31.2	30.7	33.4	49.3
10 s	1.51	1.49	1.51	1.63	32.4	32.1	33.2	42.0
1 min	1.59	1.58	1.59	1.62	36.2	36.0	36.0	37.5
15 min	1.61	1.61	1.60	1.61	35.0	35.0	35.0	35.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.06	−0.02	−0.02	−0.21	3.97	3.97	4.08	2.89
1 s	0.01	0.07	0.02	−0.32	4.03	3.88	3.59	3.20
5 s	0.13	0.14	0.09	−0.05	3.97	3.95	3.48	4.85
10 s	0.22	0.23	0.22	0.09	7.29	7.48	6.88	7.43
1 min	−0.13	−0.13	−0.14	−0.14				

Table 2: Clock-lag differences, MU, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.49	0.62	0.66
s.d. of $\hat{\epsilon}_t$ in cents	2.42	2.39	2.07
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.28	0.35	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.13	0.24	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.05	0.13	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.06	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.51,0.48	−0.54,0.52
95% C.I. for $\theta_t - X_t$	—	−0.91,0.86	−1.05,0.99
99% C.I. for $\theta_t - X_t$	—	−1.39,1.39	−1.84,1.81
s.e. of $\theta_t - X_t$ in cents	—	0.42	0.67

Table 3: Residual summaries, `MU`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.10, is above two, while the Hill index of the Huber residual itself on this day is 1.71, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.72	1.71	1.44 to 2.24	1.39 to 2.31
Residual divided by δ_t	1.74	1.69	1.44 to 2.21	1.66 to 2.12
Repricing windows removed	1.73	1.73	1.39 to 2.29	1.59 to 2.29
Calm windows only	1.66	1.73		
Every 10th trade	1.71	1.81	1.31 to 2.22	1.53 to 2.26
Six-trade half-life	1.90	1.84	1.45 to 2.33	1.36 to 2.48

Table 4: Hill estimator of the tail index of the Huber filter residual, `MU`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 17.5 to 20.2 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.1 cents on 16 October 2024, 79% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

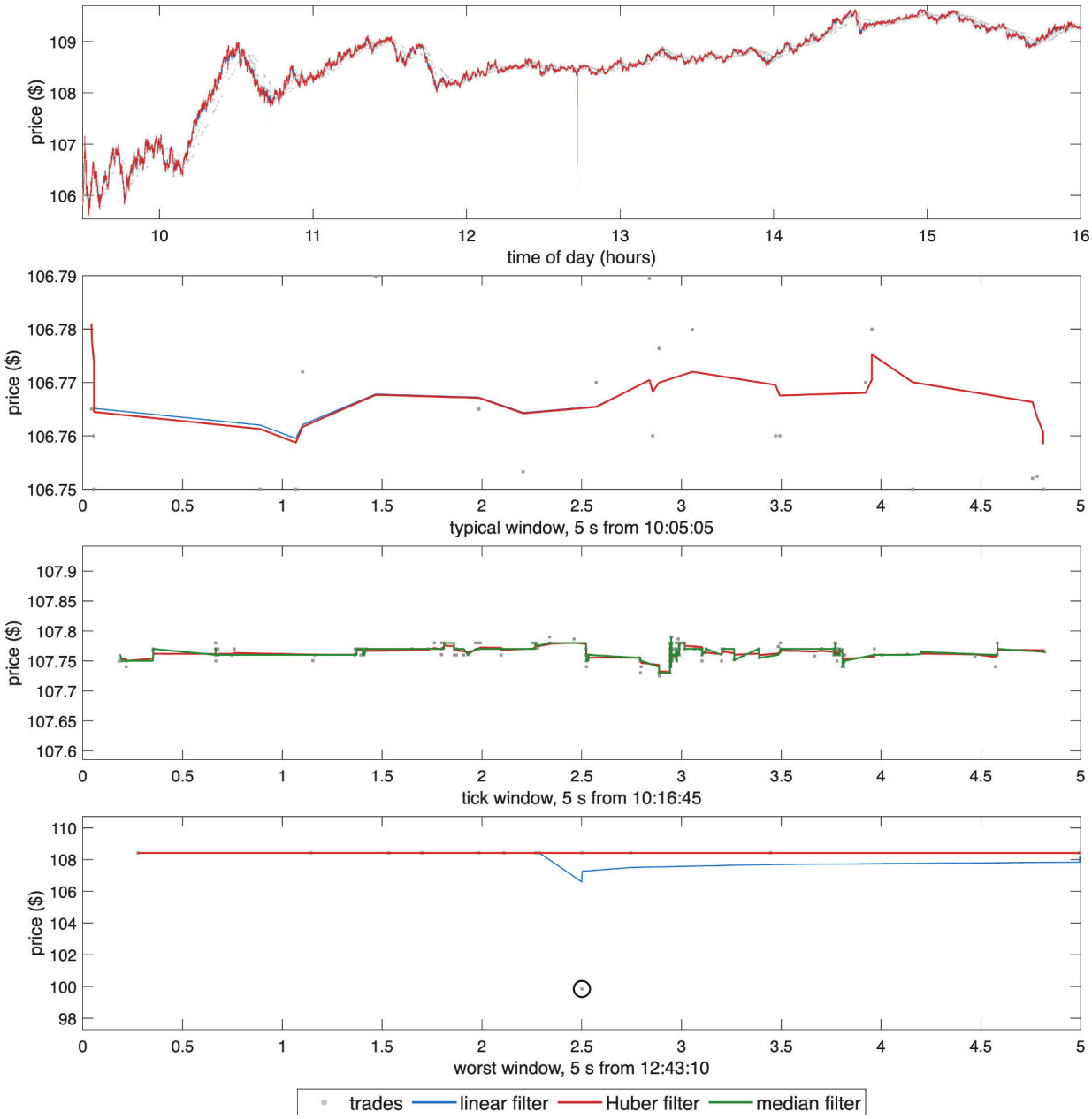

Figure 1: MU trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:05:05. 3rd: 5-second window at 10:16:45 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

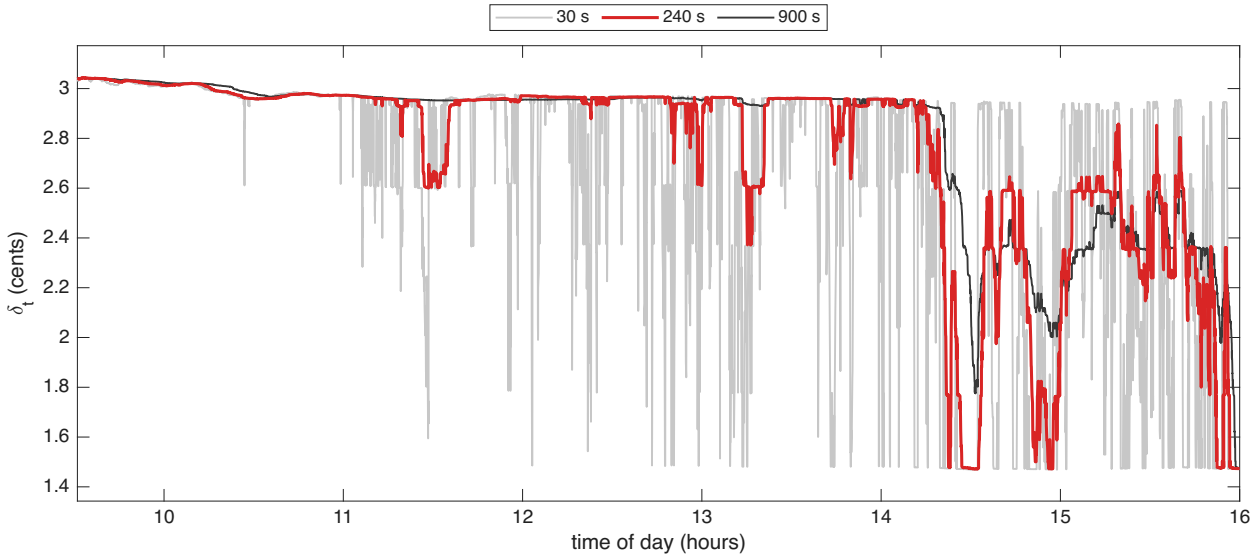

Figure 2: The Huber threshold δ_t in cents for `MU` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

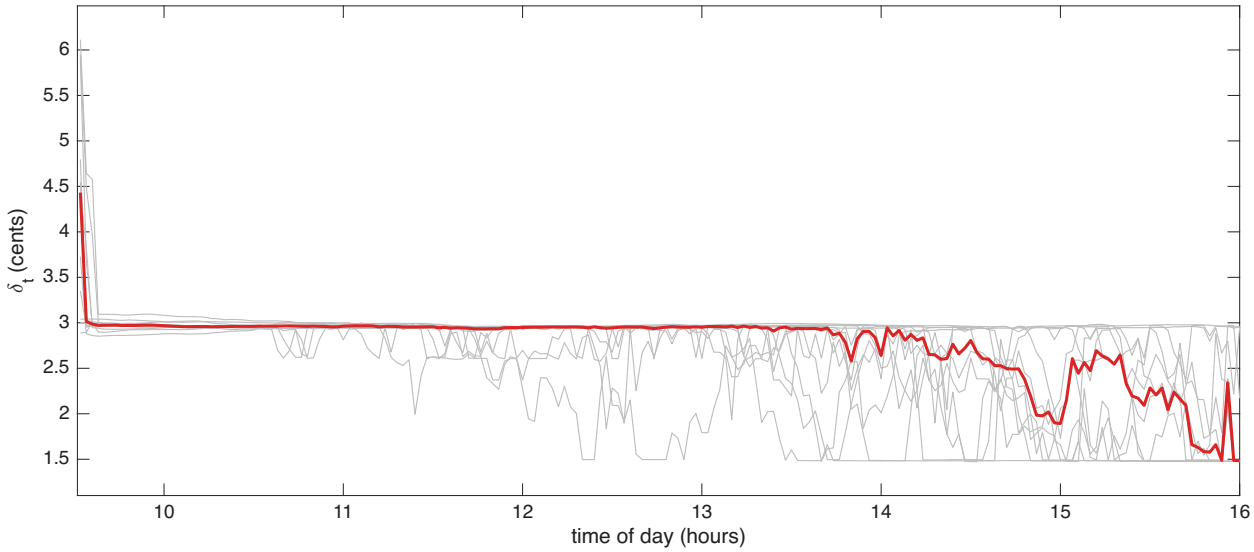

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `MU` trading days in Oct 2024 (grey lines) and their median (red line).

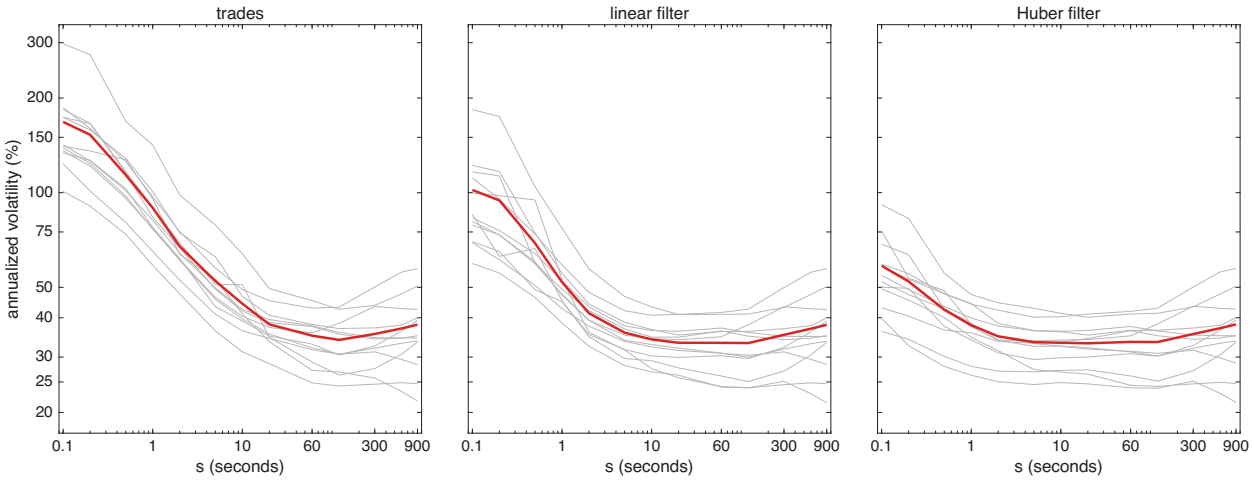

Figure 4: Volatility signature curves for `MU` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

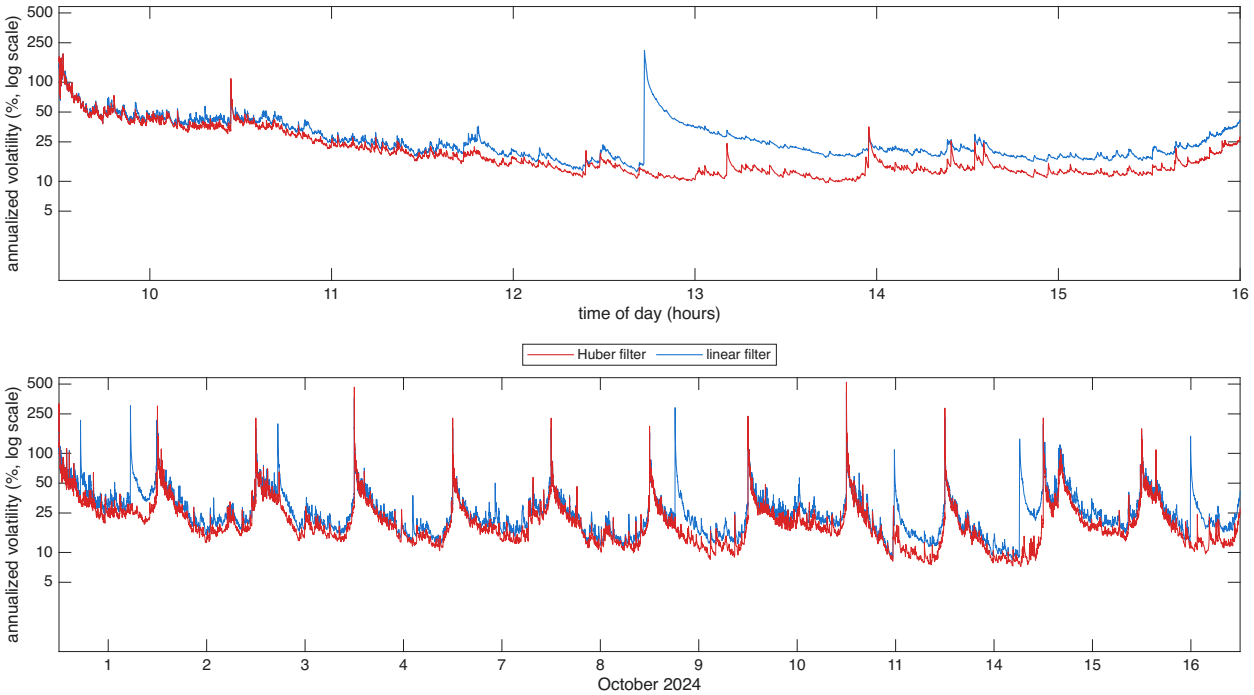

Figure 5: Filtered spot volatility for `MU` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 96$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

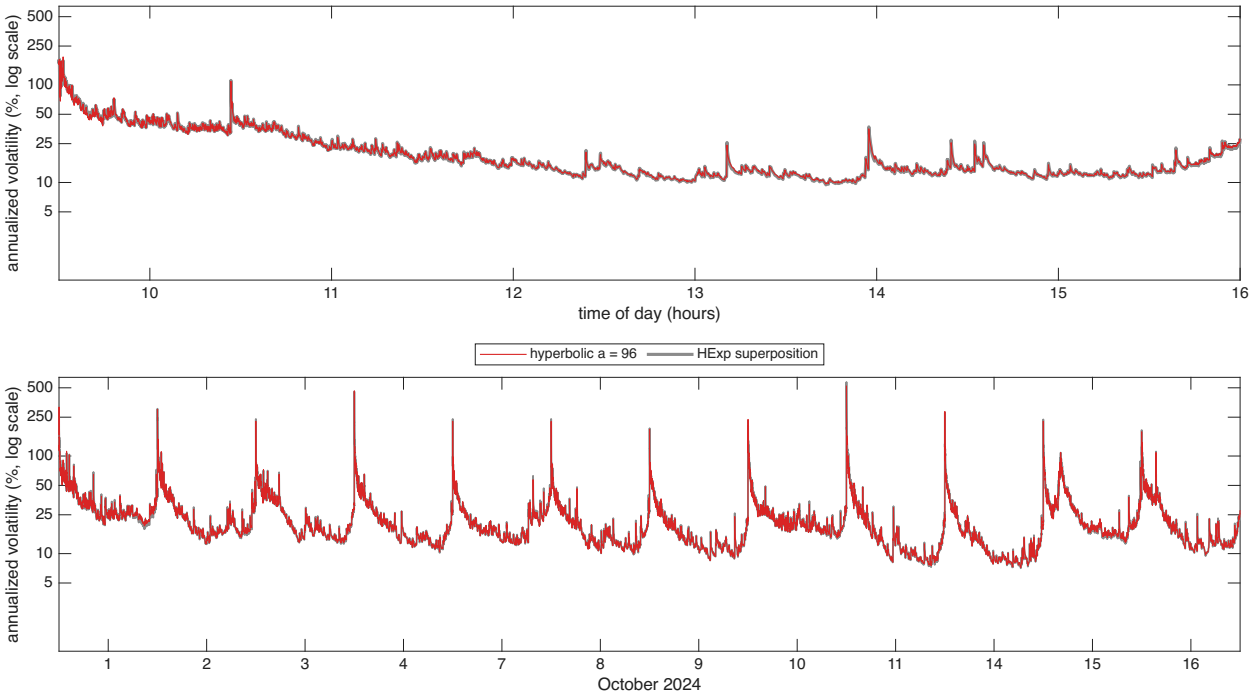

Figure 6: Hyperbolic against superposition weights for `MU` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 96$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

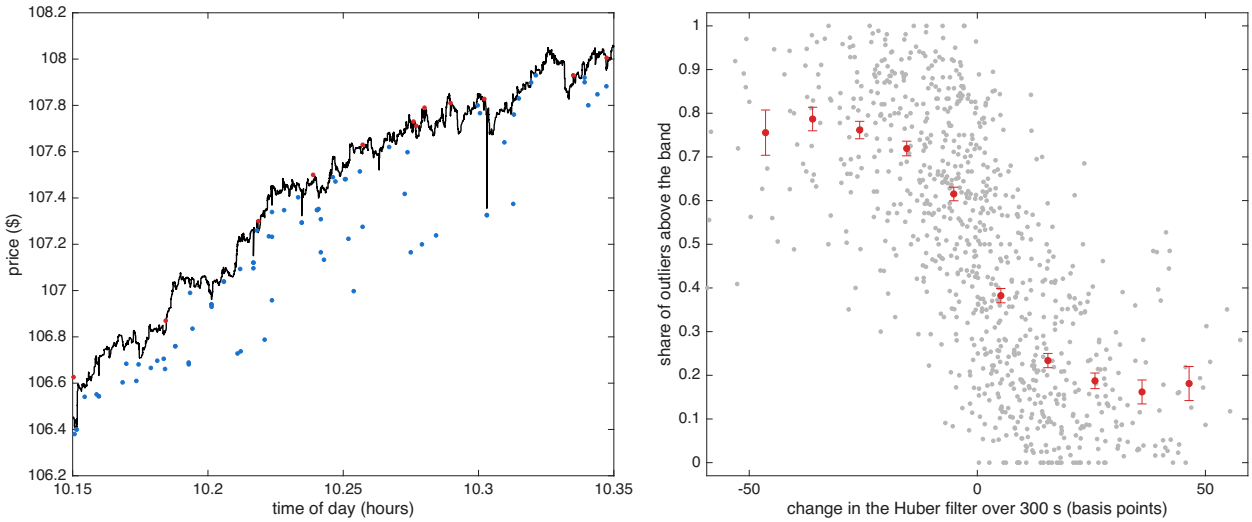

Figure 7: Outliers & robust price for `MU` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

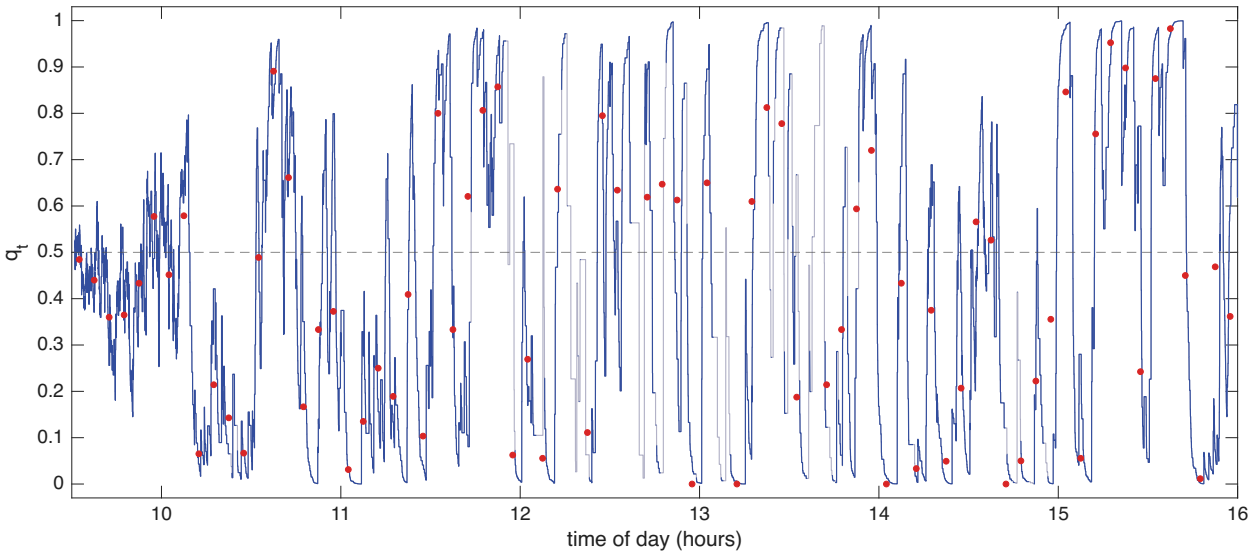

Figure 8: The directional conditional probability for `MU`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 36 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

INTC 1 to 16 October 2024

trades on 16 October 2024	260,042
average trading rate	8.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,485 (0.96%)
price half-life: run / minimizer[1]	3.0 / 3.0 trades
cost of running 3 trades	+0.12%
scale half-life: run / minimizer[2]	240 / 60 s
cost of running 240 s	+0.17%
direction half-life[3]	34 s [26, 47]
Mandelbrot a minimizing QLIKE	59 trades
a within 0.0005 of the minimum	42 to 87
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	25% $\rightarrow$ 65%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	23% $\rightarrow$ 32%
median spot volatility, Huber / linear input[5]	28.4 / 32.4 %
hyperbolic against HExp, correlation of logs	0.9833
HExp QLIKE advantage, leave-one-day-out[6]	+0.01180
15-minute signature level	24 to 48 %
Huber signature, 1 min over 1 s, 12 day average	0.98

INTC over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 13:21; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.062	0.049	0.053	0.239	0.207	0.105	0.110	0.641
2	0.057	0.055	0.060	0.194	0.184	0.114	0.115	0.480
5	0.063	0.061	0.067	0.142	0.164	0.120	0.120	0.325
10	0.065	0.063	0.068	0.114	0.145	0.117	0.118	0.244
25	0.067	0.066	0.068	0.093	0.121	0.107	0.108	0.171
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.20	0.16	0.11	−0.44	2.12	2.22	2.58	2.73
2	0.01	0.04	0.03	−0.03	2.18	1.73	2.61	2.84
5	−0.01	−0.02	−0.02	0.00	1.63	1.54	2.25	2.90
10	−0.01	−0.03	−0.02	0.00	1.41	1.47	2.07	3.10
25	0.01	0.02	0.01	0.01	2.02	1.92	2.16	3.31

Table 1: Trade-lag differences, `INTC`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.87	0.82	0.84	1.32	48.2	33.9	37.1	173.8
1 s	0.54	0.52	0.54	0.73	39.7	34.7	37.8	96.3
5 s	0.44	0.44	0.44	0.51	37.9	36.4	37.1	56.3
10 s	0.44	0.44	0.44	0.48	38.4	37.3	37.3	47.8
1 min	0.43	0.43	0.42	0.43	39.4	38.8	38.4	40.3
15 min	0.45	0.45	0.45	0.45	40.5	40.3	40.2	40.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.15	−0.09	−0.04	−0.29	2.71	2.78	3.12	4.00
1 s	−0.08	0.00	−0.01	−0.34	3.68	3.67	4.20	5.85
5 s	0.15	0.19	0.18	−0.06	10.72	6.43	5.47	5.20
10 s	0.17	0.19	0.19	0.08	15.19	13.00	12.35	12.31
1 min	−0.06	−0.06	−0.05	−0.07				

Table 2: Clock-lag differences, `INTC`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.15	0.20	0.20
s.d. of $\hat{\epsilon}_t$ in cents	0.48	0.45	0.42
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.58
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.14	0.16	0.11
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.03	0.05	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.00	0.00	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	0.01
80% C.I. for $\theta_t - X_t$	—	−0.13,0.13	−0.14,0.14
95% C.I. for $\theta_t - X_t$	—	−0.24,0.24	−0.26,0.26
99% C.I. for $\theta_t - X_t$	—	−0.37,0.37	−0.48,0.46
s.e. of $\theta_t - X_t$ in cents	—	0.12	0.15

Table 3: Residual summaries, `INTC`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.58, is above two, while the Hill index of the Huber residual itself on this day is 1.85, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.88	1.85	1.06 to 2.39	1.06 to 2.39
Residual divided by δ_t	2.10	1.86	1.24 to 2.35	1.27 to 2.51
Repricing windows removed	2.02	1.90	1.67 to 2.39	1.66 to 2.47
Calm windows only	2.11	1.92		
Every 10th trade	1.81	1.77	1.12 to 2.17	1.27 to 1.99
Six-trade half-life	1.89	1.93	1.06 to 2.54	1.04 to 2.55

Table 4: Hill estimator of the tail index of the Huber filter residual, `INTC`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 8.5 to 10.0 cents over the days and removes 0% to 7% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 4.9 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

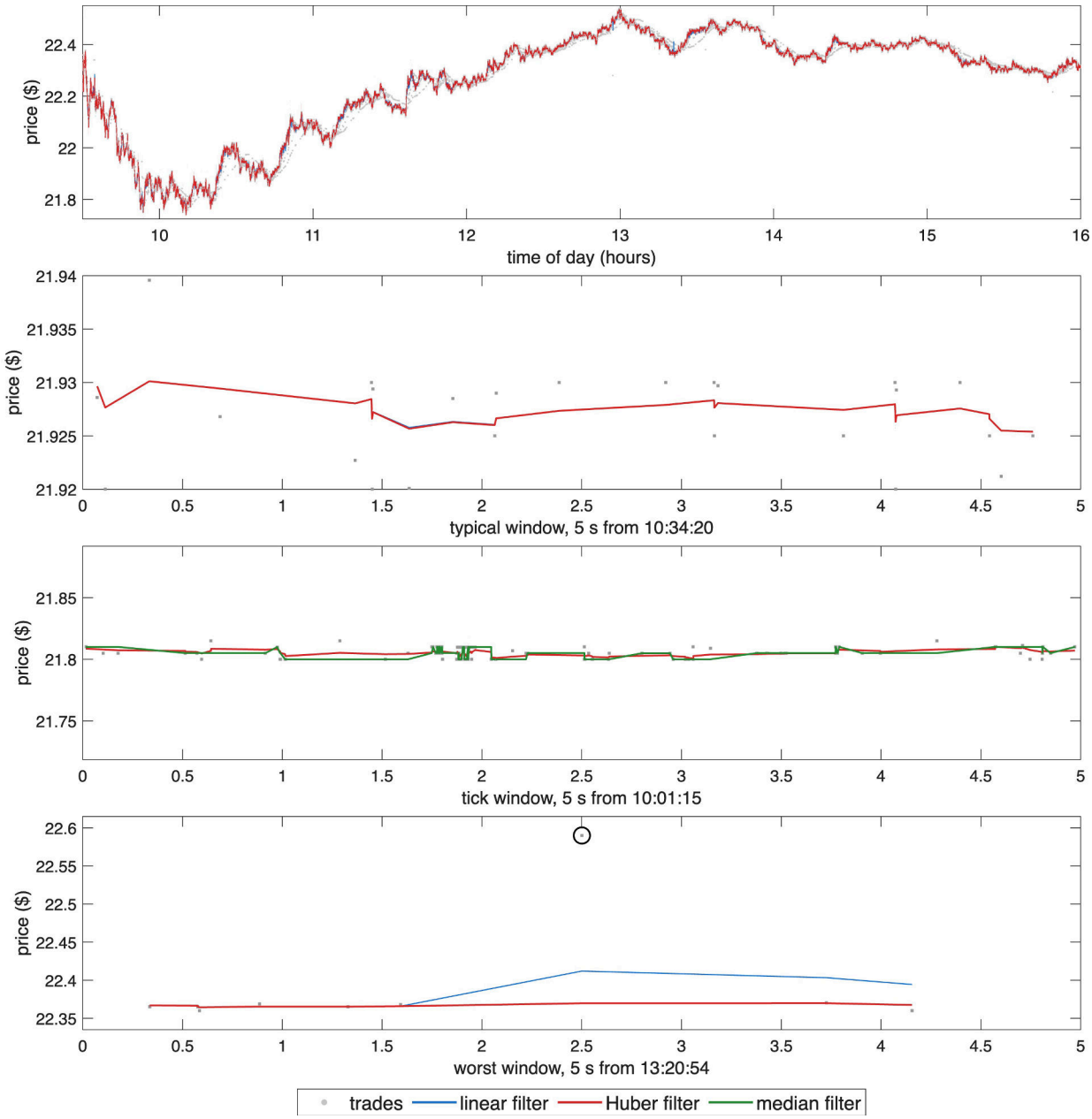

Figure 1: `INTC` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:34:20. 3rd: 5-second window at 10:01:15 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

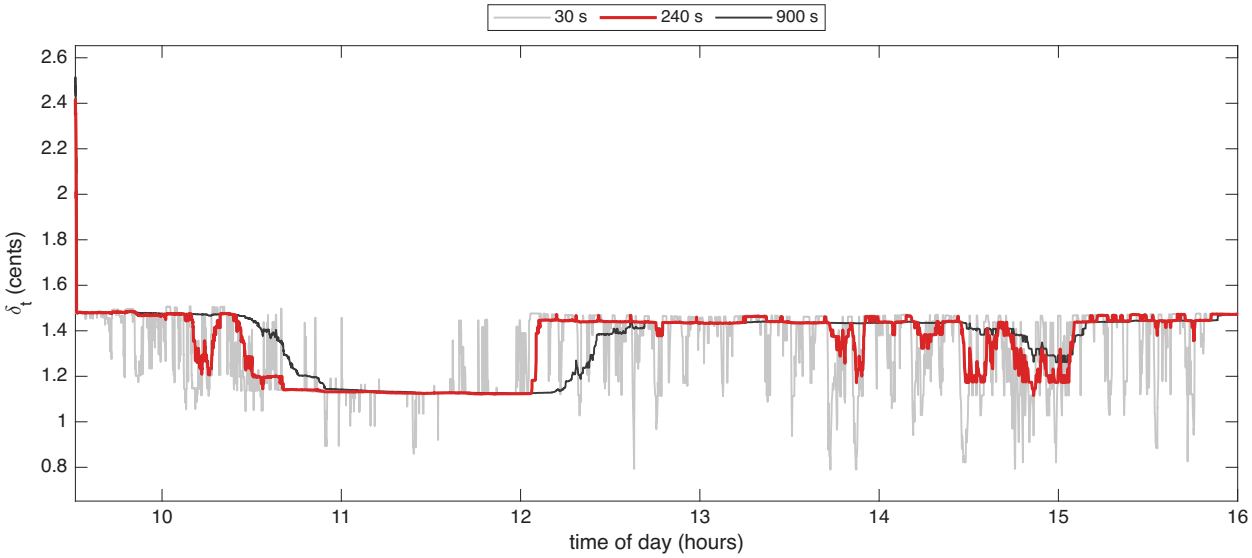

Figure 2: The Huber threshold δ_t in cents for `INTC` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

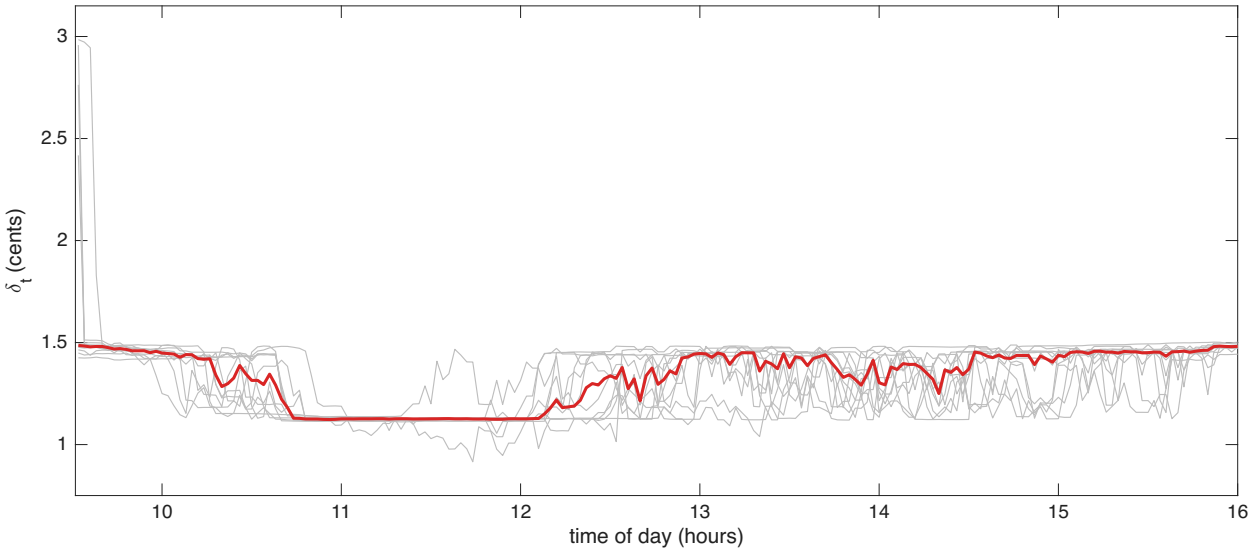

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `INTC` trading days in Oct 2024 (grey lines) and their median (red line).

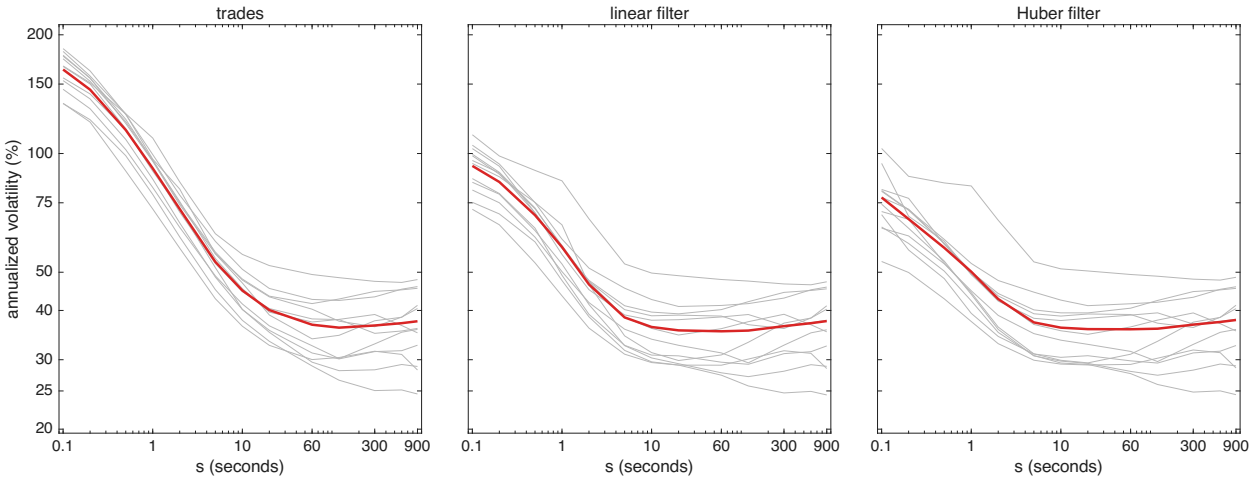

Figure 4: Volatility signature curves for `INTC` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

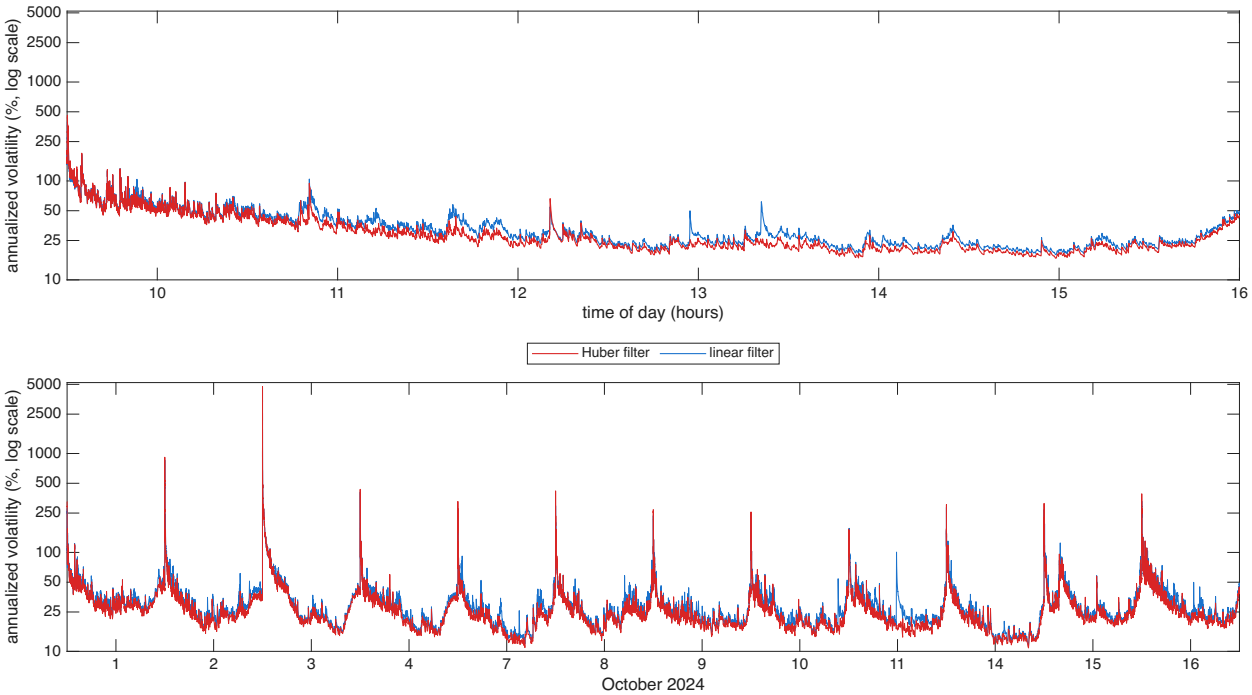

Figure 5: Filtered spot volatility for `INTC` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 59$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

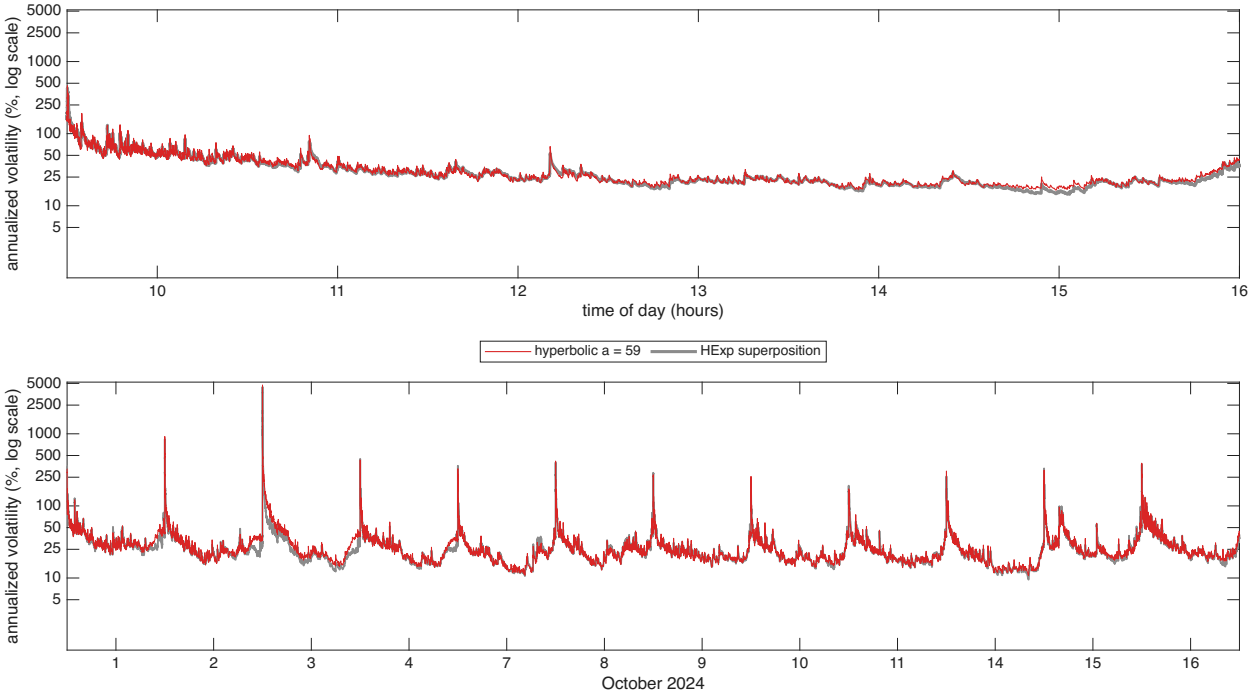

Figure 6: Hyperbolic against superposition weights for `INTC` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 59$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

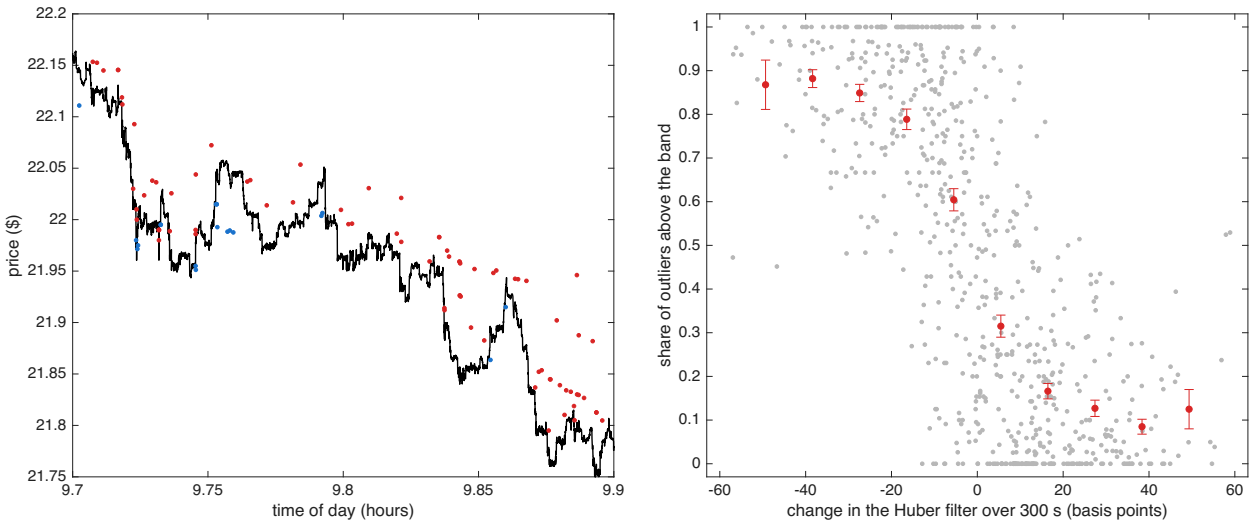

Figure 7: Outliers & robust price for `INTC` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

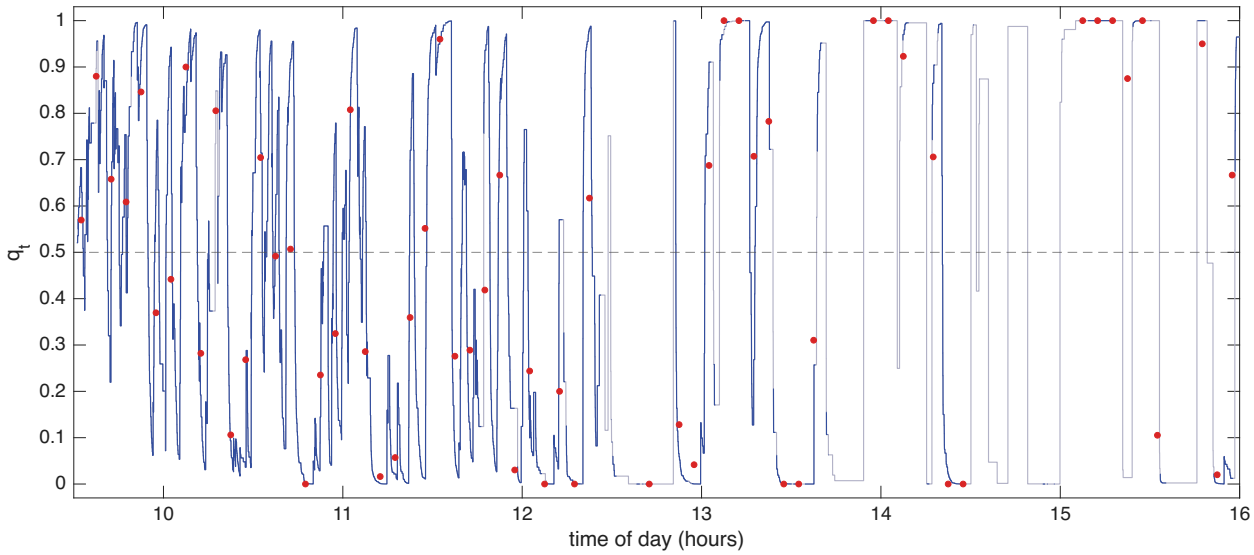

Figure 8: The directional conditional probability for INTC, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 34 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

Selecting a on single days gives values from 2 to 268 trades, against 59 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 89 trades for this stock.

The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum. On 3 October 2024, 50 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 19 per cent on any other day. Every number that day contributes to carries that burst.

Diagnostics.

META 1 to 16 October 2024

trades on 16 October 2024	226,251
average trading rate	8.7 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	9,789 (4.33%)
price half-life: run / minimizer[1]	3.0 / 3.1 trades
cost of running 3 trades	+0.14%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.01%
direction half-life[3]	14 s [12, 18]
Mandelbrot a minimizing QLIKE	111 trades
a within 0.0005 of the minimum	72 to 181
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	32% $\rightarrow$ 97%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	31% $\rightarrow$ 47%
median spot volatility, Huber / linear input[5]	18.6 / 20.7 %
hyperbolic against HExp, correlation of logs	0.9985
HExp QLIKE advantage, leave-one-day-out[6]	-0.00177
15-minute signature level	15 to 30 %
Huber signature, 1 min over 1 s, 12 day average	1.06

META over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:58; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.237	0.913	0.918	4.425	2.937	1.635	1.664	9.937
2	1.130	0.969	1.009	3.575	2.581	1.677	1.719	7.403
5	1.204	1.012	1.068	2.582	2.235	1.625	1.704	4.927
10	1.162	0.993	1.041	2.002	1.905	1.509	1.593	3.606
25	1.076	0.952	0.977	1.509	1.573	1.367	1.415	2.473
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.05	0.07	−0.45	3.42	2.65	2.68	3.01
2	0.02	−0.02	−0.01	−0.02	3.25	2.58	2.74	2.97
5	−0.03	−0.04	−0.04	0.00	3.25	2.74	2.77	2.87
10	0.00	−0.01	−0.02	0.00	3.40	3.05	2.78	2.93
25	0.00	0.00	0.00	0.00	3.72	3.70	3.27	3.00

Table 1: Trade-lag differences, `META`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	10.91	9.69	9.74	18.90	26.8	17.3	19.8	106.9
1 s	6.38	5.90	6.03	9.18	22.7	18.7	19.6	48.7
5 s	4.92	4.67	4.71	5.93	19.3	18.0	18.2	27.8
10 s	4.76	4.61	4.63	5.39	19.1	18.4	18.4	24.0
1 min	4.55	4.51	4.49	4.72	19.5	19.3	19.1	20.0
15 min	4.13	4.12	4.12	4.15	18.0	17.9	17.9	18.0
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.19	−0.15	−0.12	−0.39	4.03	4.19	4.58	3.80
1 s	−0.10	−0.03	−0.03	−0.35	5.13	5.64	6.15	4.90
5 s	0.15	0.20	0.18	−0.17	9.58	10.63	8.73	5.23
10 s	0.14	0.16	0.16	−0.01	8.71	9.24	9.97	7.14
1 min	0.03	0.04	0.04	0.01				

Table 2: Clock-lag differences, `META`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.81	3.40	3.53
s.d. of $\hat{\epsilon}_t$ in cents	7.10	6.83	6.39
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.68
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.15	0.14	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.04	0.03	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	−0.02	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	−1.98,1.92	−2.13,2.07
95% C.I. for $\theta_t - X_t$	—	−3.54,3.47	−3.93,3.84
99% C.I. for $\theta_t - X_t$	—	−5.49,5.61	−6.65,6.56
s.e. of $\theta_t - X_t$ in cents	—	1.65	2.12

Table 3: Residual summaries, `META`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.68, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.32	2.37	1.58 to 2.52	1.36 to 2.68
Residual divided by δ_t	2.23	2.19	1.86 to 2.50	1.68 to 2.82
Repricing windows removed	2.44	2.34	1.85 to 2.66	2.00 to 2.80
Calm windows only	2.49	2.32		
Every 10th trade	2.49	2.15	1.50 to 2.63	1.44 to 2.66
Six-trade half-life	2.37	2.48	1.58 to 2.66	1.35 to 2.73

Table 4: Hill estimator of the tail index of the Huber filter residual, `META`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 60.5 to 100.8 cents over the days and removes 1% to 5% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 30.3 cents on 16 October 2024, 74% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

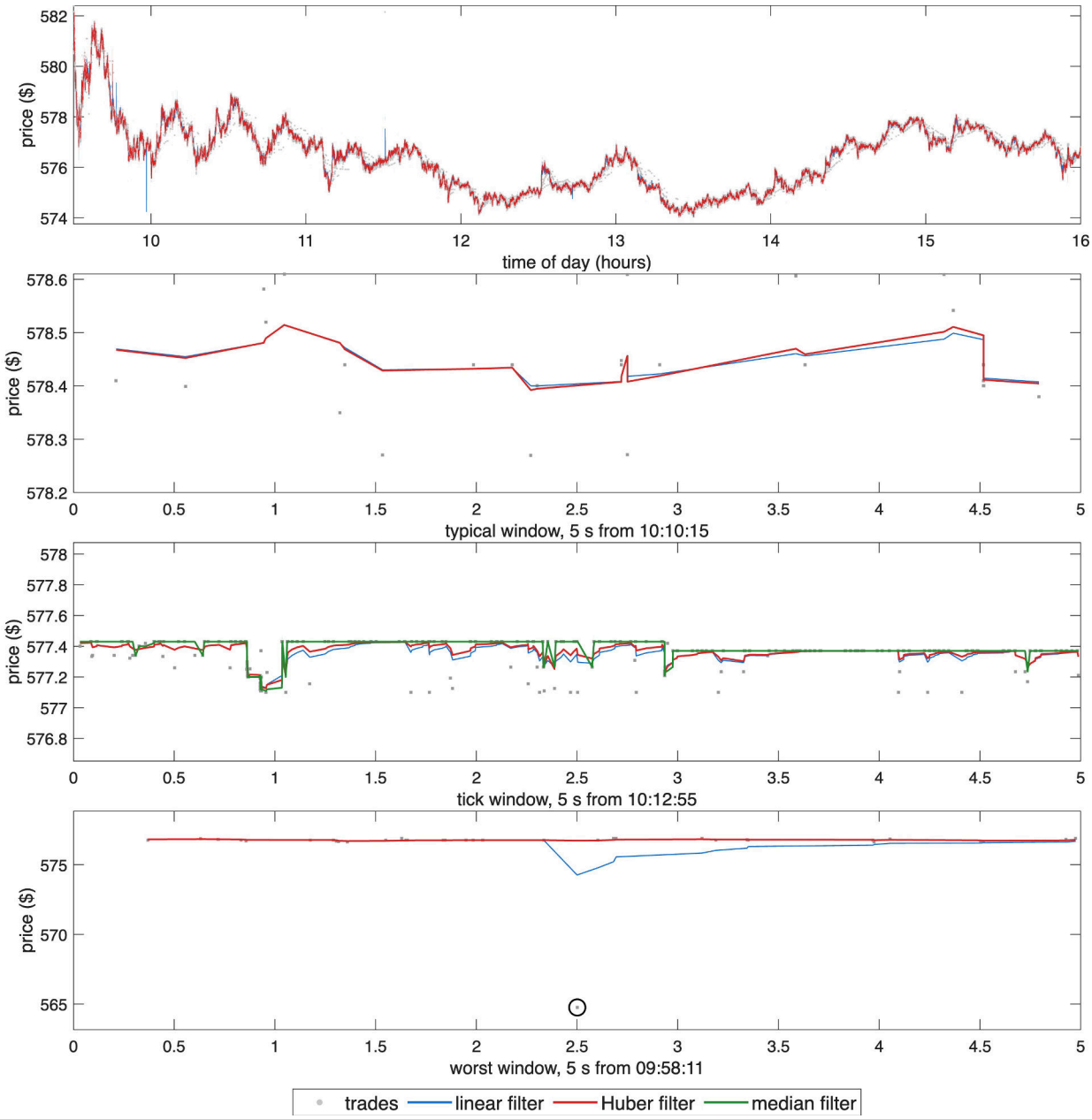

Figure 1: `META` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:10:15. 3rd: 5-second window at 10:12:55 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

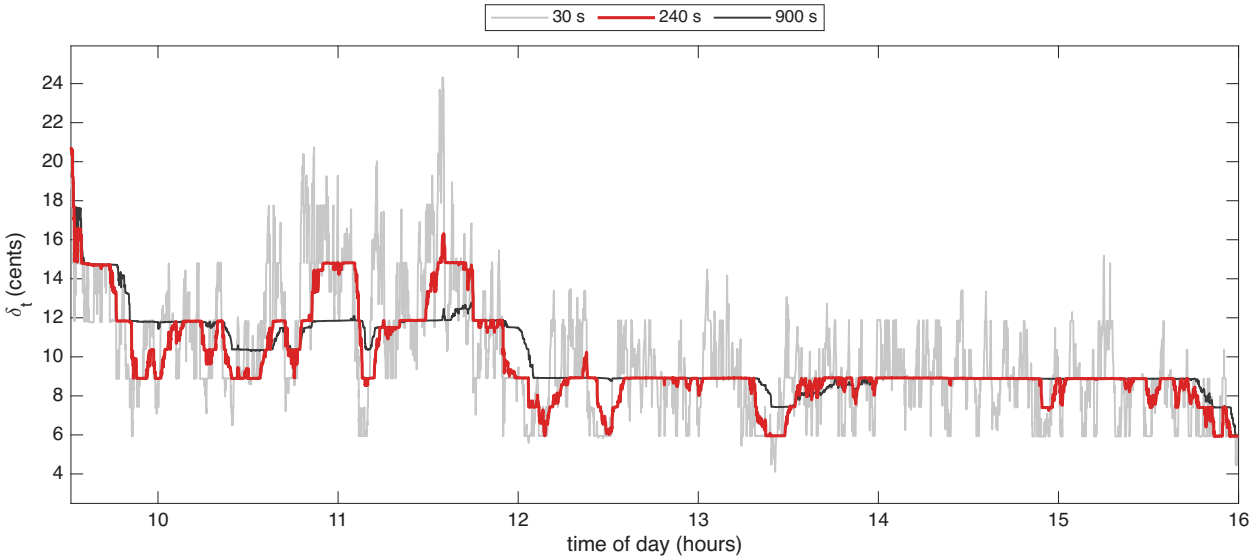

Figure 2: The Huber threshold δ_t in cents for `META` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

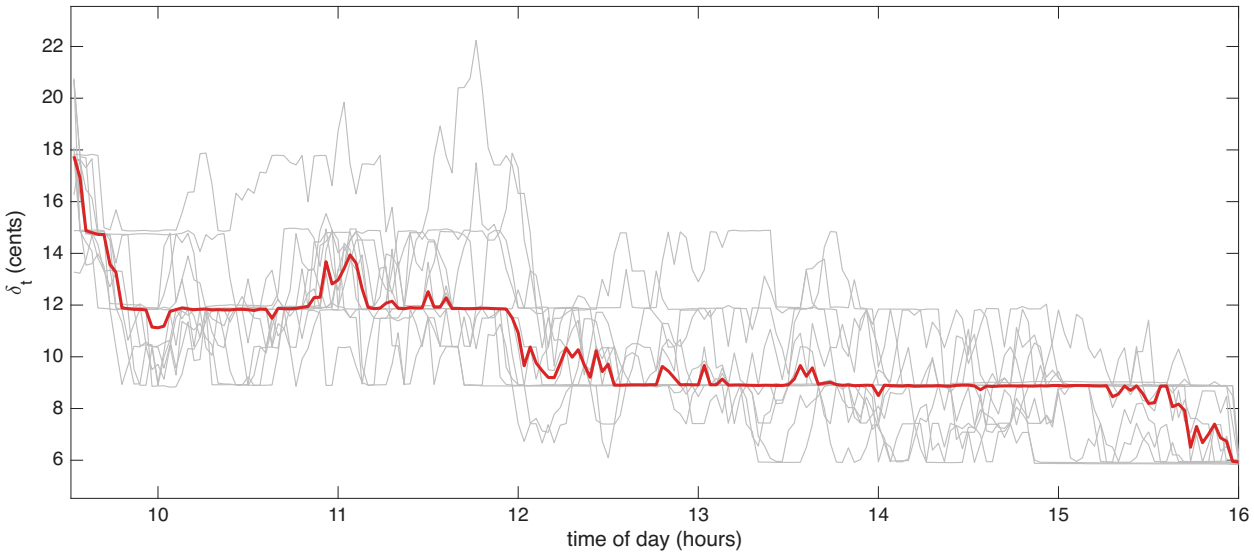

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `META` trading days in Oct 2024 (grey lines) and their median (red line).

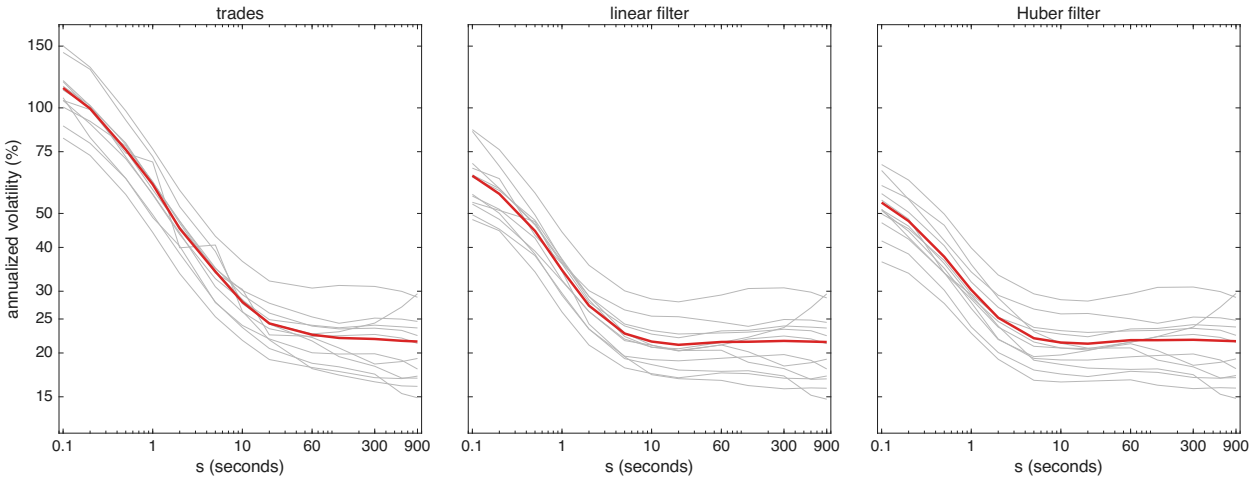

Figure 4: Volatility signature curves for `META` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

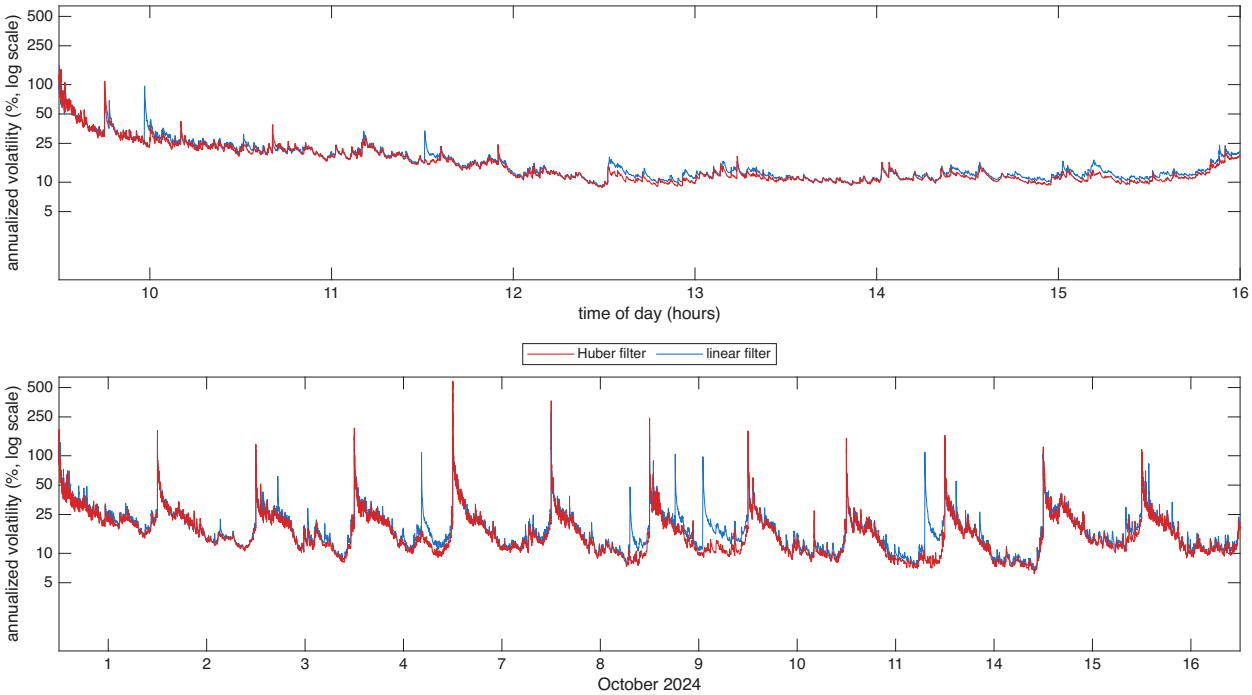

Figure 5: Filtered spot volatility for `META` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 111$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

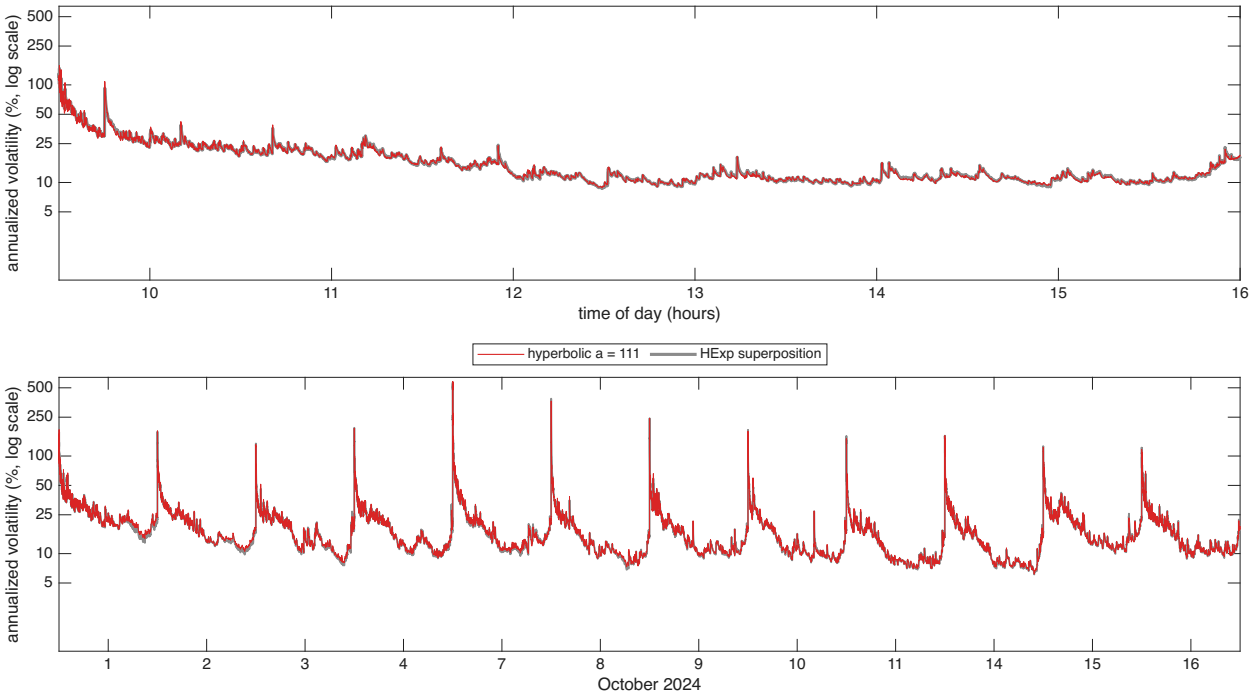

Figure 6: Hyperbolic against superposition weights for `META` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 111$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

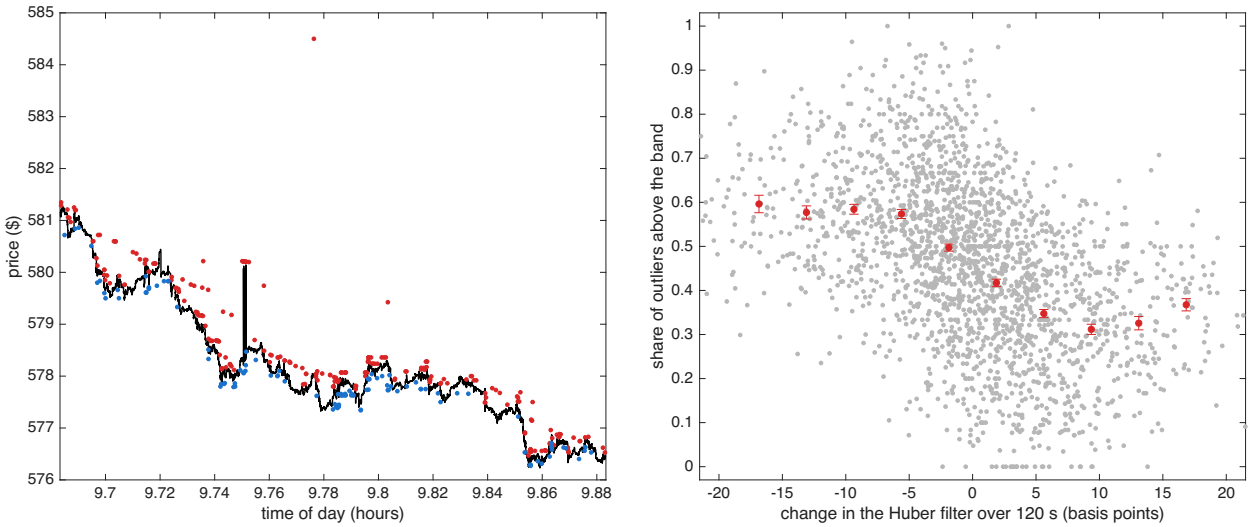

Figure 7: Outliers & robust price for `META` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

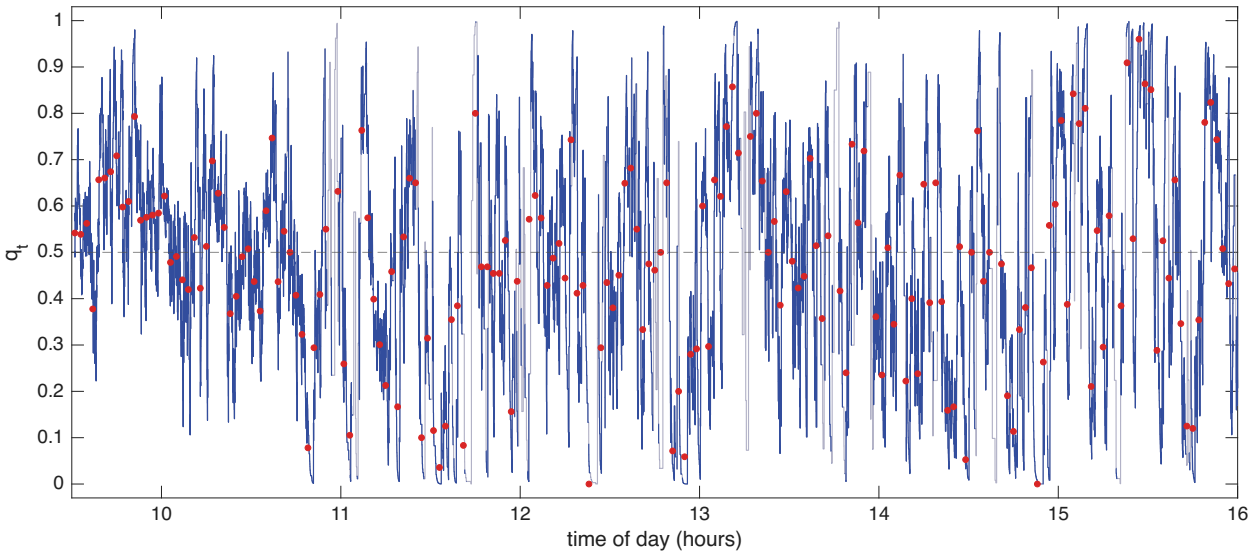

Figure 8: The directional conditional probability for `META`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 14 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

On 7 October 2024, 28 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 16 per cent on any other day. Every number that day contributes to carries that burst.

Diagnostics.

GOOG 1 to 16 October 2024

trades on 16 October 2024	155,257
average trading rate	8.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,439 (0.93%)
price half-life: run / minimizer[1]	3.0 / 1.2 trades
cost of running 3 trades	+11.54%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.03%
direction half-life[3]	82 s [60, 120]
Mandelbrot a minimizing QLIKE	1632 trades
a within 0.0005 of the minimum	1632 to 1632
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	9% $\rightarrow$ 65%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	7% $\rightarrow$ 7%
median spot volatility, Huber / linear input[5]	12.8 / 14.9 %
hyperbolic against HExp, correlation of logs	0.9984
HExp QLIKE advantage, leave-one-day-out[6]	+0.00556
15-minute signature level	14 to 35 %
Huber signature, 1 min over 1 s, 12 day average	1.33

GOOG over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:08; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.147	0.122	0.133	0.458	0.564	0.396	0.575	3.573
2	0.161	0.148	0.165	0.406	0.479	0.353	0.565	2.562
5	0.206	0.192	0.212	0.351	0.439	0.356	0.532	1.655
10	0.243	0.227	0.244	0.331	0.419	0.368	0.493	1.205
25	0.279	0.269	0.276	0.320	0.420	0.397	0.450	0.816
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.28	−0.21	−0.03	−0.49	1.47	1.39	1.89	1.94
2	0.09	0.14	−0.04	0.00	1.68	1.39	1.89	2.05
5	−0.03	−0.02	−0.02	0.00	1.70	1.55	1.90	2.18
10	−0.05	−0.06	−0.01	0.00	1.92	1.88	2.03	2.27
25	0.02	0.02	0.03	0.04	3.68	3.21	2.55	2.62

Table 1: Trade-lag differences, GOOG, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.71	1.58	1.63	2.59	11.1	8.6	20.7	131.9
1 s	1.27	1.20	1.23	1.57	12.2	11.1	19.3	68.2
5 s	1.16	1.13	1.14	1.28	14.6	14.2	16.6	36.3
10 s	1.14	1.12	1.12	1.21	15.2	14.9	16.1	30.6
1 min	1.11	1.11	1.10	1.13	15.2	15.2	15.3	18.6
15 min	1.04	1.04	1.03	1.04	15.9	15.9	15.8	16.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.15	−0.14	−0.08	−0.20	4.91	5.02	4.00	2.73
1 s	−0.05	−0.01	−0.02	−0.25	5.81	5.66	4.43	3.89
5 s	0.00	0.02	−0.02	−0.26	6.88	6.99	5.42	3.03
10 s	−0.05	−0.04	−0.04	−0.12	5.96	5.85	5.11	4.57
1 min	−0.09	−0.09	−0.09	−0.21				

Table 2: Clock-lag differences, GOOG, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.37	0.48	0.51
s.d. of $\hat{\epsilon}_t$ in cents	2.57	2.55	2.18
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	1.89
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.35	0.45	-0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.16	0.33	-0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.06	0.17	-0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.07	-0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.00	0.03
80% C.I. for $\theta_t - X_t$	—	-0.40,0.43	-0.41,0.45
95% C.I. for $\theta_t - X_t$	—	-0.70,0.77	-0.77,0.86
99% C.I. for $\theta_t - X_t$	—	-1.14,1.24	-1.48,1.61
s.e. of $\theta_t - X_t$ in cents	—	0.36	∞(0.66)

Table 3: Residual summaries, GOOG, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. does not exist in theory, the sample value is in brackets.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.53	1.35	1.39 to 1.71	1.24 to 1.65
Residual divided by δ_t	1.48	1.41	1.38 to 1.86	1.29 to 1.74
Repricing windows removed	1.51	1.43	1.43 to 1.75	1.35 to 1.63
Calm windows only	1.49	1.37		
Every 10th trade	1.42	1.35	1.32 to 1.93	1.14 to 1.66
Six-trade half-life	1.75	1.46	1.45 to 1.96	1.31 to 1.79

Table 4: Hill estimator of the tail index of the Huber filter residual, GOOG, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 14.1 to 20.2 cents over the days and removes 1% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 9.3 cents on 16 October 2024, 87% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

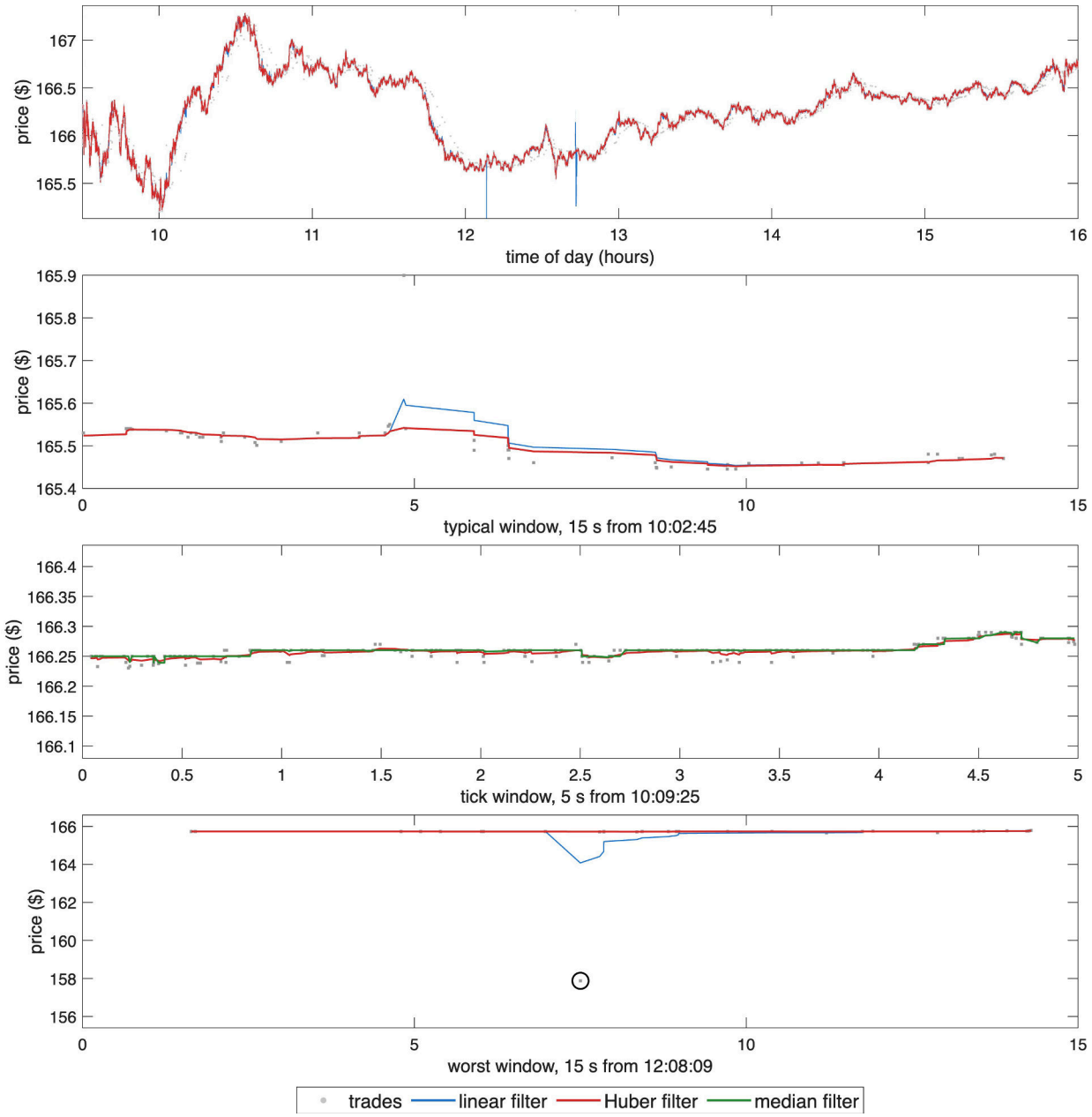

Figure 1: GOOG trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:02:45. 3rd: 5-second window at 10:09:25 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

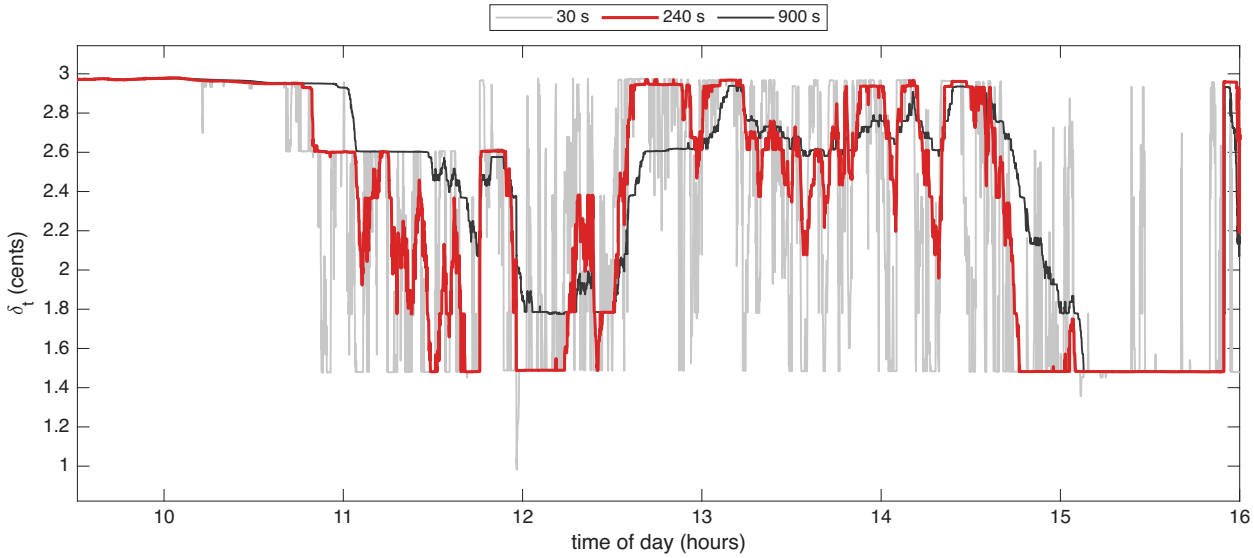

Figure 2: The Huber threshold δ_t in cents for GOOG on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

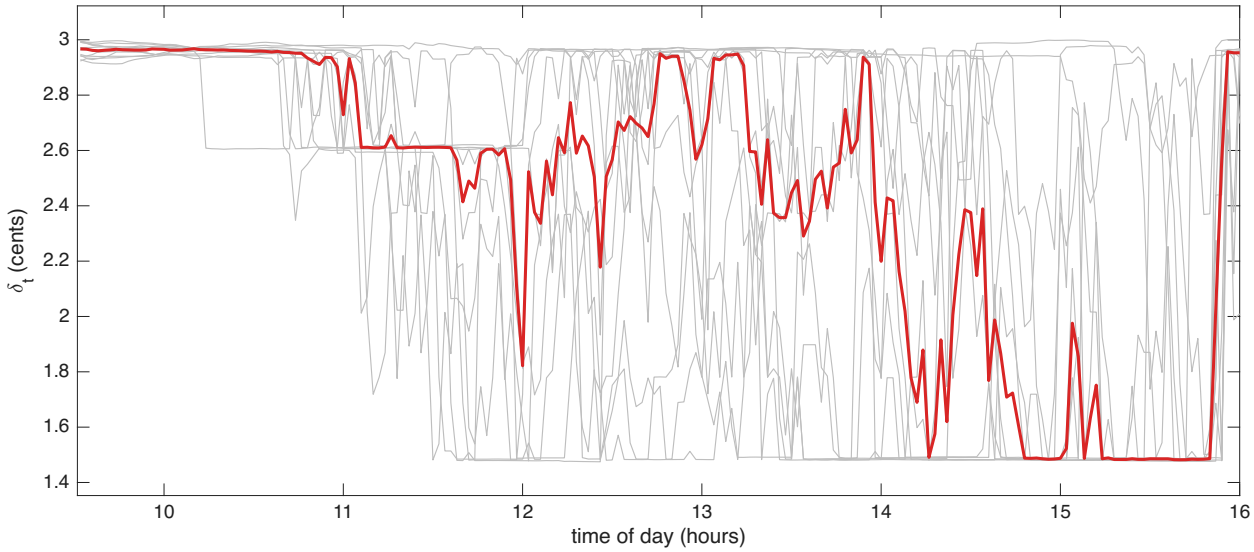

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 GOOG trading days in Oct 2024 (grey lines) and their median (red line).

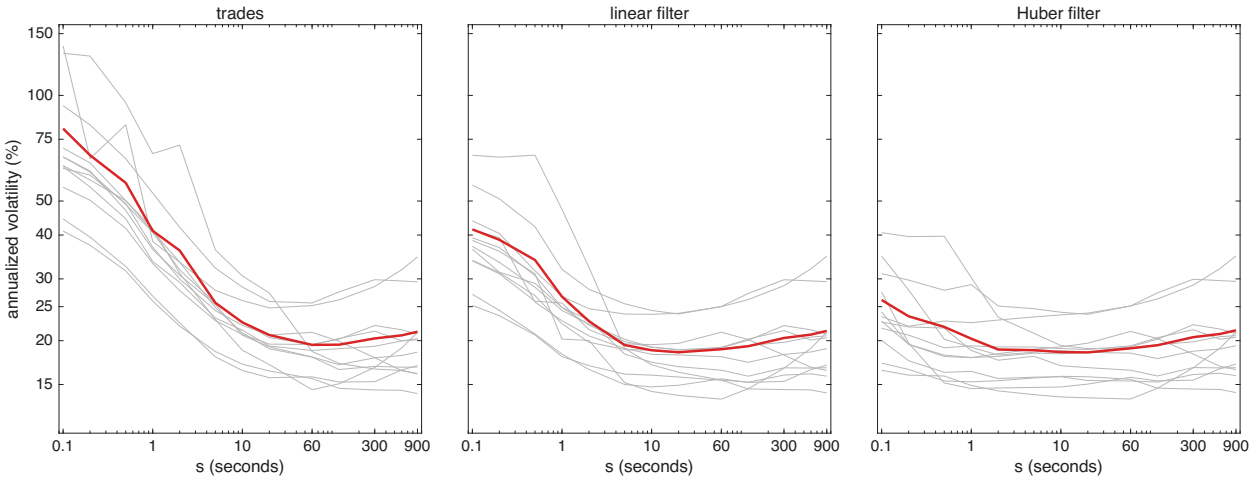

Figure 4: Volatility signature curves for GOOG over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

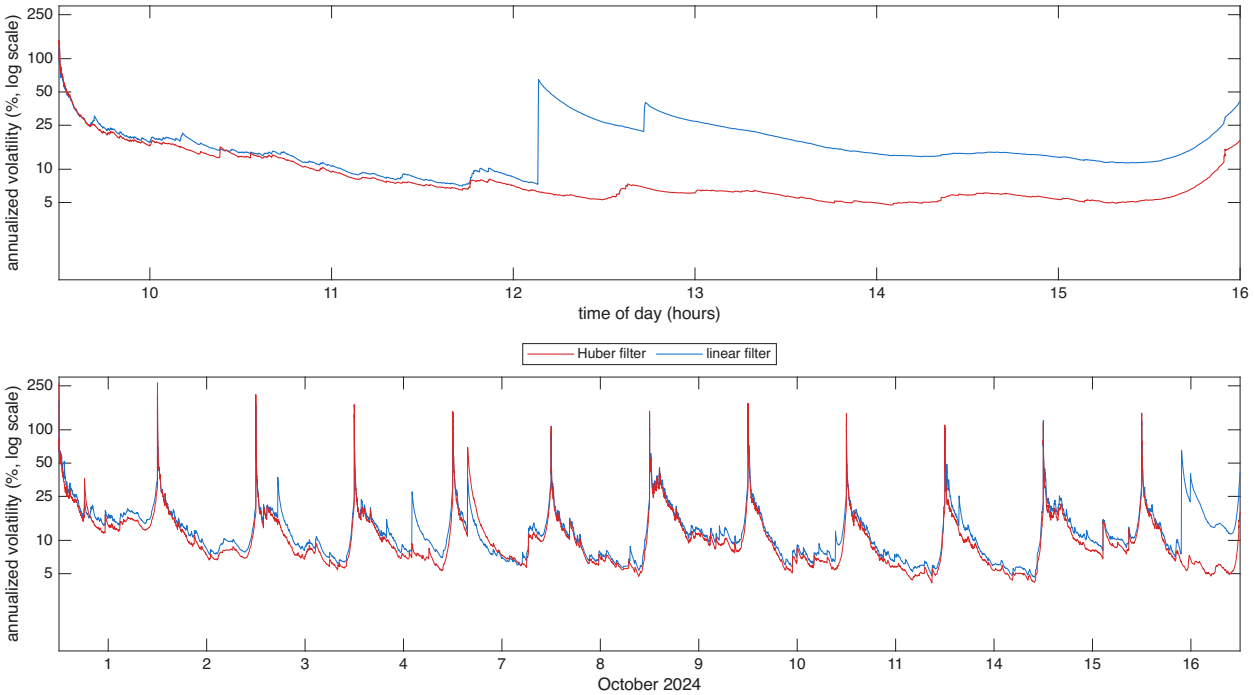

Figure 5: Filtered spot volatility for GOOG over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 1632$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

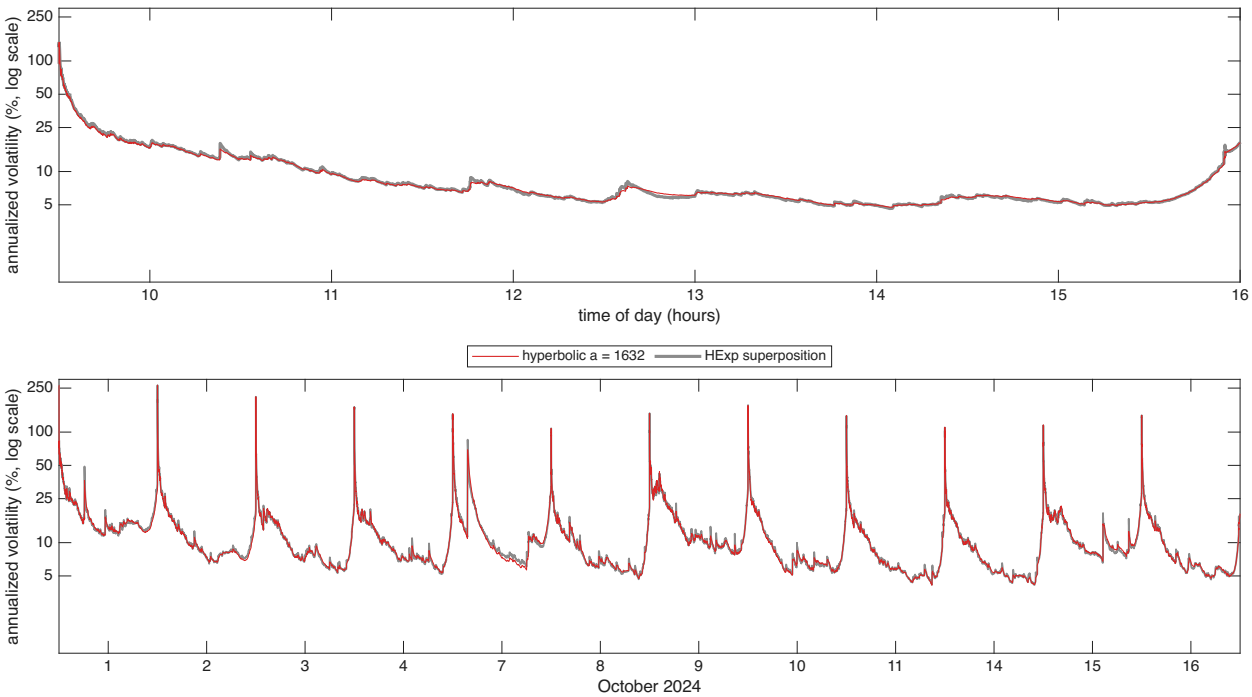

Figure 6: Hyperbolic against superposition weights for `GOOG` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 1632$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

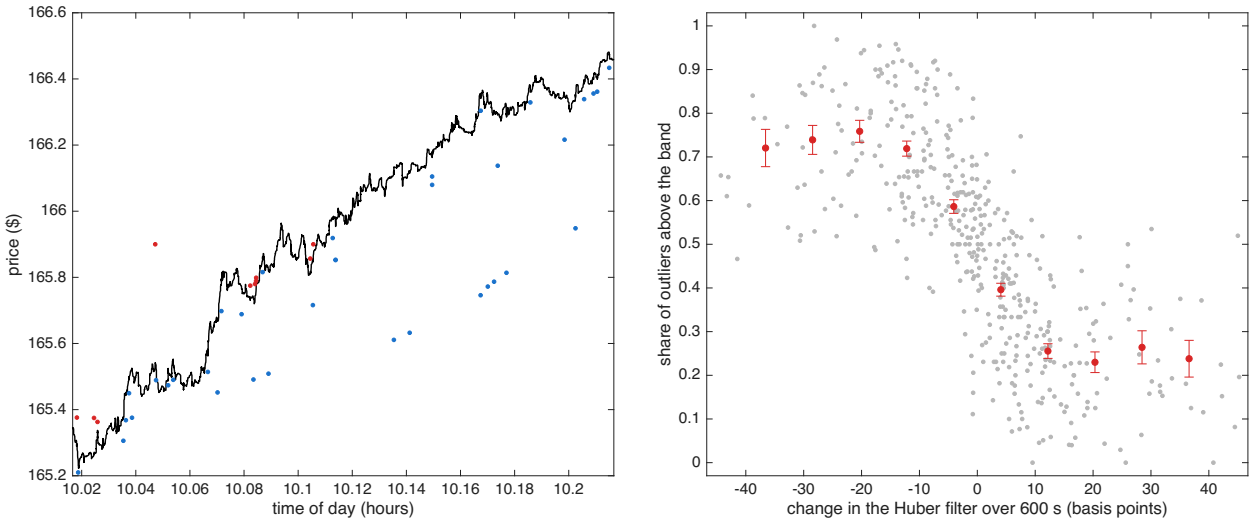

Figure 7: Outliers & robust price for `GOOG` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

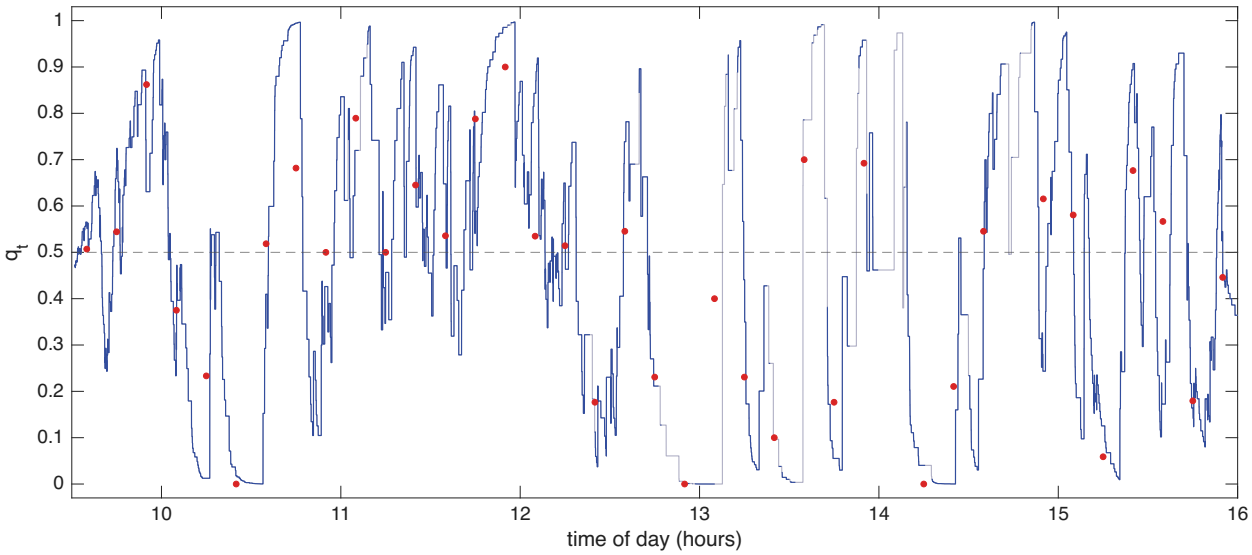

Figure 8: The directional conditional probability for GOOG, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 82 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The common half-life of 3 trades raises the k-trade-ahead Huber forecast loss by 11.5% over the minimizing half-life of 1.2 trades.
At the horizons k = 30 the half-life minimizing the k-trade-ahead Huber forecast loss is 1.1 trades, within one grid point of the end of the 1 to 5 trade grid searched.
The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 23 to 26112 trades, against 1632 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 82 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

Diagnostics.

COIN 1 to 16 October 2024

trades on 16 October 2024	178,823
average trading rate	5.8 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	8,986 (5.03%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.28%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	20 s [16, 25]
Mandelbrot a minimizing QLIKE	73 trades
a within 0.0005 of the minimum	47 to 108
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	117% $\to$ 130%
Huber $\hat{\sigma}$, before $\to$ peak	115% $\to$ 130%
median spot volatility, Huber / linear input[5]	44.5 / 48.4 %
hyperbolic against HExp, correlation of logs	0.9968
HExp QLIKE advantage, leave-one-day-out[6]	+0.00123
15-minute signature level	44 to 100 %
Huber signature, 1 min over 1 s, 12 day average	1.25

COIN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:46; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.100	0.867	0.890	4.161	2.429	1.416	1.507	8.845
2	1.037	0.942	1.005	3.379	2.164	1.485	1.583	6.589
5	1.167	1.045	1.126	2.499	1.994	1.555	1.655	4.450
10	1.218	1.106	1.171	2.031	1.844	1.579	1.658	3.348
25	1.264	1.188	1.216	1.656	1.751	1.627	1.657	2.472
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.10	0.10	−0.45	3.41	2.74	2.92	3.33
2	0.05	0.03	0.03	−0.02	3.34	2.80	3.25	3.37
5	−0.02	0.01	−0.01	0.00	3.58	3.25	3.79	3.48
10	0.00	0.00	0.00	0.01	4.05	3.88	4.26	3.50
25	0.00	0.00	0.00	0.00	5.26	5.17	5.71	3.65

Table 1: Trade-lag differences, `COIN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	10.22	9.45	9.79	17.88	57.9	40.2	44.8	230.8
1 s	6.60	6.22	6.41	9.38	56.8	49.9	52.6	135.2
5 s	5.57	5.40	5.45	6.60	58.8	56.8	57.1	83.2
10 s	5.41	5.30	5.30	6.02	61.5	60.4	60.0	73.7
1 min	5.15	5.12	5.10	5.25	63.8	63.5	63.0	65.3
15 min	5.43	5.43	5.43	5.45	68.9	68.9	68.9	69.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.02	0.00	−0.22	4.01	4.13	4.45	6.58
1 s	0.04	0.11	0.07	−0.30	5.89	5.98	6.58	4.26
5 s	0.14	0.17	0.16	−0.08	9.87	9.50	10.23	6.32
10 s	0.08	0.09	0.09	−0.08	7.87	8.66	9.43	11.67
1 min	−0.06	−0.05	−0.05	−0.07				

Table 2: Clock-lag differences, `COIN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.73	3.26	3.42
s.d. of $\hat{\epsilon}_t$ in cents	6.30	6.09	5.79
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.91
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.19	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.09	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.02	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.01	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−2.18,2.11	−2.40,2.34
95% C.I. for $\theta_t - X_t$	—	−3.78,3.66	−4.31,4.13
99% C.I. for $\theta_t - X_t$	—	−5.48,5.55	−6.91,6.62
s.e. of $\theta_t - X_t$ in cents	—	1.74	2.14

Table 3: Residual summaries, `COIN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.91, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.25	2.38	1.90 to 2.98	1.93 to 2.69
Residual divided by δ_t	1.96	2.17	1.85 to 2.73	1.89 to 2.64
Repricing windows removed	2.26	2.42	1.89 to 2.97	2.00 to 2.68
Calm windows only	2.23	2.34		
Every 10th trade	1.90	2.28	1.85 to 2.61	1.64 to 3.04
Six-trade half-life	2.34	2.49	1.98 to 3.29	1.98 to 2.85

Table 4: Hill estimator of the tail index of the Huber filter residual, `COIN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 40.5 to 67.2 cents over the days and removes 0% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 30.2 cents on 16 October 2024, 68% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

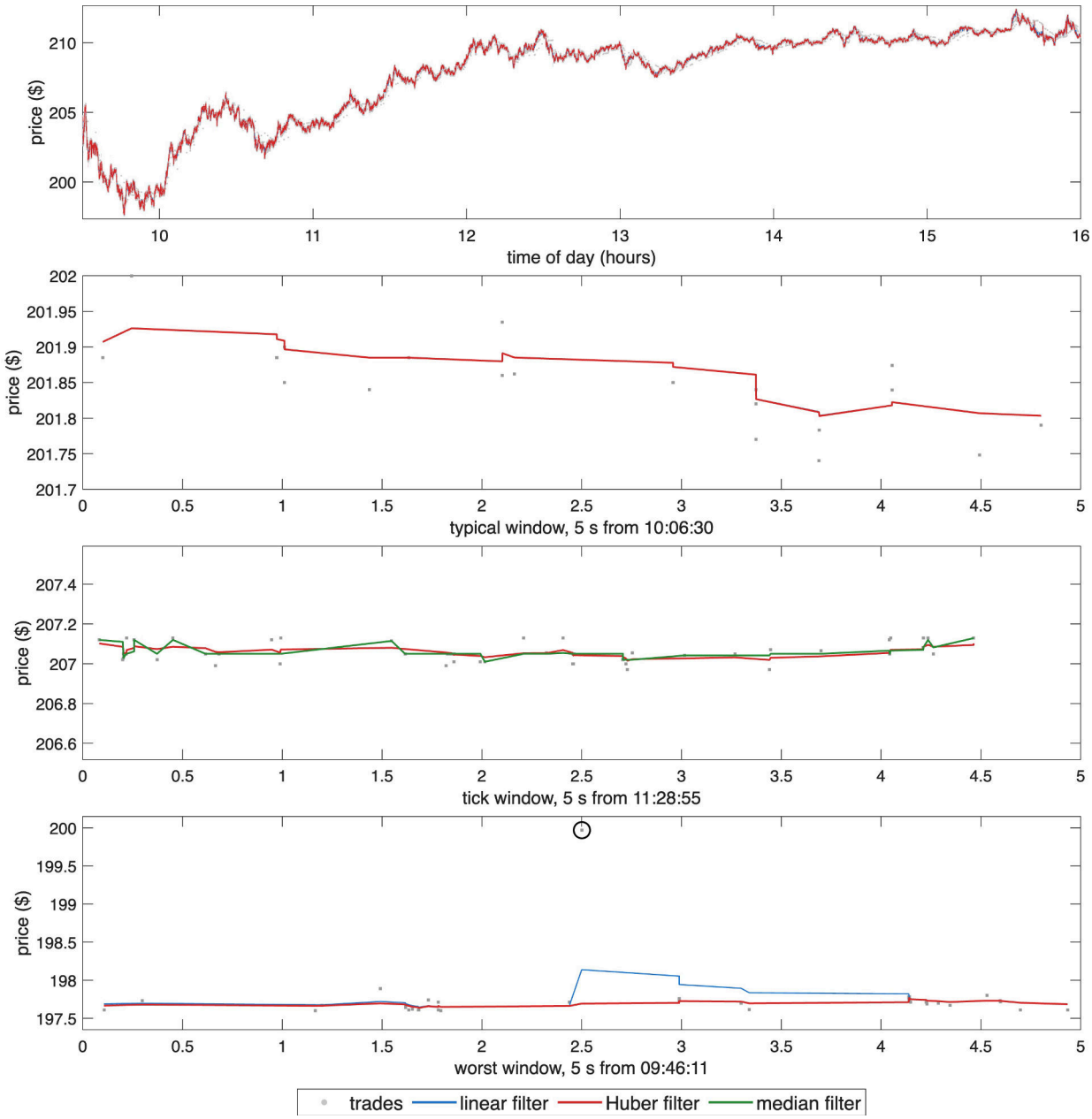

Figure 1: `COIN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:06:30. 3rd: 5-second window at 11:28:55 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

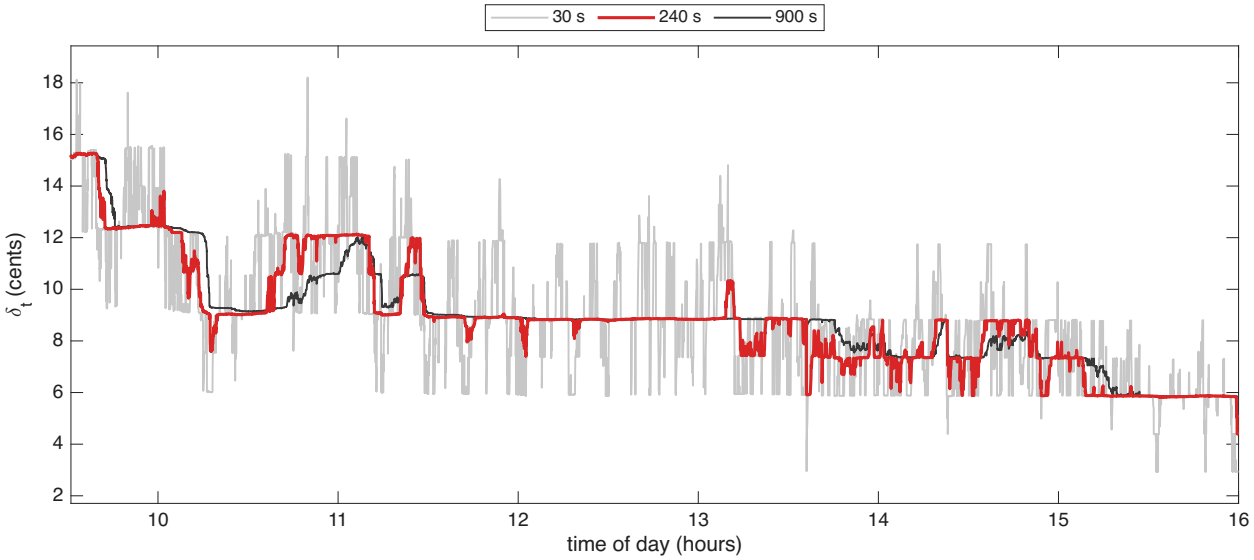

Figure 2: The Huber threshold δ_t in cents for `COIN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

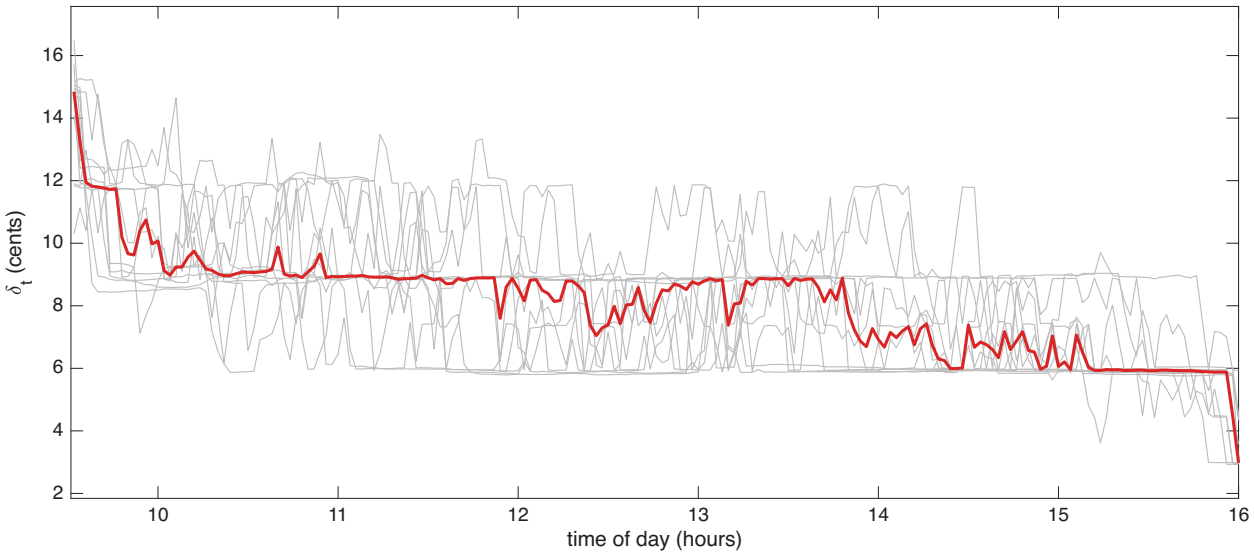

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `COIN` trading days in Oct 2024 (grey lines) and their median (red line).

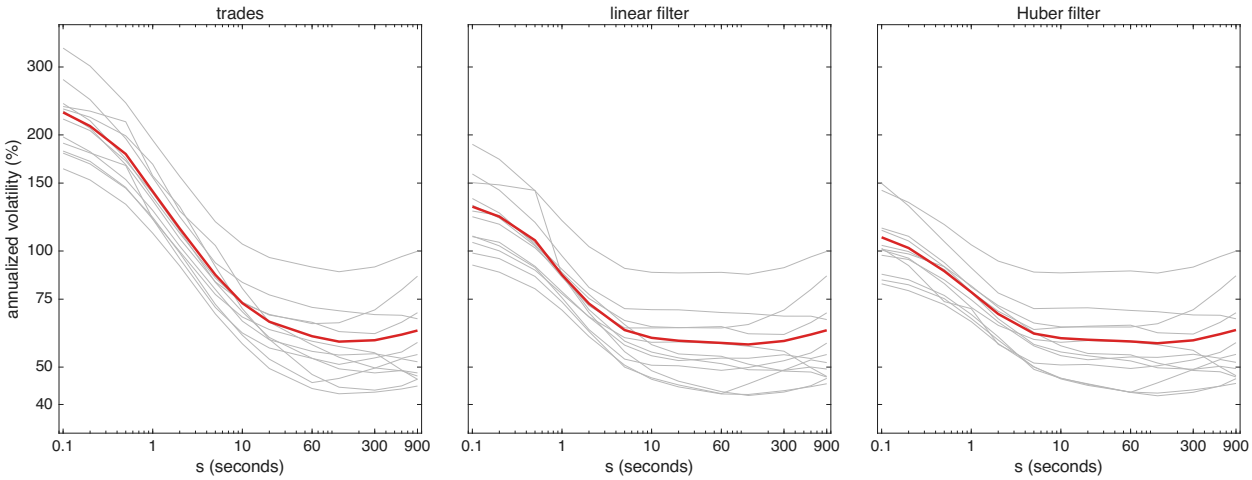

Figure 4: Volatility signature curves for `COIN` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

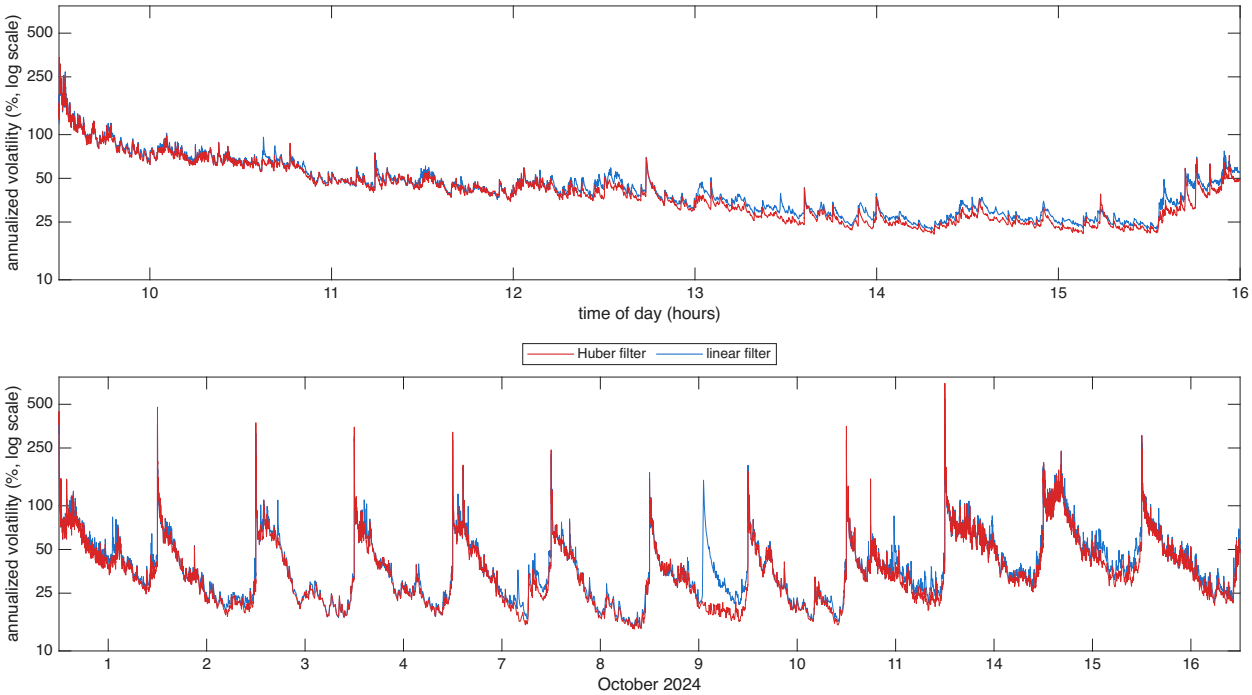

Figure 5: Filtered spot volatility for `COIN` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 73$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

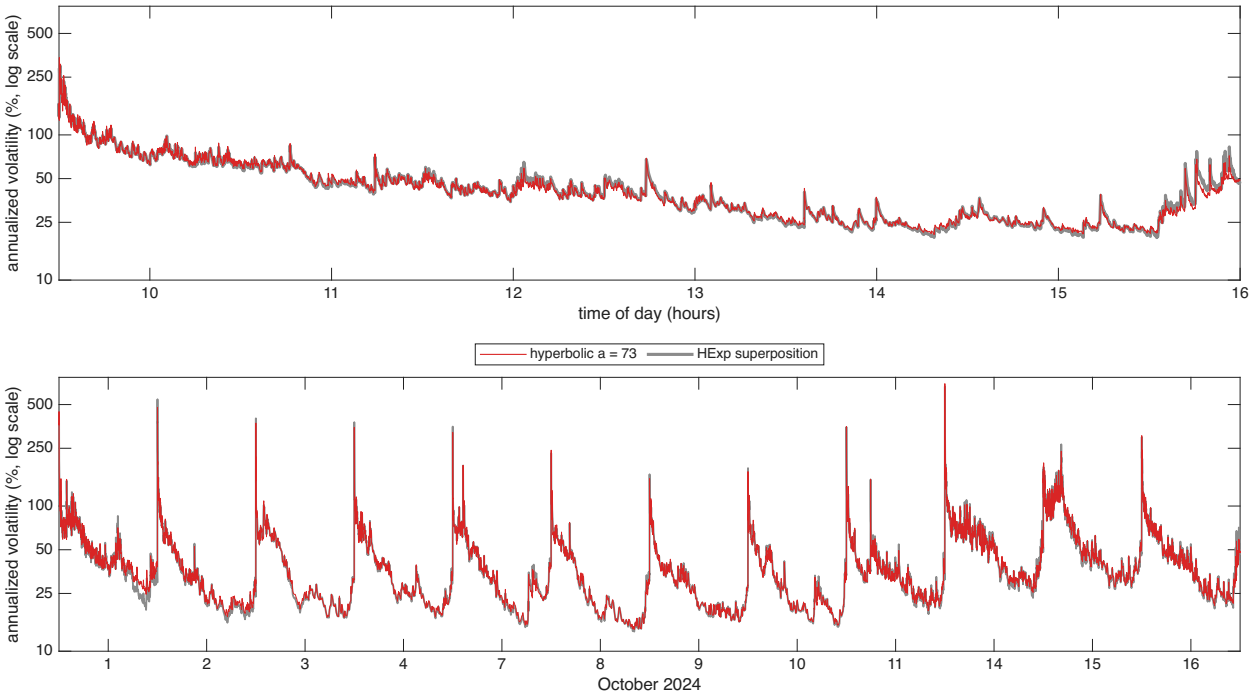

Figure 6: Hyperbolic against superposition weights for `COIN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 73$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

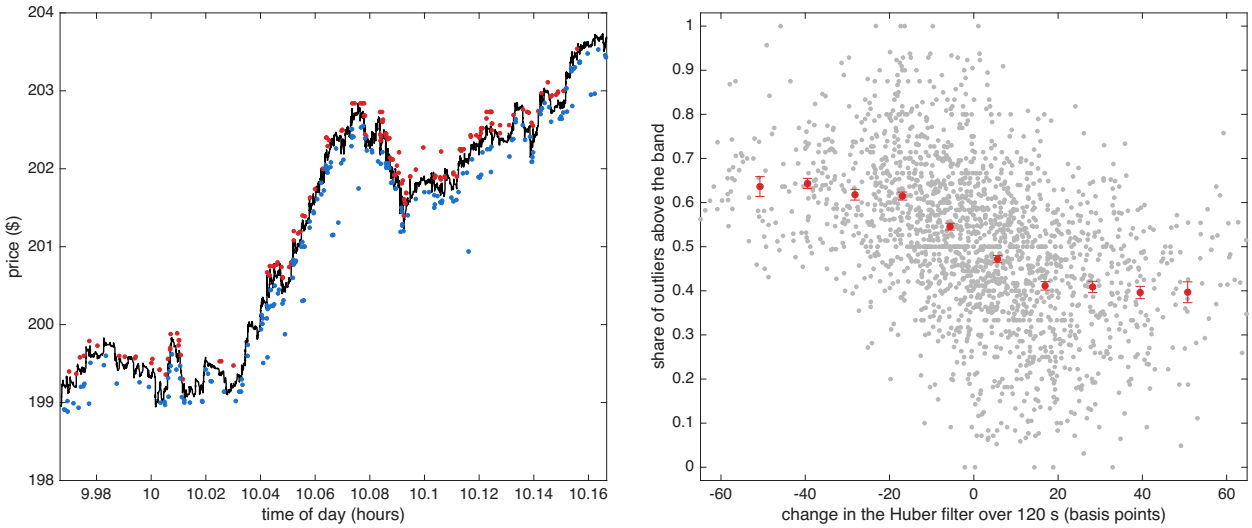

Figure 7: Outliers & robust price for `COIN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

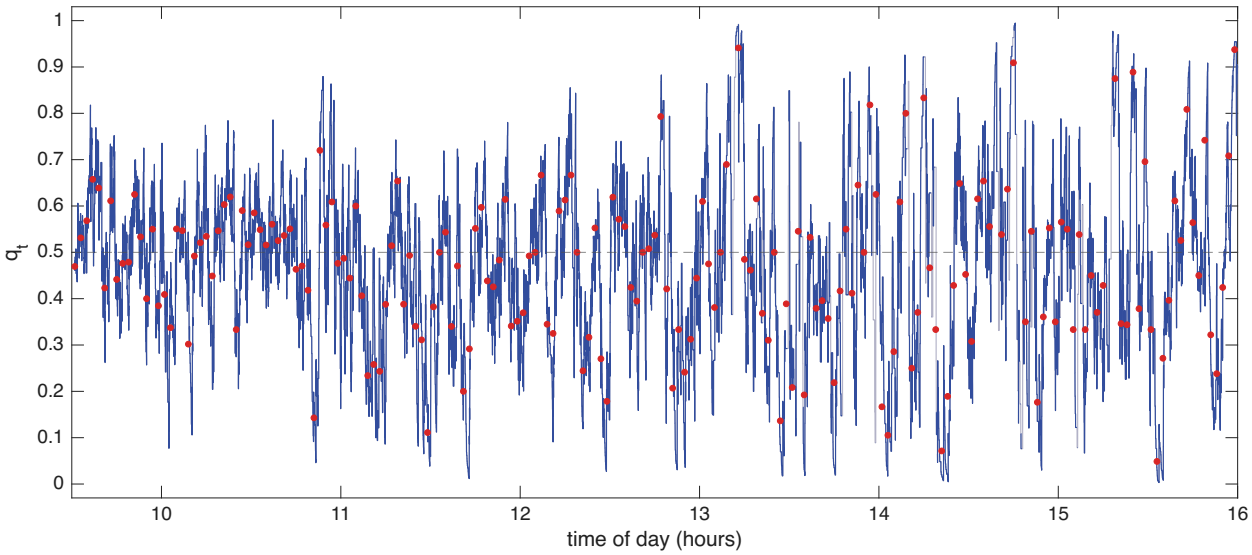

Figure 8: The directional conditional probability for COIN, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 20 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

The largest Huber residual on 16 October 2024 occurs at 09:32, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:46.

Diagnostics.

MRVL 1 to 16 October 2024

trades on 16 October 2024	178,140
average trading rate	5.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	3,745 (2.10%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+3.90%
scale half-life: run / minimizer[2]	240 / 60 s
cost of running 240 s	+0.07%
direction half-life[3]	31 s [25, 39]
Mandelbrot a minimizing QLIKE	93 trades
a within 0.0005 of the minimum	60 to 144
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	34% $\to$ 422%
Huber $\hat{\sigma}$, before $\to$ peak	17% $\to$ 31%
median spot volatility, Huber / linear input[5]	23.1 / 29.3 %
hyperbolic against HExp, correlation of logs	0.9981
HExp QLIKE advantage, leave-one-day-out[6]	-0.00099
15-minute signature level	25 to 76 %
Huber signature, 1 min over 1 s, 12 day average	1.41

MRVL over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:53; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.185	0.157	0.187	0.726	0.648	0.436	0.860	5.231
2	0.199	0.187	0.228	0.616	0.598	0.445	0.838	3.743
5	0.249	0.233	0.279	0.501	0.541	0.444	0.773	2.416
10	0.287	0.269	0.307	0.445	0.502	0.442	0.686	1.739
25	0.328	0.314	0.332	0.407	0.486	0.457	0.574	1.157
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.15	0.04	−0.05	−0.49	1.73	1.40	2.28	2.75
2	−0.03	−0.03	−0.05	−0.01	1.76	1.46	2.35	2.78
5	−0.01	0.00	−0.04	0.00	2.17	1.88	2.24	2.88
10	0.00	0.01	−0.01	0.00	2.64	2.43	2.25	2.90
25	0.00	0.00	0.00	0.00	3.48	3.51	2.74	2.96

Table 1: Trade-lag differences, `MRVL`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.79	1.68	1.85	3.31	32.9	25.7	42.8	216.3
1 s	1.30	1.24	1.36	2.03	33.4	30.1	45.0	133.4
5 s	1.17	1.14	1.19	1.48	38.8	37.6	45.2	75.5
10 s	1.14	1.12	1.14	1.33	40.3	39.5	41.1	60.8
1 min	1.17	1.16	1.16	1.21	41.9	41.6	41.1	44.8
15 min	1.04	1.04	1.04	1.04	38.0	38.0	37.8	38.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.05	−0.04	0.00	−0.13	3.92	3.72	3.58	3.04
1 s	−0.10	−0.06	−0.10	−0.38	4.18	4.07	3.40	3.13
5 s	−0.02	0.01	−0.04	−0.27	6.09	6.17	4.29	3.44
10 s	0.11	0.13	0.11	−0.13	6.59	6.95	4.43	4.21
1 min	−0.04	−0.04	−0.03	−0.07				

Table 2: Clock-lag differences, `MRVL`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.53	0.66	0.72
s.d. of $\hat{\epsilon}_t$ in cents	3.76	3.72	3.17
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.26
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.29	0.35	−0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.12	0.23	−0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.11	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.06	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.49,0.50	−0.54,0.57
95% C.I. for $\theta_t - X_t$	—	−0.91,0.93	−1.12,1.13
99% C.I. for $\theta_t - X_t$	—	−1.41,1.48	−2.14,2.11
s.e. of $\theta_t - X_t$ in cents	—	0.43	0.92

Table 3: Residual summaries, `MRVL`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.26, is above two, while the Hill index of the Huber residual itself on this day is 1.94, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.08	1.94	1.75 to 2.33	1.78 to 2.51
Residual divided by δ_t	1.96	2.11	1.74 to 2.65	1.83 to 2.64
Repricing windows removed	2.11	1.96	1.75 to 2.33	1.79 to 2.50
Calm windows only	1.92	1.93		
Every 10th trade	1.96	1.90	1.60 to 2.12	1.46 to 2.29
Six-trade half-life	2.24	2.05	1.85 to 2.51	1.93 to 2.67

Table 4: Hill estimator of the tail index of the Huber filter residual, `MRVL`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 15.5 to 20.2 cents over the days and removes 0% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.1 cents on 16 October 2024, 84% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

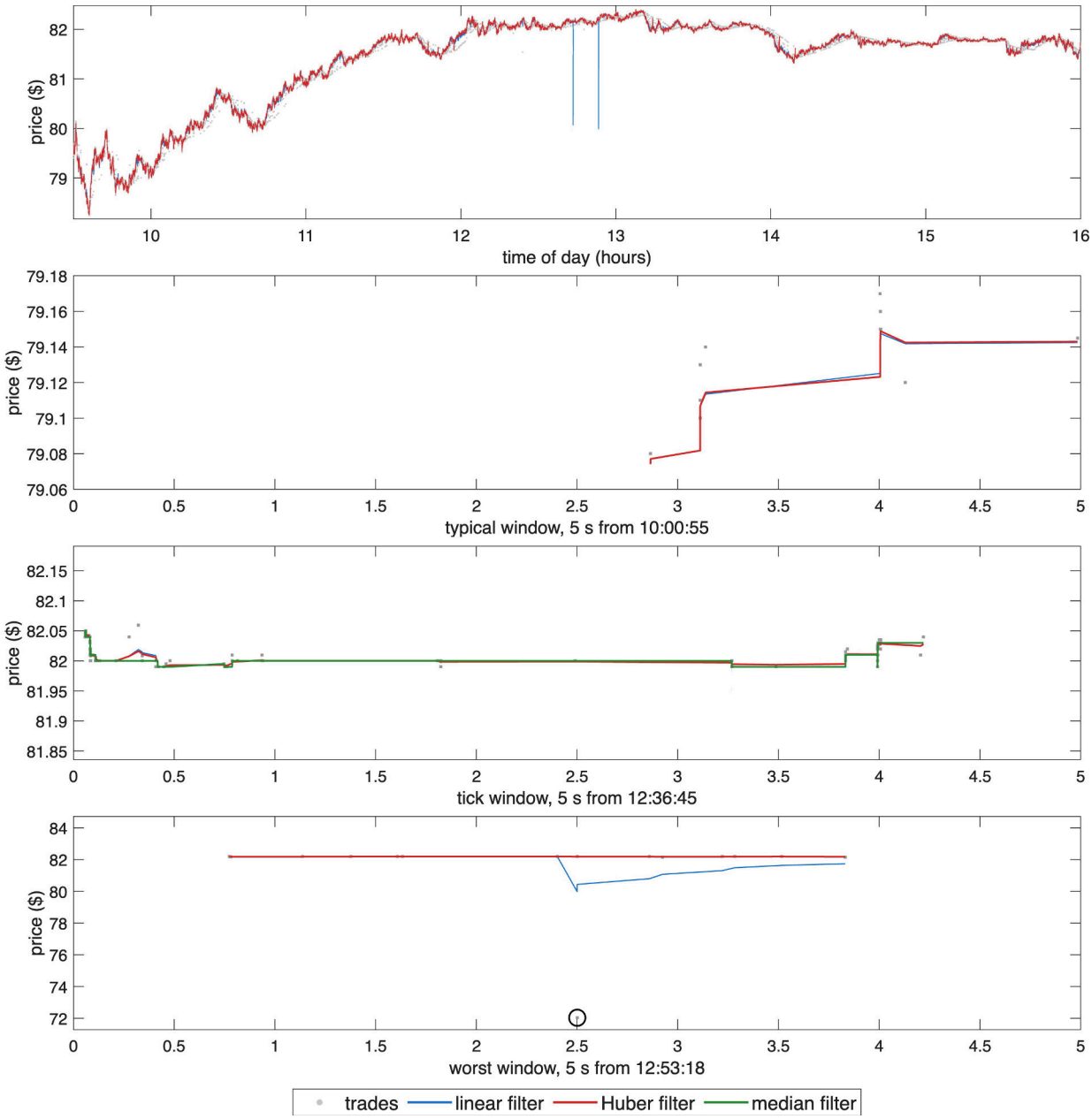

Figure 1: `MRVL` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:00:55. 3rd: 5-second window at 12:36:45 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

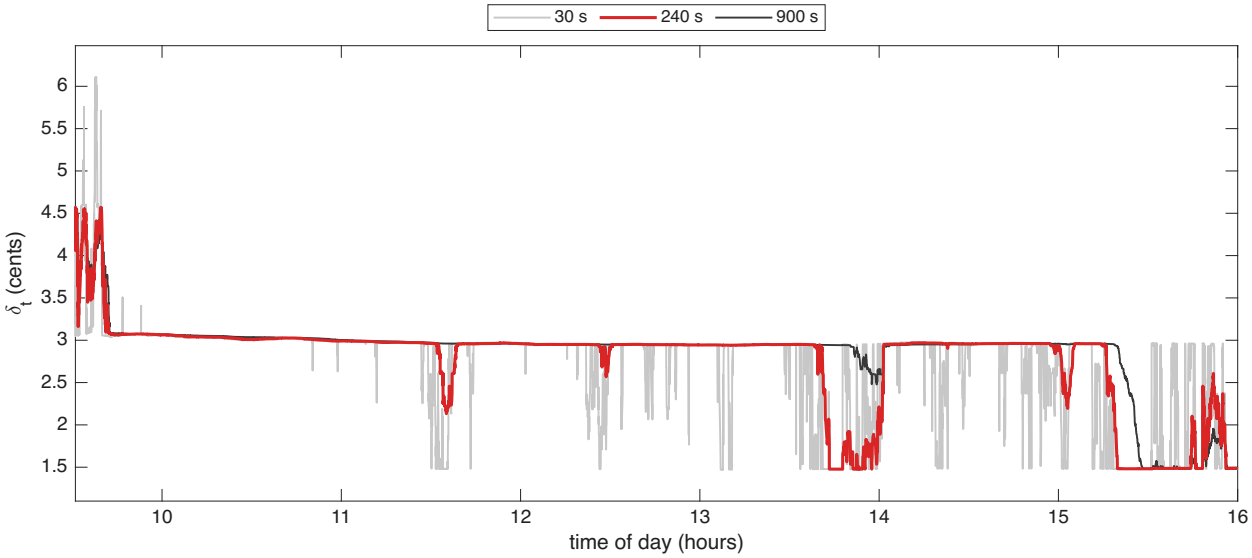

Figure 2: The Huber threshold δ_t in cents for `MRVL` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

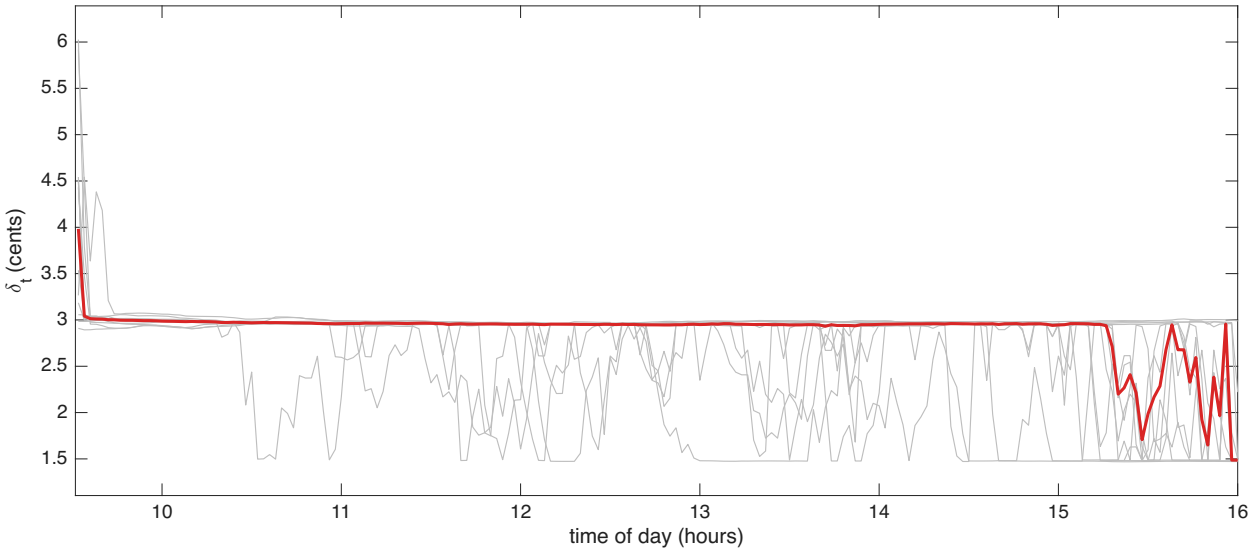

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `MRVL` trading days in Oct 2024 (grey lines) and their median (red line).

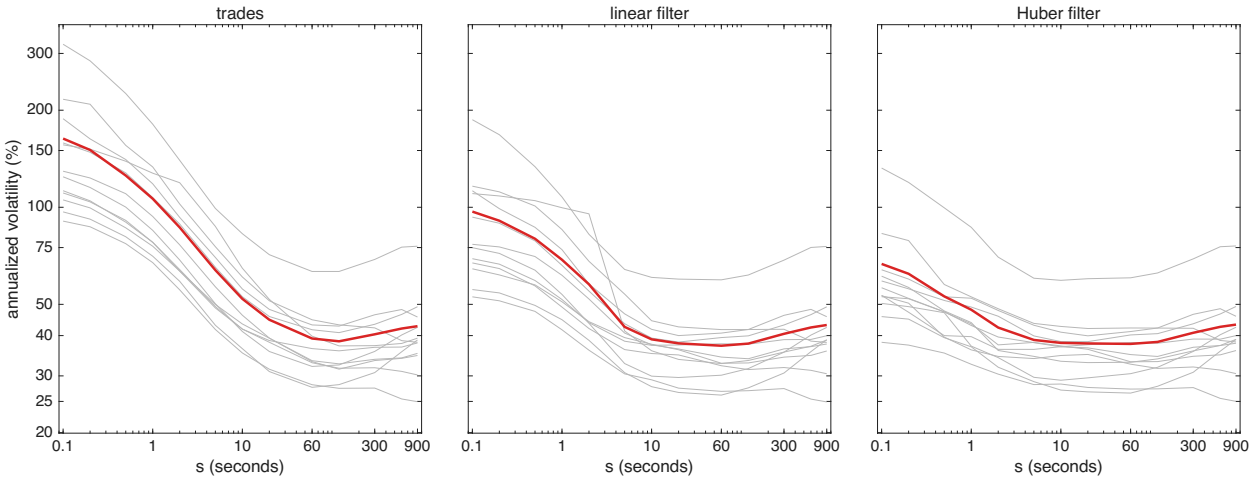

Figure 4: Volatility signature curves for `MRVL` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

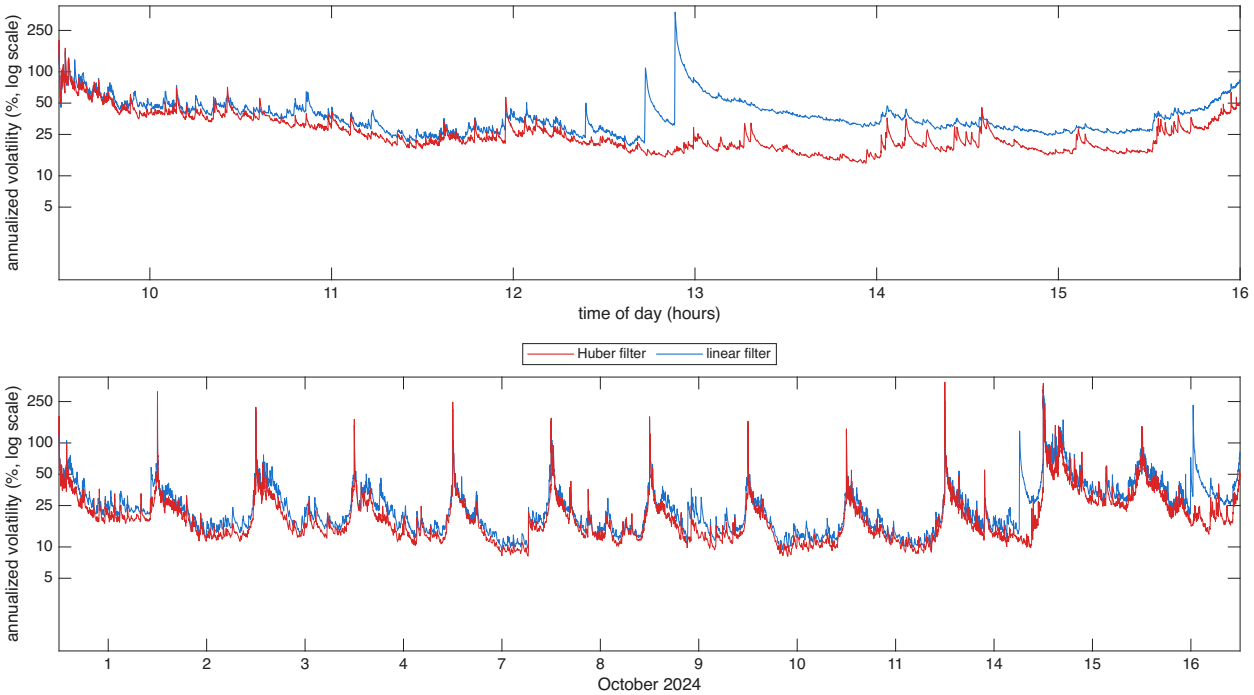

Figure 5: Filtered spot volatility for `MRVL` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 93$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

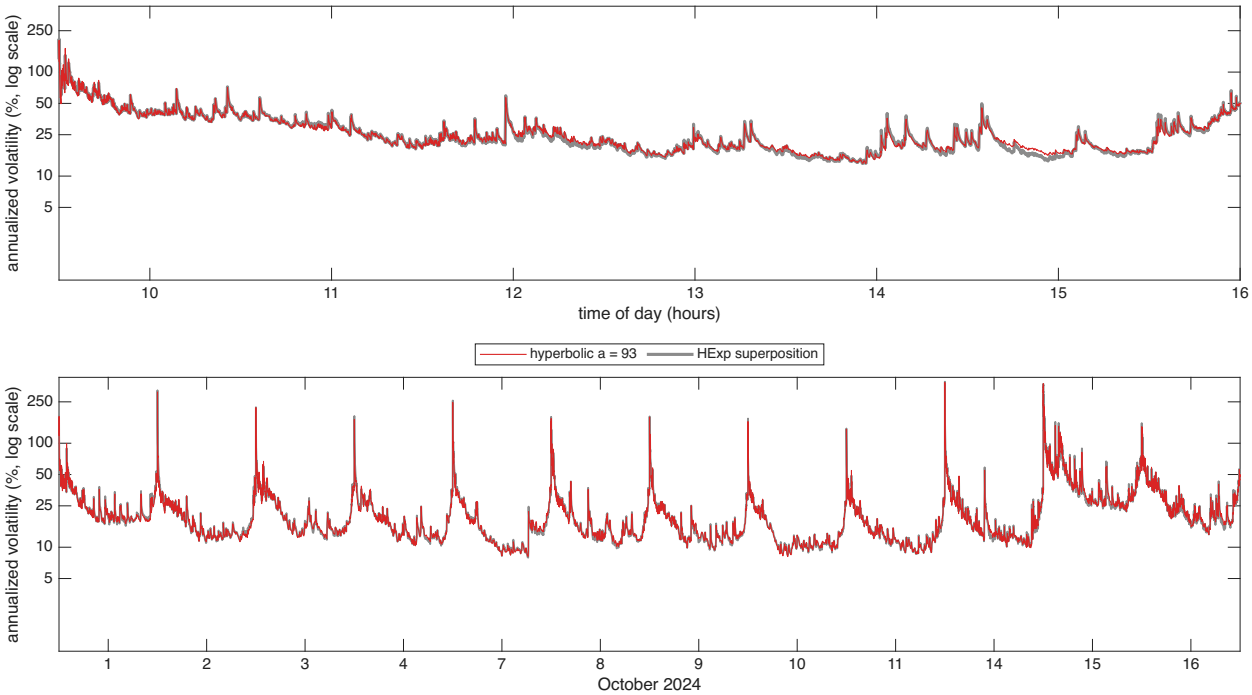

Figure 6: Hyperbolic against superposition weights for `MRVL` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 93$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

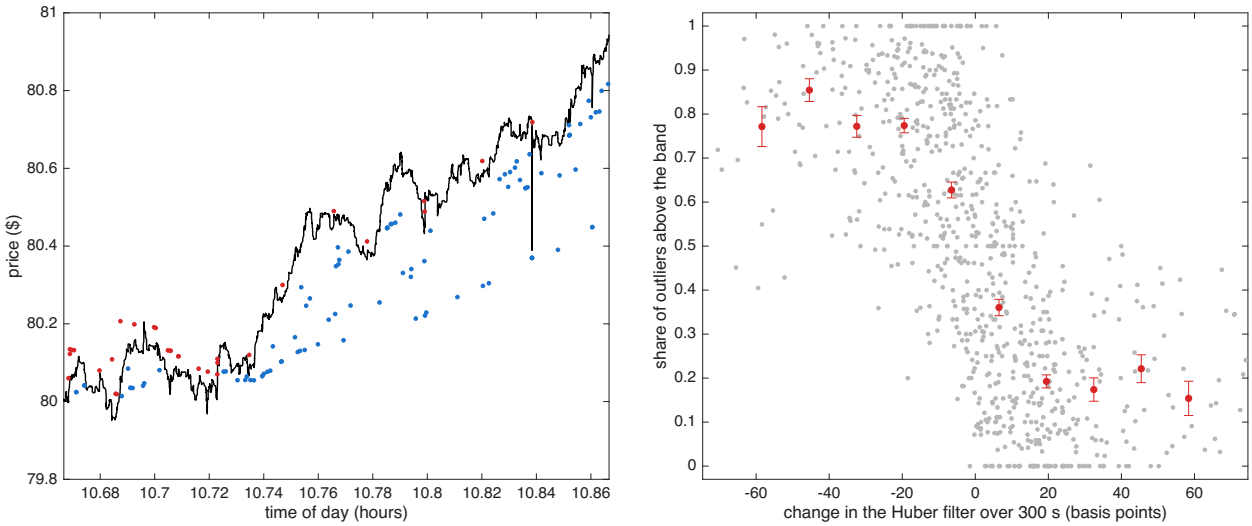

Figure 7: Outliers & robust price for `MRVL` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

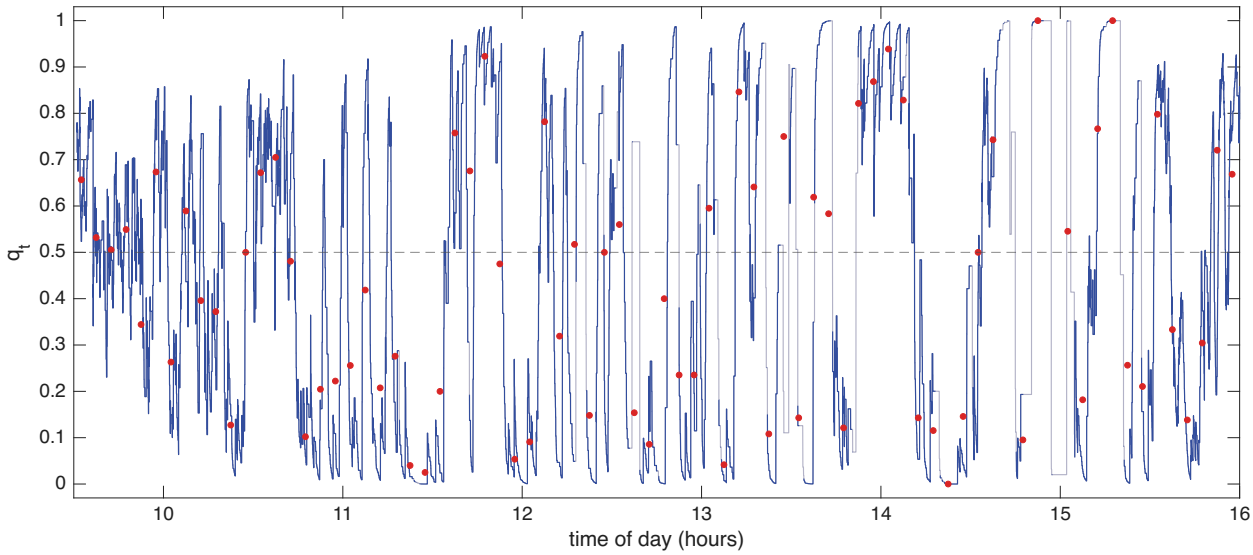

Figure 8: The directional conditional probability for MRVL, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 31 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

RIVN 1 to 16 October 2024

trades on 16 October 2024	65,638
average trading rate	4.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	451 (0.69%)
price half-life: run / minimizer[1]	3.0 / 3.2 trades
cost of running 3 trades	+0.19%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.00%
direction half-life[3]	141 s [107, 182]
Mandelbrot a minimizing QLIKE	52 trades
a within 0.0005 of the minimum	39 to 73
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	19% $\to$ 37%
Huber $\hat{\sigma}$, before $\to$ peak	18% $\to$ 36%
median spot volatility, Huber / linear input[5]	38.7 / 43.2 %
hyperbolic against HExp, correlation of logs	0.9979
HExp QLIKE advantage, leave-one-day-out[6]	-0.00416
15-minute signature level	29 to 114 %
Huber signature, 1 min over 1 s, 12 day average	1.22

RIVN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:28; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.059	0.045	0.045	0.222	0.155	0.072	0.075	0.442
2	0.052	0.049	0.049	0.176	0.130	0.076	0.078	0.333
5	0.055	0.051	0.052	0.125	0.112	0.077	0.079	0.223
10	0.053	0.050	0.051	0.097	0.097	0.074	0.076	0.164
25	0.050	0.048	0.049	0.074	0.081	0.070	0.070	0.115
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.29	0.10	0.09	−0.43	3.60	6.81	5.31	6.50
2	0.05	0.00	−0.01	−0.05	3.85	4.22	4.63	6.71
5	−0.02	−0.04	−0.04	−0.02	4.42	4.48	4.88	6.65
10	−0.01	−0.01	−0.01	0.00	4.34	5.16	5.49	6.90
25	0.00	0.01	0.01	0.00	6.31	5.45	5.50	7.02

Table 1: Trade-lag differences, `RIVN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.36	0.33	0.33	0.65	47.6	25.7	26.8	156.1
1 s	0.23	0.22	0.22	0.38	40.0	26.9	27.7	113.9
5 s	0.16	0.15	0.16	0.22	35.2	29.7	30.2	66.4
10 s	0.15	0.14	0.14	0.19	33.3	29.8	30.0	51.3
1 min	0.13	0.13	0.13	0.14	31.9	30.6	30.5	35.1
15 min	0.12	0.12	0.12	0.12	29.1	28.9	28.8	29.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.06	0.07	0.07	−0.09	6.84	7.19	7.14	9.64
1 s	−0.20	−0.12	−0.12	−0.38	8.42	8.91	8.73	9.35
5 s	−0.06	0.04	0.03	−0.29	7.89	7.44	7.10	9.38
10 s	0.10	0.15	0.14	−0.16	7.41	8.19	8.15	11.50
1 min	−0.11	−0.09	−0.09	−0.14				

Table 2: Clock-lag differences, `RIVN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.14	0.17	0.17
s.d. of $\hat{\epsilon}_t$ in cents	0.32	0.29	0.29
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	5.36
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.09	0.10	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.01	0.00	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	−0.02	−0.03	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	−0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.01	0.01
80% C.I. for $\theta_t - X_t$	—	−0.11,0.11	−0.11,0.11
95% C.I. for $\theta_t - X_t$	—	−0.20,0.20	−0.20,0.21
99% C.I. for $\theta_t - X_t$	—	−0.32,0.31	−0.33,0.32
s.e. of $\theta_t - X_t$ in cents	—	0.10	0.10

Table 3: Residual summaries, `RIVN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 5.36, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.05	3.85	1.83 to 3.05	1.96 to 3.85
Residual divided by δ_t	2.95	3.99	1.83 to 3.15	1.88 to 3.99
Repricing windows removed	3.05	3.85	1.83 to 3.05	1.93 to 3.85
Calm windows only	3.01	3.81		
Every 10th trade	2.31	2.71	1.83 to 2.79	1.58 to 2.71
Six-trade half-life	3.30	4.14	1.88 to 3.30	2.13 to 4.14

Table 4: Hill estimator of the tail index of the Huber filter residual, `RIVN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 7.7 to 9.9 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 3.9 cents on 16 October 2024, 99% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

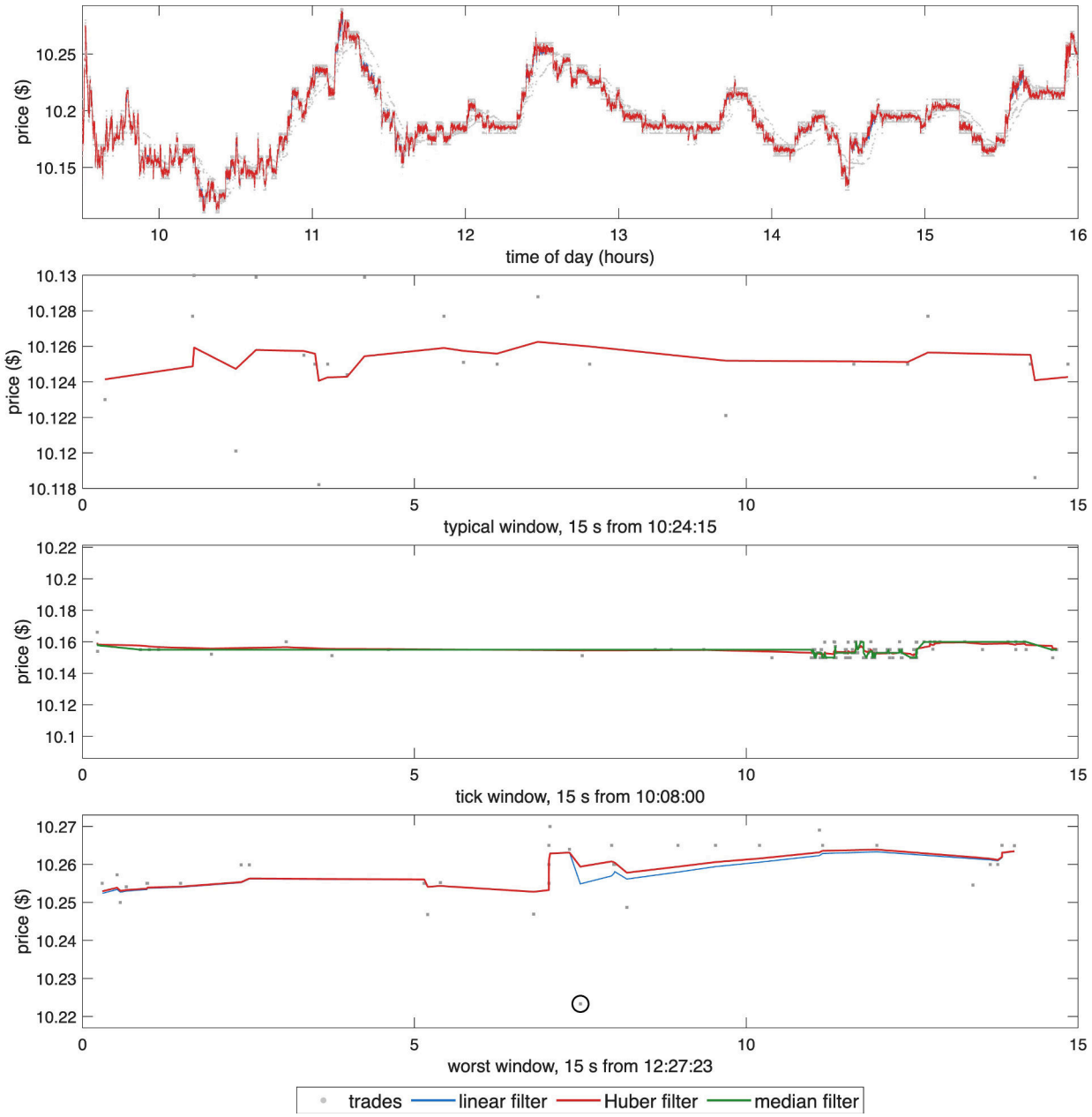

Figure 1: `RIVN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:24:15. 3rd: 15-second window at 10:08:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

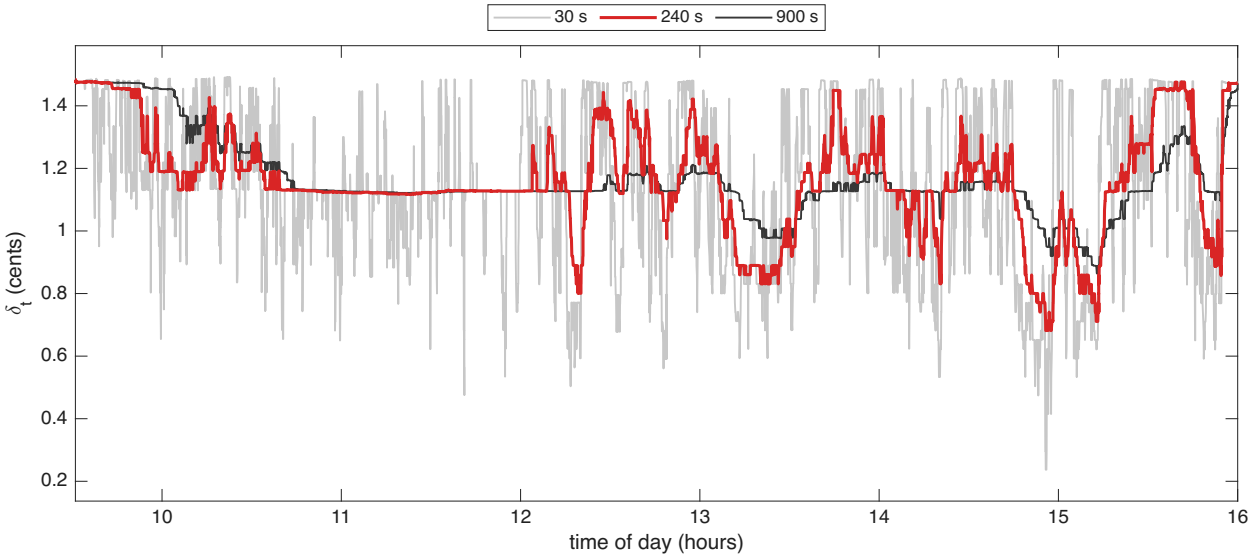

Figure 2: The Huber threshold δ_t in cents for `RIVN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

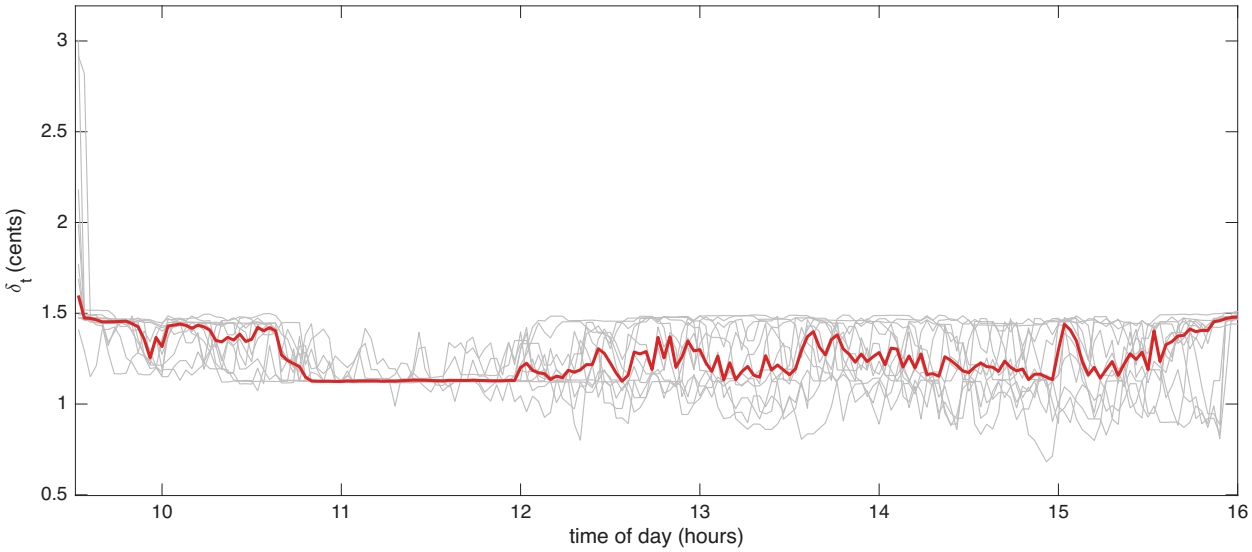

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `RIVN` trading days in Oct 2024 (grey lines) and their median (red line).

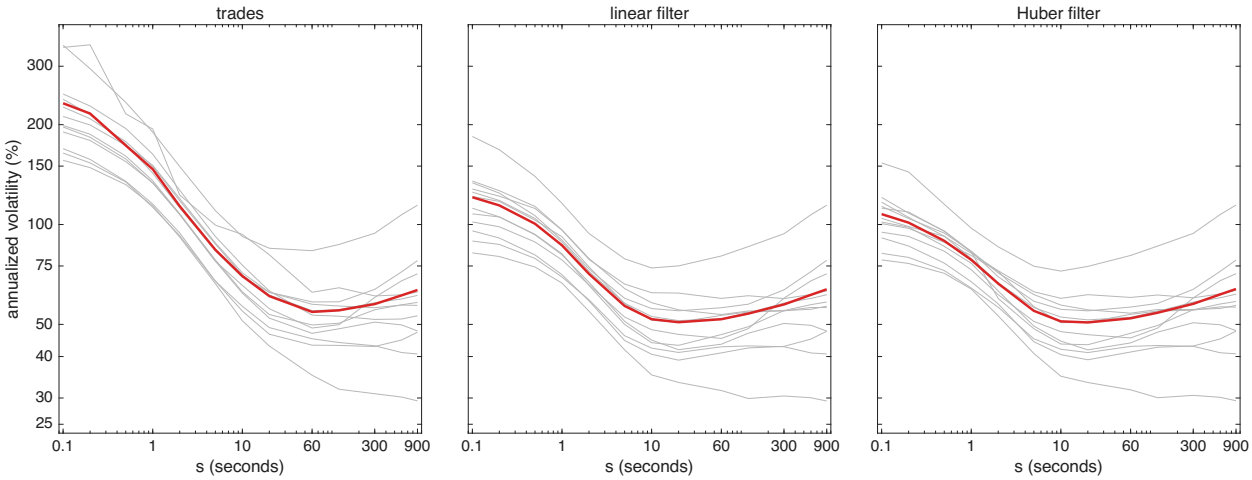

Figure 4: Volatility signature curves for `RIVN` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

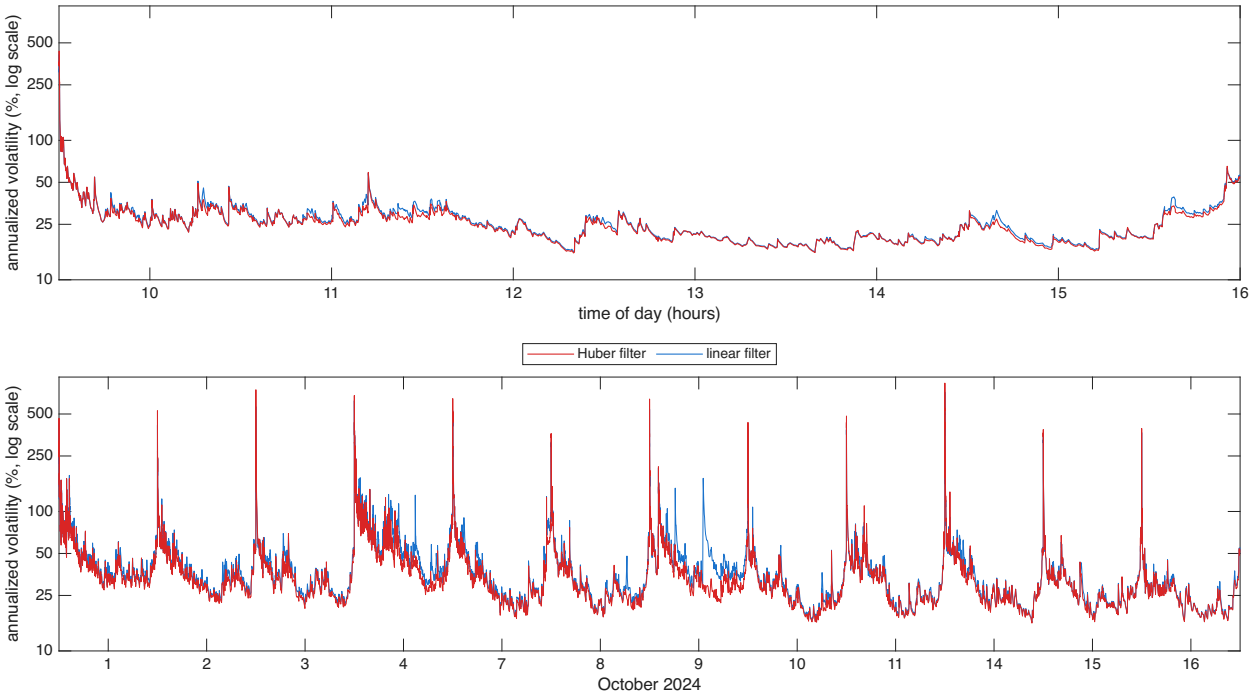

Figure 5: Filtered spot volatility for `RIVN` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 52$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

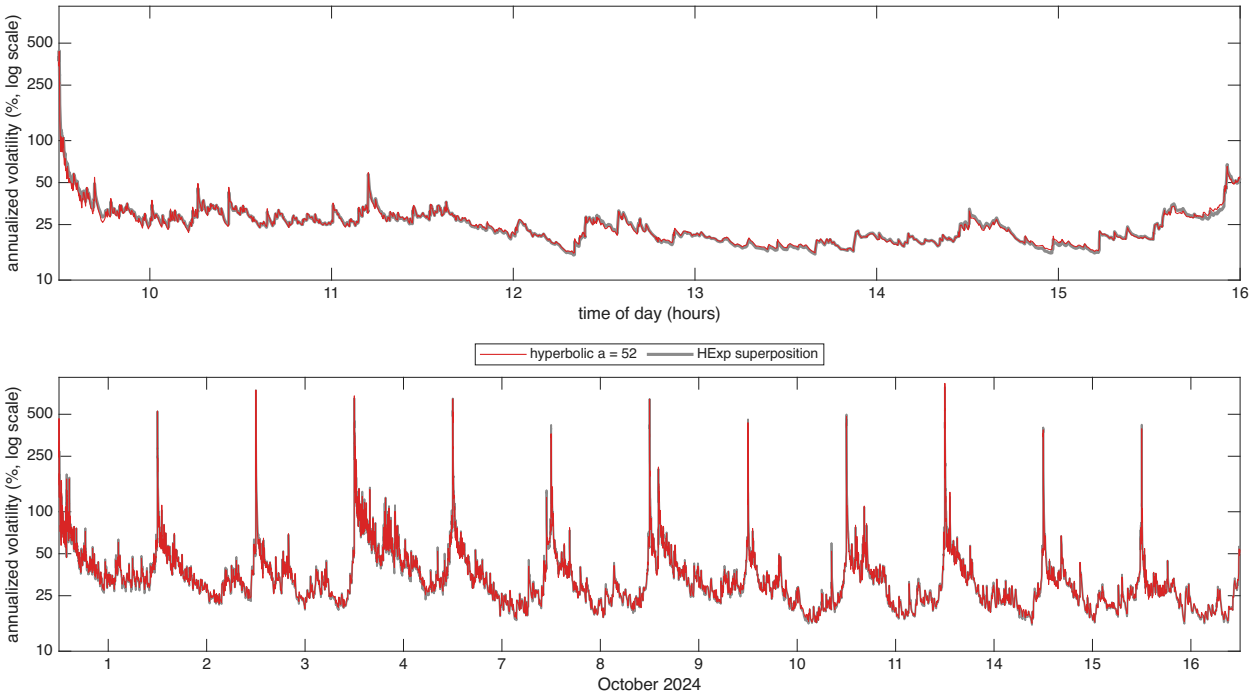

Figure 6: Hyperbolic against superposition weights for `RIVN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 52$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

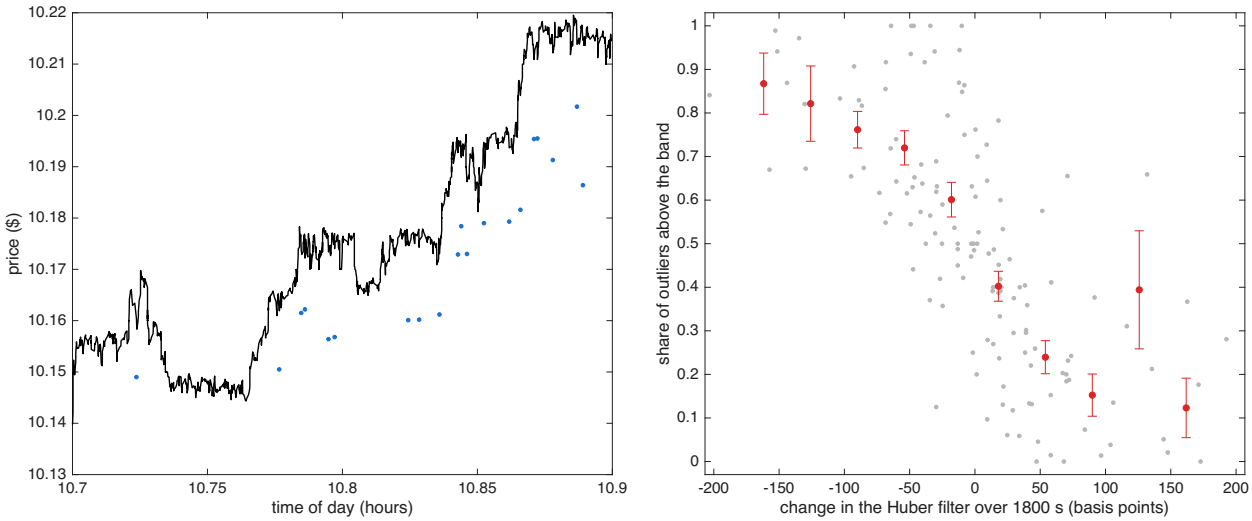

Figure 7: Outliers & robust price for `RIVN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

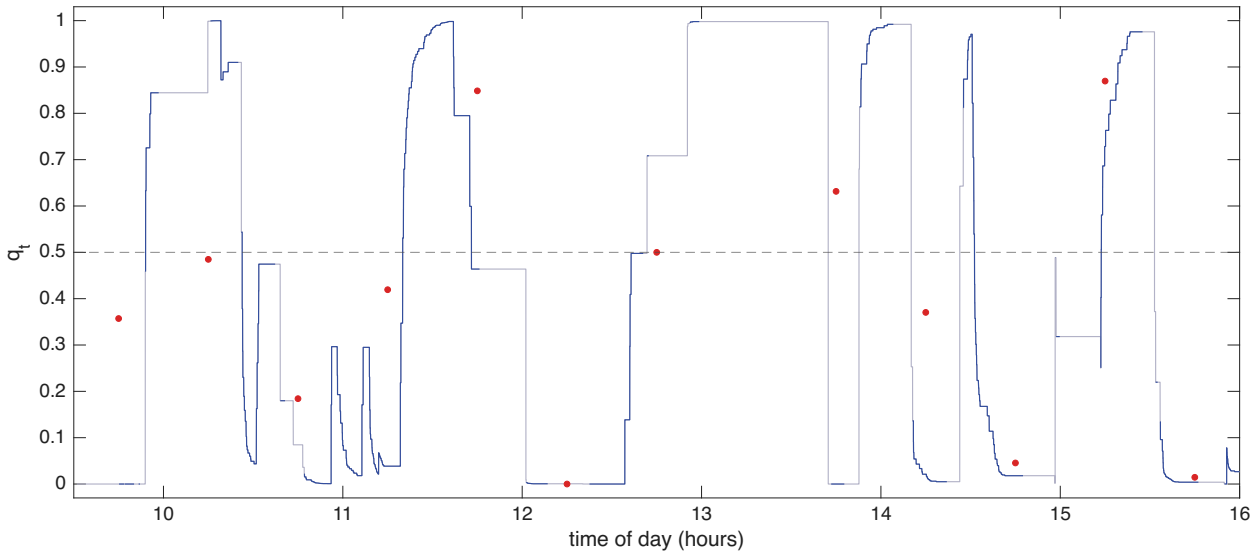

Figure 8: The directional conditional probability for `RIVN`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 141 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

On 14 October 2024, 34 per cent of the day’s largest one per cent of residuals fall inside its first minute, against at most 19 per cent on any other day. Every number that day contributes to carries that burst.

Diagnostics.

HOOD 1 to 16 October 2024

trades on 16 October 2024	86,176
average trading rate	4.5 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,122 (1.30%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.39%
scale half-life: run / minimizer[2]	240 / 90 s
cost of running 240 s	+0.16%
direction half-life[3]	76 s [61, 96]
Mandelbrot a minimizing QLIKE	54 trades
a within 0.0005 of the minimum	39 to 76
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	29% $\rightarrow$ 52%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	26% $\rightarrow$ 31%
median spot volatility, Huber / linear input[5]	29.0 / 34.3 %
hyperbolic against HExp, correlation of logs	0.9982
HExp QLIKE advantage, leave-one-day-out[6]	-0.00376
15-minute signature level	35 to 93 %
Huber signature, 1 min over 1 s, 12 day average	1.37

HOOD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:37; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.074	0.066	0.072	0.281	0.227	0.112	0.140	0.806
2	0.078	0.078	0.087	0.238	0.206	0.125	0.150	0.603
5	0.097	0.096	0.106	0.191	0.193	0.145	0.164	0.410
10	0.112	0.111	0.119	0.172	0.190	0.163	0.174	0.320
25	0.131	0.130	0.133	0.162	0.196	0.185	0.189	0.255
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.18	0.23	0.14	−0.44	3.26	3.11	2.77	3.32
2	0.02	0.14	0.07	−0.03	2.59	2.80	3.37	3.40
5	0.01	0.08	0.03	0.00	3.36	3.15	4.37	3.49
10	0.01	0.04	0.02	0.00	3.86	3.80	4.73	3.65
25	0.01	−0.01	−0.01	0.00	4.81	4.79	4.94	3.98

Table 1: Trade-lag differences, HOOD, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.60	0.55	0.59	1.10	26.5	17.4	22.2	122.7
1 s	0.44	0.42	0.44	0.72	27.5	22.6	25.6	90.0
5 s	0.37	0.36	0.37	0.49	32.7	30.9	32.1	57.7
10 s	0.37	0.36	0.37	0.45	34.4	33.3	33.7	47.6
1 min	0.37	0.36	0.36	0.39	36.3	35.8	35.5	37.9
15 min	0.36	0.36	0.36	0.38	35.1	35.0	34.9	34.9
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.02	0.02	0.01	−0.15	4.40	4.94	5.18	4.33
1 s	−0.11	−0.06	−0.11	−0.46	5.54	5.25	5.29	4.18
5 s	0.01	0.06	0.04	−0.27	9.98	7.92	7.74	7.11
10 s	0.01	0.03	0.01	−0.22	11.05	10.48	10.57	6.67
1 min	−0.01	0.00	0.00	−0.04				

Table 2: Clock-lag differences, HOOD, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.20	0.27	0.28
s.d. of $\hat{\epsilon}_t$ in cents	0.61	0.57	0.54
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.75
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.27	0.31	0.14
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.11	0.20	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.10	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.05	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.22,0.23	−0.23,0.25
95% C.I. for $\theta_t - X_t$	—	−0.37,0.41	−0.41,0.45
99% C.I. for $\theta_t - X_t$	—	−0.63,0.67	−0.74,0.76
s.e. of $\theta_t - X_t$ in cents	—	0.19	0.22

Table 3: Residual summaries, HOOD, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.75, is above two, while the Hill index of the Huber residual itself on this day is 1.91, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.91	2.29	1.58 to 2.52	1.76 to 2.75
Residual divided by δ_t	1.83	2.22	1.62 to 2.54	1.71 to 2.76
Repricing windows removed	1.91	2.29	1.66 to 2.51	1.75 to 2.82
Calm windows only	1.87	2.32		
Every 10th trade	2.00	1.57	1.43 to 2.12	1.51 to 2.55
Six-trade half-life	2.13	2.48	1.63 to 2.76	1.90 to 2.92

Table 4: Hill estimator of the tail index of the Huber filter residual, HOOD, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 10.1 to 10.1 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 5.0 cents on 16 October 2024, 97% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

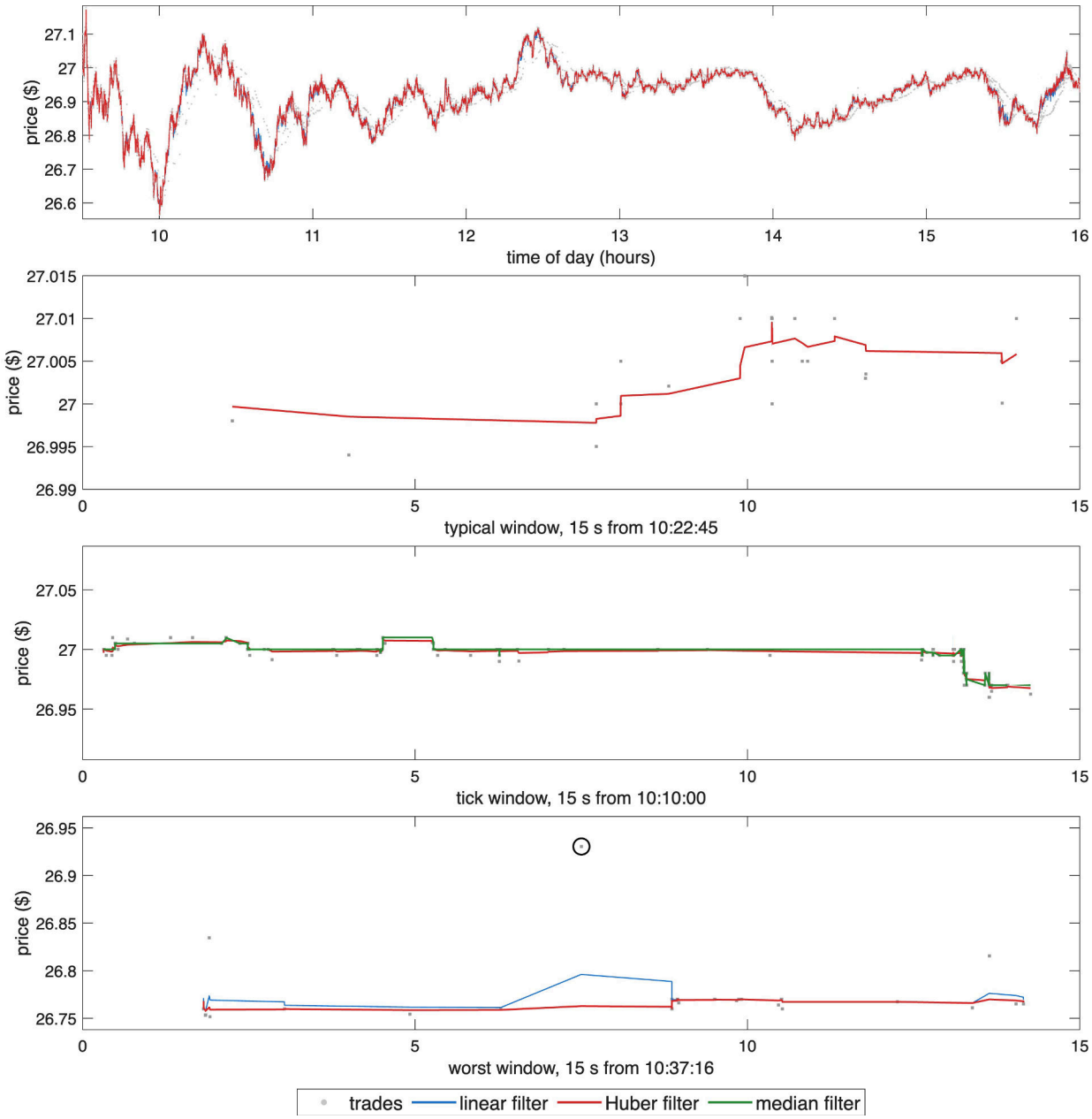

Figure 1: HOOD trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:22:45. 3rd: 15-second window at 10:10:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

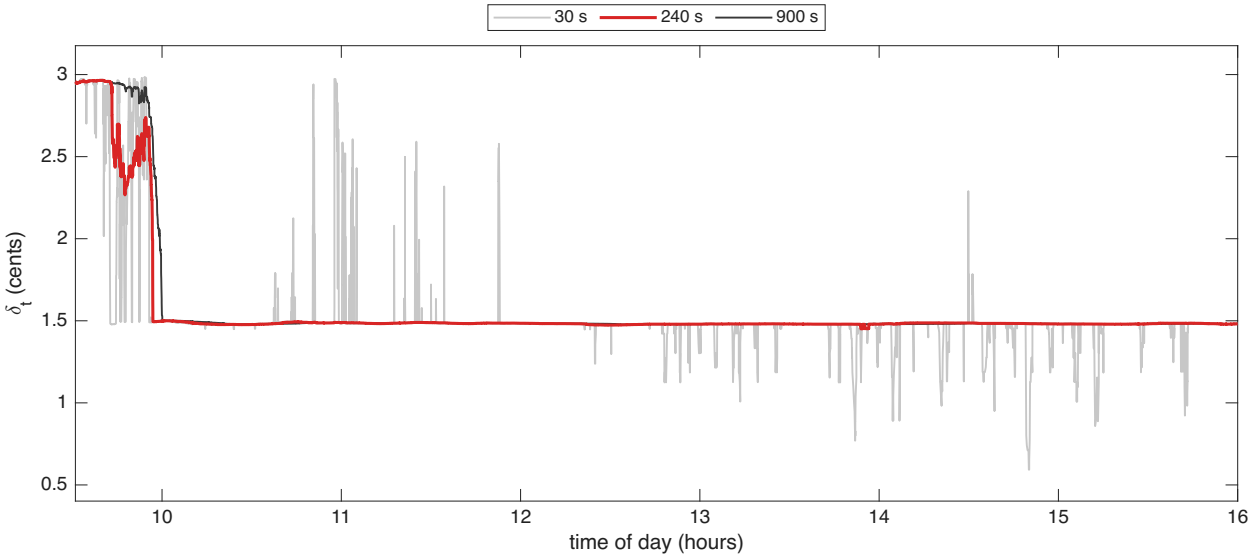

Figure 2: The Huber threshold δ_t in cents for HOOD on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

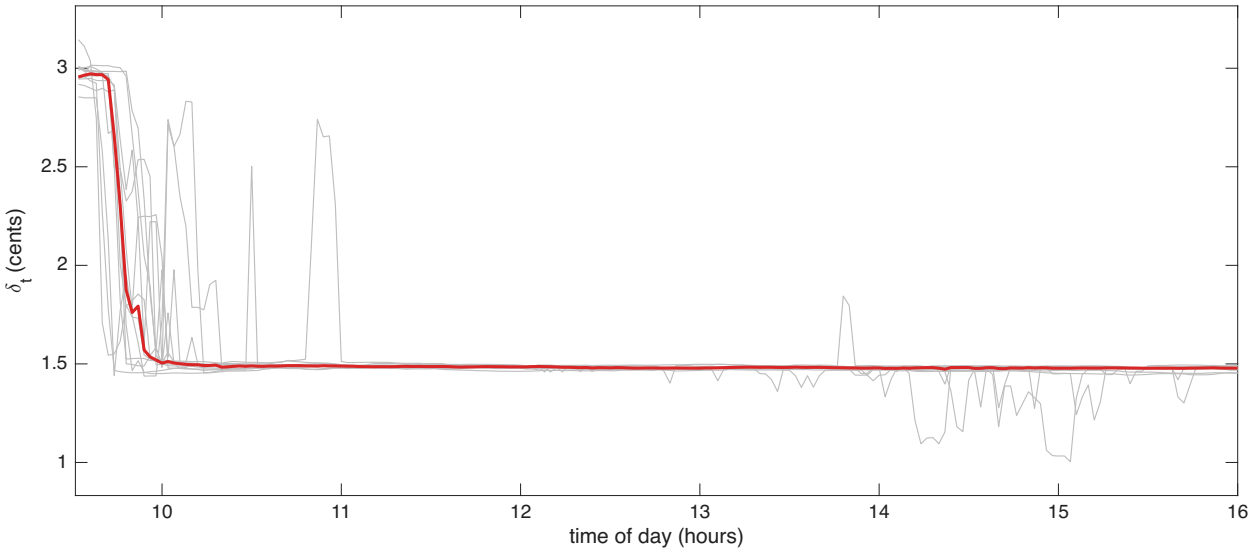

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 HOOD trading days in Oct 2024 (grey lines) and their median (red line).

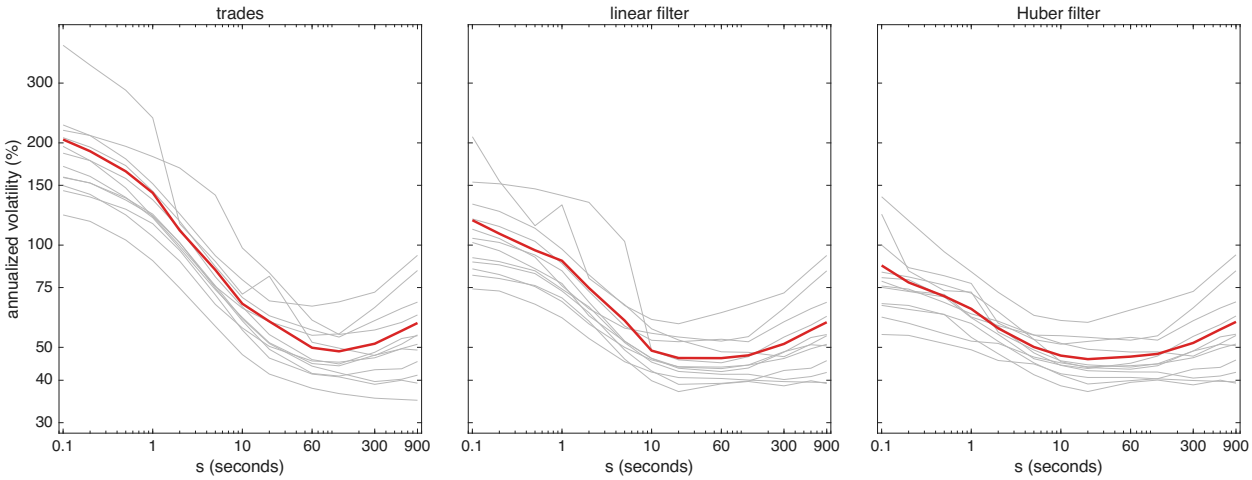

Figure 4: Volatility signature curves for HOOD over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

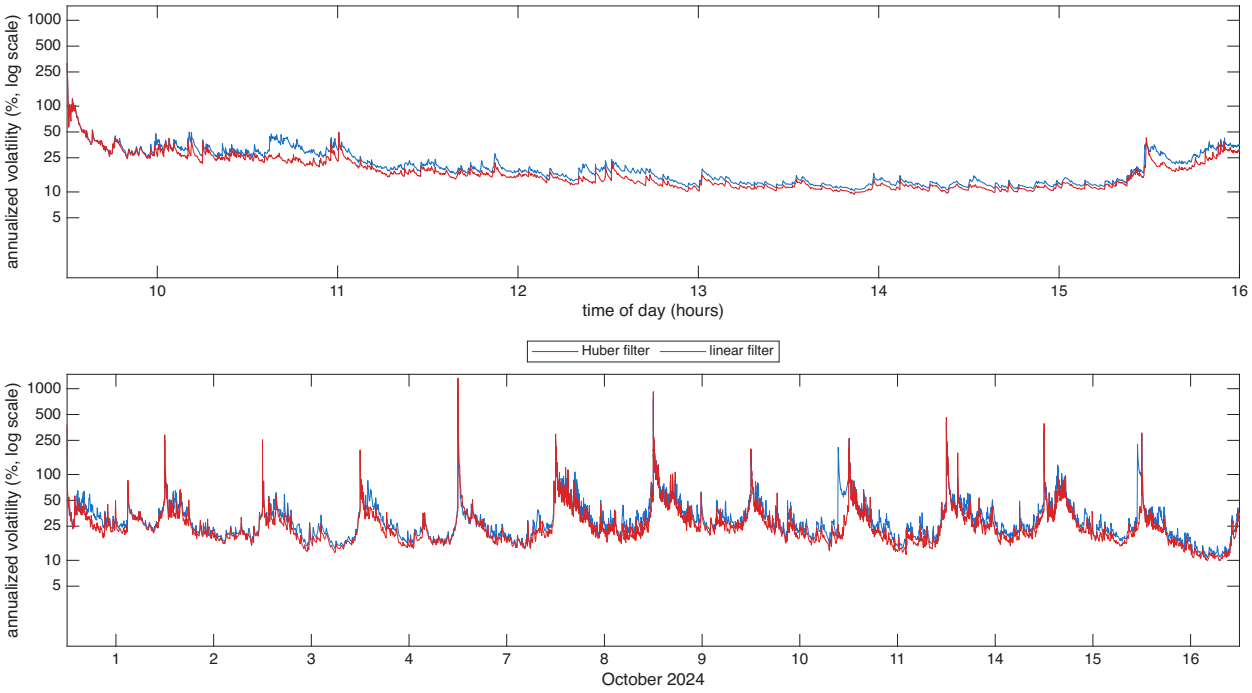

Figure 5: Filtered spot volatility for HOOD over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 54$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

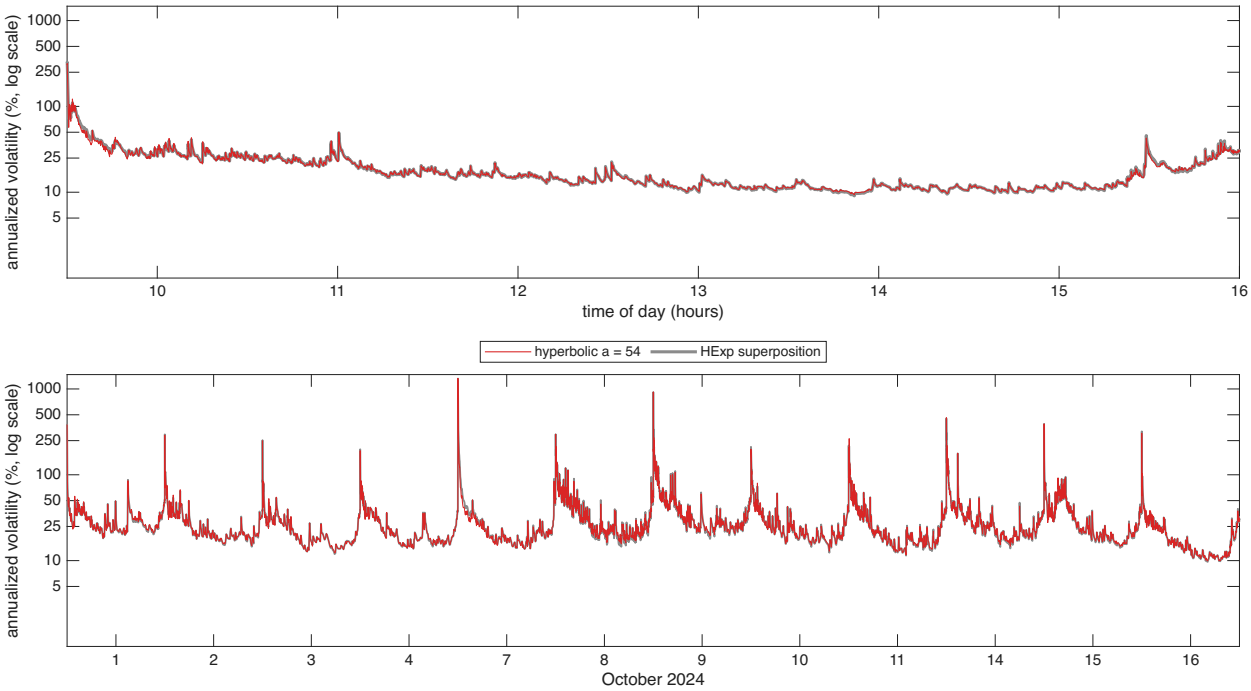

Figure 6: Hyperbolic against superposition weights for HOOD over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 54$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

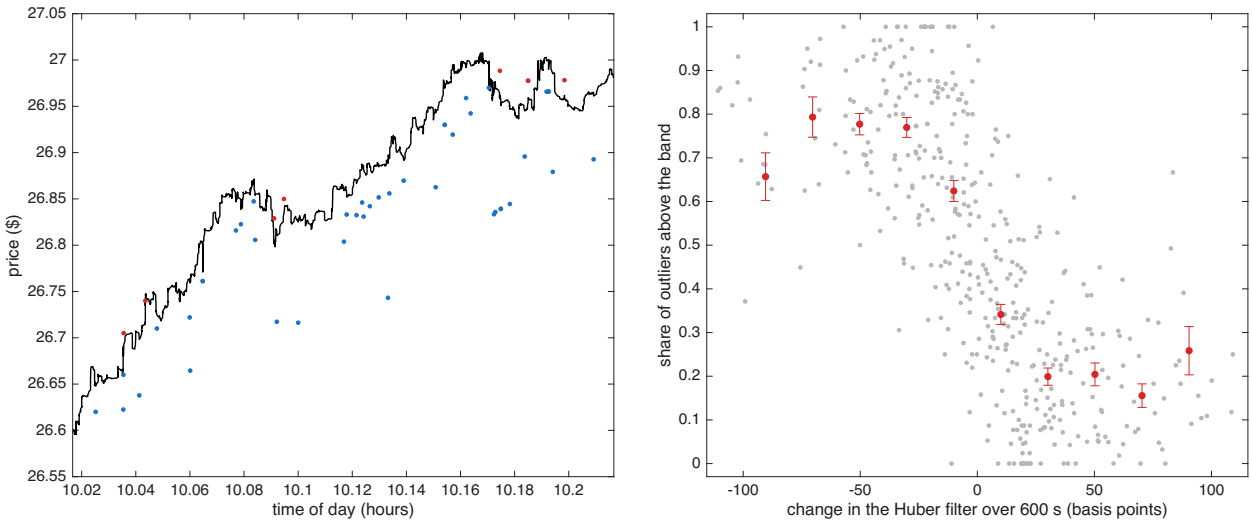

Figure 7: Outliers & robust price for HOOD on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

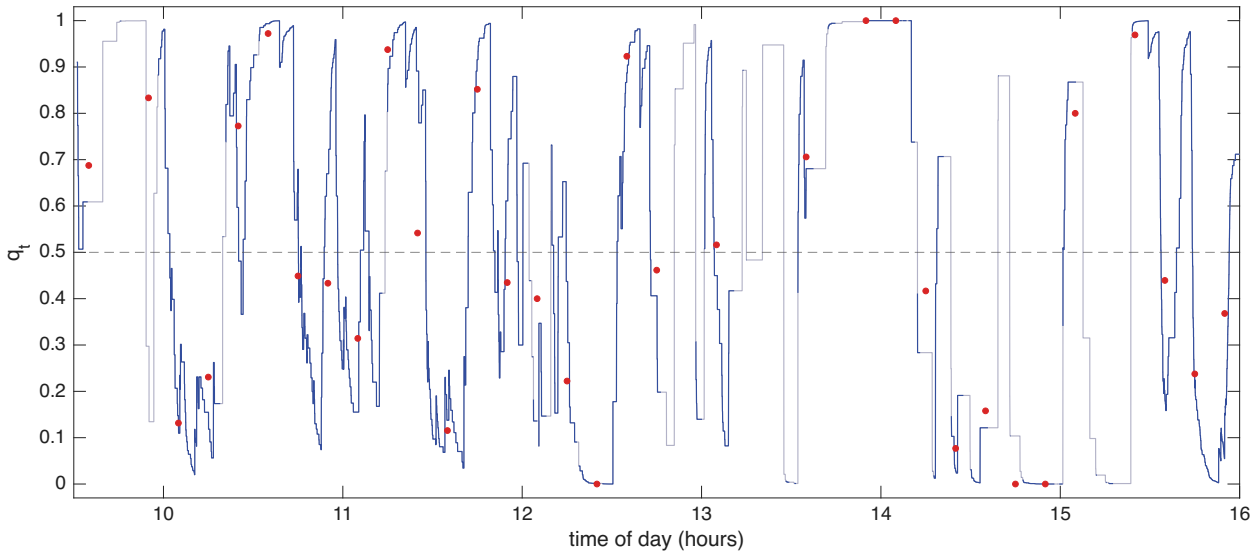

Figure 8: The directional conditional probability for HOOD, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 76 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 4 to 816 trades, against 54 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 45 trades for this stock.

Diagnostics.

CELH 1 to 16 October 2024

trades on 16 October 2024	84,533
average trading rate	4.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,832 (2.17%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+1.52%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.08%
direction half-life[3]	56 s [46, 60]
Mandelbrot a minimizing QLIKE	45 trades
a within 0.0005 of the minimum	33 to 60
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	28% $\rightarrow$ 313%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	25% $\rightarrow$ 94%
median spot volatility, Huber / linear input[5]	40.0 / 49.4 %
hyperbolic against HExp, correlation of logs	0.9972
HExp QLIKE advantage, leave-one-day-out[6]	-0.00174
15-minute signature level	35 to 106 %
Huber signature, 1 min over 1 s, 12 day average	1.28

CELH over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:14; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.237	0.186	0.199	0.881	0.583	0.299	0.391	2.326
2	0.229	0.209	0.230	0.721	0.513	0.323	0.405	1.709
5	0.264	0.241	0.266	0.546	0.473	0.357	0.418	1.139
10	0.288	0.267	0.287	0.462	0.454	0.386	0.426	0.864
25	0.321	0.305	0.313	0.403	0.464	0.434	0.447	0.652
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.17	0.07	−0.46	3.45	2.91	2.41	2.64
2	0.05	0.09	0.02	−0.01	3.67	3.25	2.85	2.68
5	0.00	0.04	0.01	0.01	4.60	4.29	3.77	2.71
10	0.00	0.03	0.00	−0.01	4.81	4.73	4.00	2.77
25	0.00	0.01	0.00	0.00	5.48	5.38	4.72	3.09

Table 1: Trade-lag differences, `CELH`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.05	1.92	2.02	3.31	53.8	35.4	49.0	280.8
1 s	1.39	1.32	1.38	2.13	55.0	47.8	58.2	198.9
5 s	1.12	1.09	1.13	1.42	64.3	61.5	65.3	123.1
10 s	1.08	1.06	1.08	1.26	67.1	65.4	67.4	102.2
1 min	1.10	1.10	1.10	1.15	71.0	70.4	69.8	75.6
15 min	1.18	1.18	1.18	1.19	79.5	79.3	79.2	79.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.07	0.13	0.15	−0.07	7.41	7.65	7.58	3.43
1 s	0.03	0.09	0.04	−0.43	8.60	8.36	7.69	5.42
5 s	0.07	0.10	0.08	−0.19	11.01	11.11	10.06	6.68
10 s	0.19	0.21	0.20	−0.01	10.99	9.79	9.70	9.70
1 min	−0.03	−0.02	−0.02	−0.04				

Table 2: Clock-lag differences, `CELH`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.59	0.73	0.77
s.d. of $\hat{\epsilon}_t$ in cents	1.68	1.62	1.50
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.40
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.23	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.13	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.06	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.03	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.52,0.52	−0.56,0.56
95% C.I. for $\theta_t - X_t$	—	−0.93,0.97	−1.05,1.08
99% C.I. for $\theta_t - X_t$	—	−1.46,1.54	−1.76,1.80
s.e. of $\theta_t - X_t$ in cents	—	0.45	0.55

Table 3: Residual summaries, `CELH`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.40, is above two, while the Hill index of the Huber residual itself on this day is 1.92, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.92	2.28	1.61 to 2.94	1.85 to 2.69
Residual divided by δ_t	1.88	2.41	1.64 to 2.55	1.86 to 3.00
Repricing windows removed	1.92	2.28	1.59 to 2.94	1.82 to 2.69
Calm windows only	1.85	2.20		
Every 10th trade	1.73	2.15	1.38 to 2.51	1.26 to 2.42
Six-trade half-life	2.09	2.47	1.73 to 3.13	1.95 to 2.93

Table 4: Hill estimator of the tail index of the Huber filter residual, `CELH`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 10.1 to 20.1 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.1 cents on 16 October 2024, 84% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

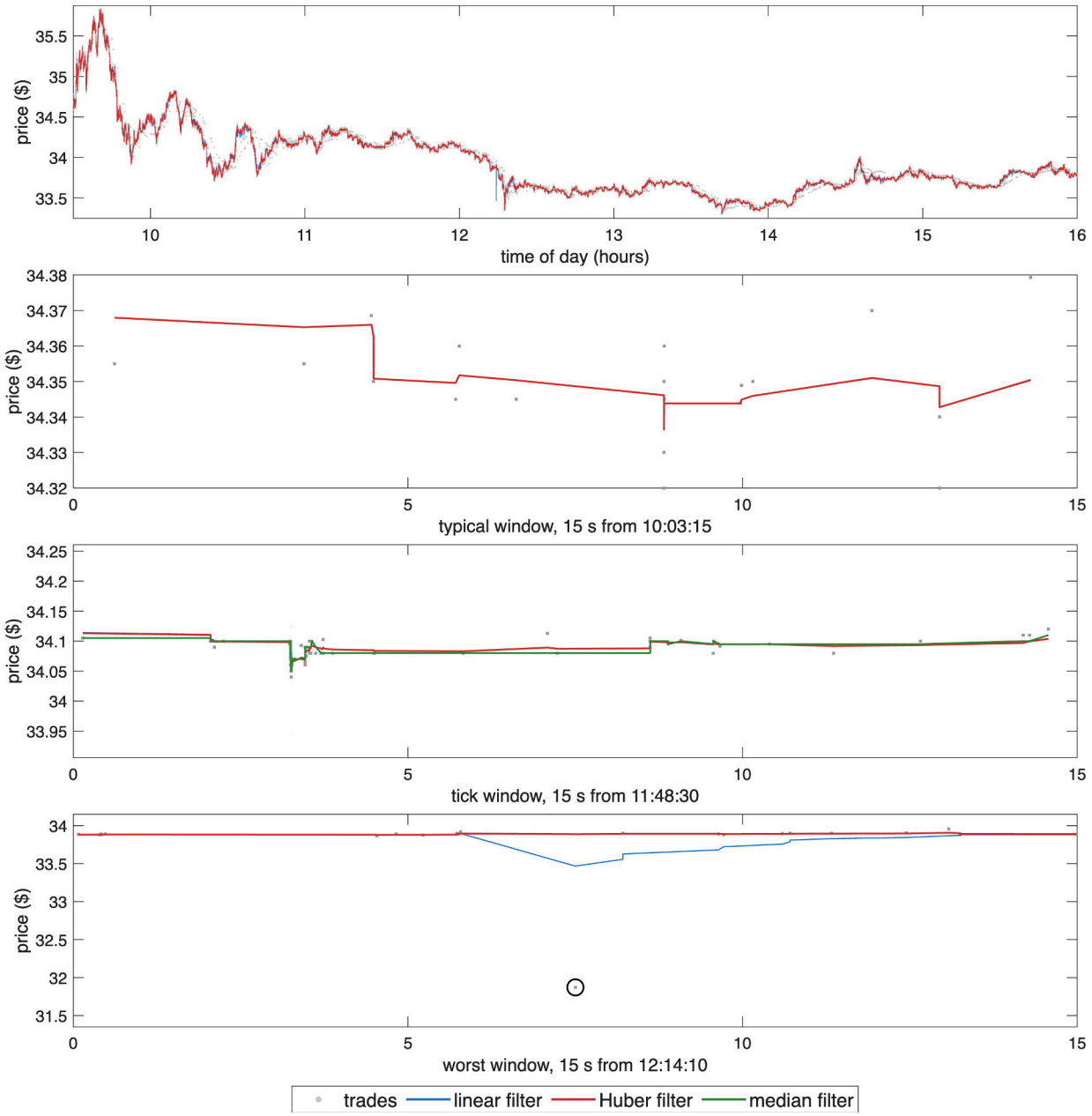

Figure 1: `CELH` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:03:15. 3rd: 15-second window at 11:48:30 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

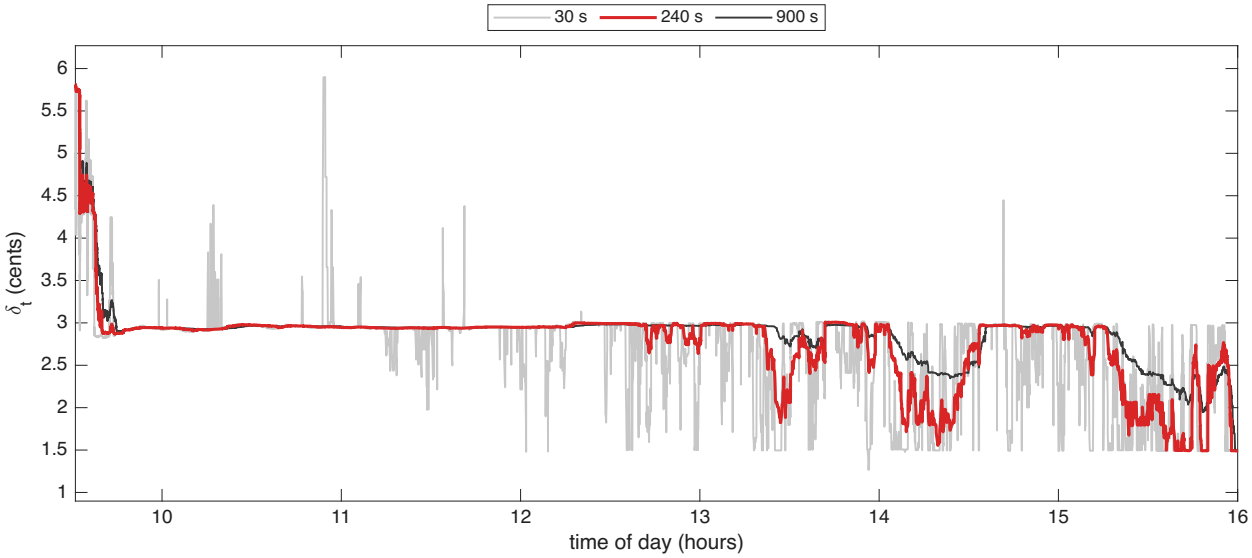

Figure 2: The Huber threshold δ_t in cents for `CELH` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

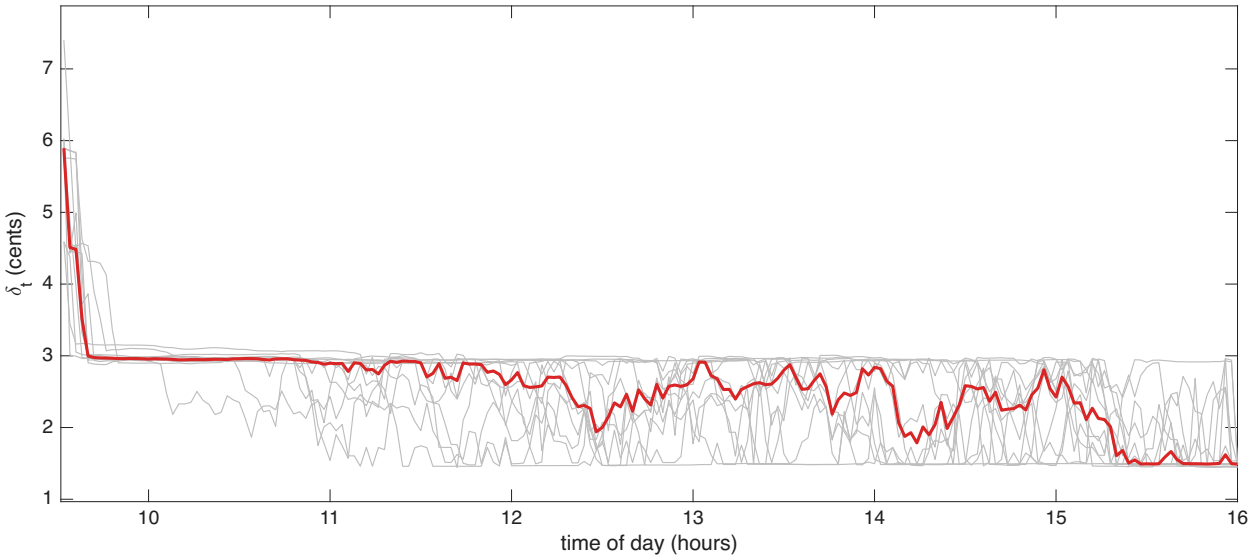

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `CELH` trading days in Oct 2024 (grey lines) and their median (red line).

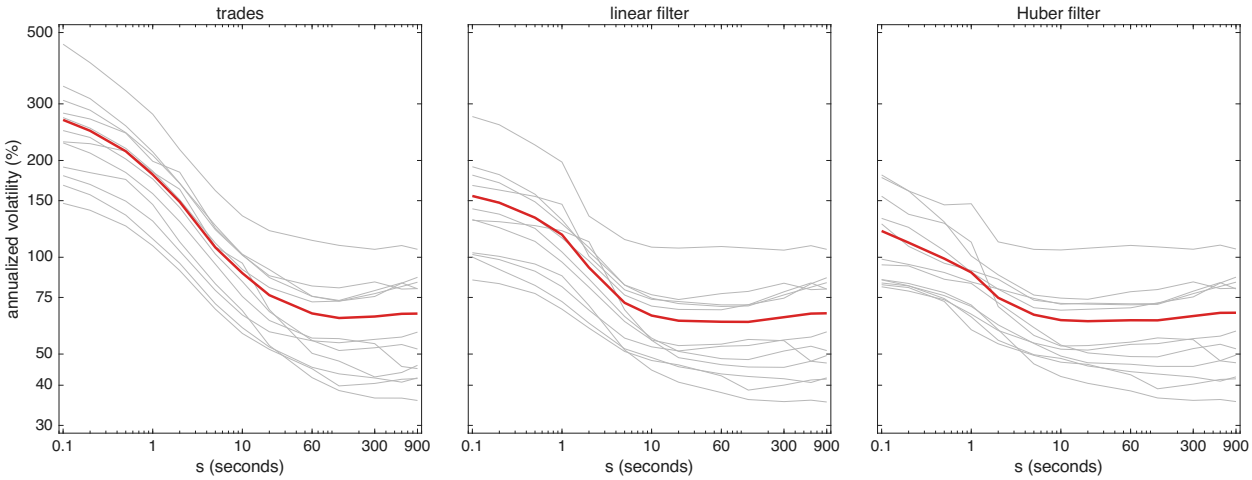

Figure 4: Volatility signature curves for `CELH` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

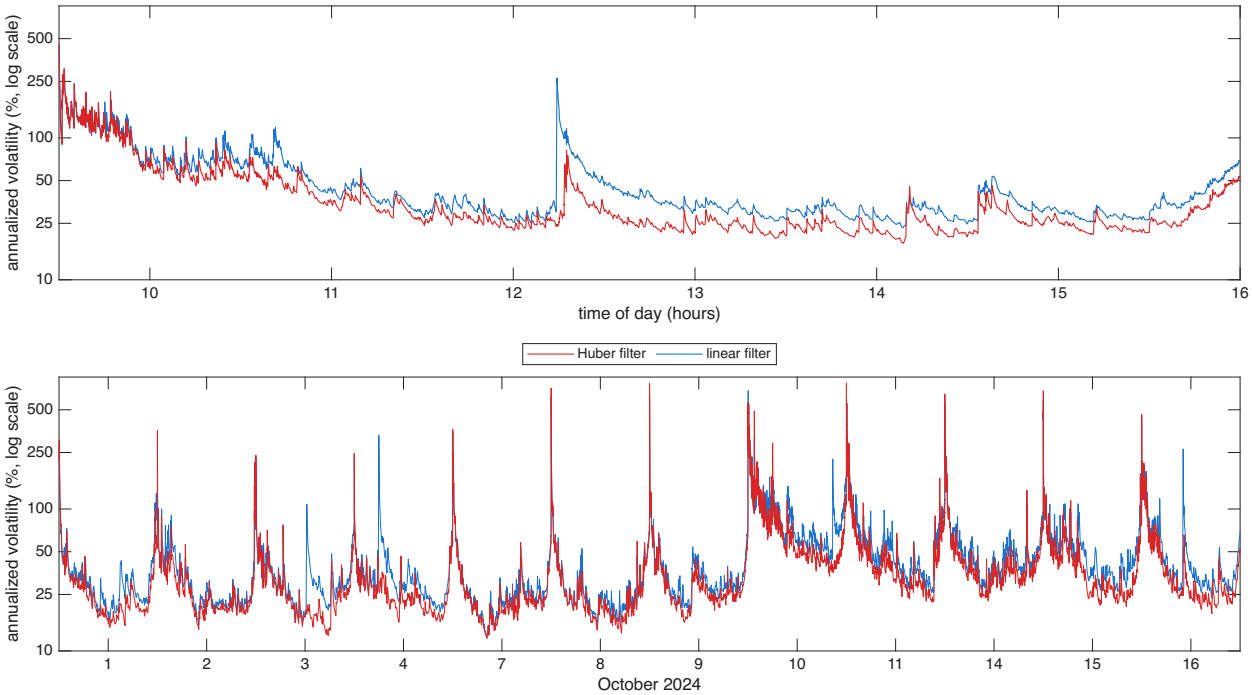

Figure 5: Filtered spot volatility for `CELH` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 45$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

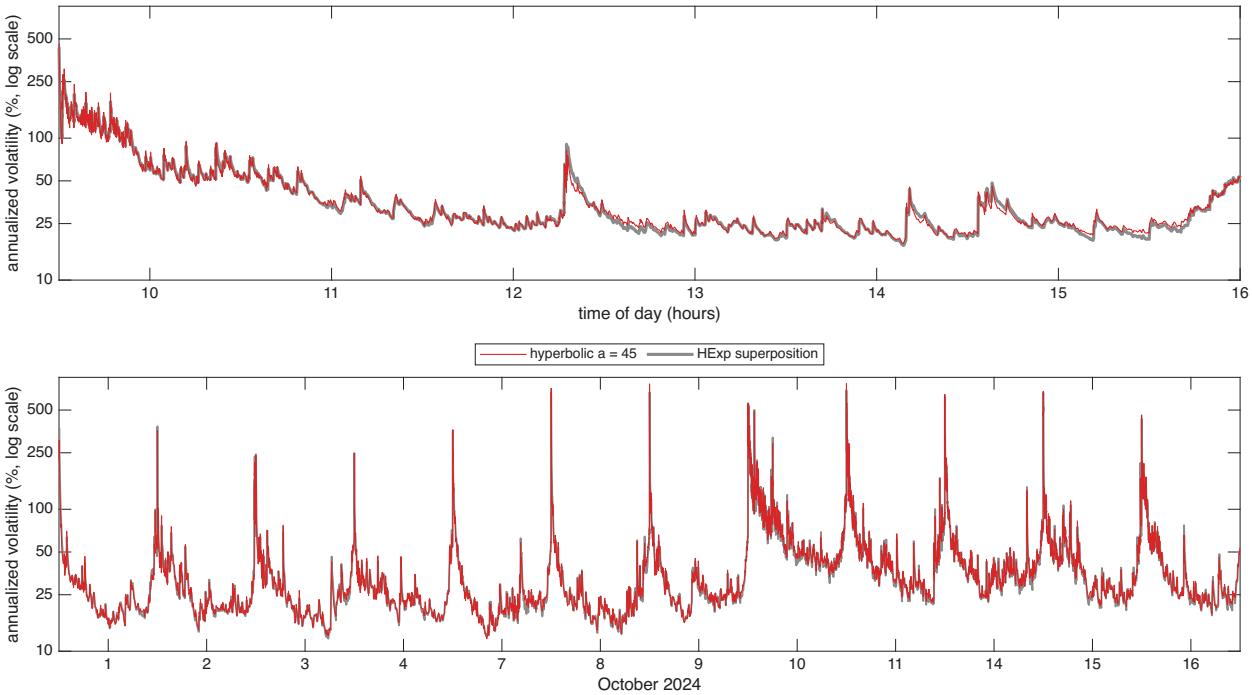

Figure 6: Hyperbolic against superposition weights for `CELH` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 45$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

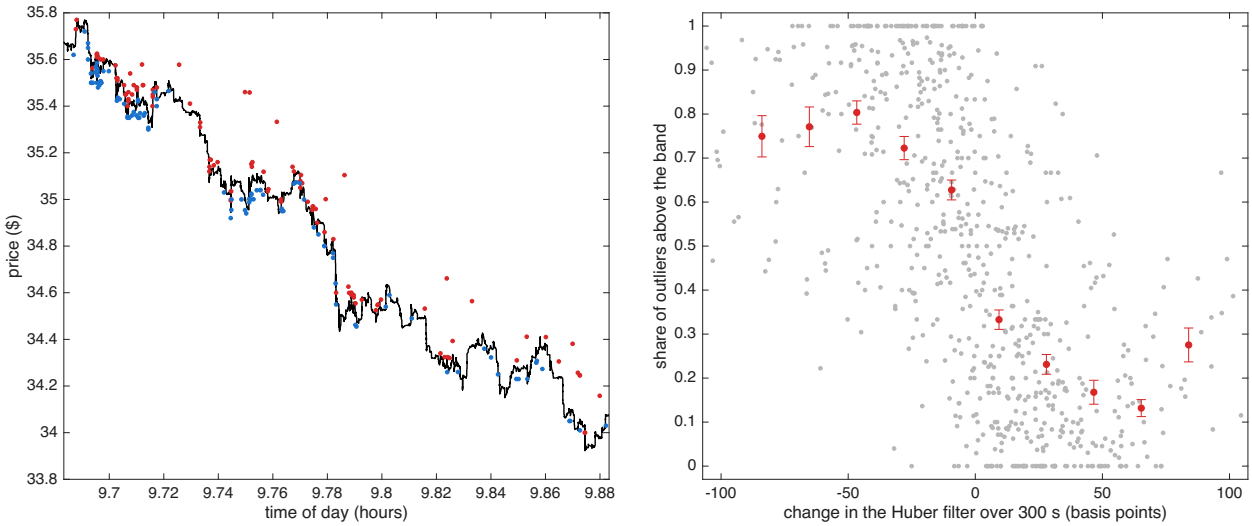

Figure 7: Outliers & robust price for `CELH` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

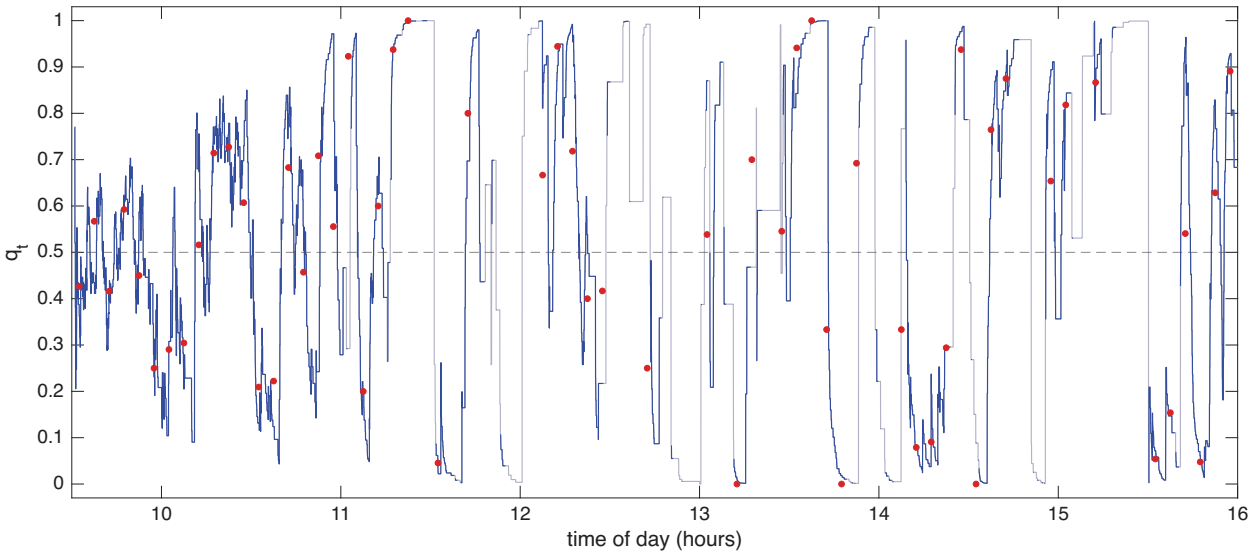

Figure 8: The directional conditional probability for CELH, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 56 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

QCOM 1 to 16 October 2024

trades on 16 October 2024	99,429
average trading rate	4.3 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	3,866 (3.89%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.26%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.01%
direction half-life[3]	33 s [28, 41]
Mandelbrot a minimizing QLIKE	39 trades
a within 0.0005 of the minimum	29 to 57
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	38% $\to$ 81%
Huber $\hat{\sigma}$, before $\to$ peak	35% $\to$ 47%
median spot volatility, Huber / linear input[5]	20.7 / 23.8 %
hyperbolic against HExp, correlation of logs	0.9931
HExp QLIKE advantage, leave-one-day-out[6]	+0.00149
15-minute signature level	18 to 52 %
Huber signature, 1 min over 1 s, 12 day average	1.36

QCOM over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:13; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.541	0.427	0.462	1.989	1.680	1.155	0.926	5.337
2	0.536	0.483	0.540	1.652	1.490	1.122	0.986	4.011
5	0.638	0.572	0.632	1.278	1.371	1.133	1.054	2.746
10	0.709	0.643	0.685	1.094	1.275	1.116	1.066	2.095
25	0.765	0.722	0.735	0.947	1.135	1.058	1.039	1.533
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	−0.06	0.13	−0.44	2.06	1.58	2.32	2.33
2	0.02	0.03	0.05	−0.02	1.90	1.59	2.37	2.32
5	0.02	0.04	−0.01	0.00	1.63	1.54	2.24	2.31
10	−0.04	−0.06	−0.02	0.00	1.66	1.64	2.38	2.36
25	0.00	0.00	0.00	0.01	2.13	2.22	3.14	2.52

Table 1: Trade-lag differences, `QCOM`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	4.30	3.87	3.92	7.37	31.3	23.9	24.9	128.1
1 s	2.84	2.63	2.73	4.29	28.1	24.5	26.1	82.7
5 s	2.43	2.33	2.37	3.00	30.2	28.9	29.8	49.8
10 s	2.33	2.27	2.27	2.68	31.3	30.5	30.4	42.0
1 min	2.31	2.29	2.27	2.36	33.7	33.4	32.7	34.8
15 min	3.14	3.14	3.13	3.14	47.5	47.4	47.2	47.5
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.21	−0.18	−0.04	−0.31	2.23	2.39	2.70	2.97
1 s	−0.05	0.00	0.08	−0.30	3.33	3.28	3.19	4.79
5 s	0.01	0.03	−0.01	−0.25	8.74	8.66	7.48	6.79
10 s	0.08	0.10	0.10	−0.15	9.42	9.35	8.69	5.80
1 min	0.19	0.19	0.21	0.17				

Table 2: Clock-lag differences, `QCOM`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.38	1.67	1.77
s.d. of $\hat{\epsilon}_t$ in cents	4.00	3.88	3.56
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.33
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.25	0.27	0.13
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.11	0.17	0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.05	0.08	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.03	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−1.20,1.14	−1.34,1.25
95% C.I. for $\theta_t - X_t$	—	−2.07,1.98	−2.43,2.28
99% C.I. for $\theta_t - X_t$	—	−3.14,2.96	−4.39,3.83
s.e. of $\theta_t - X_t$ in cents	—	0.98	1.36

Table 3: Residual summaries, `QCOM`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.33, is above two, while the Hill index of the Huber residual itself on this day is 1.73, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.73	2.02	1.73 to 3.39	1.97 to 2.88
Residual divided by δ_t	2.79	2.30	1.87 to 2.83	1.94 to 3.44
Repricing windows removed	2.13	2.04	1.93 to 3.42	1.96 to 2.87
Calm windows only	2.32	2.06		
Every 10th trade	1.36	2.27	1.36 to 2.91	1.56 to 2.71
Six-trade half-life	1.66	2.17	1.66 to 3.76	2.04 to 3.02

Table 4: Hill estimator of the tail index of the Huber filter residual, `QCOM`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 20.2 to 40.3 cents over the days and removes 0% to 4% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 15.1 cents on 16 October 2024, 77% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

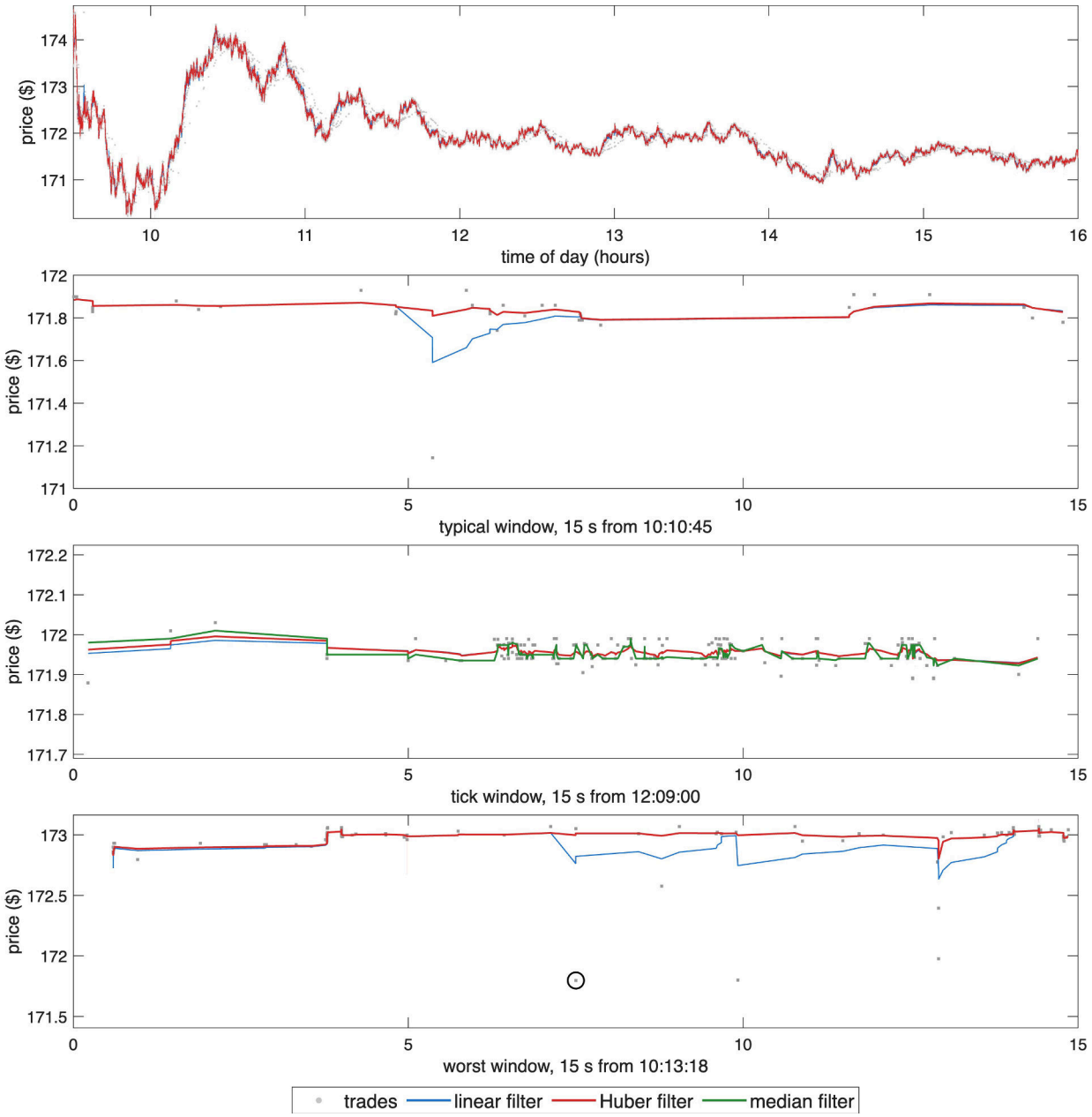

Figure 1: QCOM trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:10:45. 3rd: 15-second window at 12:09:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

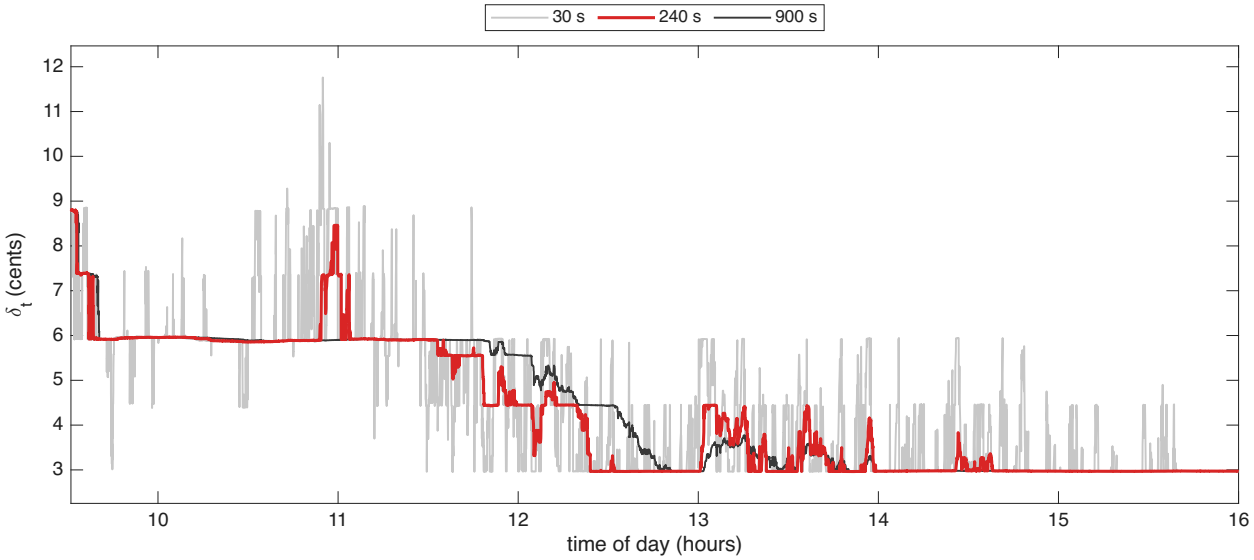

Figure 2: The Huber threshold δ_t in cents for QCOM on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

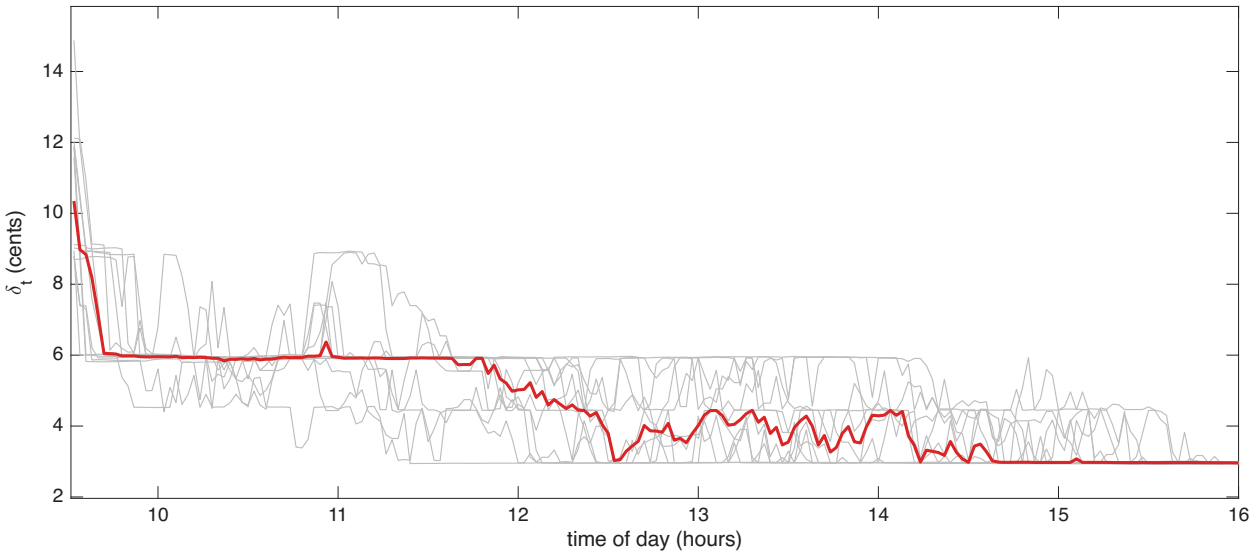

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 QCOM trading days in Oct 2024 (grey lines) and their median (red line).

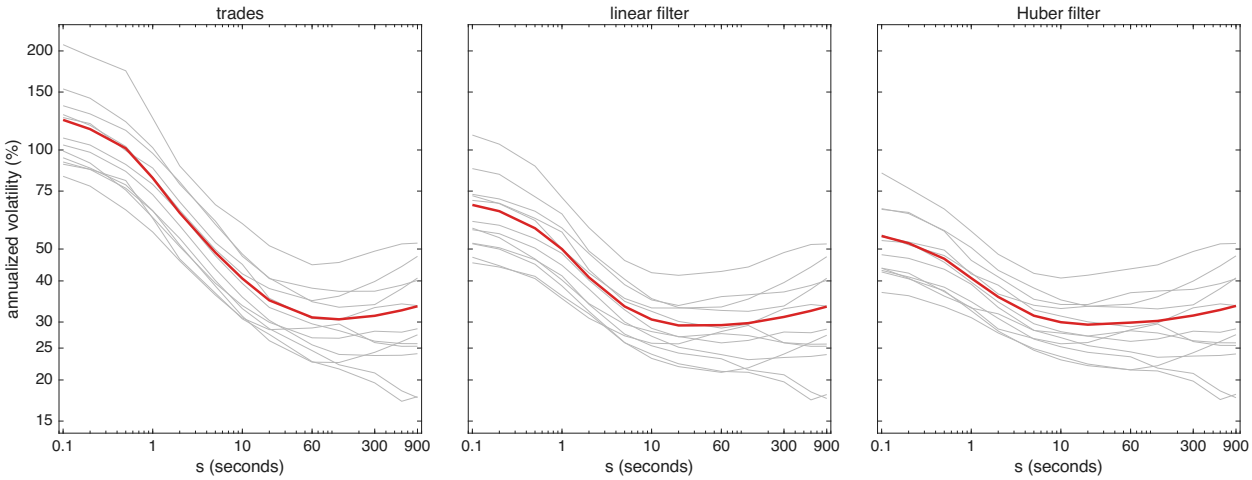

Figure 4: Volatility signature curves for QCOM over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

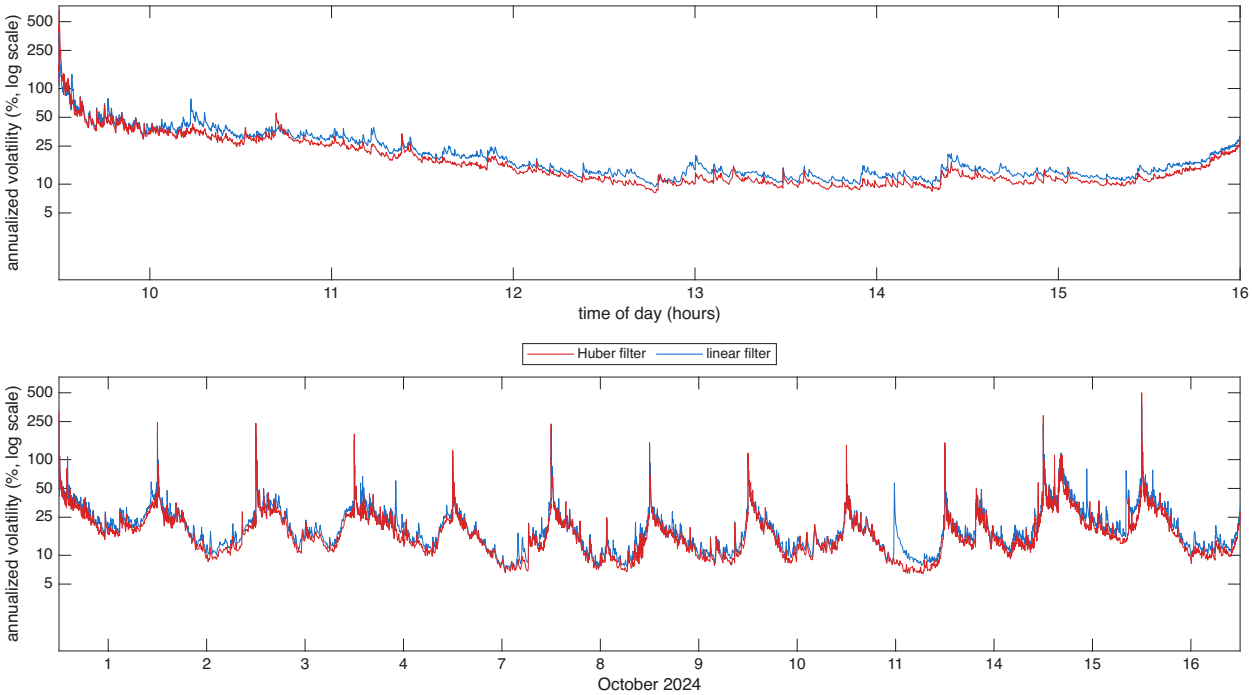

Figure 5: Filtered spot volatility for QCOM over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 39$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

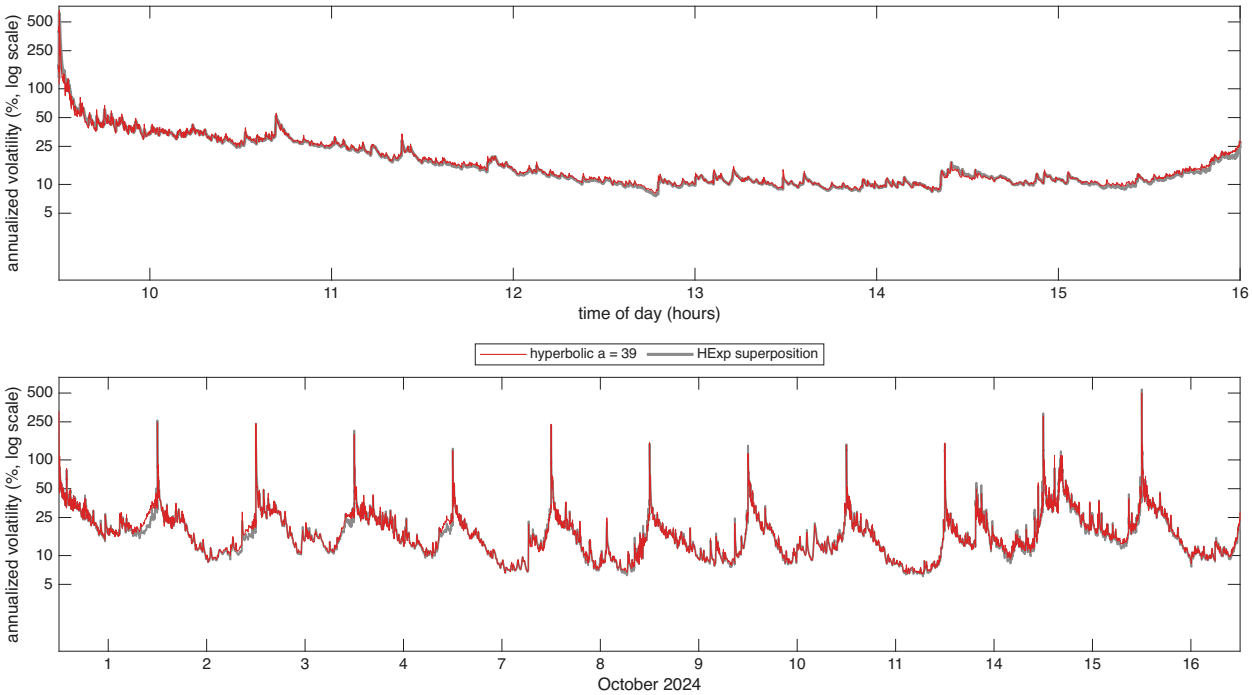

Figure 6: Hyperbolic against superposition weights for QCOM over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 39$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 30, 150, 750, 3750 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

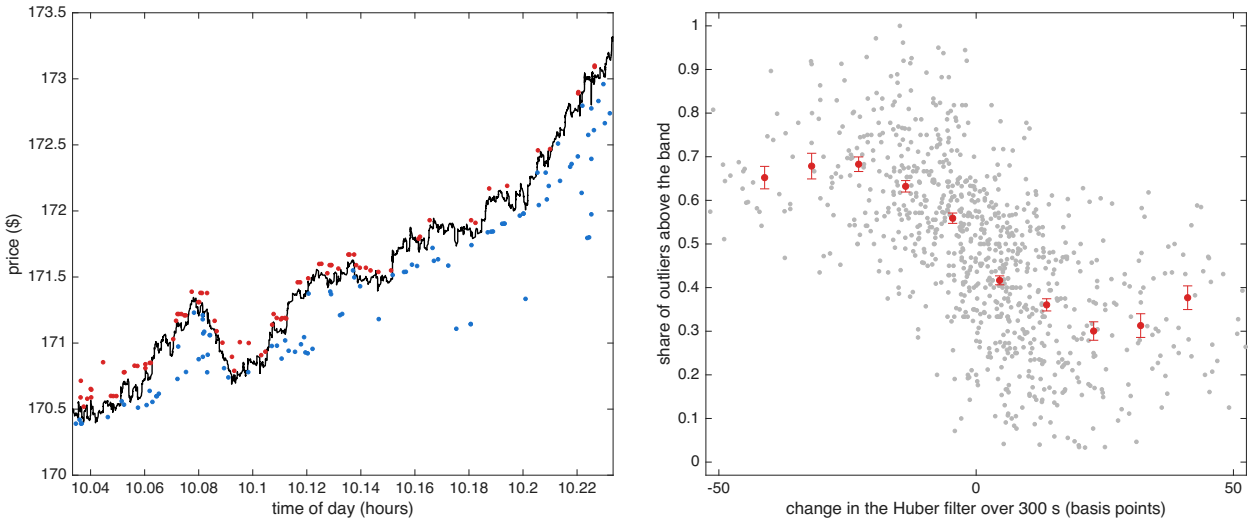

Figure 7: Outliers & robust price for QCOM on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

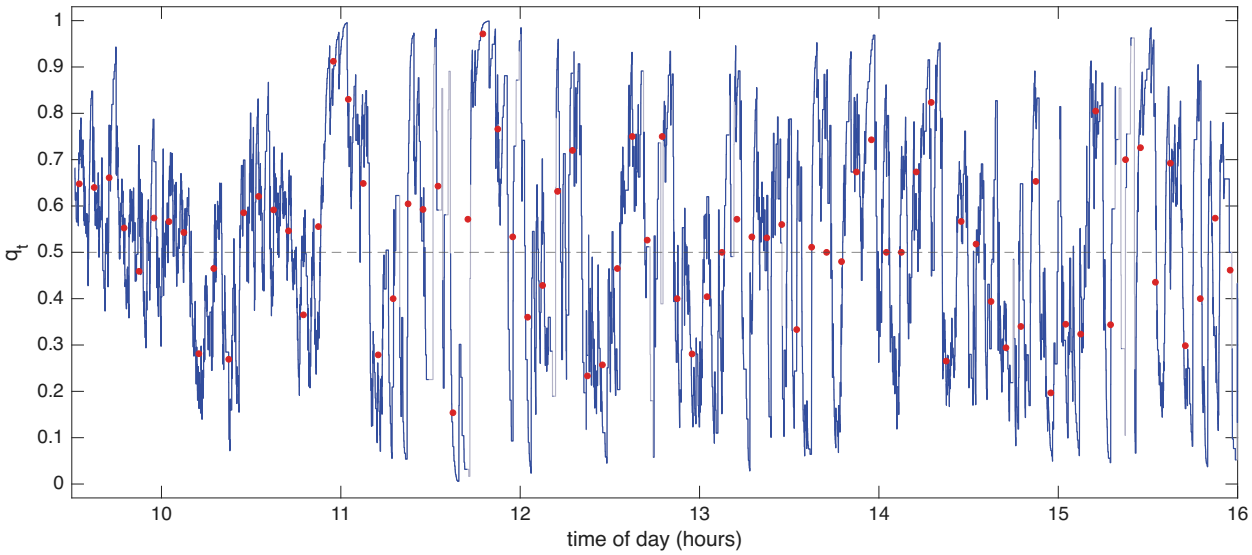

Figure 8: The directional conditional probability for QCOM, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 33 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The largest Huber residual on 16 October 2024 occurs at 09:34, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:13.

Diagnostics.

PYPL 1 to 16 October 2024

trades on 16 October 2024	74,100
average trading rate	4.3 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,236 (3.02%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+3.28%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	42 s [37, 49]
Mandelbrot a minimizing QLIKE	95 trades
a within 0.0005 of the minimum	65 to 141
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	15% $\to$ 23%
Huber $\hat{\sigma}$, before $\to$ peak	17% $\to$ 18%
median spot volatility, Huber / linear input[5]	12.5 / 15.6 %
hyperbolic against HExp, correlation of logs	0.9987
HExp QLIKE advantage, leave-one-day-out[6]	-0.00877
15-minute signature level	18 to 37 %
Huber signature, 1 min over 1 s, 12 day average	1.52

PYPL over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:53; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.122	0.111	0.124	0.525	0.576	0.399	0.265	1.519
2	0.119	0.123	0.145	0.430	0.446	0.333	0.284	1.147
5	0.144	0.144	0.167	0.330	0.408	0.336	0.302	0.790
10	0.164	0.160	0.179	0.282	0.371	0.326	0.305	0.598
25	0.194	0.187	0.193	0.250	0.363	0.338	0.314	0.451
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.40	−0.30	0.15	−0.43	1.36	1.26	2.74	2.55
2	0.19	0.23	0.06	−0.02	1.54	1.38	2.49	2.55
5	−0.06	−0.06	−0.02	−0.02	1.38	1.34	2.12	2.57
10	0.05	0.05	−0.01	−0.03	1.54	1.49	2.17	3.07
25	0.02	0.02	−0.01	0.00	3.71	2.95	2.76	3.99

Table 1: Trade-lag differences, PYPL, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.83	0.83	0.96	1.73	20.0	15.6	13.8	69.5
1 s	0.65	0.60	0.63	1.06	16.0	13.4	12.9	48.1
5 s	0.47	0.46	0.48	0.67	14.7	13.8	14.4	28.7
10 s	0.47	0.46	0.47	0.59	15.6	14.9	14.8	23.4
1 min	0.48	0.48	0.47	0.51	17.2	16.8	16.4	18.2
15 min	0.47	0.47	0.47	0.47	17.8	17.7	17.4	17.9
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.10	0.00	−0.04	−0.18	2.85	3.32	2.70	3.97
1 s	−0.51	−0.46	−0.35	−0.45	9.00	5.43	4.28	5.36
5 s	−0.08	−0.05	−0.06	−0.32	8.36	8.43	10.58	7.60
10 s	0.03	0.06	0.06	−0.21	9.69	9.04	8.60	8.39
1 min	0.02	0.03	0.05	−0.05				

Table 2: Clock-lag differences, PYPL, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.36	0.44	0.47
s.d. of $\hat{\epsilon}_t$ in cents	1.17	1.09	1.02
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.73
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.23	0.15
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.14	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.06	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.04	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.29,0.30	−0.33,0.34
95% C.I. for $\theta_t - X_t$	—	−0.56,0.55	−0.67,0.65
99% C.I. for $\theta_t - X_t$	—	−0.94,0.91	−1.25,1.25
s.e. of $\theta_t - X_t$ in cents	—	0.27	0.39

Table 3: Residual summaries, PYPL, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.73, is above two, while the Hill index of the Huber residual itself on this day is 1.93, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.93	2.48	1.66 to 2.36	1.71 to 2.58
Residual divided by δ_t	2.65	3.44	1.95 to 2.93	1.86 to 3.44
Repricing windows removed	2.47	3.08	1.80 to 2.47	1.71 to 3.08
Calm windows only	2.44	3.06		
Every 10th trade	1.54	2.07	1.34 to 2.03	1.38 to 2.51
Six-trade half-life	1.91	2.26	1.82 to 2.43	1.86 to 2.66

Table 4: Hill estimator of the tail index of the Huber filter residual, PYPL, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 10.1 to 19.9 cents over the days and removes 0% to 4% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 5.0 cents on 16 October 2024, 92% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

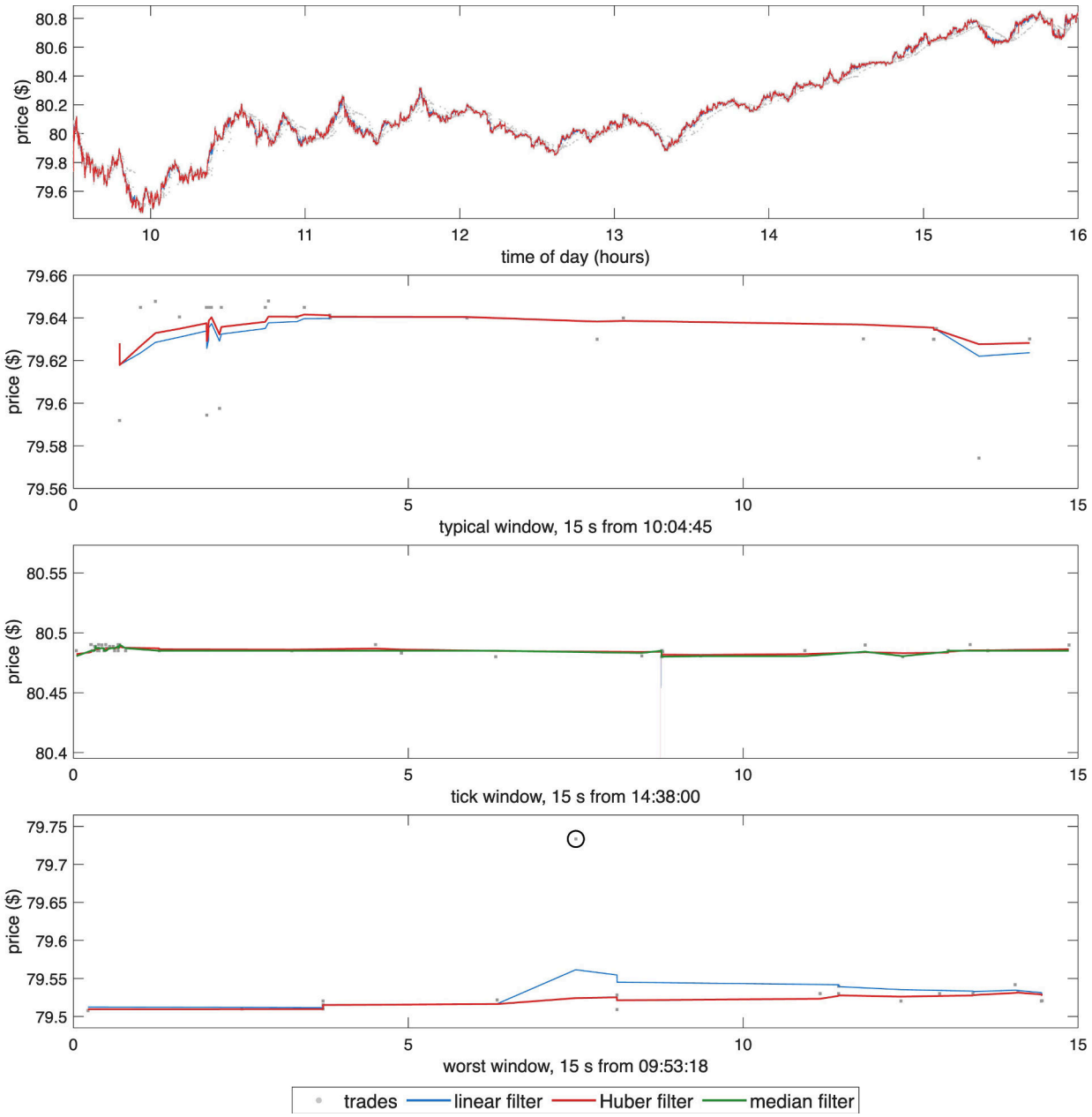

Figure 1: PYPL trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:04:45. 3rd: 15-second window at 14:38:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

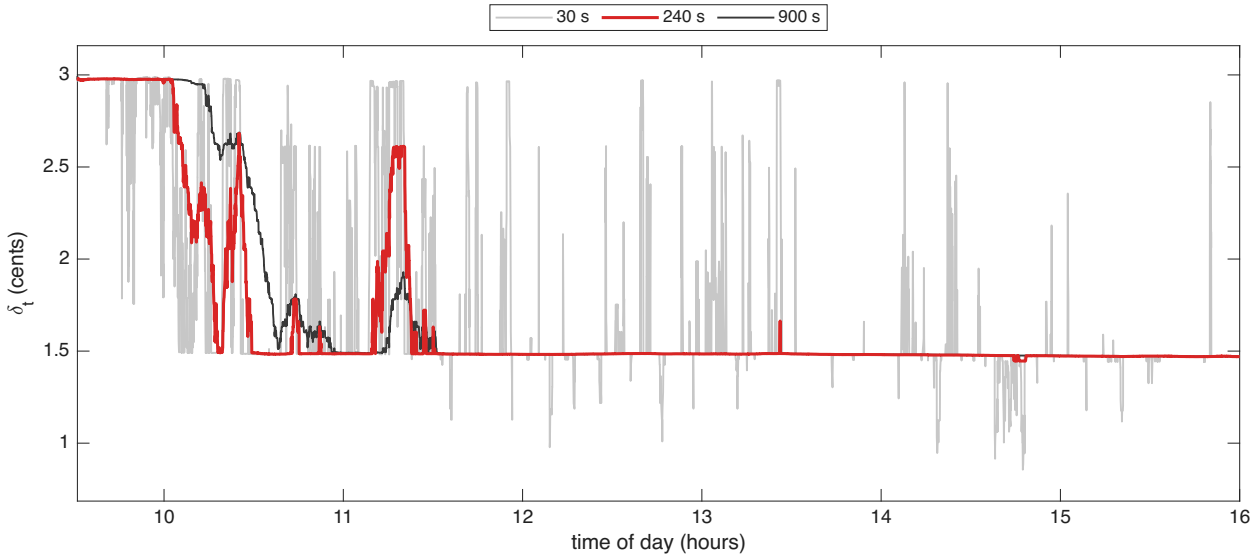

Figure 2: The Huber threshold δ_t in cents for PYPL on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

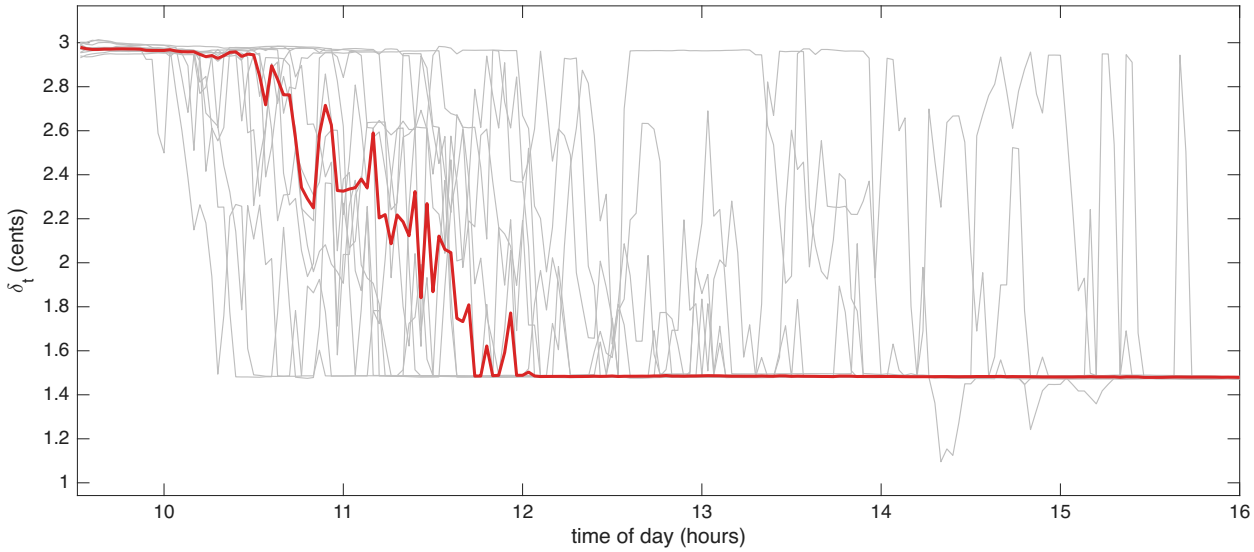

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 PYPL trading days in Oct 2024 (grey lines) and their median (red line).

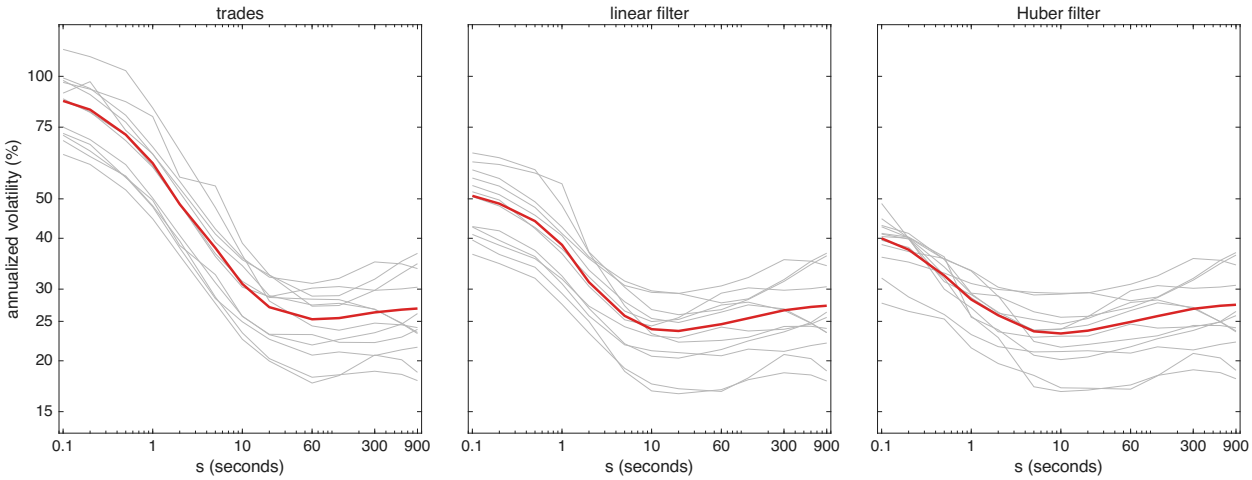

Figure 4: Volatility signature curves for PYPL over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

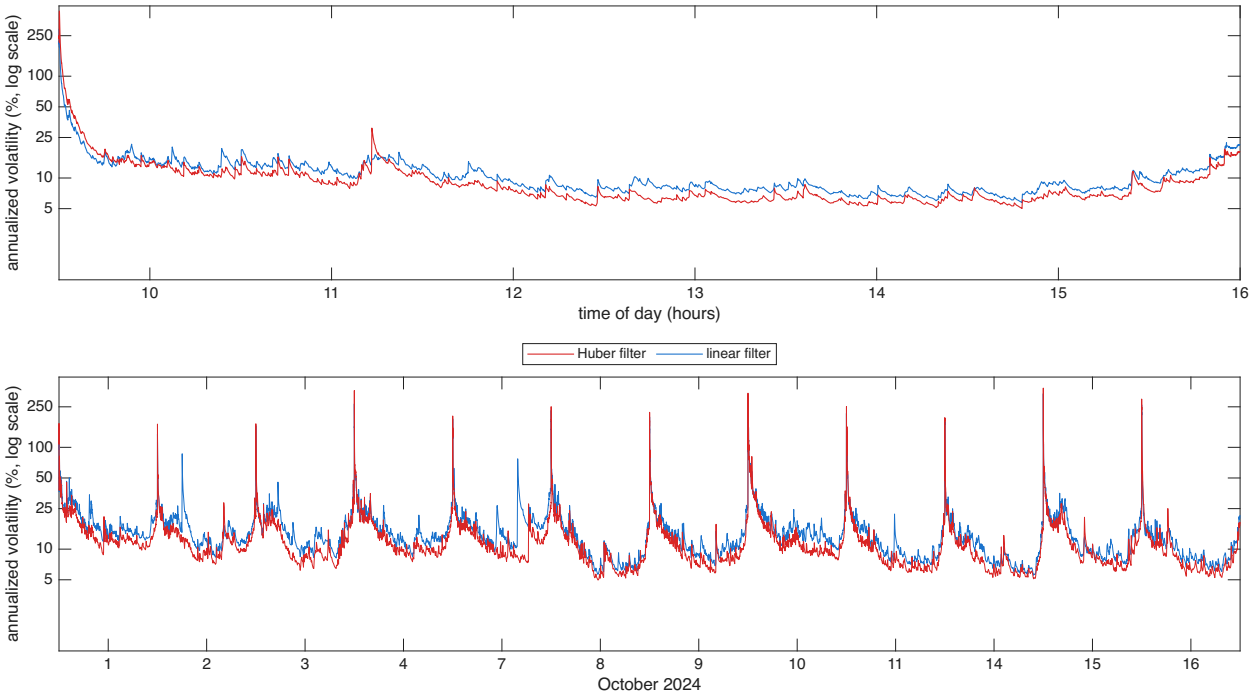

Figure 5: Filtered spot volatility for PYPL over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 95$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

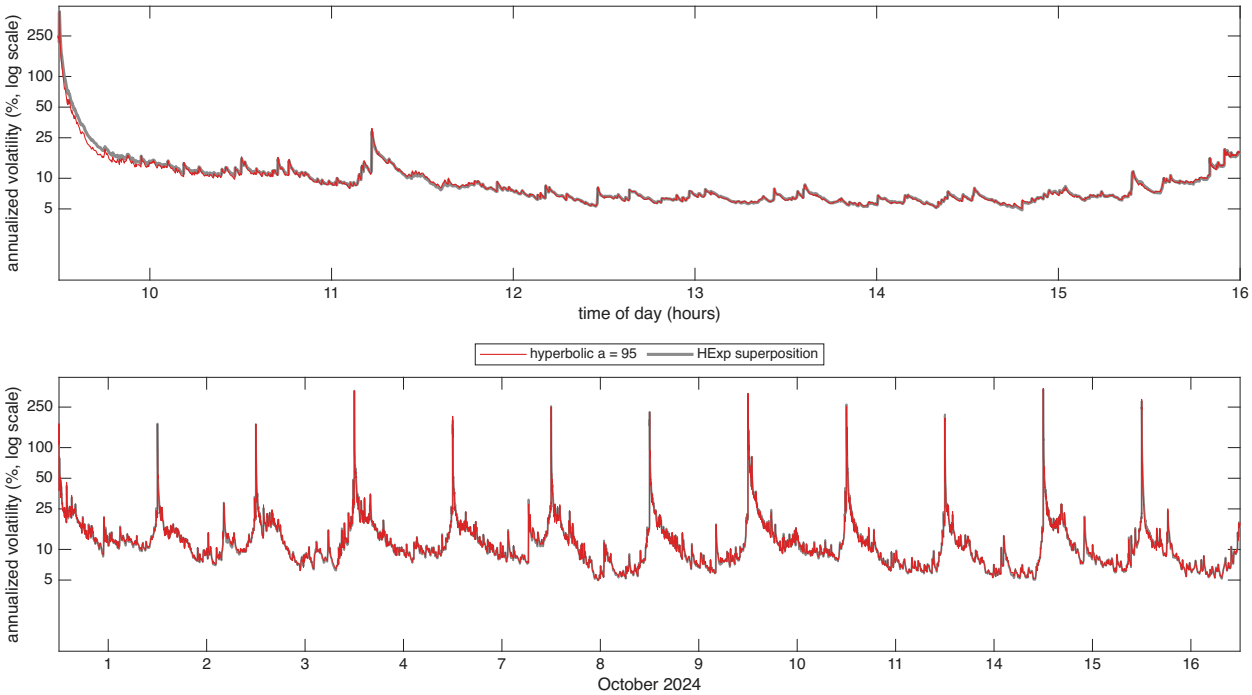

Figure 6: Hyperbolic against superposition weights for PYPL over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 95$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

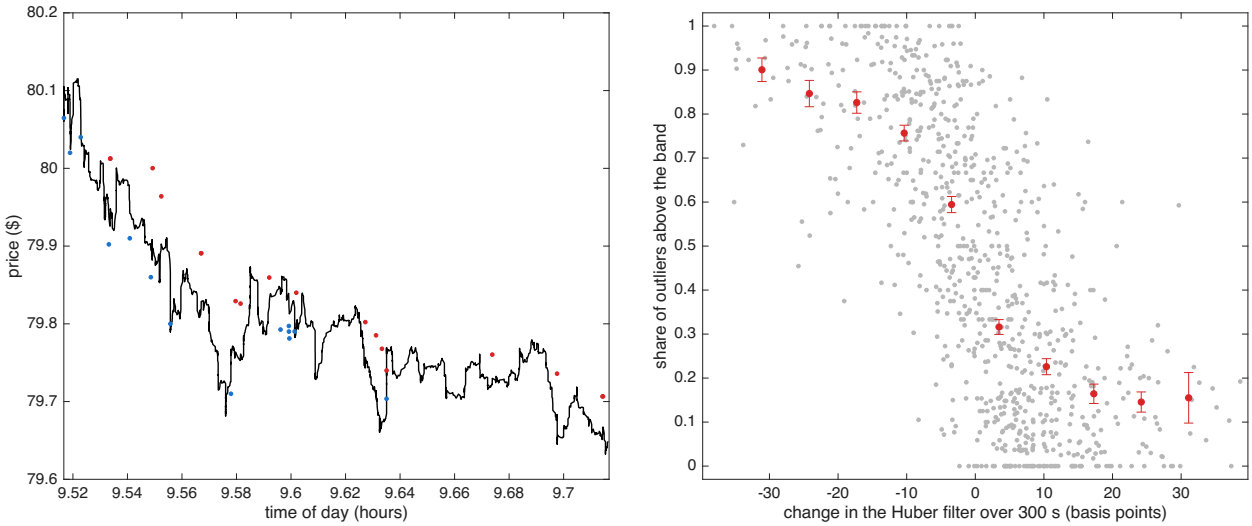

Figure 7: Outliers & robust price for PYPL on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

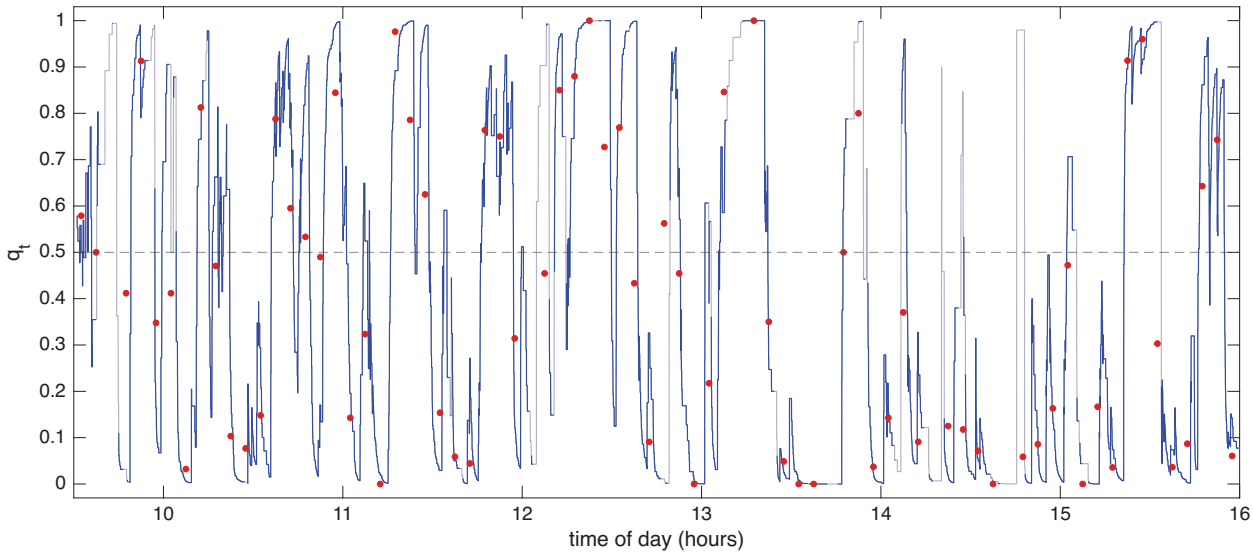

Figure 8: The directional conditional probability for PYPL, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 42 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:53.

Diagnostics.

CSCO 1 to 16 October 2024

trades on 16 October 2024	173,638
average trading rate	4.0 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,382 (1.37%)
price half-life: run / minimizer[1]	3.0 / 2.4 trades
cost of running 3 trades	+0.58%
scale half-life: run / minimizer[2]	240 / 120 s
cost of running 240 s	+0.05%
direction half-life[3]	34 s [29, 42]
Mandelbrot a minimizing QLIKE	42 trades
a within 0.0005 of the minimum	30 to 56
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	16% $\rightarrow$ 268%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	7% $\rightarrow$ 12%
median spot volatility, Huber / linear input[5]	9.4 / 12.0 %
hyperbolic against HExp, correlation of logs	0.9968
HExp QLIKE advantage, leave-one-day-out[6]	-0.00434
15-minute signature level	9 to 26 %
Huber signature, 1 min over 1 s, 12 day average	1.35

CSCO over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:57; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.072	0.062	0.073	0.310	0.266	0.140	0.292	1.811
2	0.072	0.072	0.087	0.254	0.236	0.148	0.288	1.308
5	0.084	0.083	0.100	0.193	0.209	0.156	0.267	0.845
10	0.092	0.090	0.105	0.161	0.192	0.160	0.238	0.609
25	0.101	0.099	0.108	0.138	0.178	0.163	0.201	0.405
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.13	−0.03	−0.48	2.08	2.01	2.03	2.16
2	0.00	0.01	−0.05	−0.01	1.95	1.96	1.99	2.19
5	−0.01	0.02	−0.04	0.00	2.30	2.07	1.84	2.21
10	−0.01	0.00	−0.02	−0.01	2.31	2.26	1.94	2.26
25	0.00	0.01	0.00	0.00	2.99	3.03	2.31	2.36

Table 1: Trade-lag differences, `CSCO`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.07	1.01	1.10	1.70	19.3	14.2	33.6	204.1
1 s	0.70	0.68	0.73	0.96	18.5	16.5	29.0	144.6
5 s	0.53	0.53	0.56	0.69	19.9	19.3	25.7	55.4
10 s	0.50	0.50	0.51	0.59	20.5	20.1	24.1	50.5
1 min	0.51	0.51	0.51	0.53	23.2	22.9	22.9	27.4
15 min	0.58	0.58	0.58	0.59	26.0	25.9	25.8	26.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.09	−0.01	−0.08	−0.45	4.87	4.60	3.88	2.57
1 s	−0.21	−0.13	−0.12	−0.54	6.83	6.90	4.40	2.68
5 s	0.06	0.08	−0.11	−0.51	7.70	7.35	3.94	3.54
10 s	0.09	0.11	0.03	−0.26	6.69	6.52	4.49	2.55
1 min	0.21	0.21	0.22	0.13				

Table 2: Clock-lag differences, `CSCO`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.21	0.26	0.28
s.d. of $\hat{\epsilon}_t$ in cents	1.32	1.29	1.11
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.23	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.13	−0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.04	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.17,0.18	−0.18,0.20
95% C.I. for $\theta_t - X_t$	—	−0.36,0.36	−0.42,0.42
99% C.I. for $\theta_t - X_t$	—	−0.67,0.67	−0.90,0.84
s.e. of $\theta_t - X_t$ in cents	—	0.17	0.33

Table 3: Residual summaries, `CSCO`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.03, is above two, while the Hill index of the Huber residual itself on this day is 1.56, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.02	1.56	1.22 to 2.77	1.56 to 2.82
Residual divided by δ_t	2.29	1.57	1.26 to 2.59	1.57 to 2.79
Repricing windows removed	2.08	1.51	1.22 to 2.77	1.51 to 2.86
Calm windows only	2.07	1.61		
Every 10th trade	1.75	1.57	1.28 to 2.32	1.57 to 2.67
Six-trade half-life	2.16	1.66	1.30 to 3.01	1.66 to 3.15

Table 4: Hill estimator of the tail index of the Huber filter residual, `CSCO`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 10.1 to 10.1 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 5.0 cents on 16 October 2024, 86% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

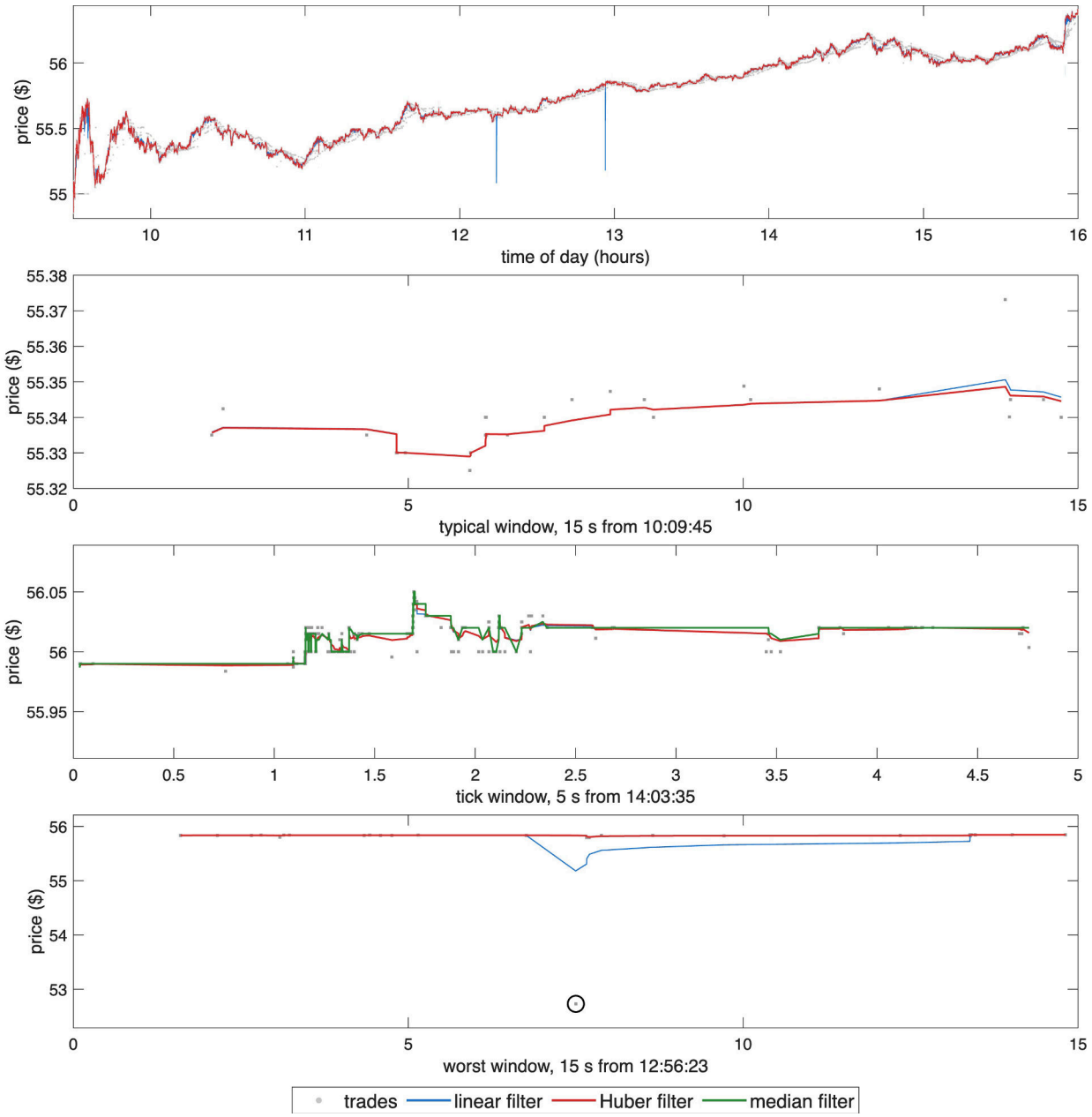

Figure 1: CSCO trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:09:45. 3rd: 5-second window at 14:03:35 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

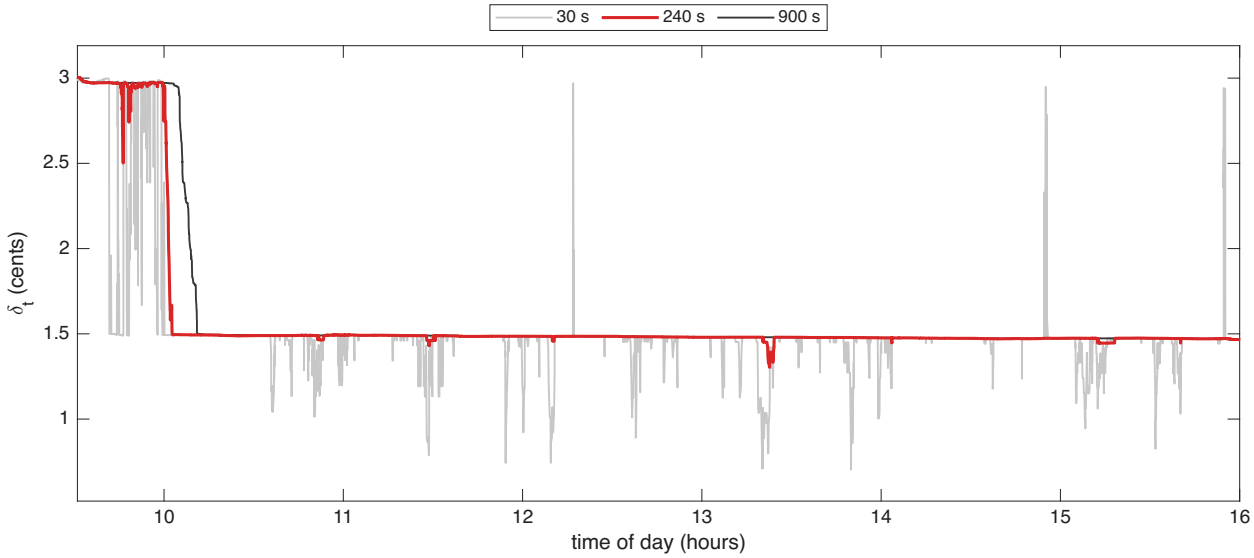

Figure 2: The Huber threshold δ_t in cents for `CSCO` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

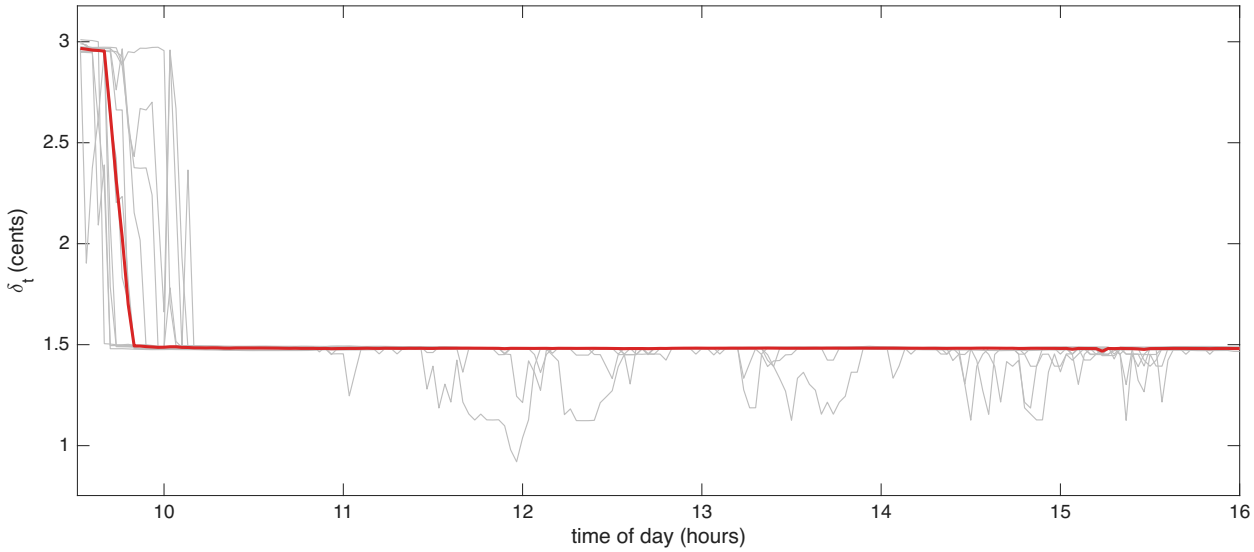

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `CSCO` trading days in Oct 2024 (grey lines) and their median (red line).

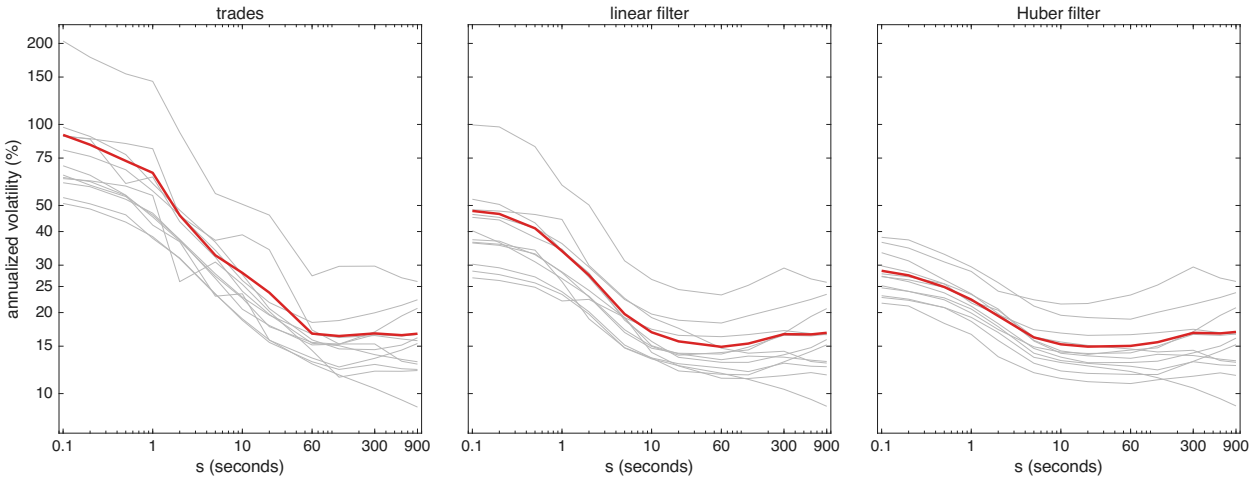

Figure 4: Volatility signature curves for CSCO over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

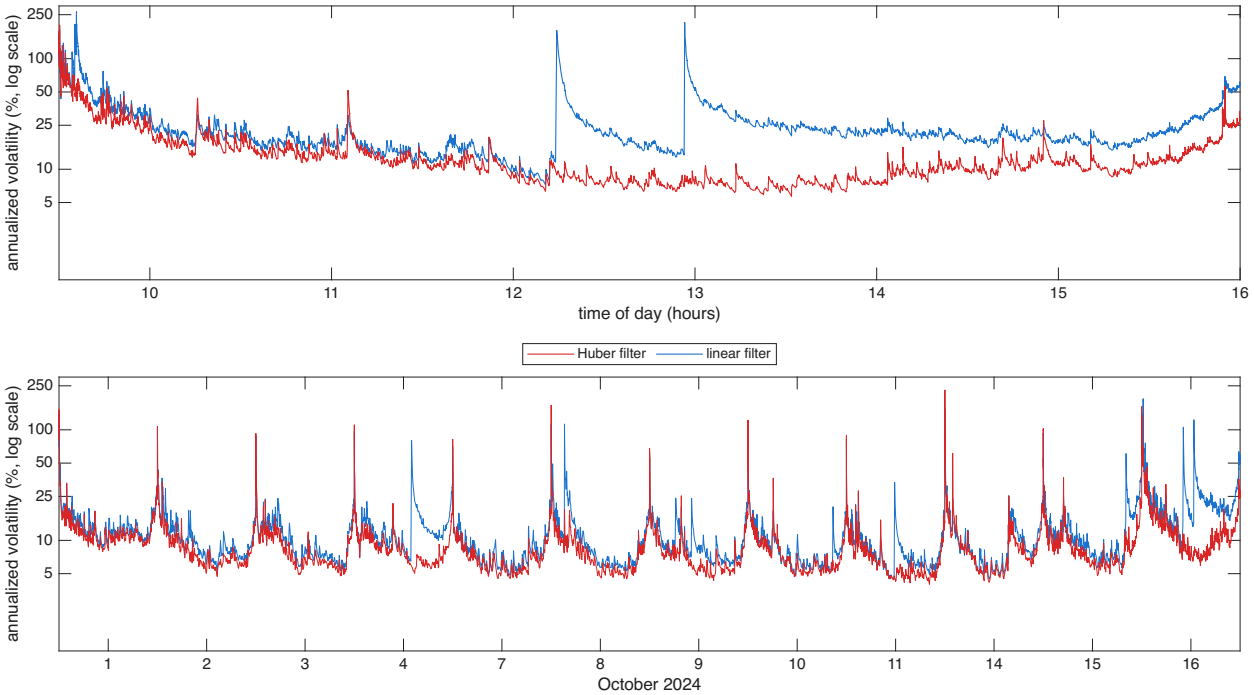

Figure 5: Filtered spot volatility for CSCO over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 42$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

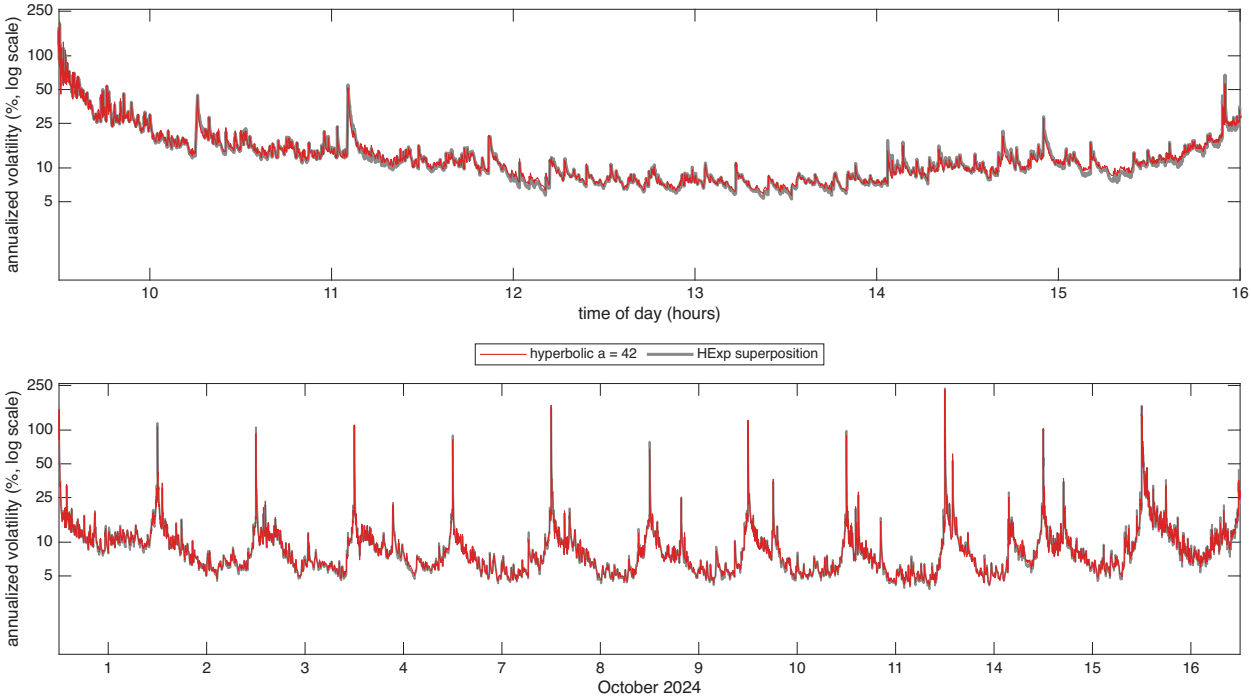

Figure 6: Hyperbolic against superposition weights for `CSCO` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 42$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

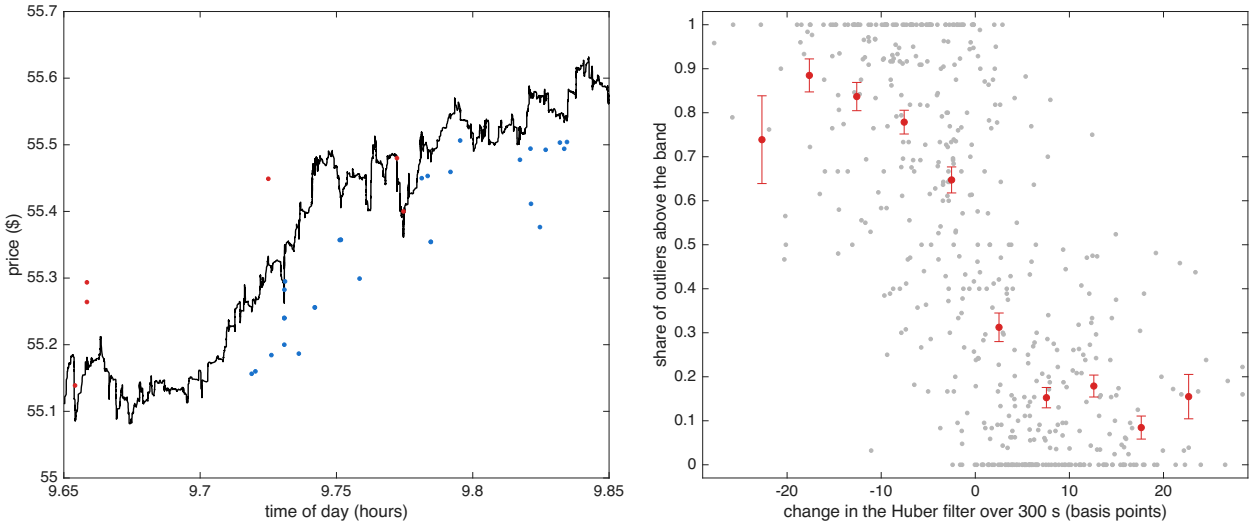

Figure 7: Outliers & robust price for `CSCO` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

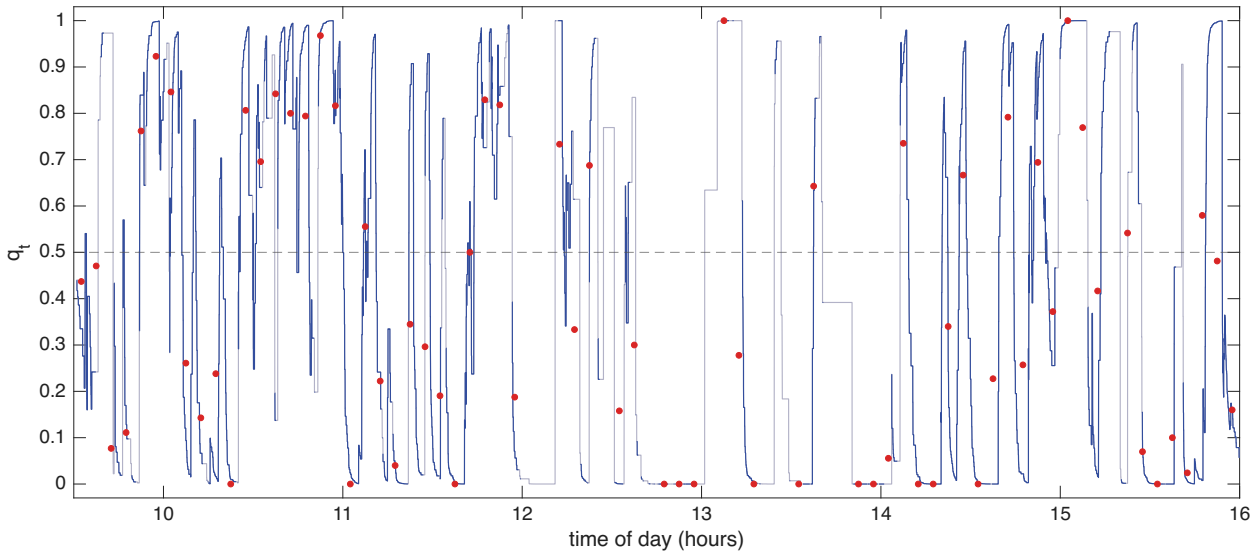

Figure 8: The directional conditional probability for CSCO, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 34 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

Diagnostics.

PEP 1 to 16 October 2024

trades on 16 October 2024	74,962
average trading rate	4.0 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,583 (2.11%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+1.96%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.03%
direction half-life[3]	70 s [58, 88]
Mandelbrot a minimizing QLIKE	148 trades
a within 0.0005 of the minimum	100 to 204
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	4% $\to$ 59%
Huber $\hat{\sigma}$, before $\to$ peak	4% $\to$ 4%
median spot volatility, Huber / linear input[5]	8.9 / 11.8 %
hyperbolic against HExp, correlation of logs	0.9981
HExp QLIKE advantage, leave-one-day-out[6]	-0.00163
15-minute signature level	10 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.24

PEP over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:57; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.212	0.173	0.193	0.794	0.535	0.289	0.478	2.892
2	0.220	0.204	0.232	0.671	0.495	0.325	0.487	2.111
5	0.271	0.249	0.280	0.530	0.483	0.380	0.492	1.387
10	0.311	0.288	0.310	0.469	0.487	0.427	0.495	1.042
25	0.350	0.334	0.342	0.424	0.509	0.483	0.508	0.770
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.14	0.26	0.04	−0.47	2.64	2.45	2.68	3.11
2	0.05	0.16	0.00	−0.02	2.68	2.44	2.86	3.17
5	0.00	0.08	0.00	0.00	3.16	2.95	3.09	3.31
10	0.00	0.03	0.00	0.00	3.80	3.59	2.85	3.34
25	0.00	0.00	0.00	0.01	3.97	3.88	3.13	3.60

Table 1: Trade-lag differences, PEP, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.41	1.26	1.33	2.60	8.8	5.9	11.0	57.3
1 s	1.01	0.94	0.99	1.61	8.9	7.6	12.0	48.8
5 s	0.86	0.82	0.85	1.13	11.3	10.8	11.8	25.7
10 s	0.86	0.84	0.85	1.03	12.1	11.8	12.3	22.4
1 min	0.83	0.82	0.82	0.87	12.3	12.2	12.1	14.3
15 min	0.64	0.63	0.63	0.65	9.8	9.8	9.8	10.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.08	−0.05	−0.07	−0.19	5.64	4.66	3.64	4.31
1 s	−0.02	−0.01	−0.03	−0.28	4.92	4.97	4.24	4.58
5 s	0.07	0.10	0.07	−0.24	5.86	5.58	4.85	4.86
10 s	−0.01	0.01	0.00	−0.16	7.57	7.85	6.95	5.75
1 min	−0.15	−0.14	−0.14	−0.17				

Table 2: Clock-lag differences, PEP, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.57	0.71	0.74
s.d. of $\hat{\epsilon}_t$ in cents	2.10	2.06	1.83
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.68
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.26	0.32	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.11	0.21	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.05	0.11	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.05	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	−0.55,0.52	−0.59,0.56
95% C.I. for $\theta_t - X_t$	—	−0.98,0.98	−1.09,1.09
99% C.I. for $\theta_t - X_t$	—	−1.57,1.70	−1.96,2.09
s.e. of $\theta_t - X_t$ in cents	—	0.48	0.64

Table 3: Residual summaries, PEP, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.68, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.13	2.49	1.50 to 2.48	1.70 to 2.54
Residual divided by δ_t	2.09	2.51	1.74 to 2.54	1.63 to 3.05
Repricing windows removed	2.08	2.48	1.69 to 2.45	1.70 to 2.69
Calm windows only	2.02	2.30		
Every 10th trade	1.77	2.21	1.50 to 2.16	1.60 to 2.49
Six-trade half-life	2.35	2.73	1.50 to 2.68	1.88 to 2.73

Table 4: Hill estimator of the tail index of the Huber filter residual, PEP, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.0 to 20.2 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 10.1 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

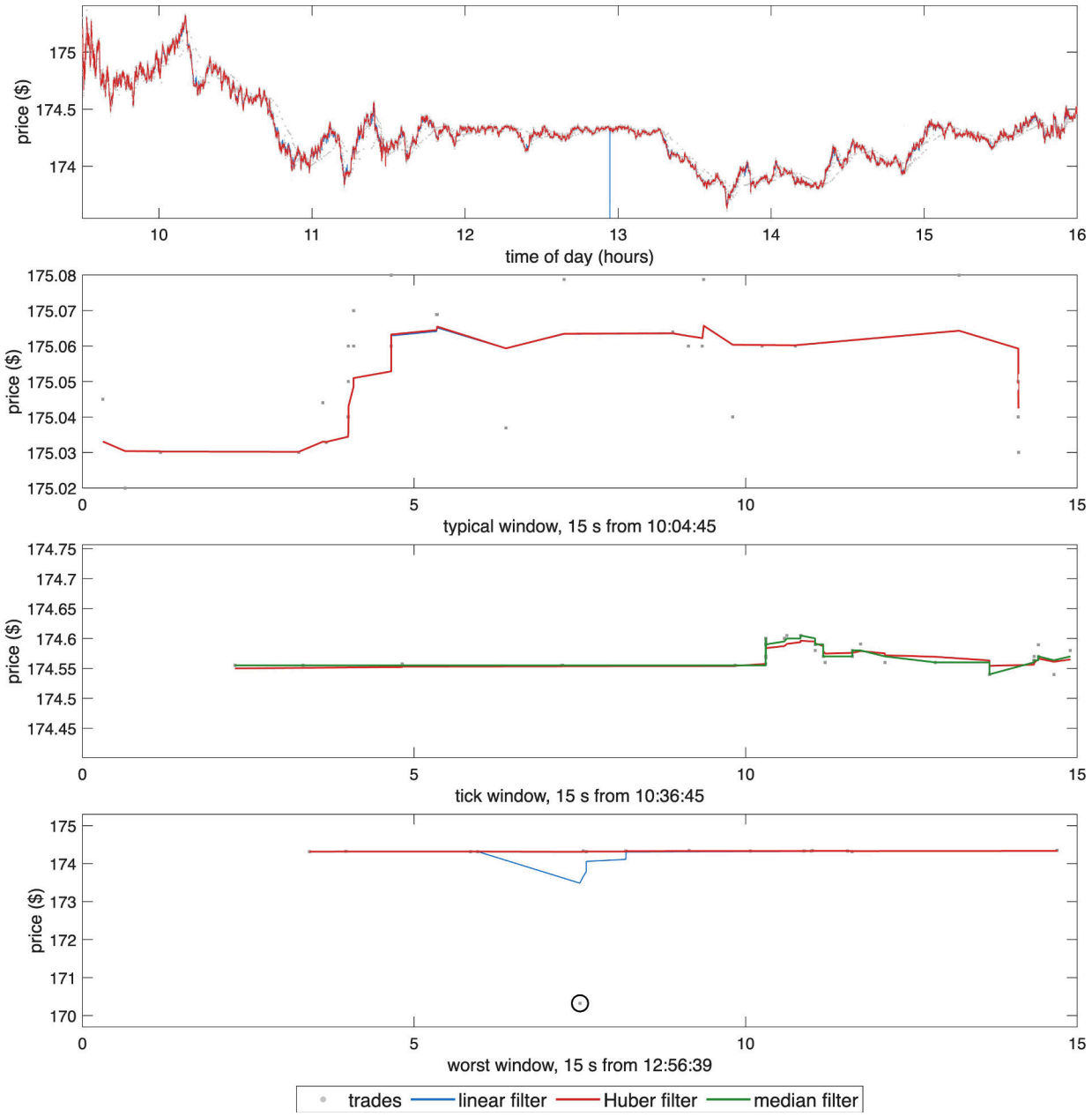

Figure 1: PEP trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:04:45. 3rd: 15-second window at 10:36:45 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

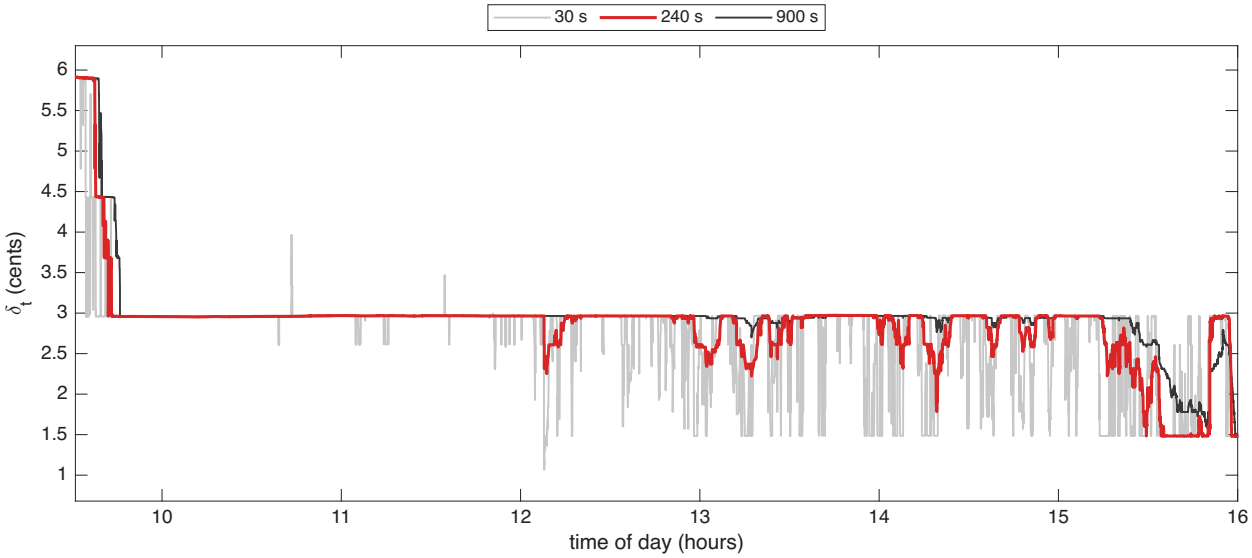

Figure 2: The Huber threshold δ_t in cents for `PEP` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

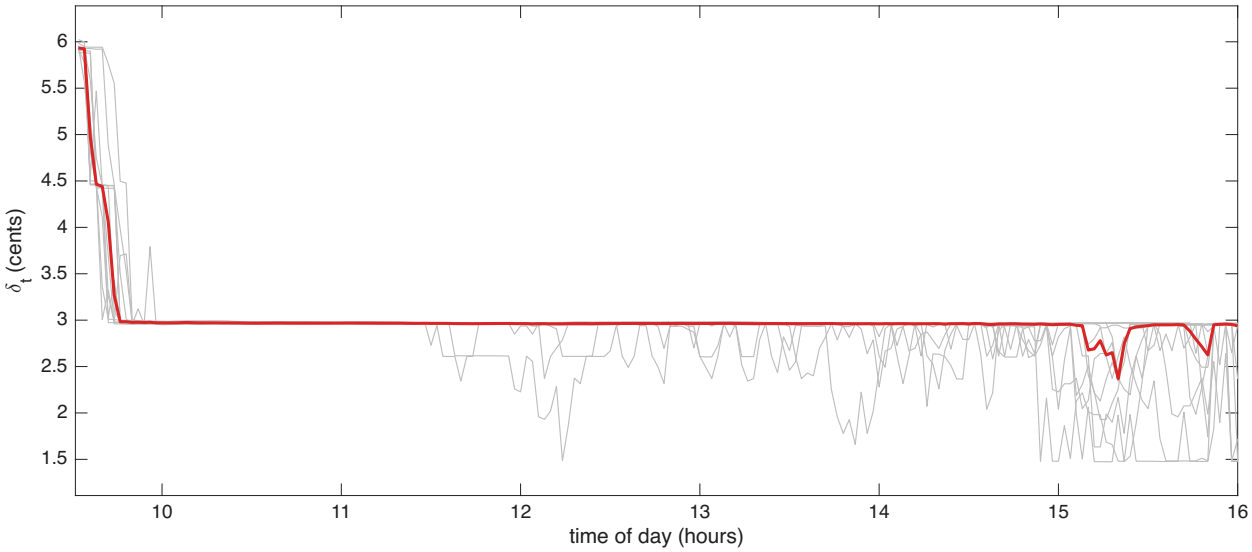

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `PEP` trading days in Oct 2024 (grey lines) and their median (red line).

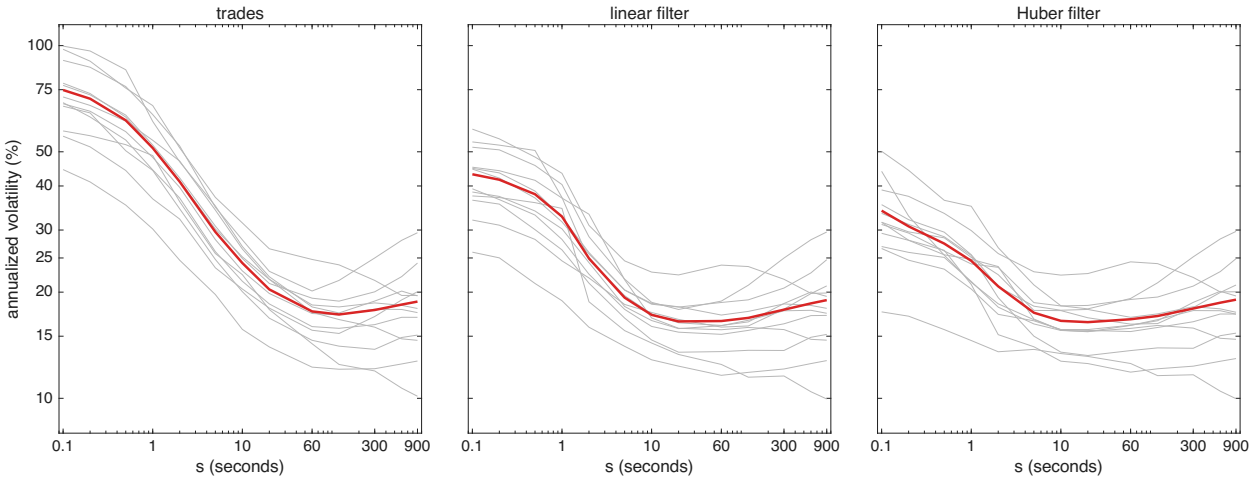

Figure 4: Volatility signature curves for PEP over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

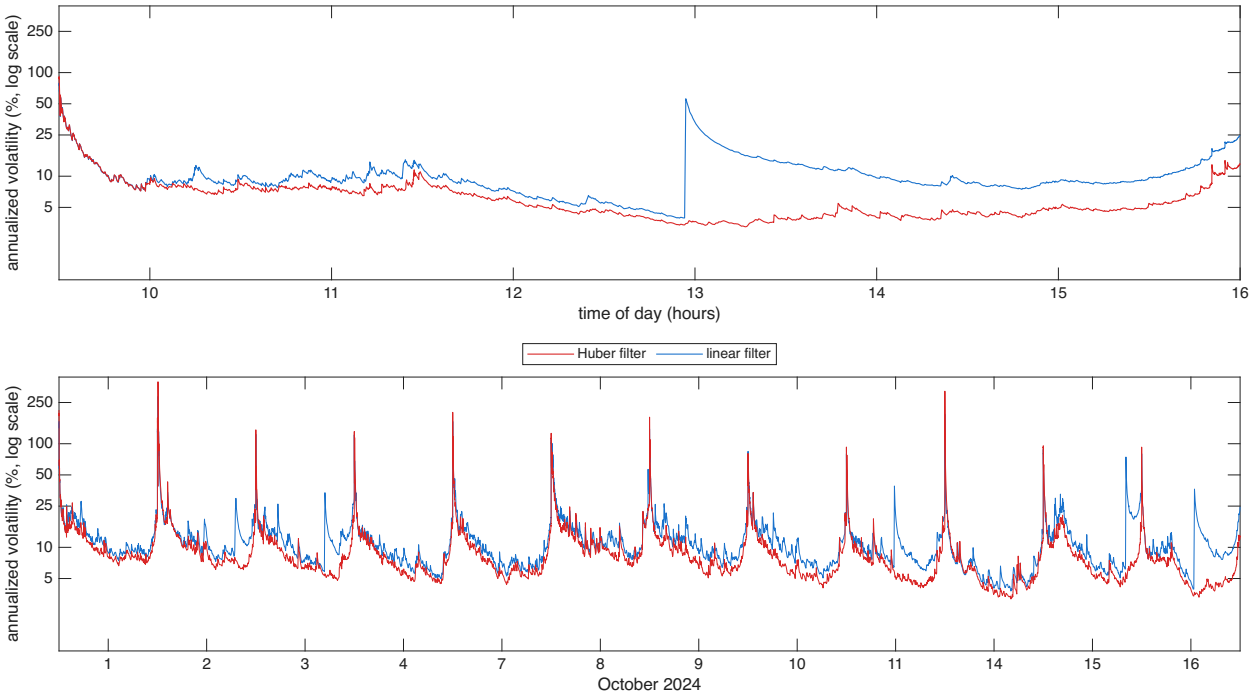

Figure 5: Filtered spot volatility for PEP over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 148$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

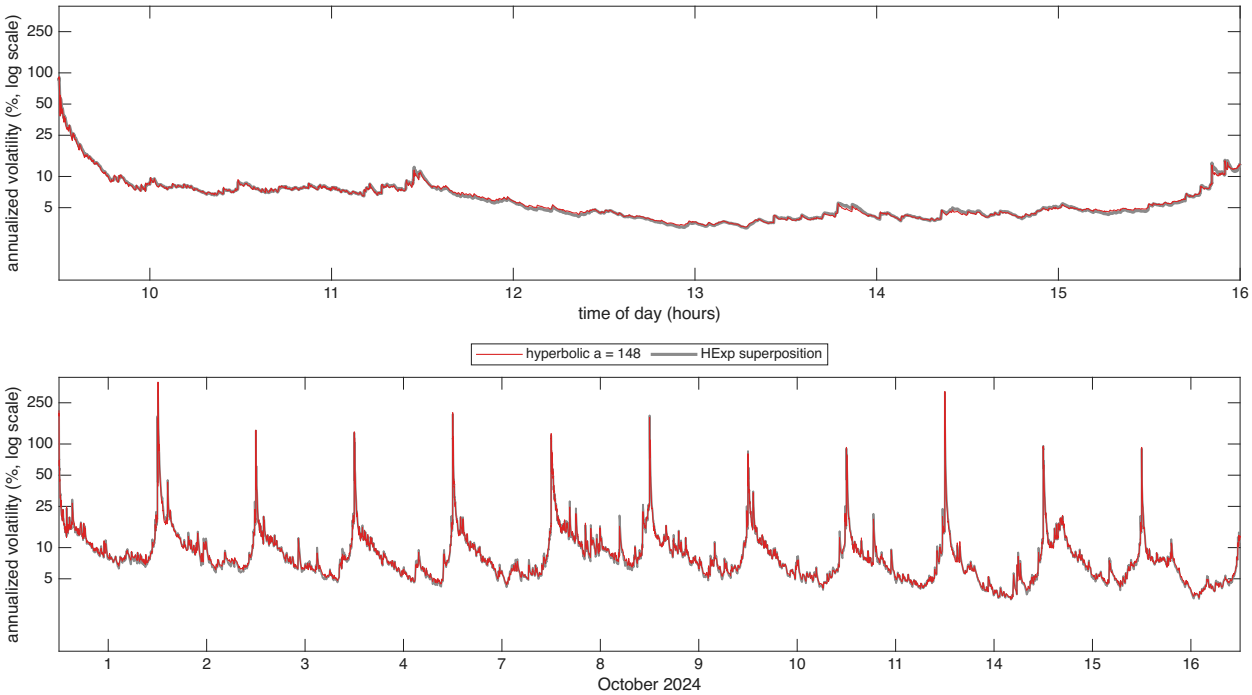

Figure 6: Hyperbolic against superposition weights for `PEP` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 148$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

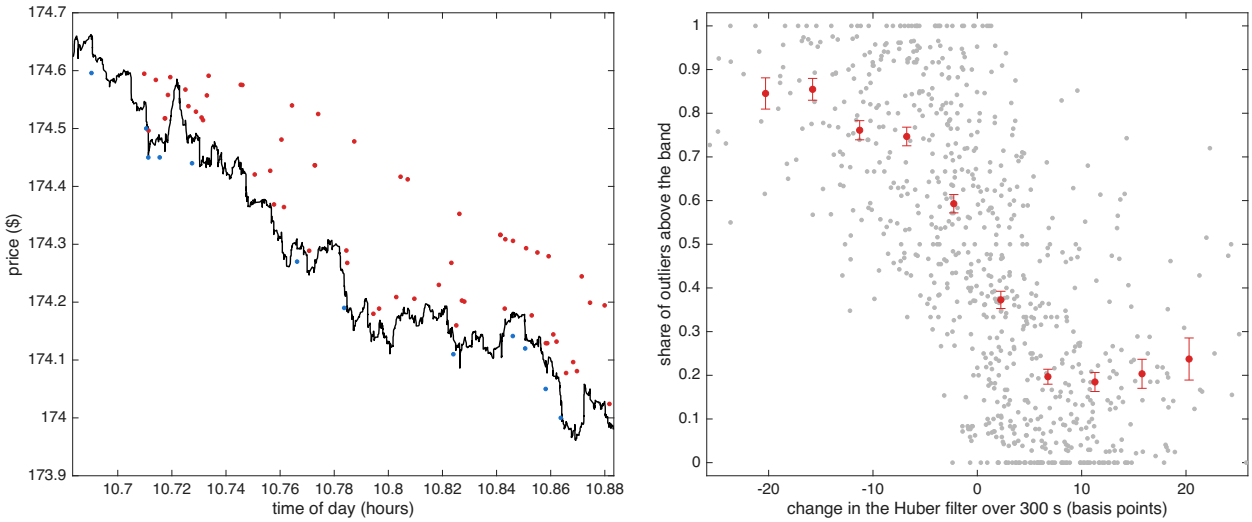

Figure 7: Outliers & robust price for `PEP` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

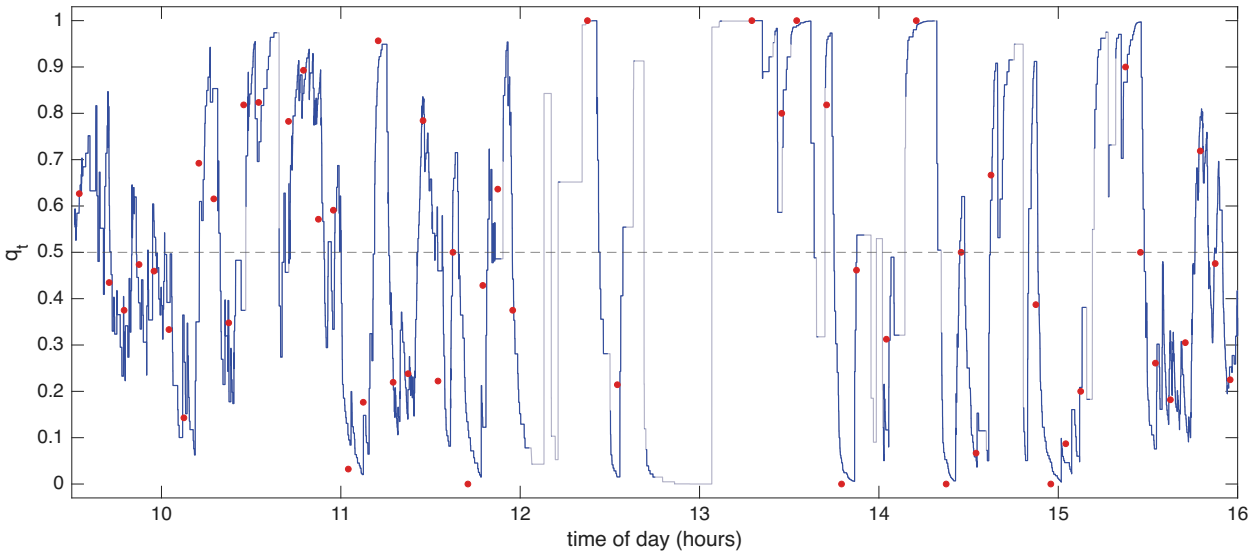

Figure 8: The directional conditional probability for PEP, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 70 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 20 to 26112 trades, against 148 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 40 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

Diagnostics.

AMAT 1 to 16 October 2024

trades on 16 October 2024	181,209
average trading rate	3.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	6,827 (3.77%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.80%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.01%
direction half-life[3]	21 s [17, 26]
Mandelbrot a minimizing QLIKE	30 trades
a within 0.0005 of the minimum	21 to 42
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	17% $\rightarrow$ 434%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	16% $\rightarrow$ 26%
median spot volatility, Huber / linear input[5]	24.4 / 29.1 %
hyperbolic against HExp, correlation of logs	0.9833
HExp QLIKE advantage, leave-one-day-out[6]	+0.00201
15-minute signature level	20 to 103 %
Huber signature, 1 min over 1 s, 12 day average	1.35

`AMAT` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.510	0.403	0.426	1.872	1.759	1.204	1.415	8.954
2	0.496	0.449	0.491	1.549	1.551	1.172	1.413	6.490
5	0.573	0.516	0.564	1.175	1.437	1.179	1.357	4.226
10	0.617	0.563	0.600	0.975	1.321	1.148	1.263	3.074
25	0.663	0.622	0.637	0.829	1.157	1.064	1.103	2.084
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.22	−0.05	0.00	−0.47	2.20	1.80	1.90	1.98
2	0.05	0.03	−0.03	−0.01	2.10	1.79	1.85	1.91
5	−0.05	−0.06	−0.03	0.00	1.83	1.66	1.72	1.89
10	0.01	0.01	−0.01	0.00	1.66	1.62	1.70	1.89
25	0.00	0.00	0.00	0.00	1.77	1.79	1.97	1.93

Table 1: Trade-lag differences, `AMAT`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	5.68	5.27	5.40	8.38	39.9	33.2	43.0	266.0
1 s	3.86	3.65	3.74	5.20	34.5	31.1	42.4	201.9
5 s	3.19	3.09	3.14	3.66	36.9	35.8	44.5	105.4
10 s	3.03	2.96	2.98	3.39	39.0	38.4	44.2	97.0
1 min	3.23	3.21	3.21	3.34	44.5	44.2	44.8	63.0
15 min	4.05	4.04	4.03	4.06	56.5	56.4	56.2	56.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.15	−0.14	−0.11	−0.36	1.95	2.08	2.95	2.45
1 s	−0.21	−0.16	−0.10	−0.36	2.16	2.22	2.58	2.64
5 s	−0.03	−0.02	−0.08	−0.17	3.41	3.42	3.01	3.72
10 s	0.08	0.09	0.03	−0.42	4.85	4.76	3.78	3.31
1 min	0.05	0.05	0.04	−0.19				

Table 2: Clock-lag differences, `AMAT`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.27	1.54	1.64
s.d. of $\hat{\epsilon}_t$ in cents	6.47	6.38	5.56
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	1.91
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.24	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.13	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.03	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−1.03,1.00	−1.14,1.10
95% C.I. for $\theta_t - X_t$	—	−1.94,1.98	−2.23,2.26
99% C.I. for $\theta_t - X_t$	—	−3.10,3.17	−3.97,4.04
s.e. of $\theta_t - X_t$ in cents	—	0.91	∞(1.73)

Table 3: Residual summaries, `AMAT`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. does not exist in theory, the sample value is in brackets.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.73	1.83	1.60 to 3.23	1.83 to 3.21
Residual divided by δ_t	2.02	1.98	1.63 to 3.39	1.95 to 3.03
Repricing windows removed	2.00	2.03	1.68 to 3.38	2.03 to 3.22
Calm windows only	1.98	1.99		
Every 10th trade	1.96	1.55	1.53 to 3.72	1.55 to 2.59
Six-trade half-life	1.71	1.91	1.66 to 3.50	1.91 to 3.45

Table 4: Hill estimator of the tail index of the Huber filter residual, `AMAT`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 30.1 to 60.4 cents over the days and removes 0% to 9% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 15.0 cents on 16 October 2024, 66% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

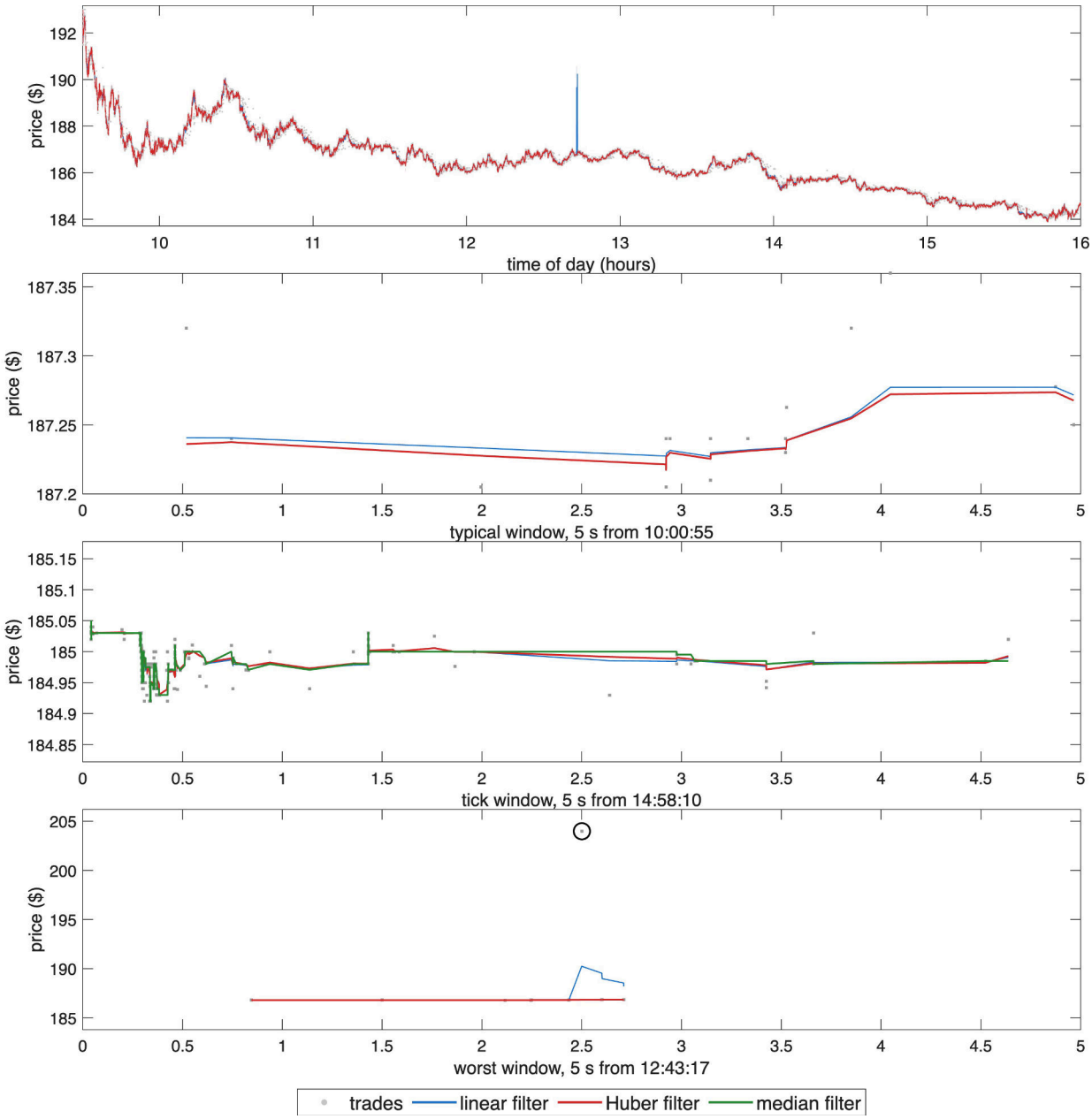

Figure 1: `AMAT` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 5-second window at the day's median trade count, from 10:00:55. 3rd: 5-second window at 14:58:10 in which the price reverses direction most often, with the median filter m_t added. Bottom: 5-second window around the largest Huber residual after 09:45, circled.

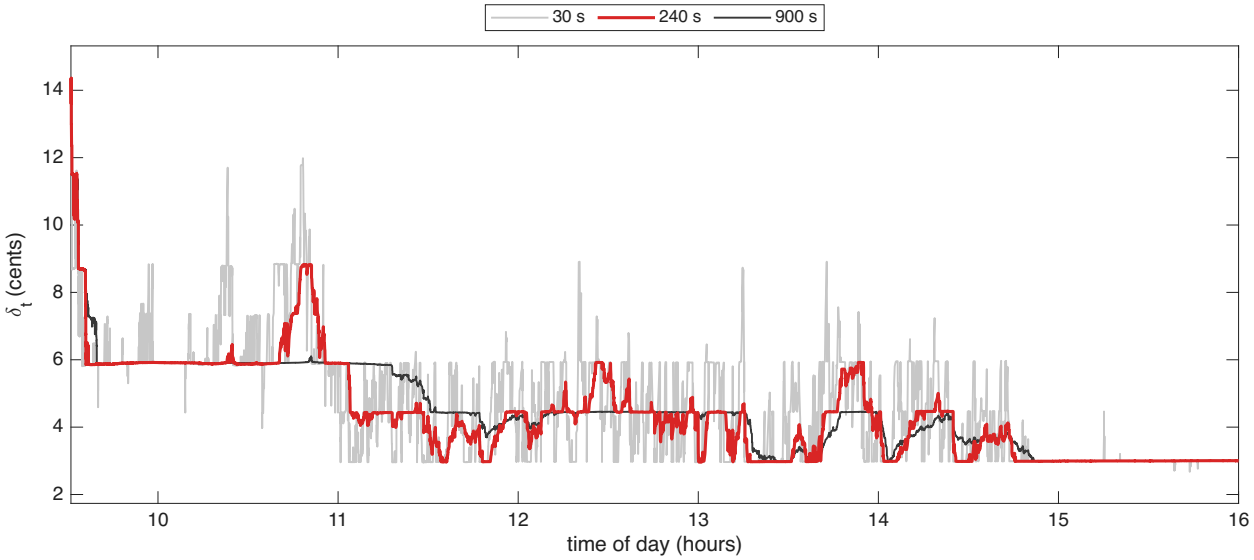

Figure 2: The Huber threshold δ_t in cents for `AMAT` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

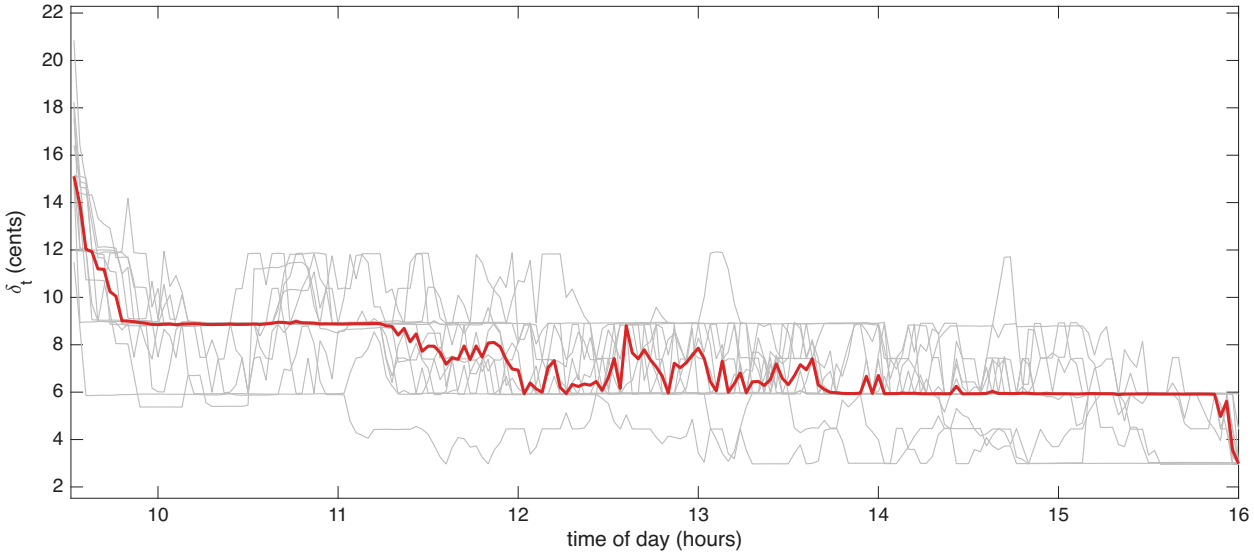

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AMAT` trading days in Oct 2024 (grey lines) and their median (red line).

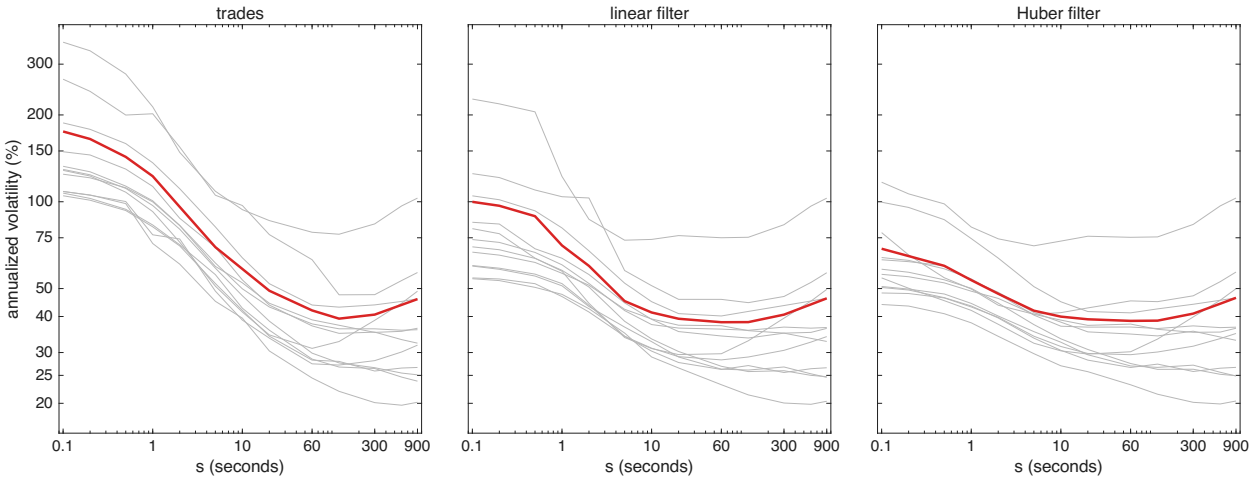

Figure 4: Volatility signature curves for `AMAT` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

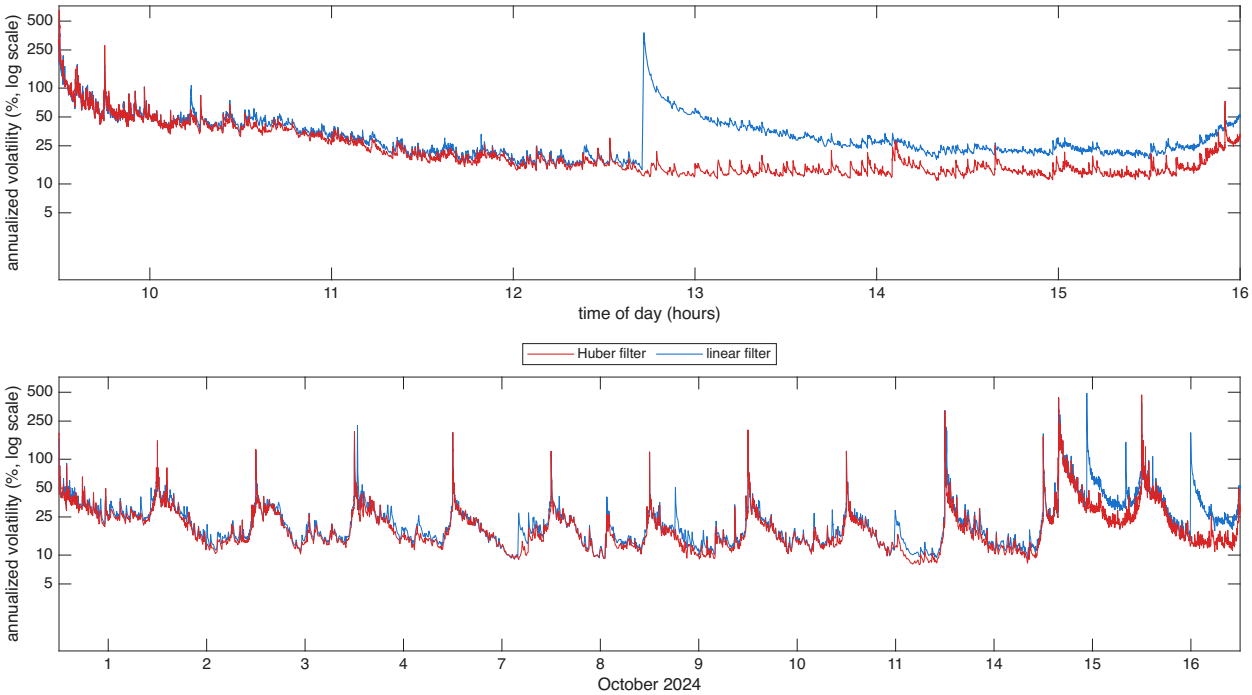

Figure 5: Filtered spot volatility for `AMAT` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 30$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

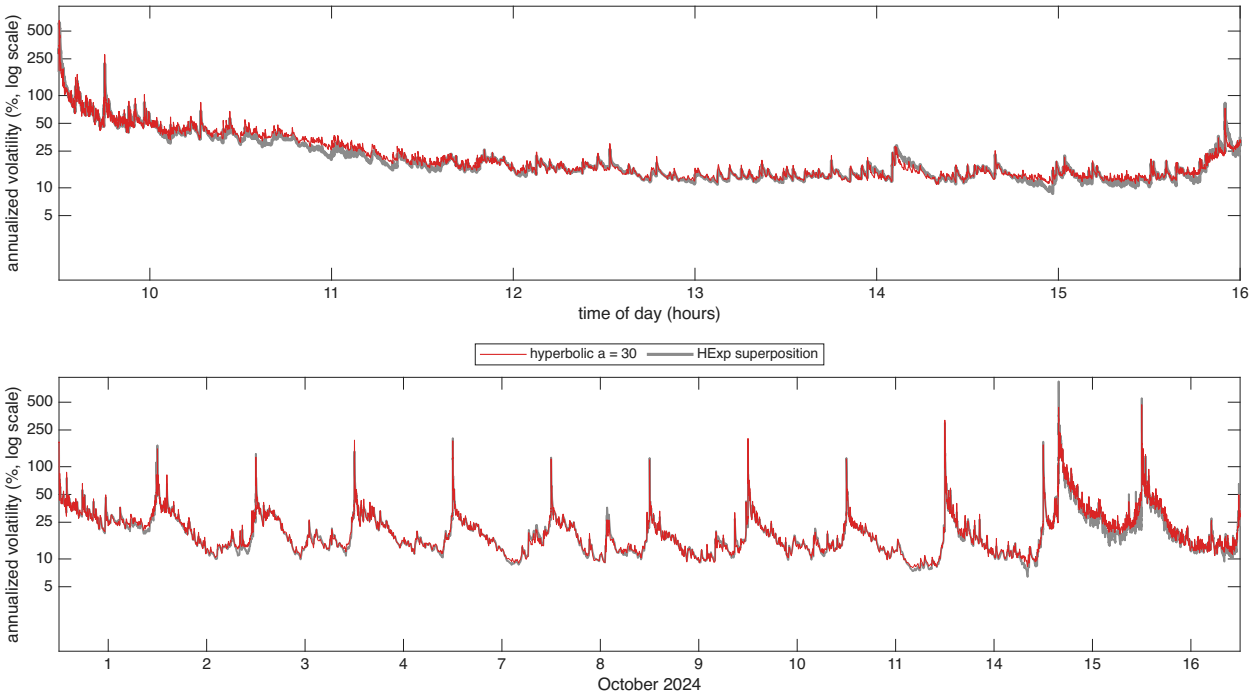

Figure 6: Hyperbolic against superposition weights for `AMAT` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 30$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

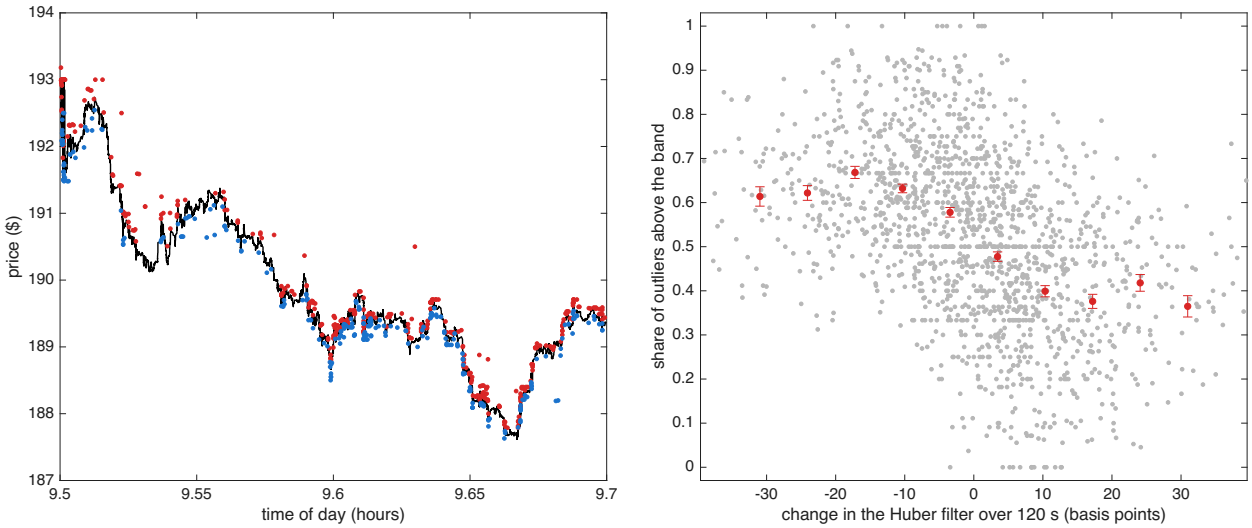

Figure 7: Outliers & robust price for `AMAT` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

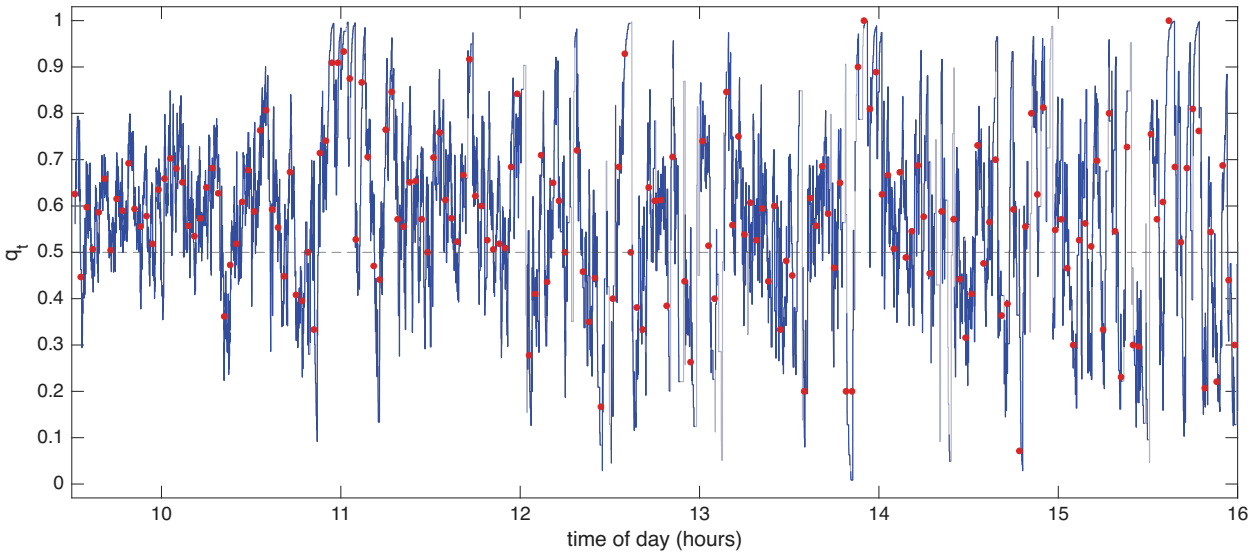

Figure 8: The directional conditional probability for AMAT, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 21 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

Selecting a on single days gives values from 9 to 408 trades, against 30 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 39 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

Diagnostics.

SBUX 1 to 16 October 2024

trades on 16 October 2024	96,945
average trading rate	3.7 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,558 (1.61%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+1.98%
scale half-life: run / minimizer[2]	240 / 60 s
cost of running 240 s	+0.13%
direction half-life[3]	57 s [47, 60]
Mandelbrot a minimizing QLIKE	53 trades
a within 0.0005 of the minimum	38 to 78
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	12% $\to$ 82%
Huber $\hat{\sigma}$, before $\to$ peak	11% $\to$ 10%
median spot volatility, Huber / linear input[5]	11.0 / 12.8 %
hyperbolic against HExp, correlation of logs	0.9984
HExp QLIKE advantage, leave-one-day-out[6]	-0.00436
15-minute signature level	16 to 28 %
Huber signature, 1 min over 1 s, 12 day average	1.39

SBUX over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:57; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.134	0.119	0.137	0.549	0.431	0.237	0.357	2.176
2	0.142	0.140	0.165	0.463	0.387	0.249	0.364	1.598
5	0.175	0.169	0.199	0.364	0.362	0.276	0.361	1.042
10	0.203	0.194	0.217	0.320	0.354	0.303	0.357	0.778
25	0.239	0.231	0.242	0.297	0.364	0.341	0.363	0.568
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.19	0.11	0.04	−0.46	2.11	1.94	2.33	2.46
2	0.03	0.10	−0.01	−0.02	2.26	2.01	2.50	2.51
5	0.02	0.06	0.00	0.01	2.62	2.39	2.84	2.59
10	0.01	0.03	0.00	0.00	2.94	2.62	2.74	2.59
25	0.00	0.01	0.00	−0.01	3.02	3.09	2.87	2.74

Table 1: Trade-lag differences, `SBUX`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.07	1.00	1.09	2.10	14.8	10.4	17.4	105.0
1 s	0.75	0.73	0.80	1.25	15.2	12.8	16.9	58.6
5 s	0.70	0.68	0.71	0.90	18.6	17.6	19.3	37.5
10 s	0.70	0.69	0.72	0.86	20.3	19.7	20.4	31.4
1 min	0.75	0.75	0.75	0.80	22.7	22.4	22.1	24.1
15 min	0.90	0.91	0.91	0.92	28.3	28.2	28.1	28.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.00	−0.02	−0.19	2.79	3.13	2.99	2.96
1 s	0.01	0.04	−0.03	−0.36	4.26	3.62	3.30	3.30
5 s	0.09	0.16	0.13	−0.18	3.64	4.37	4.47	3.65
10 s	0.07	0.10	0.06	−0.20	4.30	4.33	4.35	3.70
1 min	0.15	0.16	0.16	0.11				

Table 2: Clock-lag differences, `SBUX`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.39	0.49	0.53
s.d. of $\hat{\epsilon}_t$ in cents	1.60	1.54	1.38
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.32
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.26	0.28	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.17	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.10	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.06	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.37,0.37	−0.40,0.41
95% C.I. for $\theta_t - X_t$	—	−0.68,0.73	−0.78,0.84
99% C.I. for $\theta_t - X_t$	—	−1.13,1.19	−1.54,1.51
s.e. of $\theta_t - X_t$ in cents	—	0.34	0.47

Table 3: Residual summaries, `SBUX`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.32, is above two, while the Hill index of the Huber residual itself on this day is 1.65, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.65	2.04	1.61 to 2.53	1.65 to 2.72
Residual divided by δ_t	1.68	2.12	1.68 to 2.59	1.84 to 3.00
Repricing windows removed	1.64	2.02	1.64 to 2.56	1.85 to 2.72
Calm windows only	1.53	1.99		
Every 10th trade	1.61	1.81	1.61 to 2.24	1.54 to 2.49
Six-trade half-life	1.82	2.19	1.57 to 2.72	1.60 to 2.90

Table 4: Hill estimator of the tail index of the Huber filter residual, `SBUX`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 17.7 to 20.2 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 8.9 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

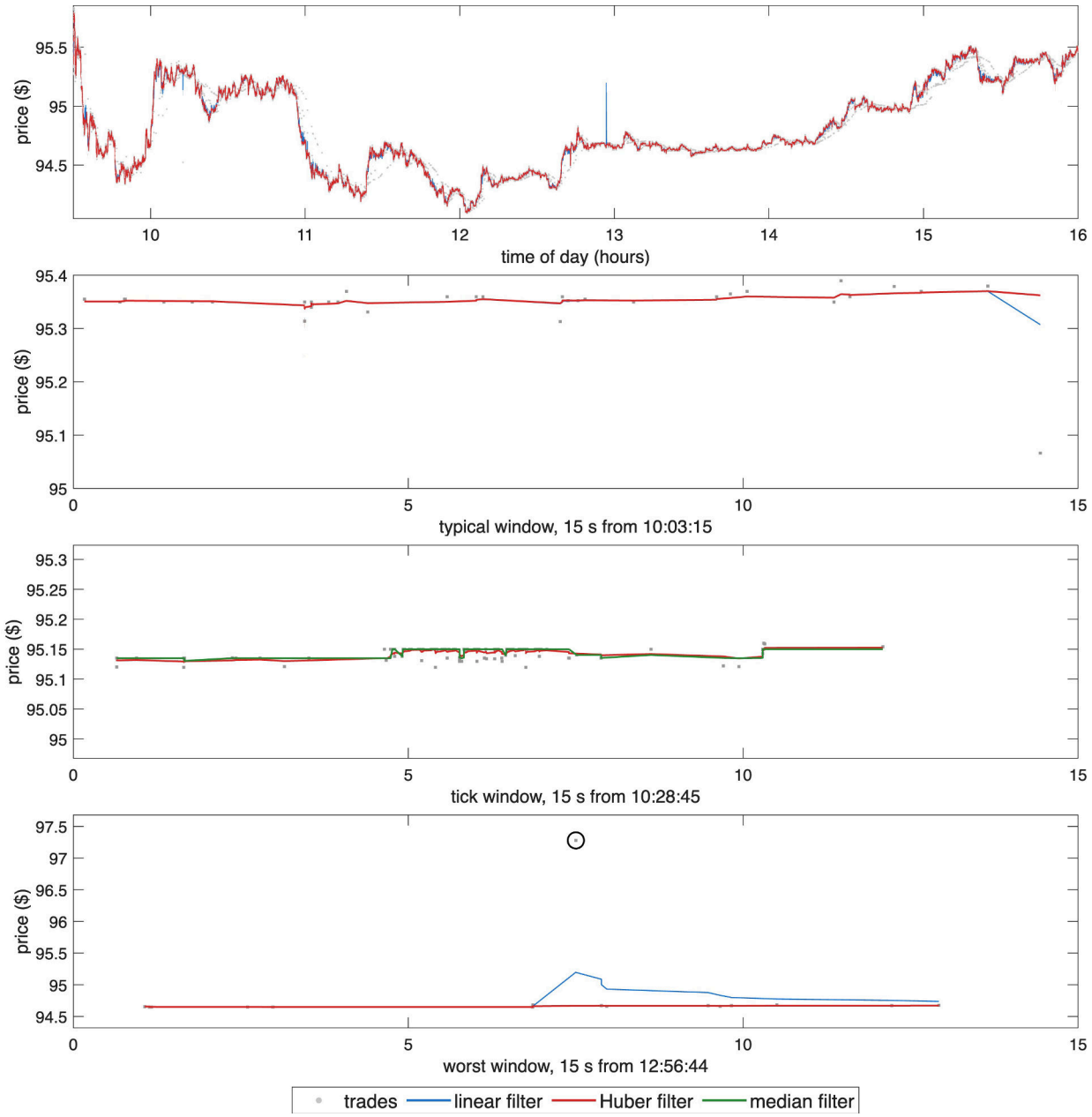

Figure 1: `SBUX` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:03:15. 3rd: 15-second window at 10:28:45 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

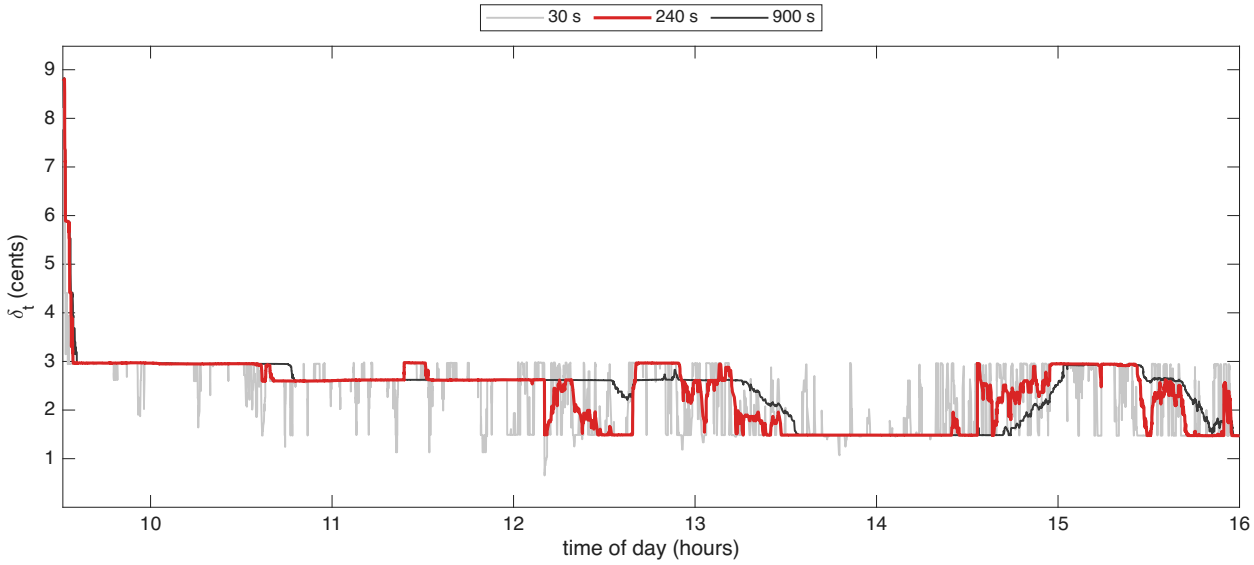

Figure 2: The Huber threshold δ_t in cents for `SBUX` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

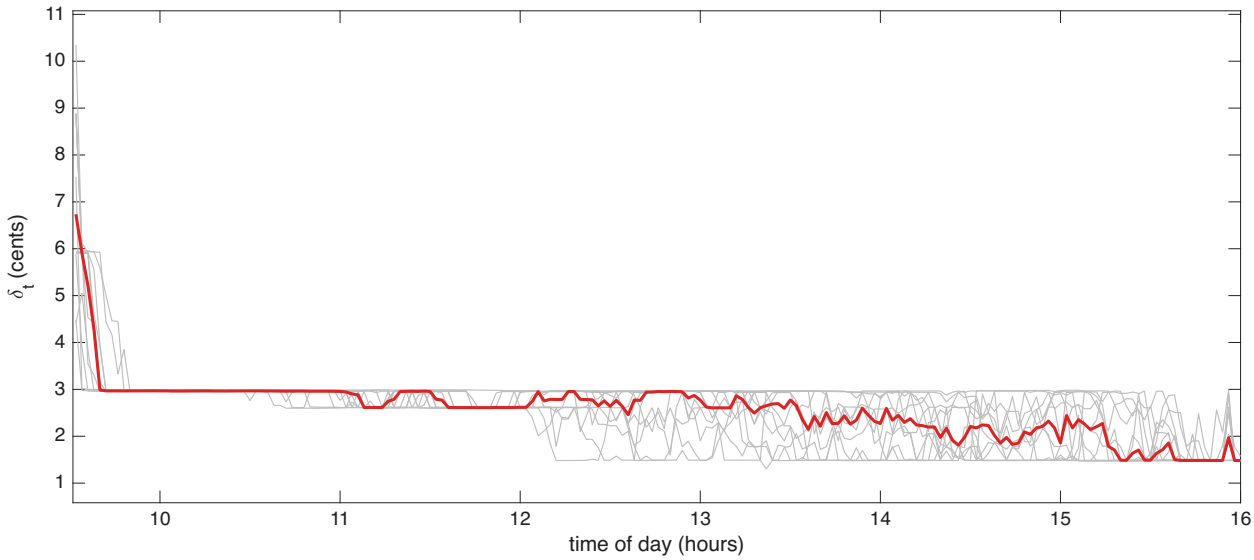

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `SBUX` trading days in Oct 2024 (grey lines) and their median (red line).

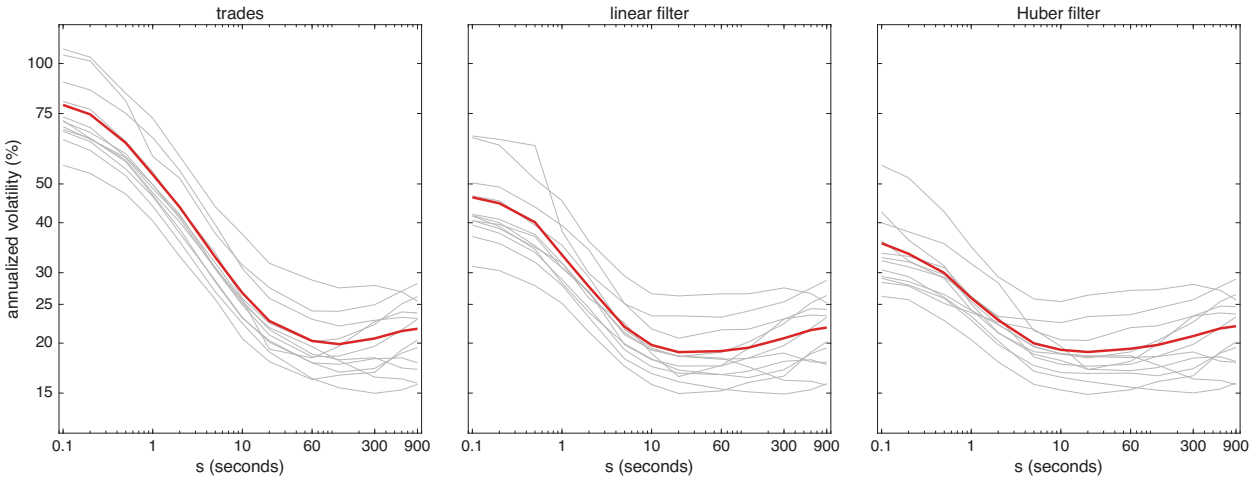

Figure 4: Volatility signature curves for `SBUX` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

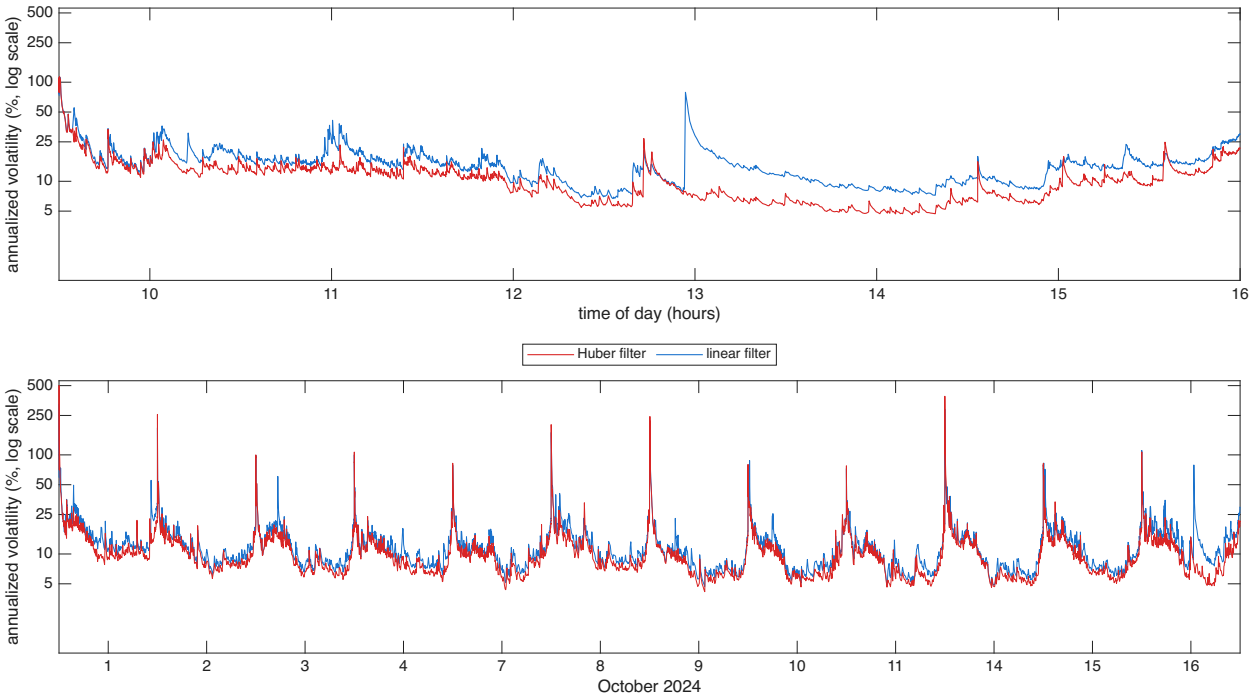

Figure 5: Filtered spot volatility for `SBUX` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 53$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

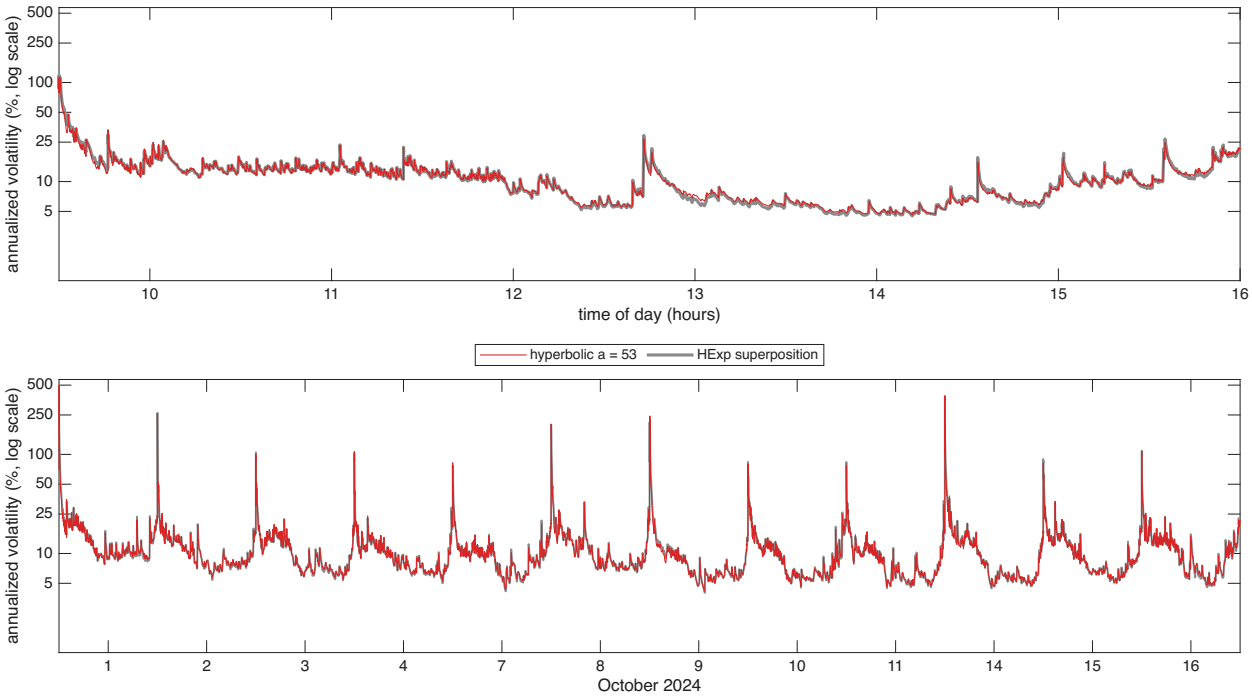

Figure 6: Hyperbolic against superposition weights for `SBUX` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 53$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

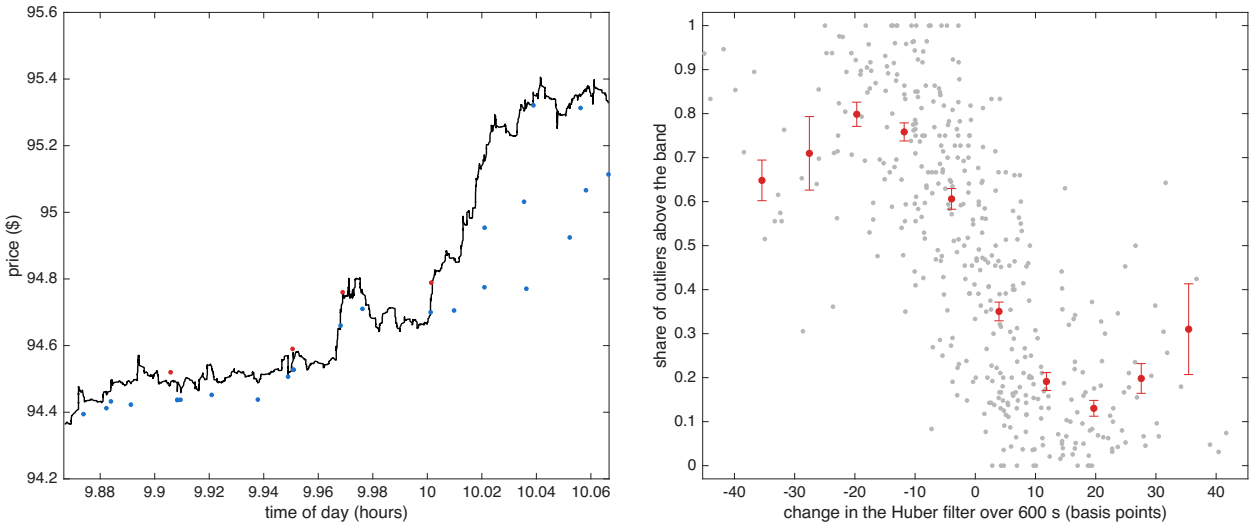

Figure 7: Outliers & robust price for `SBUX` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

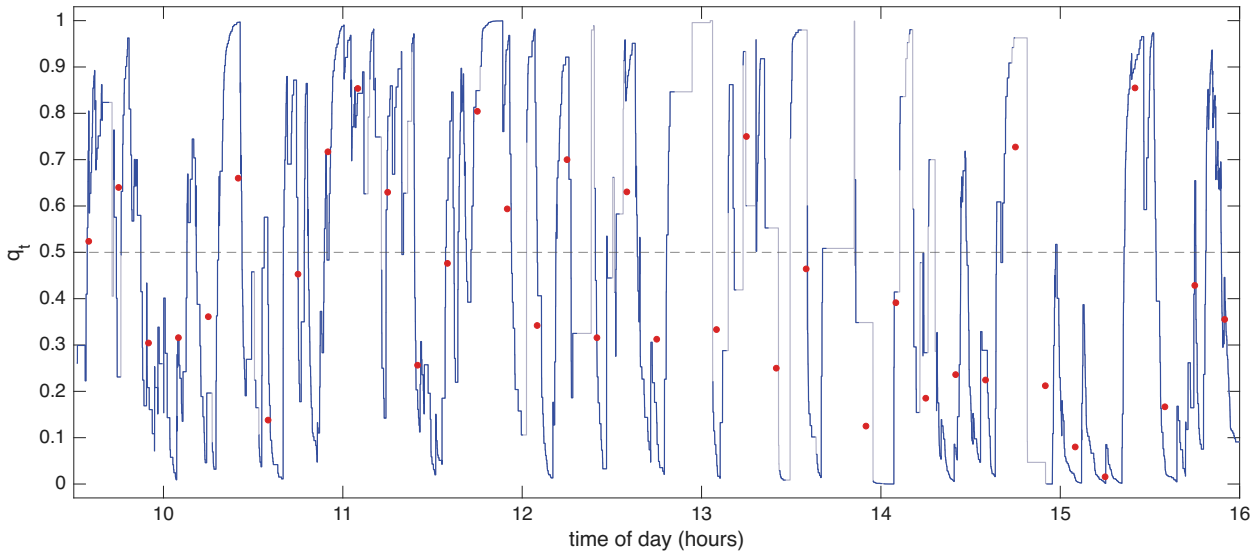

Figure 8: The directional conditional probability for SBUX, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 57 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

Selecting a on single days gives values from 13 to 408 trades, against 53 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 37 trades for this stock.

On 14 October 2024, 31 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 23 per cent on any other day. Every number that day contributes to carries that burst.

Diagnostics.

NFLX 1 to 16 October 2024

trades on 16 October 2024	87,455
average trading rate	3.7 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	4,507 (5.15%)
price half-life: run / minimizer[1]	3.0 / 3.4 trades
cost of running 3 trades	+0.16%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	30 s [25, 36]
Mandelbrot a minimizing QLIKE	91 trades
a within 0.0005 of the minimum	59 to 155
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	20% $\to$ 32%
Huber $\hat{\sigma}$, before $\to$ peak	20% $\to$ 40%
median spot volatility, Huber / linear input[5]	18.2 / 19.5 %
hyperbolic against HExp, correlation of logs	0.9993
HExp QLIKE advantage, leave-one-day-out[6]	-0.00049
15-minute signature level	15 to 49 %
Huber signature, 1 min over 1 s, 12 day average	1.07

NFLX over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:51; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.223	1.710	1.682	7.949	5.748	2.745	2.725	15.957
2	2.033	1.827	1.849	6.462	4.910	2.898	2.864	12.030
5	2.194	1.910	1.948	4.660	4.321	2.971	2.929	8.066
10	2.184	1.907	1.930	3.652	3.812	2.898	2.847	5.998
25	2.038	1.815	1.811	2.787	3.121	2.633	2.597	4.227
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.27	0.11	0.11	−0.43	4.14	3.57	3.68	4.10
2	0.07	0.01	0.01	−0.03	4.29	3.73	4.17	4.17
5	−0.02	−0.02	−0.02	0.00	4.34	4.04	4.47	4.20
10	−0.01	−0.02	−0.01	0.00	5.55	4.68	4.84	4.57
25	0.00	−0.01	−0.01	−0.01	5.02	4.65	4.52	4.34

Table 1: Trade-lag differences, `NFLX`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	14.75	13.11	13.02	26.02	29.0	15.6	15.9	91.1
1 s	8.99	8.30	8.32	13.82	23.7	16.5	16.5	58.0
5 s	6.43	6.04	6.04	8.38	19.8	16.8	16.7	33.4
10 s	5.65	5.35	5.31	7.04	18.3	16.5	16.4	26.7
1 min	4.71	4.64	4.63	5.11	16.6	16.2	16.1	18.5
15 min	4.18	4.20	4.20	4.31	14.8	14.7	14.7	15.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.09	−0.03	−0.03	−0.21	6.43	6.60	6.56	4.42
1 s	−0.13	−0.03	−0.03	−0.38	5.51	7.00	7.01	5.45
5 s	−0.13	−0.05	−0.06	−0.28	6.75	6.48	6.05	5.91
10 s	−0.11	−0.05	−0.04	−0.26	6.36	6.50	6.60	5.39
1 min	0.02	0.04	0.05	−0.05				

Table 2: Clock-lag differences, `NFLX`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	5.10	6.26	6.46
s.d. of $\hat{\epsilon}_t$ in cents	11.53	10.64	10.47
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.69
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.18	0.14	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.03	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−3.74,3.60	−3.98,3.85
95% C.I. for $\theta_t - X_t$	—	−7.16,6.91	−7.68,7.27
99% C.I. for $\theta_t - X_t$	—	−11.66,12.44	−12.85,12.80
s.e. of $\theta_t - X_t$ in cents	—	3.31	3.71

Table 3: Residual summaries, `NFLX`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.69, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.97	3.33	2.15 to 3.50	2.28 to 3.84
Residual divided by δ_t	3.44	2.98	2.38 to 3.91	2.42 to 3.38
Repricing windows removed	3.13	3.33	2.80 to 3.47	2.95 to 3.90
Calm windows only	3.24	3.29		
Every 10th trade	2.93	3.75	1.99 to 3.24	2.51 to 3.75
Six-trade half-life	3.11	3.55	2.15 to 3.73	2.08 to 4.19

Table 4: Hill estimator of the tail index of the Huber filter residual, `NFLX`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 104.8 to 171.4 cents over the days and removes 0% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 52.4 cents on 16 October 2024, 87% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

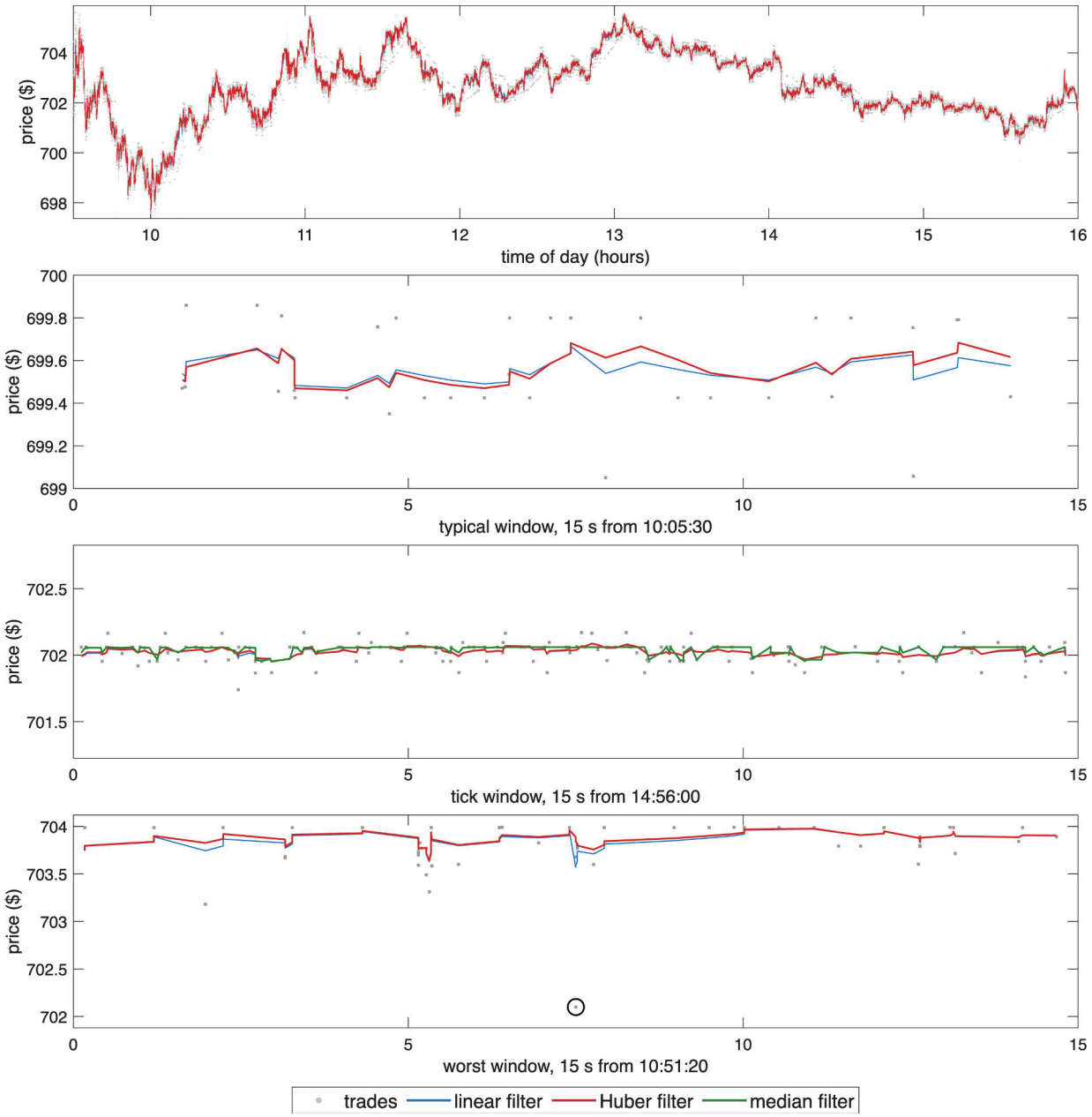

Figure 1: `NFLX` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:05:30. 3rd: 15-second window at 14:56:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

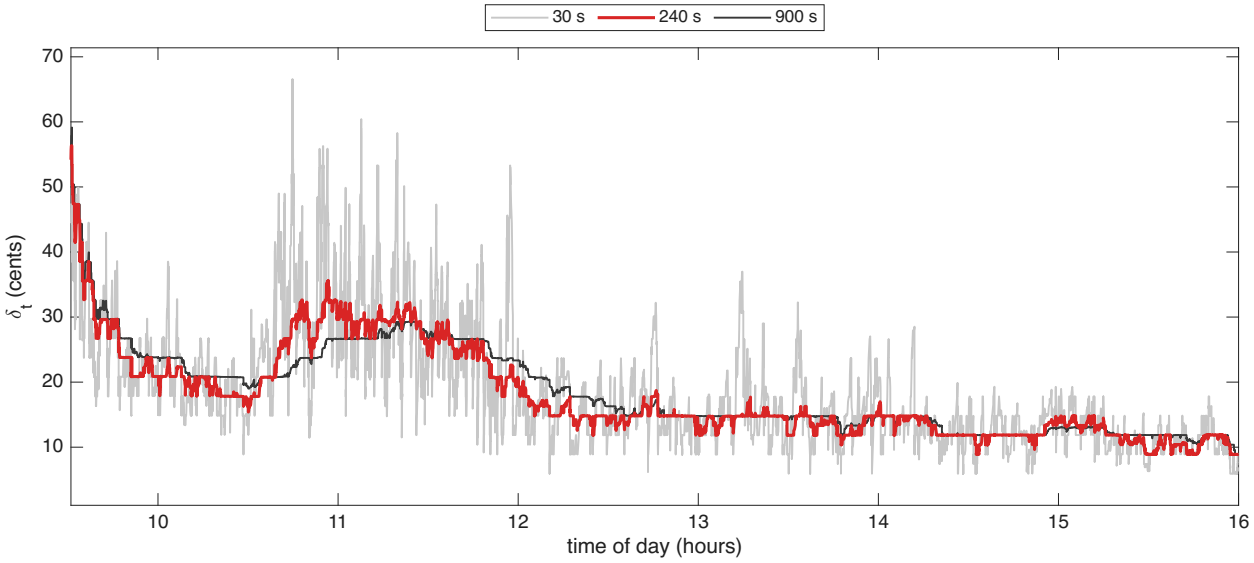

Figure 2: The Huber threshold δ_t in cents for `NFLX` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

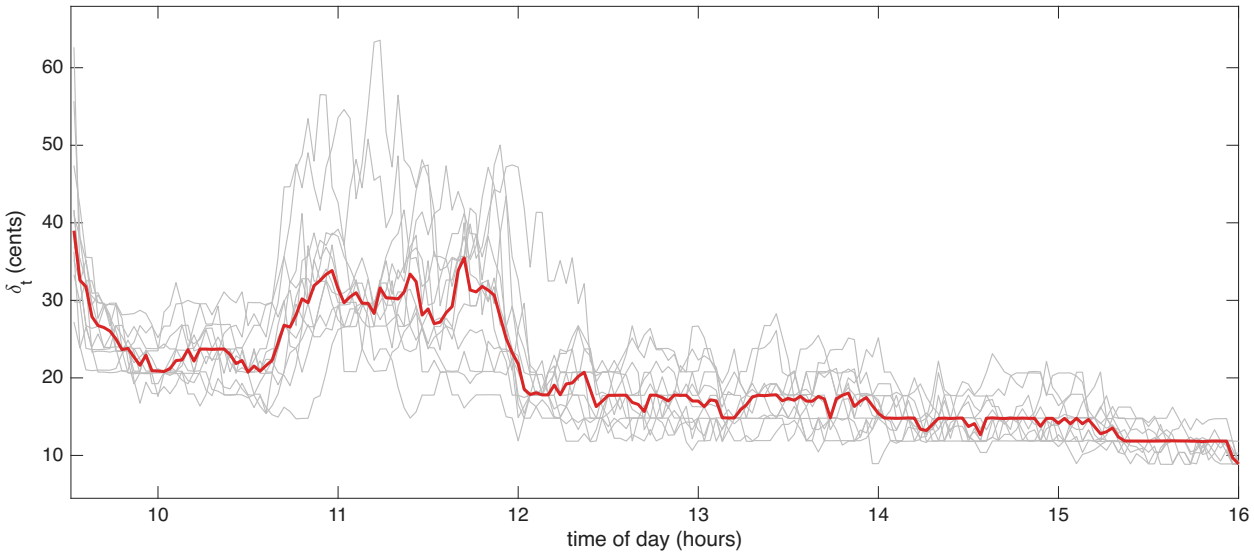

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `NFLX` trading days in Oct 2024 (grey lines) and their median (red line).

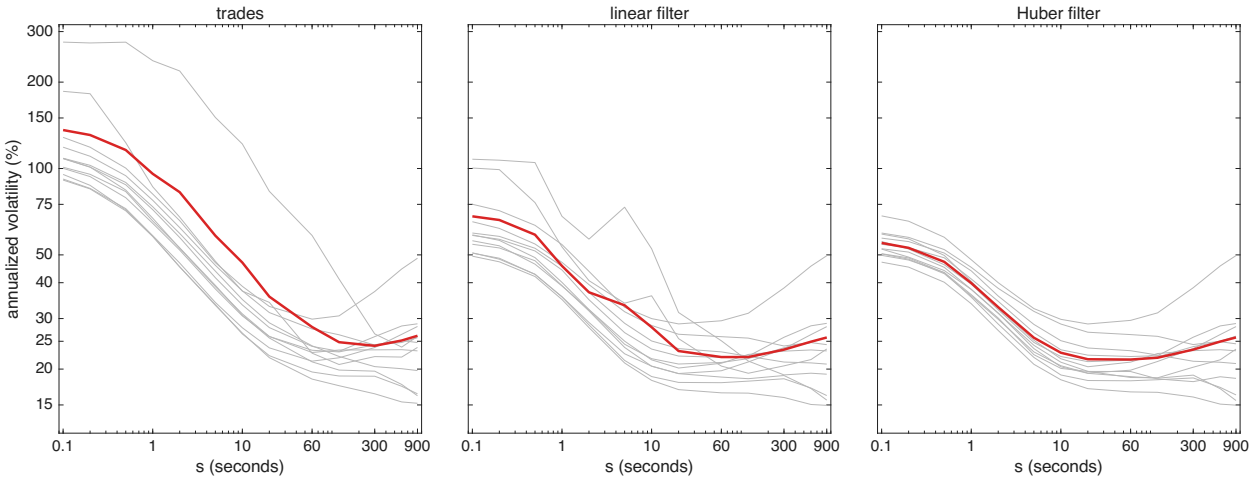

Figure 4: Volatility signature curves for `NFLX` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

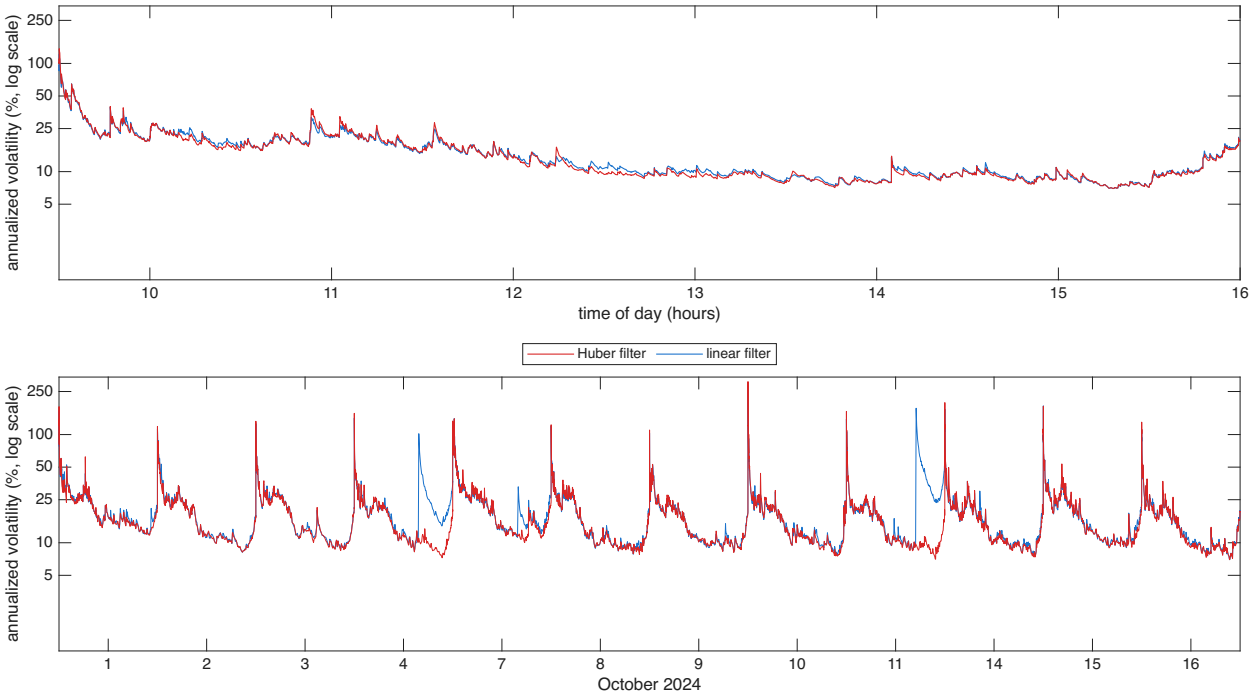

Figure 5: Filtered spot volatility for `NFLX` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 91$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

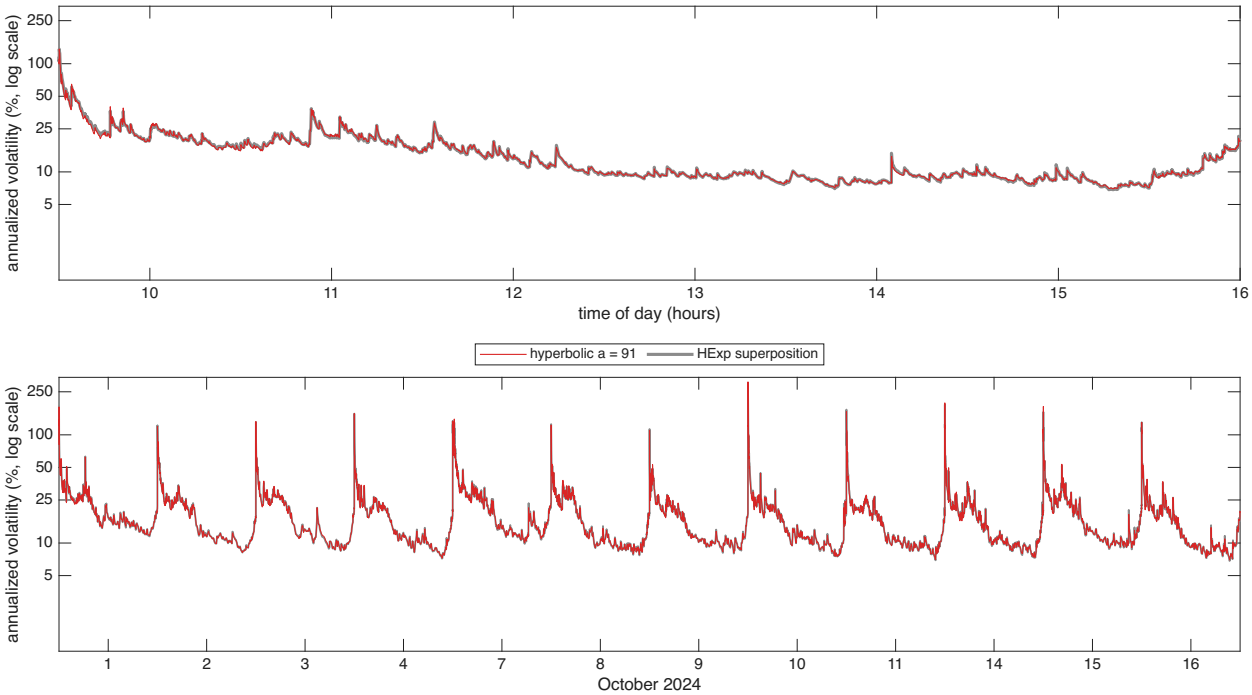

Figure 6: Hyperbolic against superposition weights for `NFLX` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 91$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

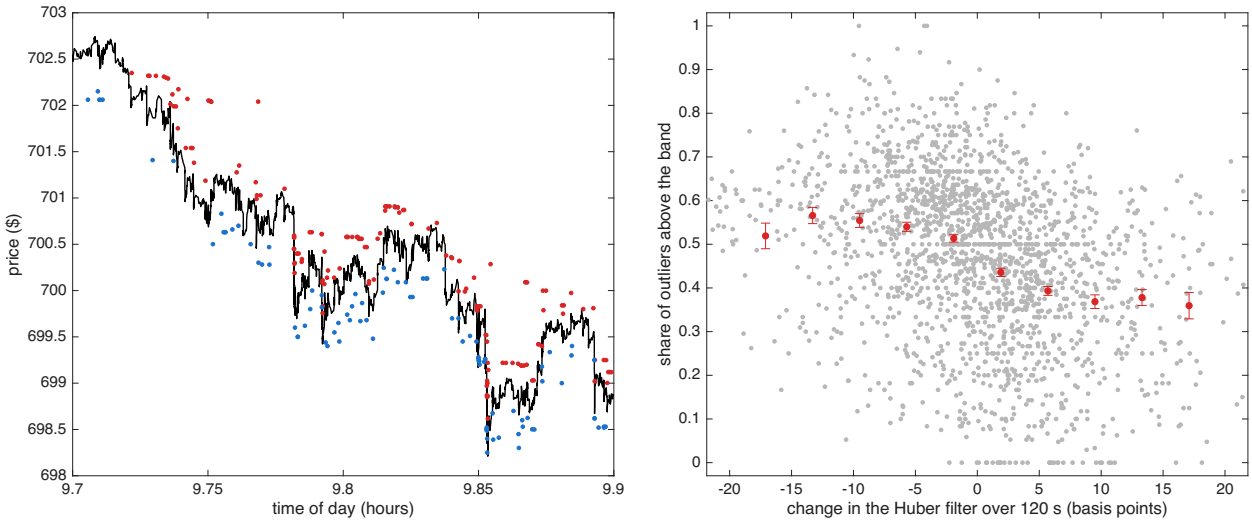

Figure 7: Outliers & robust price for `NFLX` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

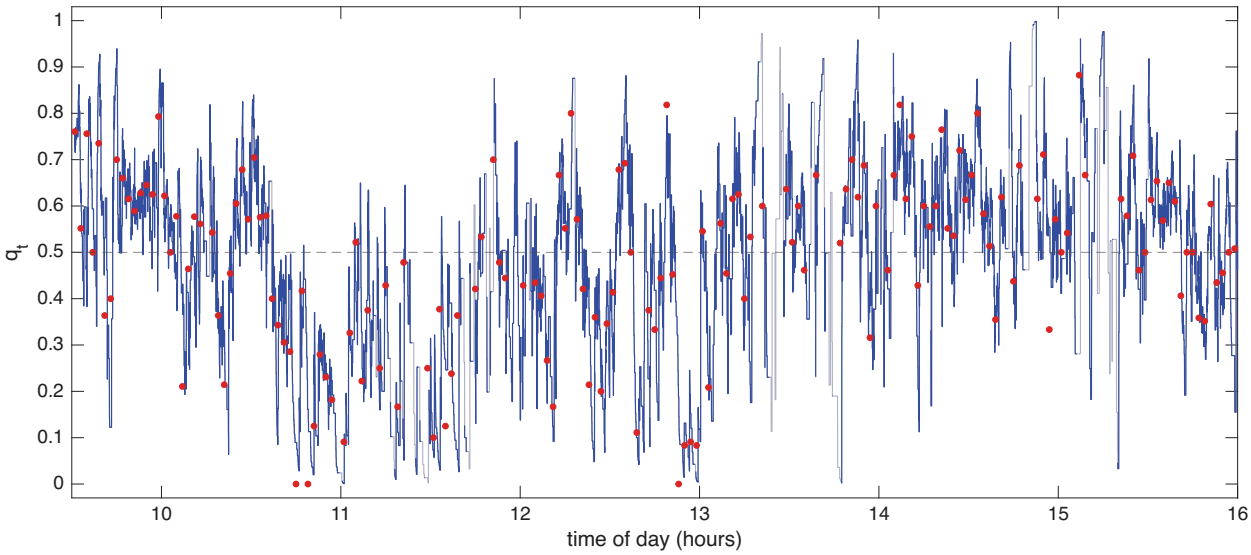

Figure 8: The directional conditional probability for `NFLX`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 30 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:51.

Diagnostics.

CRWD 1 to 16 October 2024

trades on 16 October 2024	68,326
average trading rate	3.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	3,660 (5.36%)
price half-life: run / minimizer[1]	3.0 / 2.7 trades
cost of running 3 trades	+0.25%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	37 s [31, 47]
Mandelbrot a minimizing QLIKE	102 trades
a within 0.0005 of the minimum	65 to 163
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	56% $\to$ 69%
Huber $\hat{\sigma}$, before $\to$ peak	57% $\to$ 69%
median spot volatility, Huber / linear input[5]	22.6 / 24.9 %
hyperbolic against HExp, correlation of logs	0.9987
HExp QLIKE advantage, leave-one-day-out[6]	+0.00109
15-minute signature level	22 to 67 %
Huber signature, 1 min over 1 s, 12 day average	1.25

CRWD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:45; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.343	1.120	1.102	5.494	3.035	1.768	1.790	10.664
2	1.229	1.198	1.220	4.337	2.647	1.829	1.853	7.859
5	1.370	1.283	1.321	3.135	2.421	1.899	1.919	5.289
10	1.414	1.317	1.345	2.505	2.217	1.897	1.913	3.963
25	1.419	1.336	1.338	1.965	2.022	1.869	1.867	2.876
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.24	0.07	0.07	−0.46	3.67	3.48	3.28	3.37
2	0.06	0.02	0.02	−0.02	3.52	3.64	3.88	3.57
5	−0.01	−0.01	−0.02	−0.02	3.73	3.69	4.56	3.84
10	0.00	−0.01	−0.01	0.00	4.47	4.23	5.28	3.92
25	−0.01	−0.01	−0.01	0.00	5.56	6.09	6.56	4.17

Table 1: Trade-lag differences, `CRWD`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	8.85	8.02	7.98	15.89	32.4	21.2	22.1	128.6
1 s	5.72	5.29	5.28	9.21	29.1	23.3	23.8	87.7
5 s	4.13	3.91	3.90	5.50	28.4	26.5	26.6	51.4
10 s	3.78	3.59	3.57	4.72	28.0	26.9	26.7	41.8
1 min	3.44	3.36	3.34	3.69	30.4	30.1	29.8	32.9
15 min	3.15	3.12	3.11	3.20	26.7	26.7	26.6	27.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.02	0.04	−0.16	7.11	7.25	7.53	3.77
1 s	−0.15	−0.10	−0.08	−0.34	6.08	5.81	5.96	4.00
5 s	0.02	0.05	0.05	−0.24	7.44	7.97	8.13	4.57
10 s	0.02	0.06	0.06	−0.19	7.07	7.73	7.56	8.31
1 min	0.14	0.15	0.16	0.06				

Table 2: Clock-lag differences, `CRWD`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	3.48	4.09	4.23
s.d. of $\hat{\epsilon}_t$ in cents	7.41	7.03	6.88
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.27
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.16	0.12	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.05	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.00	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	0.00	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−2.49,2.58	−2.68,2.73
95% C.I. for $\theta_t - X_t$	—	−4.54,4.79	−5.00,5.05
99% C.I. for $\theta_t - X_t$	—	−7.08,7.66	−8.06,8.19
s.e. of $\theta_t - X_t$ in cents	—	2.17	2.47

Table 3: Residual summaries, `CRWD`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.27, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.18	2.56	2.54 to 3.46	2.54 to 3.82
Residual divided by δ_t	3.04	2.56	2.22 to 3.12	2.25 to 3.32
Repricing windows removed	3.16	2.59	2.55 to 3.45	2.52 to 3.82
Calm windows only	3.23	2.70		
Every 10th trade	3.10	2.65	2.43 to 3.45	2.10 to 3.79
Six-trade half-life	3.43	2.72	2.67 to 3.64	2.67 to 4.14

Table 4: Hill estimator of the tail index of the Huber filter residual, `CRWD`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 62.1 to 90.7 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 36.8 cents on 16 October 2024, 89% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

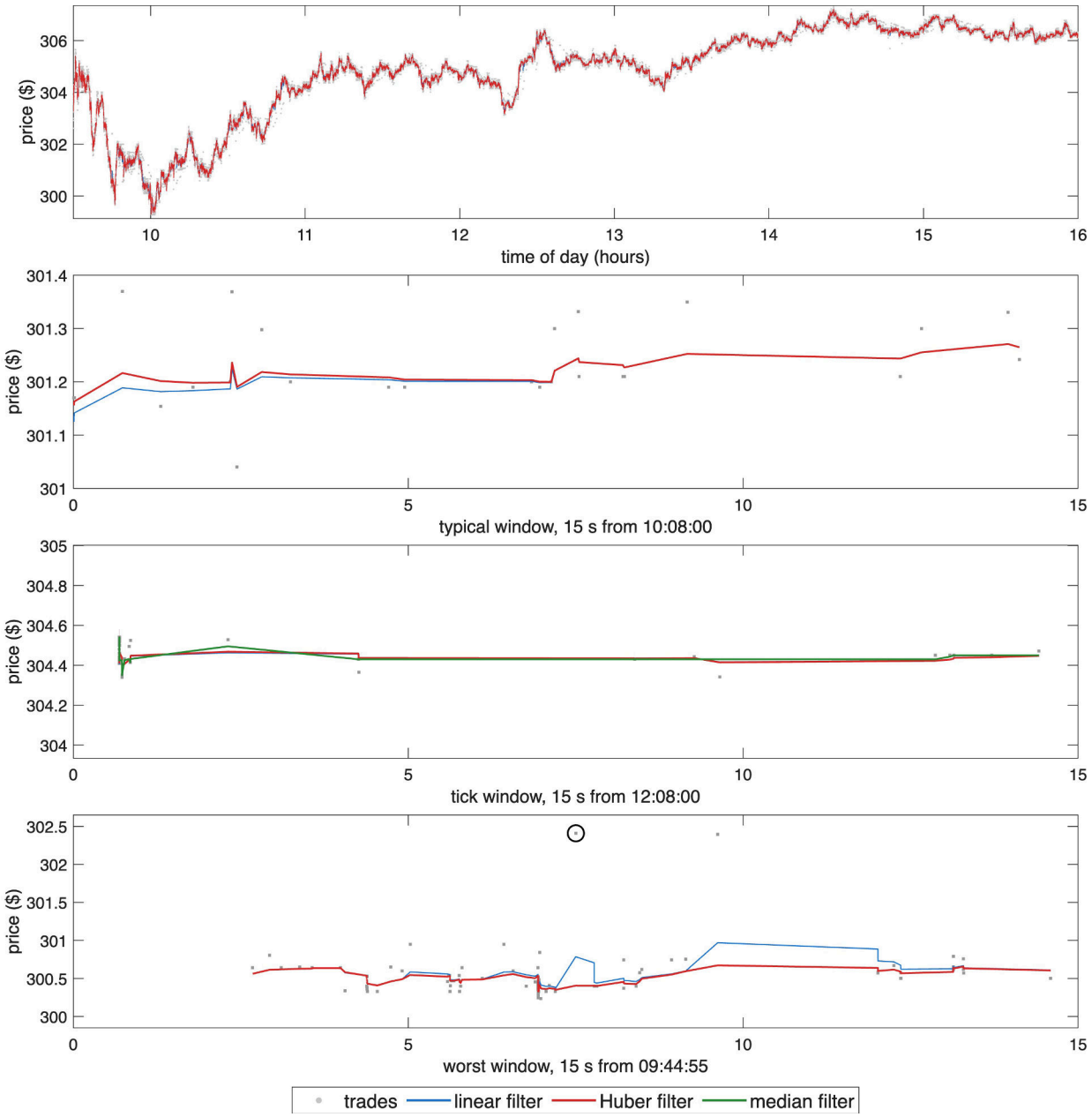

Figure 1: `CRWD` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:08:00. 3rd: 15-second window at 12:08:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

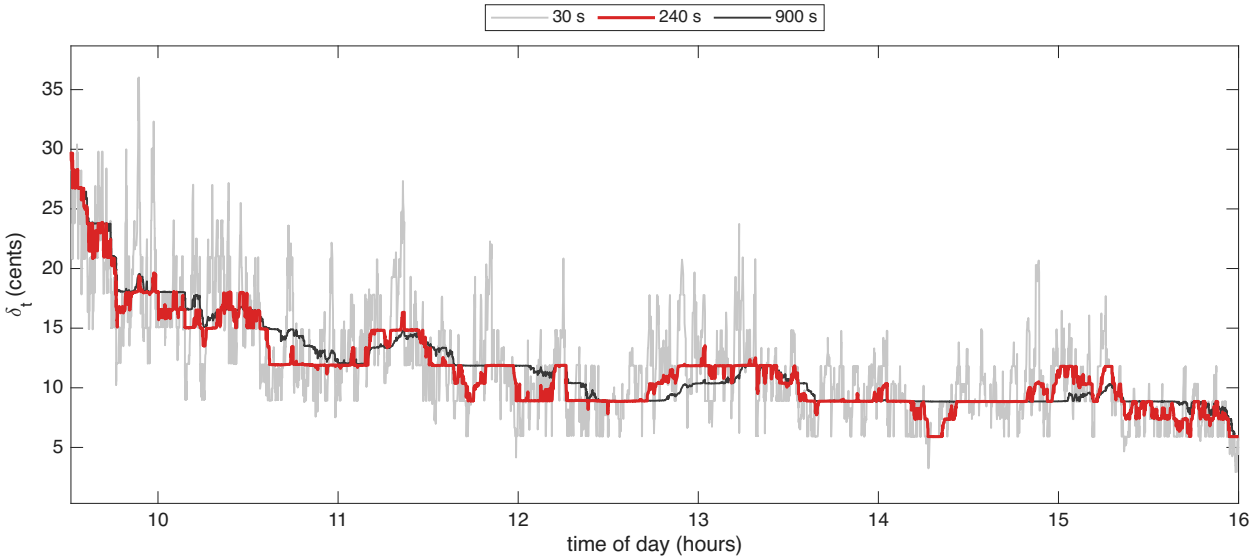

Figure 2: The Huber threshold δ_t in cents for `CRWD` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

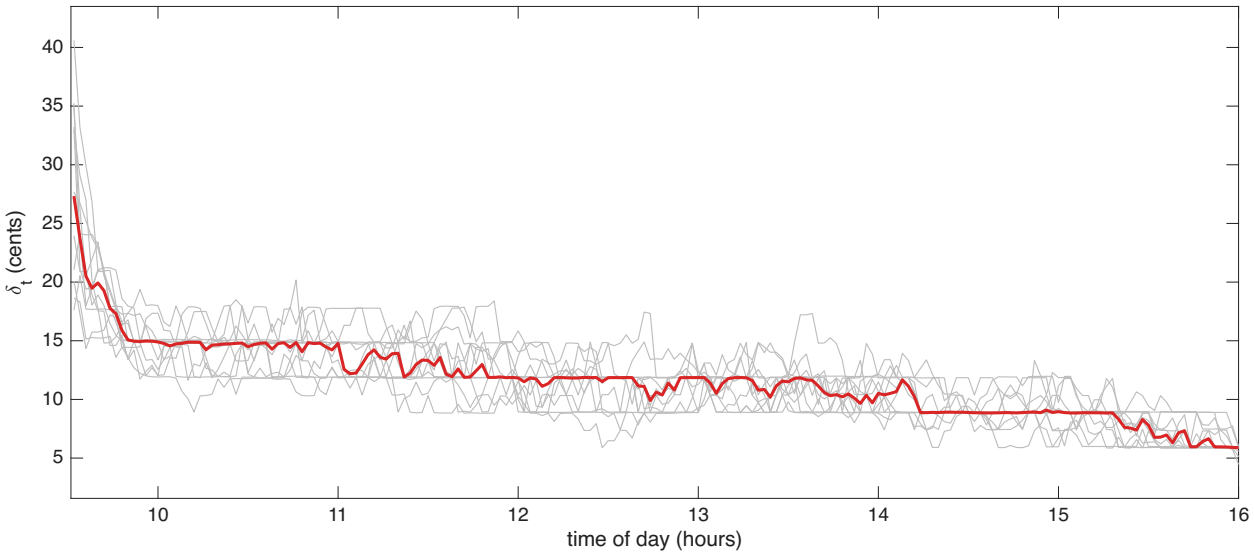

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `CRWD` trading days in Oct 2024 (grey lines) and their median (red line).

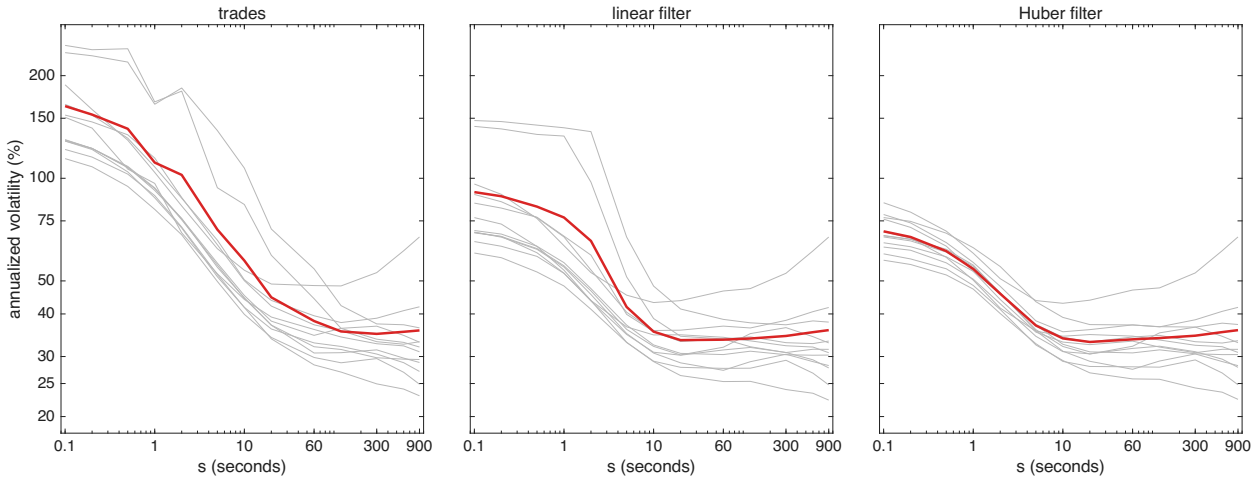

Figure 4: Volatility signature curves for `CRWD` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

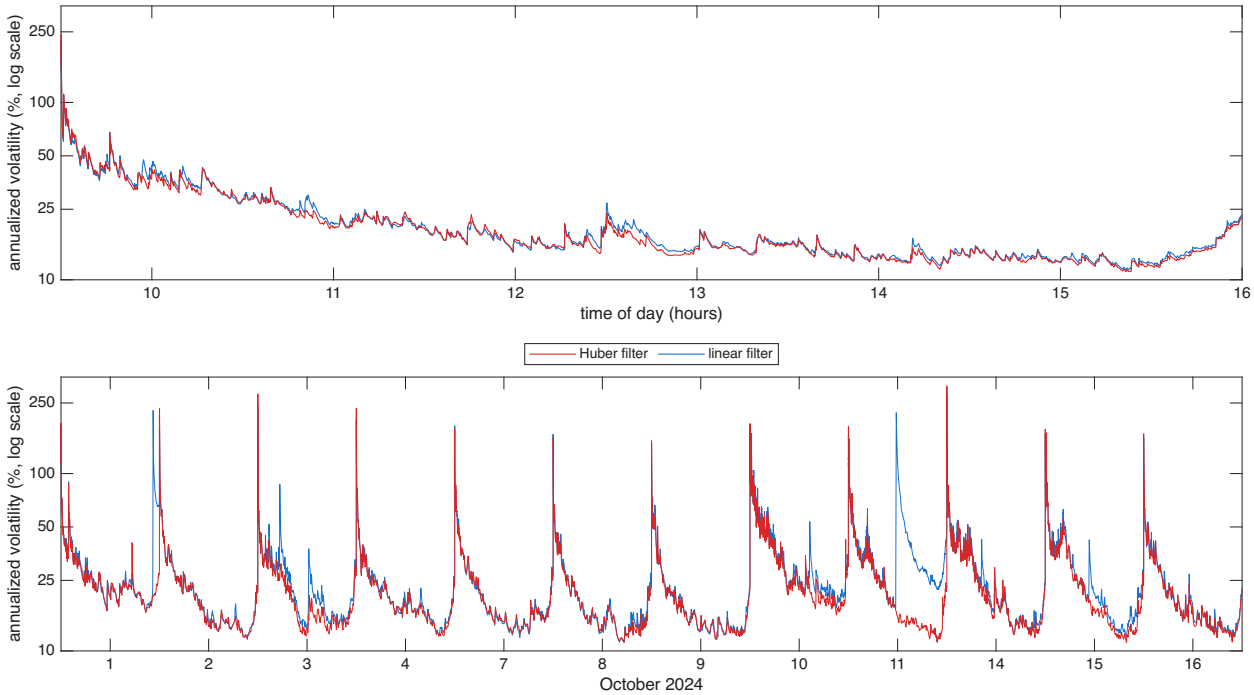

Figure 5: Filtered spot volatility for `CRWD` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 102$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

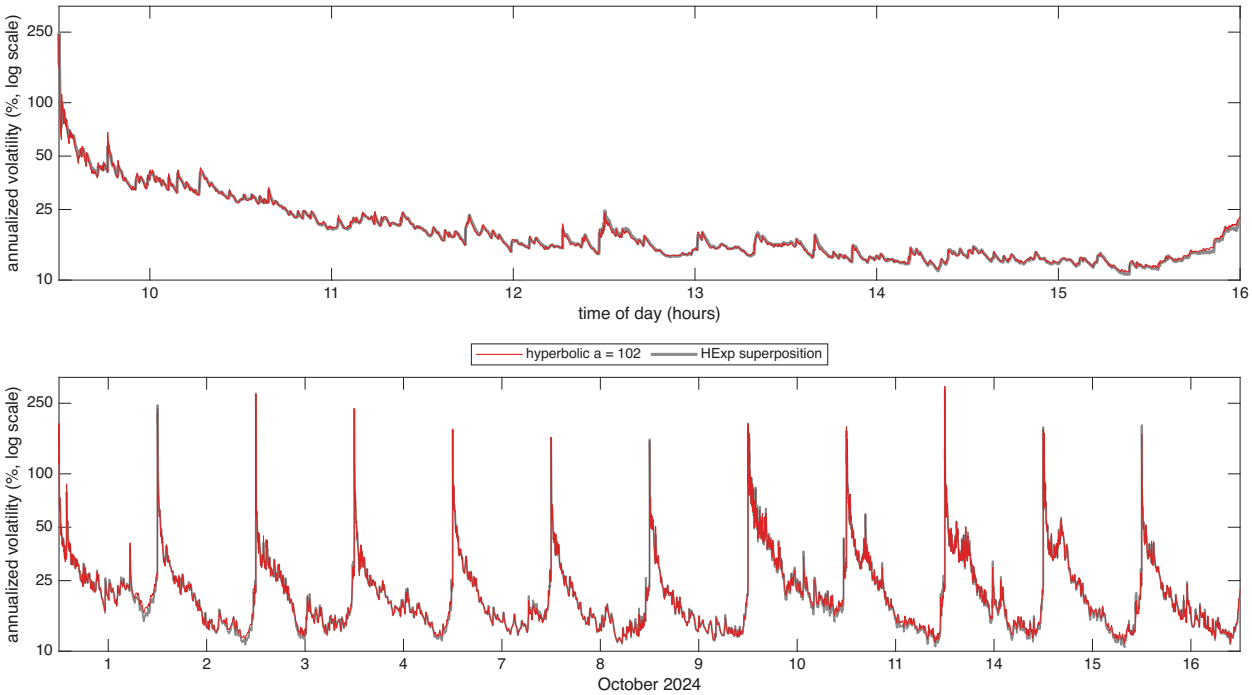

Figure 6: Hyperbolic against superposition weights for `CRWD` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 102$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

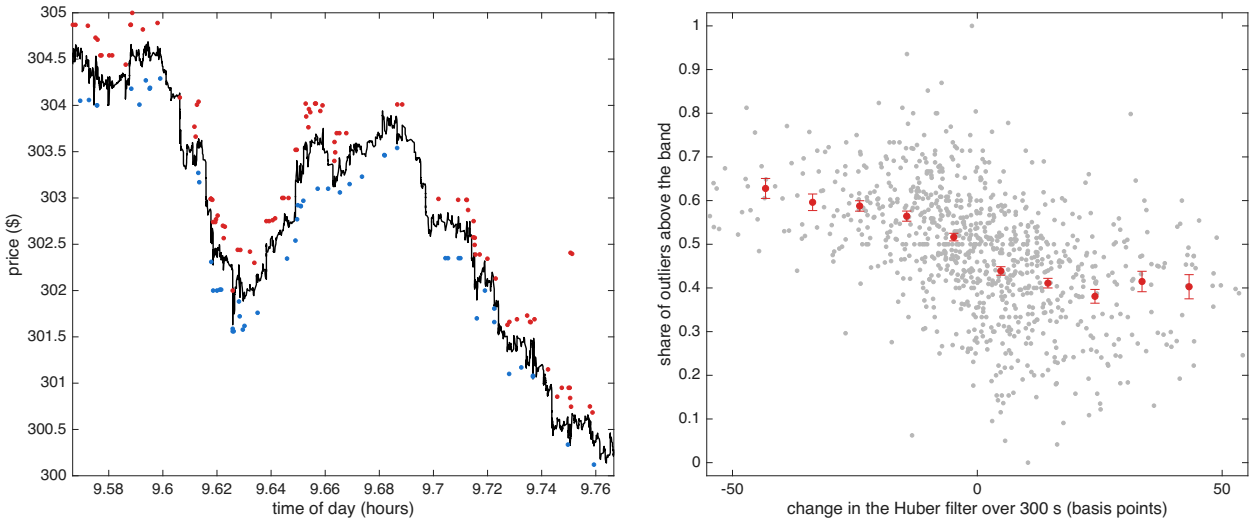

Figure 7: Outliers & robust price for `CRWD` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

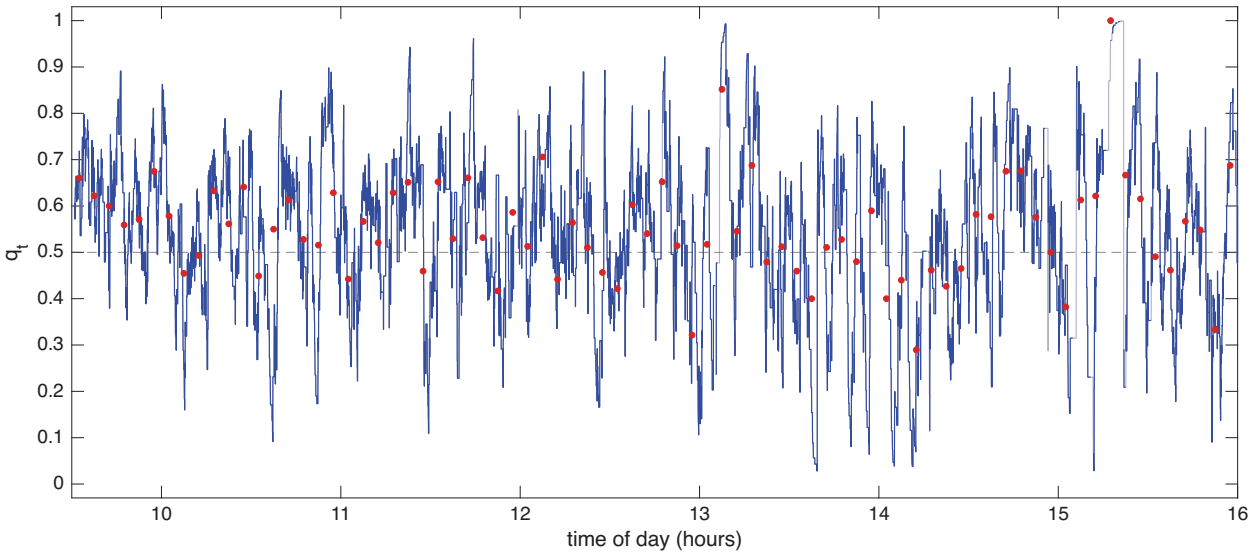

Figure 8: The directional conditional probability for `CRWD`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 37 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

ADBE 1 to 16 October 2024

trades on 16 October 2024	59,767
average trading rate	3.1 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,920 (4.89%)
price half-life: run / minimizer[1]	3.0 / 3.4 trades
cost of running 3 trades	+0.24%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.00%
direction half-life[3]	43 s [36, 52]
Mandelbrot a minimizing QLIKE	61 trades
a within 0.0005 of the minimum	40 to 95
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	17% $\rightarrow$ 24%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	18% $\rightarrow$ 24%
median spot volatility, Huber / linear input[5]	17.1 / 19.0 %
hyperbolic against HExp, correlation of logs	0.9985
HExp QLIKE advantage, leave-one-day-out[6]	-0.00061
15-minute signature level	12 to 35 %
Huber signature, 1 min over 1 s, 12 day average	1.09

ADBE over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:31; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.965	1.530	1.496	7.598	4.665	2.588	2.473	14.858
2	1.731	1.599	1.618	5.951	3.888	2.575	2.542	10.899
5	1.860	1.666	1.701	4.272	3.400	2.551	2.566	7.249
10	1.808	1.634	1.661	3.325	2.945	2.423	2.447	5.397
25	1.674	1.547	1.555	2.463	2.455	2.218	2.231	3.735
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.31	−0.01	0.06	−0.46	3.38	3.15	3.23	3.59
2	0.07	0.01	0.01	−0.01	3.40	3.19	3.19	3.54
5	−0.03	−0.02	−0.02	0.02	3.47	3.41	3.23	3.48
10	0.01	−0.02	−0.03	−0.01	3.73	3.70	3.99	3.58
25	0.00	0.01	0.01	0.01	3.94	3.63	4.03	3.55

Table 1: Trade-lag differences, `ADBE`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	10.91	9.89	9.82	19.89	29.4	18.4	17.5	100.1
1 s	7.18	6.68	6.63	11.99	25.5	18.7	17.6	70.6
5 s	5.19	4.92	4.90	7.41	21.5	18.7	18.6	41.8
10 s	4.45	4.29	4.27	5.91	19.4	17.9	17.8	32.3
1 min	3.81	3.78	3.74	4.29	18.2	17.8	17.5	20.9
15 min	3.84	3.88	3.86	4.00	20.1	20.0	19.9	20.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.02	0.06	0.04	−0.14	4.41	4.24	4.33	3.43
1 s	−0.20	−0.15	−0.12	−0.39	6.32	5.72	6.56	11.63
5 s	−0.18	−0.11	−0.11	−0.40	8.50	7.02	6.11	4.01
10 s	−0.05	−0.02	−0.03	−0.32	5.70	5.60	5.71	5.18
1 min	0.09	0.11	0.12	−0.03				

Table 2: Clock-lag differences, `ADBE`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	4.70	5.58	5.74
s.d. of $\hat{\epsilon}_t$ in cents	10.17	9.63	9.48
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.27
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.14	0.08	0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.05	0.03	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.00	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	−0.03	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.01
80% C.I. for $\theta_t - X_t$	—	−3.27,3.22	−3.47,3.39
95% C.I. for $\theta_t - X_t$	—	−5.87,5.72	−6.37,6.08
99% C.I. for $\theta_t - X_t$	—	−9.91,9.71	−11.07,10.29
s.e. of $\theta_t - X_t$ in cents	—	2.86	3.22

Table 3: Residual summaries, `ADBE`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.27, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.63	2.90	1.76 to 3.40	1.97 to 3.74
Residual divided by δ_t	3.00	2.76	2.19 to 3.45	2.15 to 3.68
Repricing windows removed	2.81	2.86	2.63 to 3.44	2.29 to 3.91
Calm windows only	3.13	2.81		
Every 10th trade	2.19	3.26	1.87 to 3.90	1.82 to 3.26
Six-trade half-life	2.60	3.01	1.62 to 3.69	1.80 to 3.91

Table 4: Hill estimator of the tail index of the Huber filter residual, `ADBE`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 76.1 to 121.0 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 54.5 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

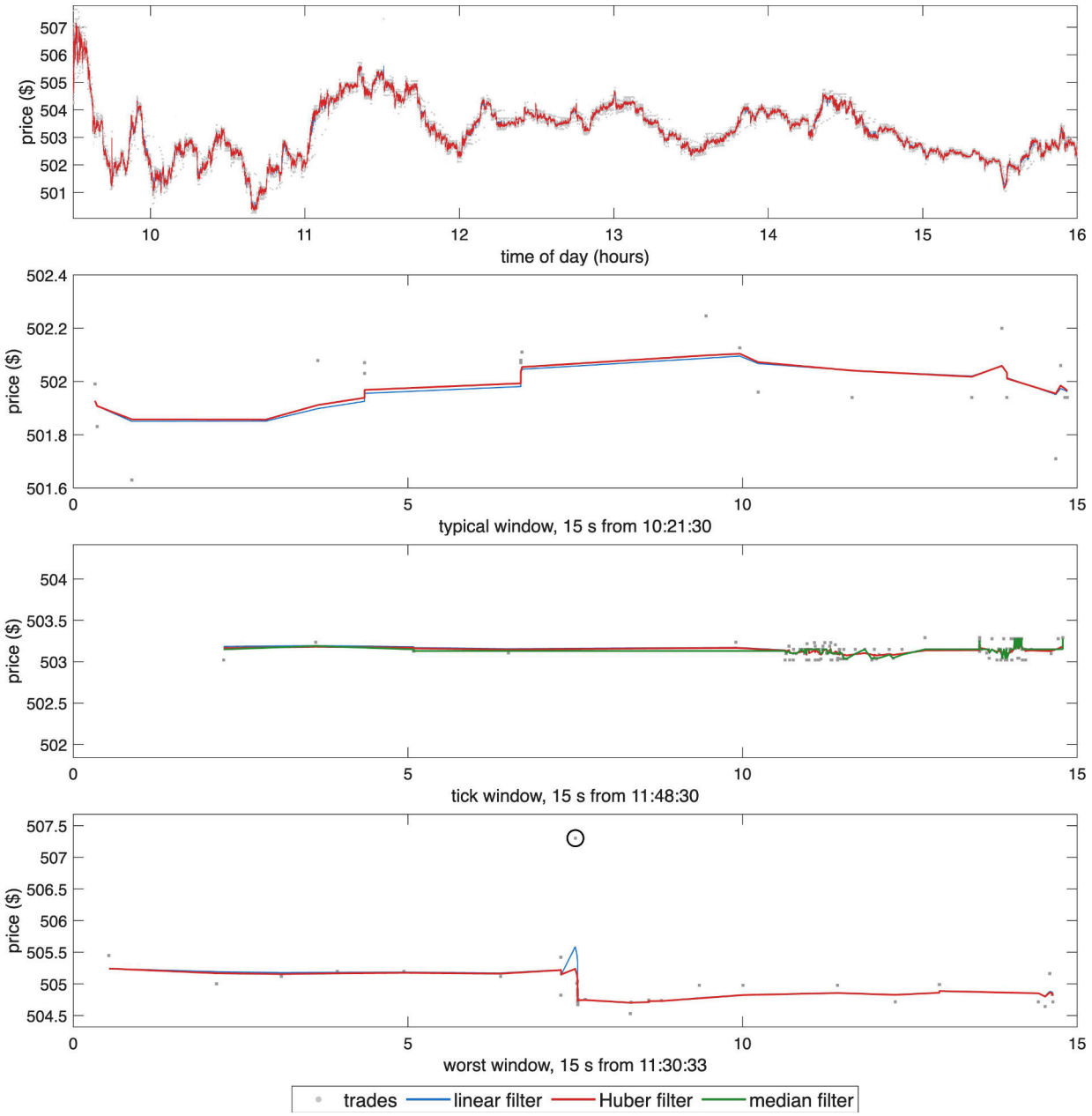

Figure 1: ADBE trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:21:30. 3rd: 15-second window at 11:48:30 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

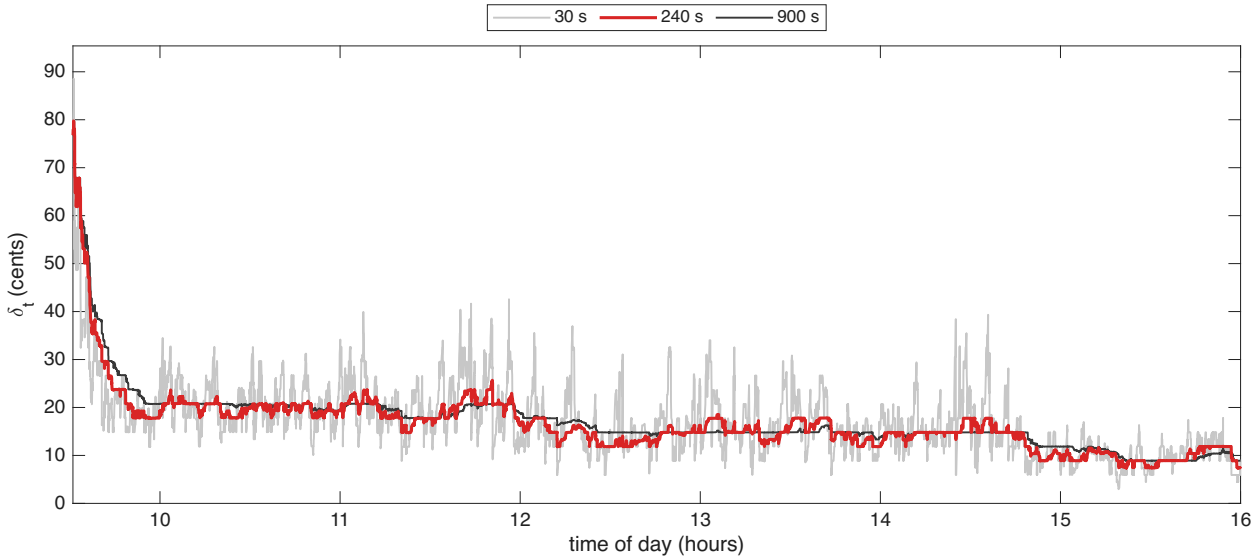

Figure 2: The Huber threshold δ_t in cents for `ADBE` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

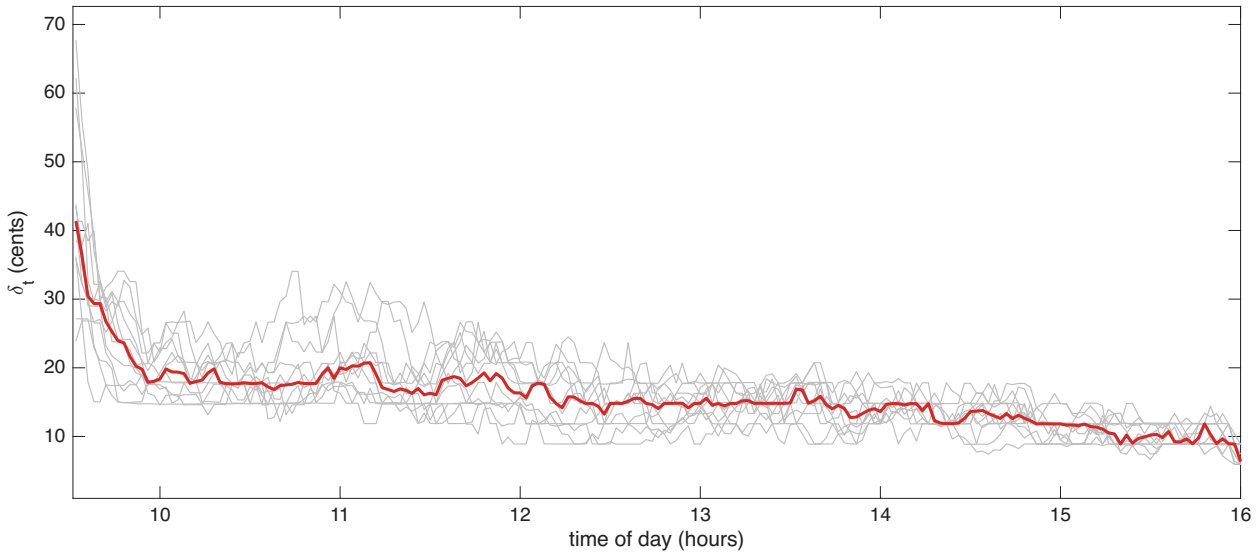

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `ADBE` trading days in Oct 2024 (grey lines) and their median (red line).

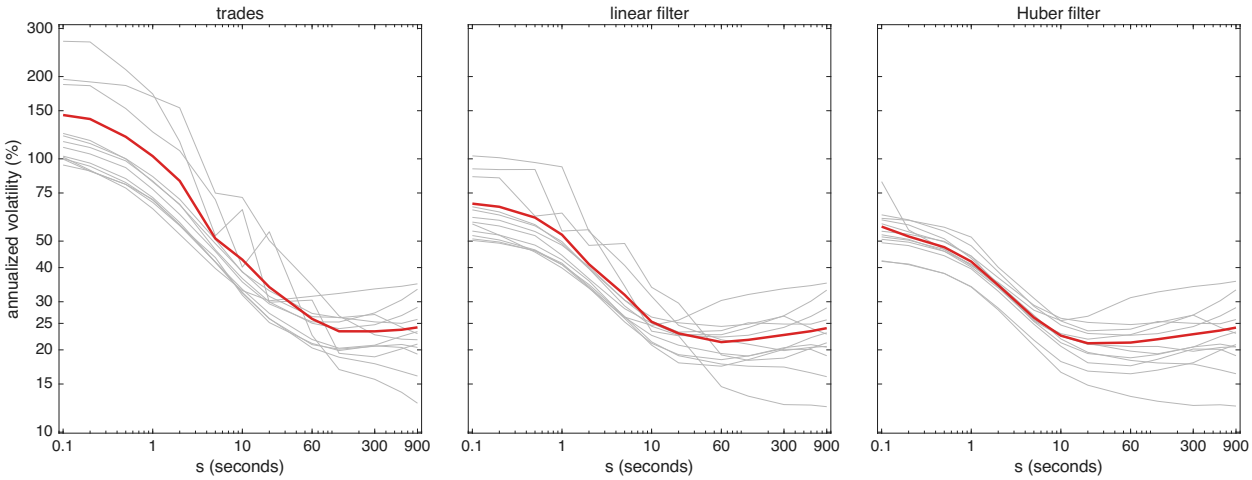

Figure 4: Volatility signature curves for `ADBE` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

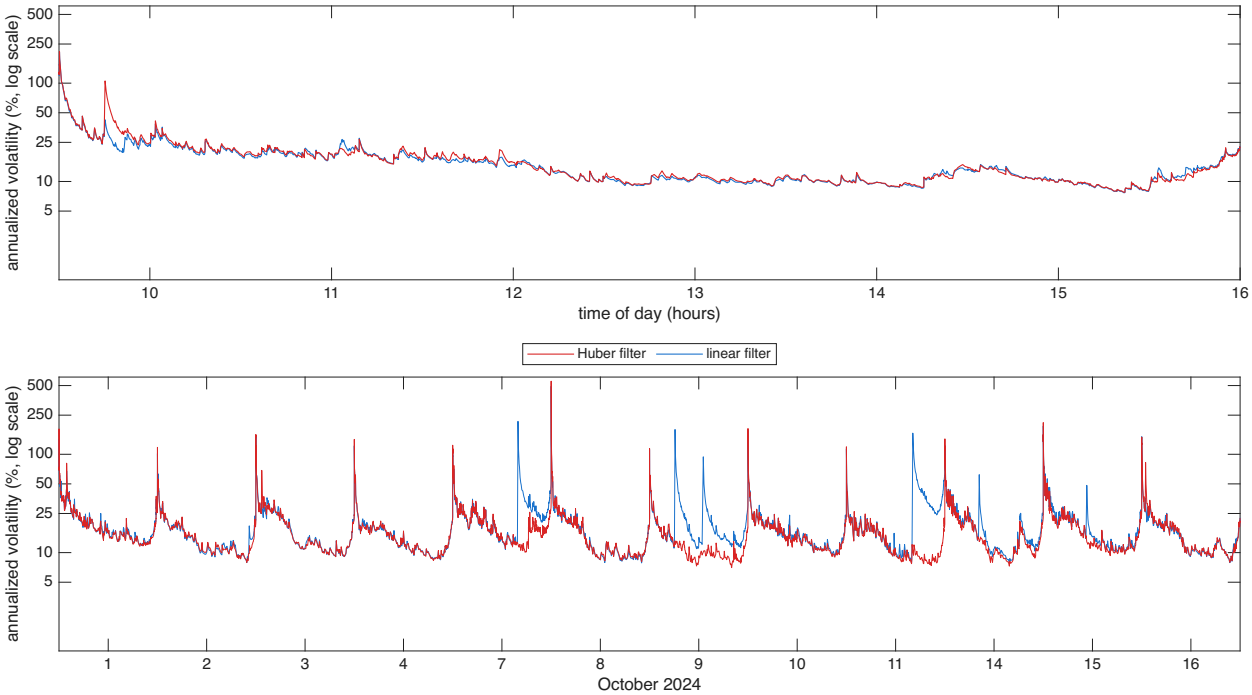

Figure 5: Filtered spot volatility for `ADBE` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 61$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

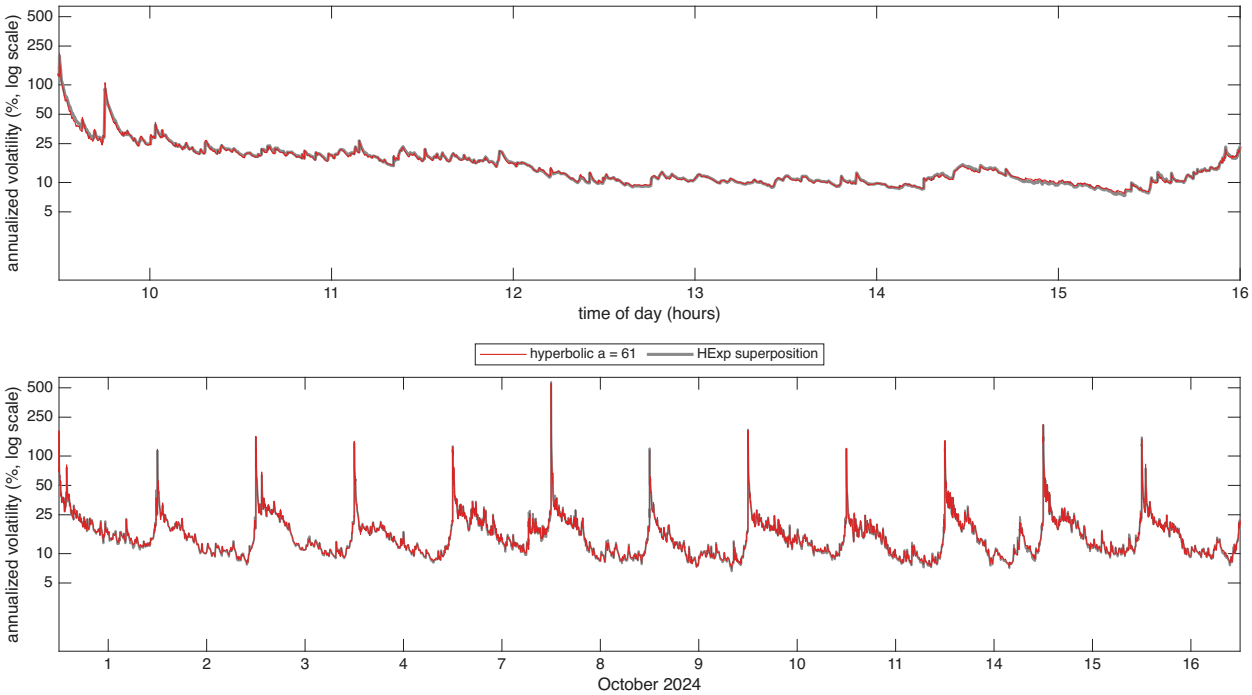

Figure 6: Hyperbolic against superposition weights for `ADBE` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 61$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

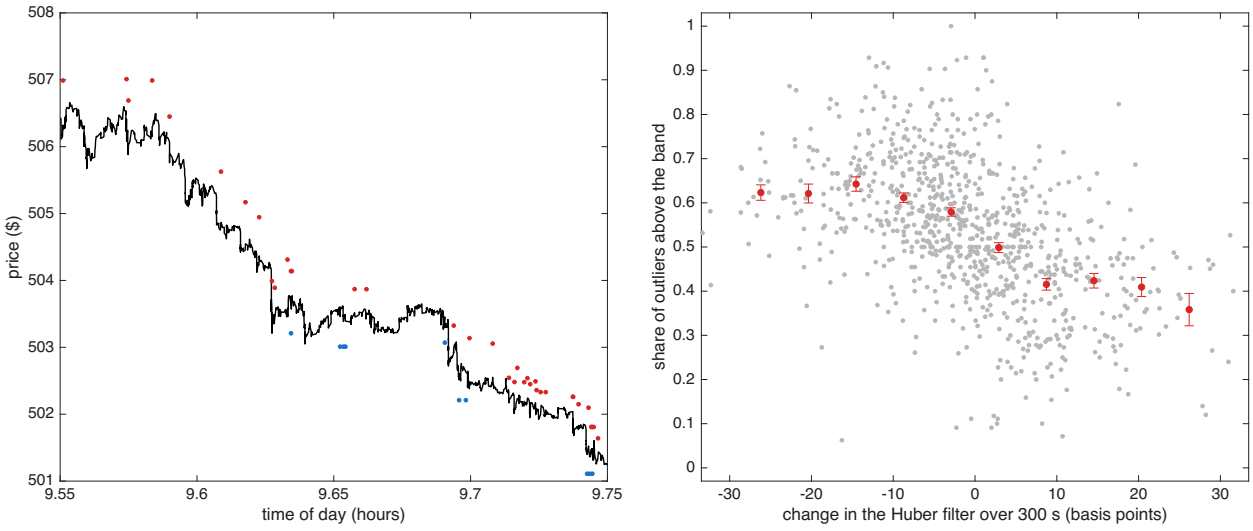

Figure 7: Outliers & robust price for `ADBE` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

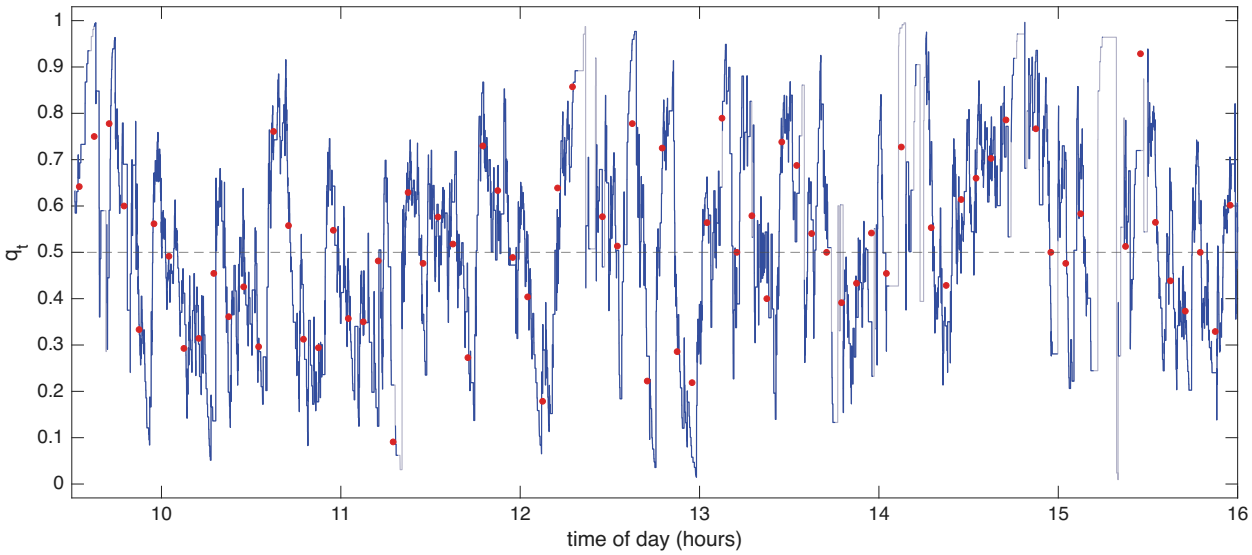

Figure 8: The directional conditional probability for ADBE, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 43 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

Selecting a on single days gives values from 19 to 408 trades, against 61 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 31 trades for this stock.

Diagnostics.

COST 1 to 16 October 2024

trades on 16 October 2024	72,911
average trading rate	3.0 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	4,085 (5.60%)
price half-life: run / minimizer[1]	3.0 / 4.0 trades
cost of running 3 trades	+0.96%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.02%
direction half-life[3]	35 s [29, 42]
Mandelbrot a minimizing QLIKE	53 trades
a within 0.0005 of the minimum	33 to 86
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	17% $\rightarrow$ 24%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	17% $\rightarrow$ 27%
median spot volatility, Huber / linear input[5]	15.5 / 15.7 %
hyperbolic against HExp, correlation of logs	0.9988
HExp QLIKE advantage, leave-one-day-out[6]	+0.00013
15-minute signature level	10 to 20 %
Huber signature, 1 min over 1 s, 12 day average	0.92

COST over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	3.290	2.379	2.276	11.834	7.346	3.620	3.402	20.757
2	2.782	2.407	2.385	9.155	6.009	3.584	3.447	15.160
5	2.856	2.393	2.409	6.341	5.062	3.463	3.387	9.969
10	2.693	2.298	2.313	4.852	4.271	3.240	3.208	7.342
25	2.383	2.097	2.097	3.545	3.420	2.880	2.862	5.060
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.33	−0.02	0.03	−0.47	4.54	3.12	3.95	4.09
2	0.08	0.01	−0.01	−0.01	4.61	3.50	4.43	4.35
5	−0.02	−0.01	−0.02	0.01	4.19	3.80	4.57	4.25
10	0.00	−0.02	−0.01	0.00	4.41	4.19	4.42	4.24
25	−0.02	−0.02	−0.01	−0.01	4.69	4.70	5.20	4.21

Table 1: Trade-lag differences, `COST`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	18.17	16.17	16.02	33.38	27.9	14.5	14.3	85.5
1 s	10.95	9.87	9.79	17.91	21.9	14.7	14.5	55.7
5 s	7.46	6.83	6.78	10.55	17.3	14.2	14.0	31.3
10 s	6.26	5.78	5.73	8.21	15.3	13.5	13.3	24.3
1 min	4.86	4.71	4.67	5.43	13.7	13.3	13.1	15.8
15 min	5.63	5.61	5.60	5.71	15.5	15.5	15.4	15.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.08	−0.05	−0.04	−0.21	8.67	9.21	8.53	5.87
1 s	−0.20	−0.01	0.02	−0.37	6.04	10.74	9.40	7.86
5 s	−0.07	0.02	0.04	−0.32	6.38	6.68	6.68	6.72
10 s	0.02	0.09	0.09	−0.25	7.58	8.15	8.04	7.27
1 min	−0.04	−0.01	−0.01	−0.13				

Table 2: Clock-lag differences, `COST`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	7.06	8.51	8.75
s.d. of $\hat{\epsilon}_t$ in cents	14.02	13.17	13.07
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.96
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.12	0.05	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.04	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−4.72,4.64	−4.98,4.86
95% C.I. for $\theta_t - X_t$	—	−8.27,8.13	−8.88,8.44
99% C.I. for $\theta_t - X_t$	—	−12.29,12.42	−13.25,13.04
s.e. of $\theta_t - X_t$ in cents	—	3.82	4.26

Table 3: Residual summaries, `COST`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.96, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.28	4.26	2.70 to 3.92	2.80 to 4.26
Residual divided by δ_t	3.78	3.89	2.74 to 4.64	2.69 to 4.63
Repricing windows removed	3.47	4.25	2.70 to 4.12	2.82 to 4.25
Calm windows only	3.78	4.43		
Every 10th trade	3.99	3.08	2.53 to 4.48	2.41 to 4.90
Six-trade half-life	3.34	4.46	2.87 to 4.13	2.93 to 4.46

Table 4: Hill estimator of the tail index of the Huber filter residual, `COST`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 121.2 to 227.8 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 81.4 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

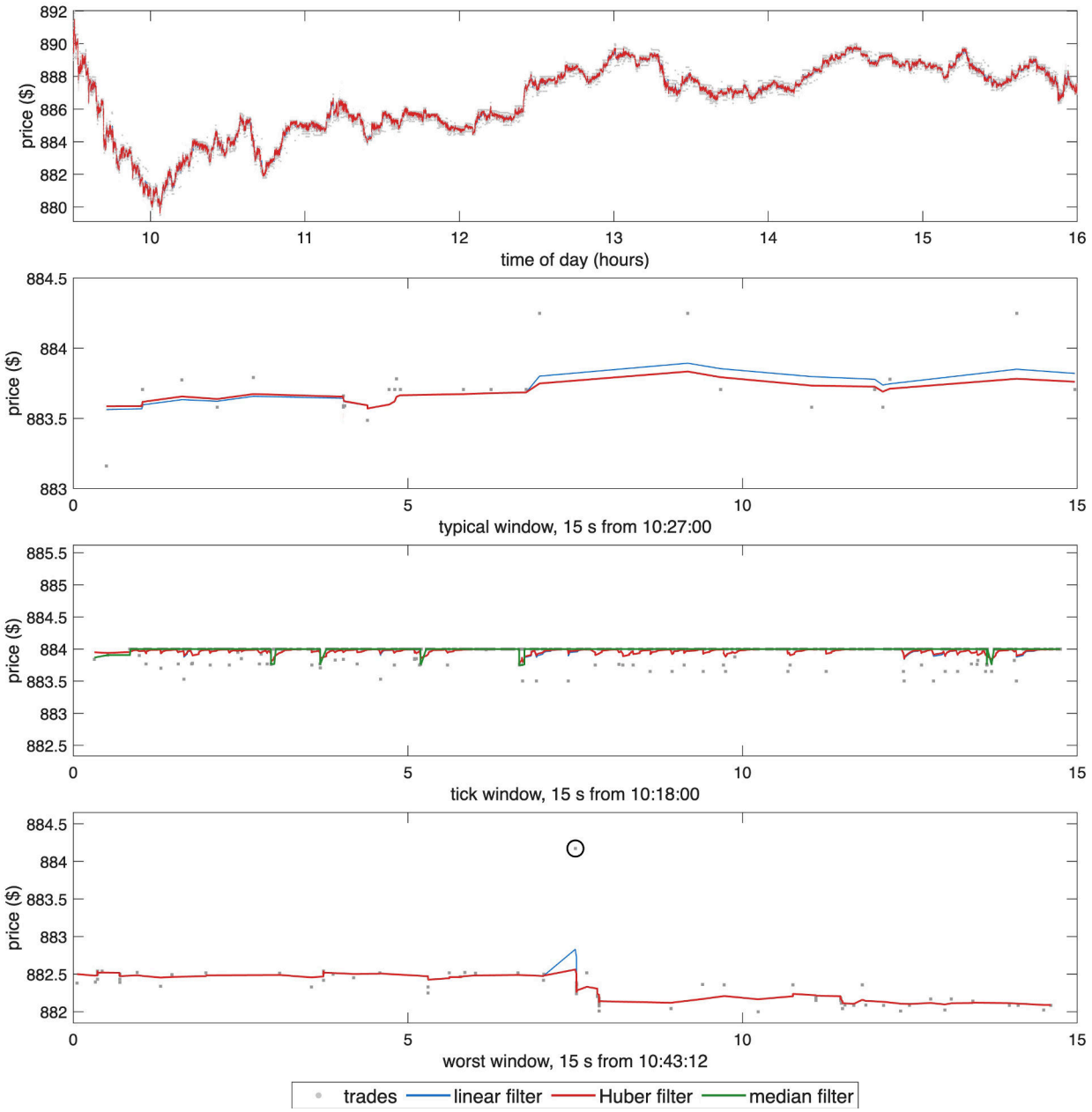

Figure 1: COST trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:27:00. 3rd: 15-second window at 10:18:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

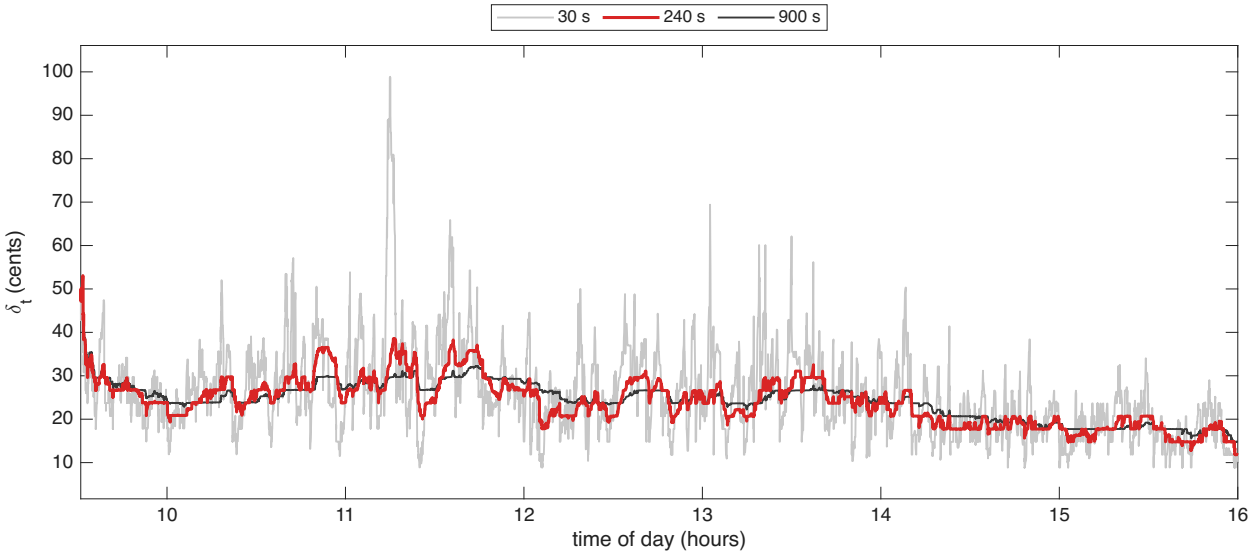

Figure 2: The Huber threshold δ_t in cents for `COST` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

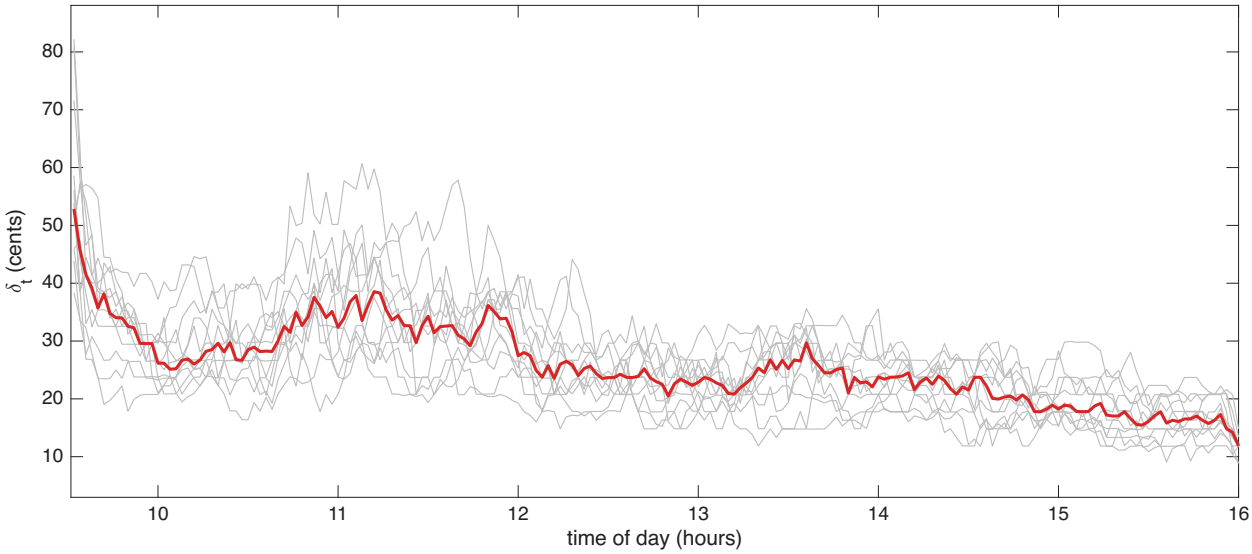

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `COST` trading days in Oct 2024 (grey lines) and their median (red line).

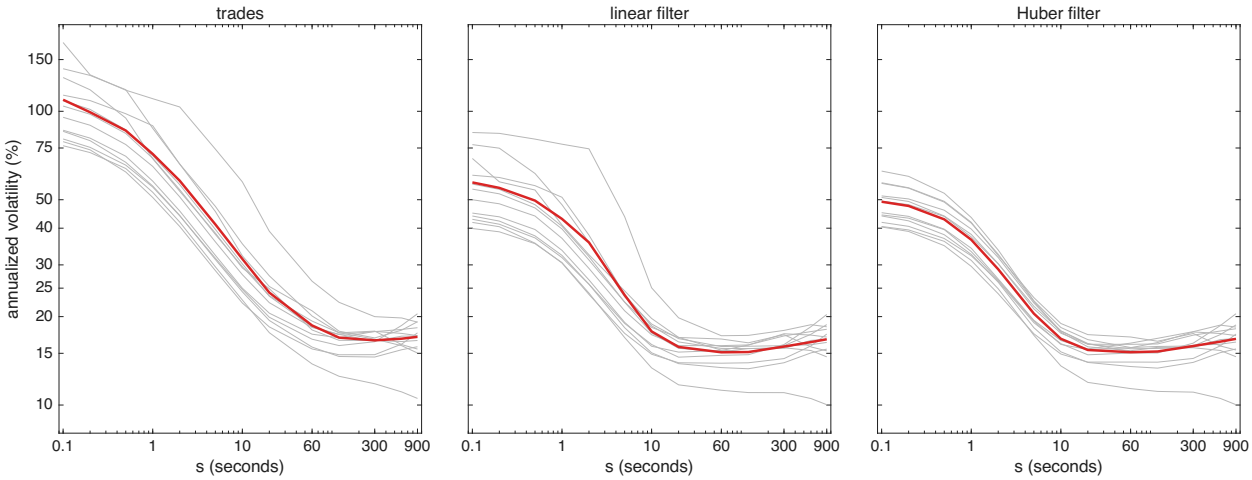

Figure 4: Volatility signature curves for `COST` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

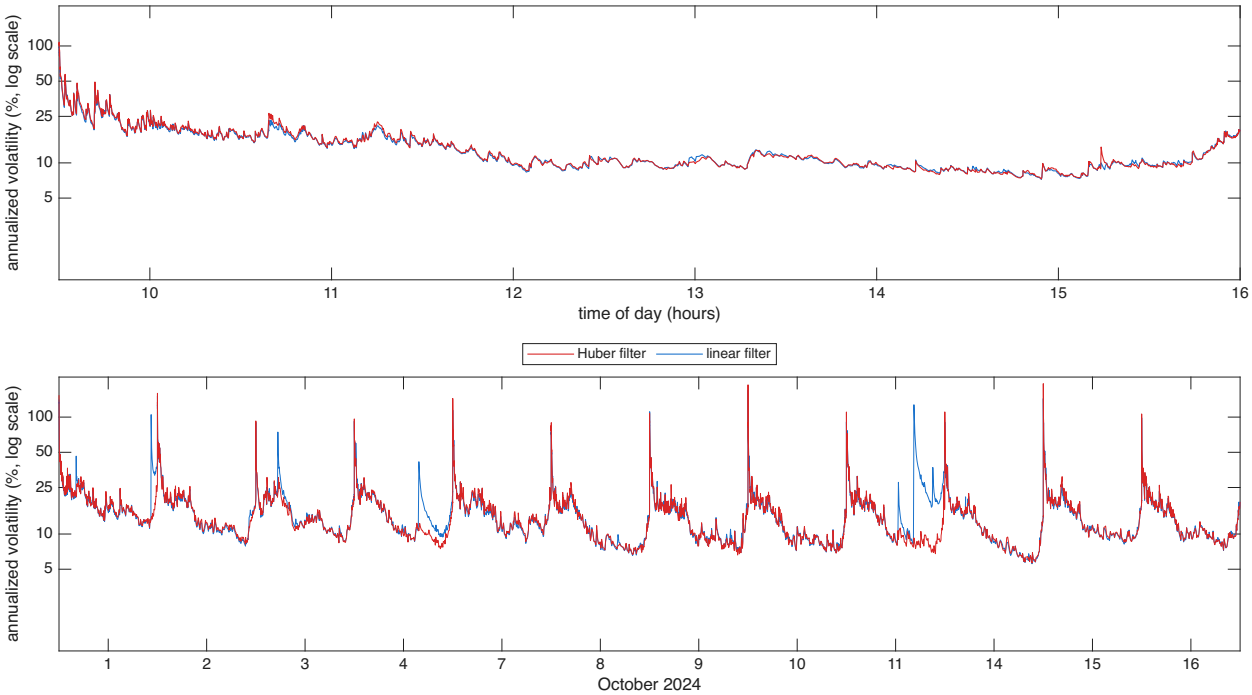

Figure 5: Filtered spot volatility for `COST` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 53$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

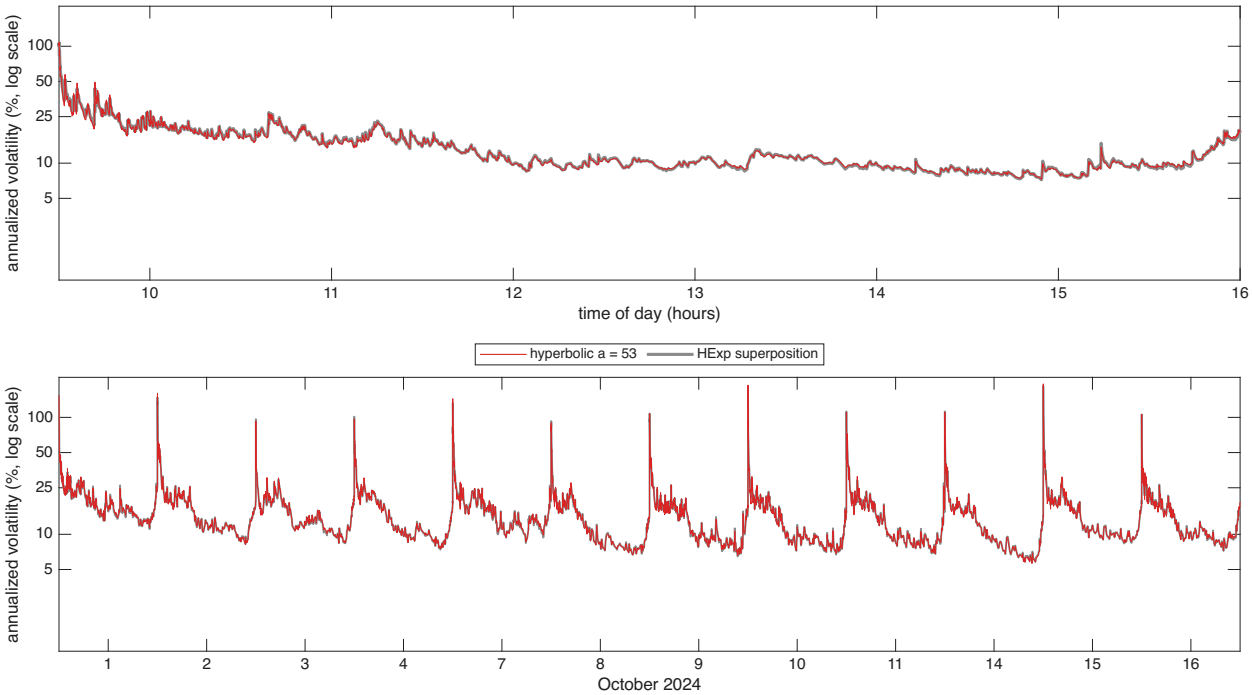

Figure 6: Hyperbolic against superposition weights for `COST` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 53$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

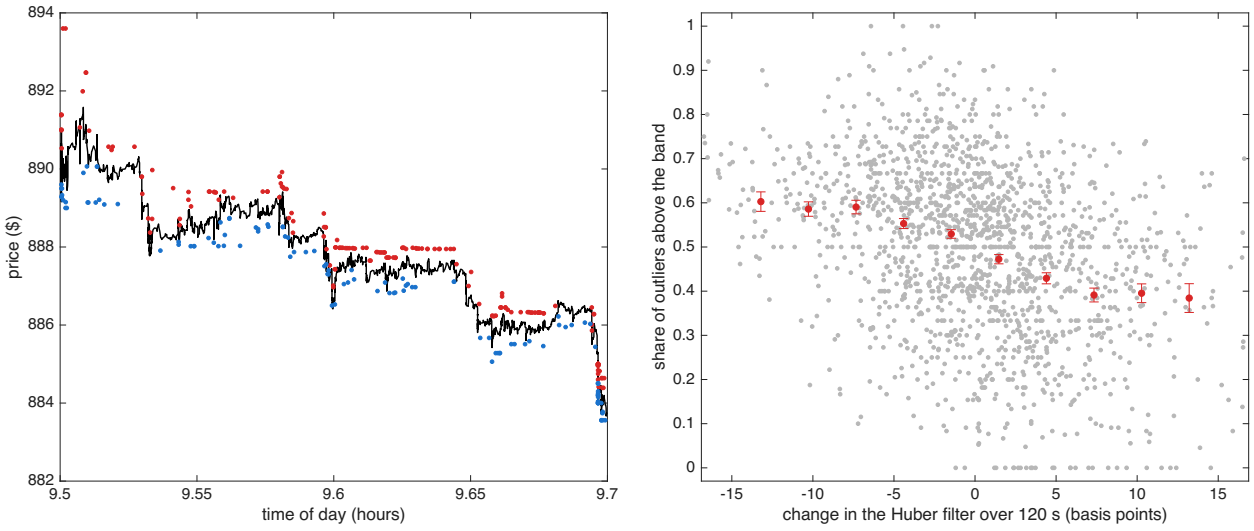

Figure 7: Outliers & robust price for `COST` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 120 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 120 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

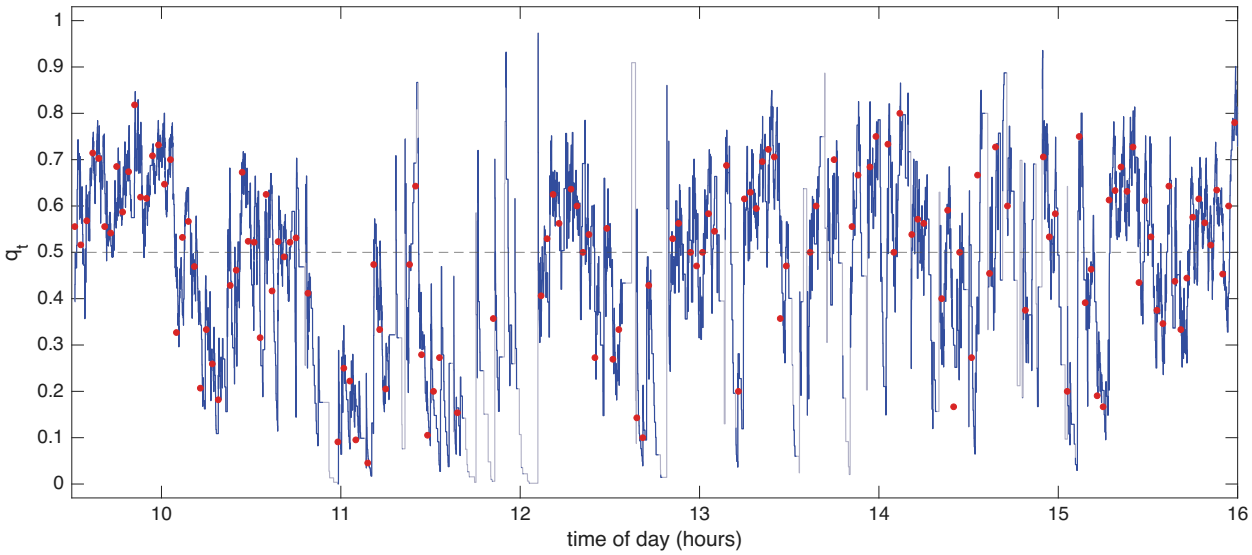

Figure 8: The directional conditional probability for COST, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 35 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 120 second bins of the previous figure.

At the horizons k = 2, 3, 5, 10, 30 and 100 the half-life minimizing the k-trade-ahead Huber forecast loss is 4, 4, 4, 5, 5 and 4 trades, within one grid point of the end of the 1 to 5 trade grid searched.
The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:43.

Diagnostics.

DKNG 1 to 16 October 2024

trades on 16 October 2024	63,023
average trading rate	2.8 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,267 (2.01%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.06%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	68 s [57, 82]
Mandelbrot a minimizing QLIKE	46 trades
a within 0.0005 of the minimum	33 to 65
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	51% $\to$ 54%
Huber $\hat{\sigma}$, before $\to$ peak	50% $\to$ 56%
median spot volatility, Huber / linear input[5]	18.0 / 21.4 %
hyperbolic against HExp, correlation of logs	0.9952
HExp QLIKE advantage, leave-one-day-out[6]	+0.00119
15-minute signature level	24 to 53 %
Huber signature, 1 min over 1 s, 12 day average	1.55

DKNG over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:45; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.101	0.090	0.098	0.412	0.345	0.190	0.195	1.123
2	0.102	0.103	0.116	0.339	0.303	0.196	0.207	0.838
5	0.121	0.121	0.135	0.264	0.271	0.208	0.224	0.572
10	0.136	0.134	0.146	0.226	0.256	0.220	0.233	0.438
25	0.159	0.155	0.160	0.204	0.264	0.248	0.251	0.344
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.07	0.13	−0.44	2.08	2.22	3.37	3.49
2	0.01	0.06	0.07	−0.01	2.44	2.25	3.39	3.49
5	0.00	0.03	0.02	0.00	2.86	2.82	3.69	3.52
10	0.02	0.04	0.02	0.01	3.28	3.44	4.17	3.71
25	0.00	0.01	0.01	0.00	5.31	5.20	5.71	4.37

Table 1: Trade-lag differences, `DKNG`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.77	0.72	0.76	1.46	24.0	16.1	18.8	103.7
1 s	0.52	0.50	0.53	0.84	23.3	18.8	20.6	79.4
5 s	0.40	0.39	0.41	0.56	25.8	24.3	25.5	52.0
10 s	0.38	0.38	0.39	0.49	28.4	27.2	27.8	43.9
1 min	0.40	0.39	0.39	0.43	31.7	31.1	30.8	33.9
15 min	0.37	0.37	0.37	0.38	30.1	30.0	29.9	30.3
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.11	−0.06	−0.05	−0.19	7.79	4.57	4.61	5.82
1 s	−0.10	−0.10	−0.10	−0.31	3.28	3.20	3.70	4.92
5 s	0.02	0.07	0.05	−0.28	4.78	5.47	5.53	5.78
10 s	0.11	0.15	0.14	−0.16	10.62	9.17	8.35	6.61
1 min	−0.02	0.00	0.00	−0.07				

Table 2: Clock-lag differences, `DKNG`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.28	0.36	0.38
s.d. of $\hat{\epsilon}_t$ in cents	0.84	0.79	0.75
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.36
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.25	0.13
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.15	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.08	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.06	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.02	0.01
80% C.I. for $\theta_t - X_t$	—	−0.27,0.25	−0.29,0.28
95% C.I. for $\theta_t - X_t$	—	−0.52,0.52	−0.59,0.60
99% C.I. for $\theta_t - X_t$	—	−0.95,0.97	−1.09,1.13
s.e. of $\theta_t - X_t$ in cents	—	0.25	0.30

Table 3: Residual summaries, DKNG, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.36, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.87	2.53	2.15 to 3.04	1.82 to 3.25
Residual divided by δ_t	3.12	2.29	2.26 to 3.65	1.95 to 3.29
Repricing windows removed	2.83	2.63	2.11 to 3.07	1.83 to 3.38
Calm windows only	2.92	2.61		
Every 10th trade	2.29	2.08	1.65 to 2.29	1.67 to 2.59
Six-trade half-life	3.12	2.73	2.37 to 3.47	1.92 to 3.62

Table 4: Hill estimator of the tail index of the Huber filter residual, DKNG, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 10.1 to 16.3 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 5.0 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

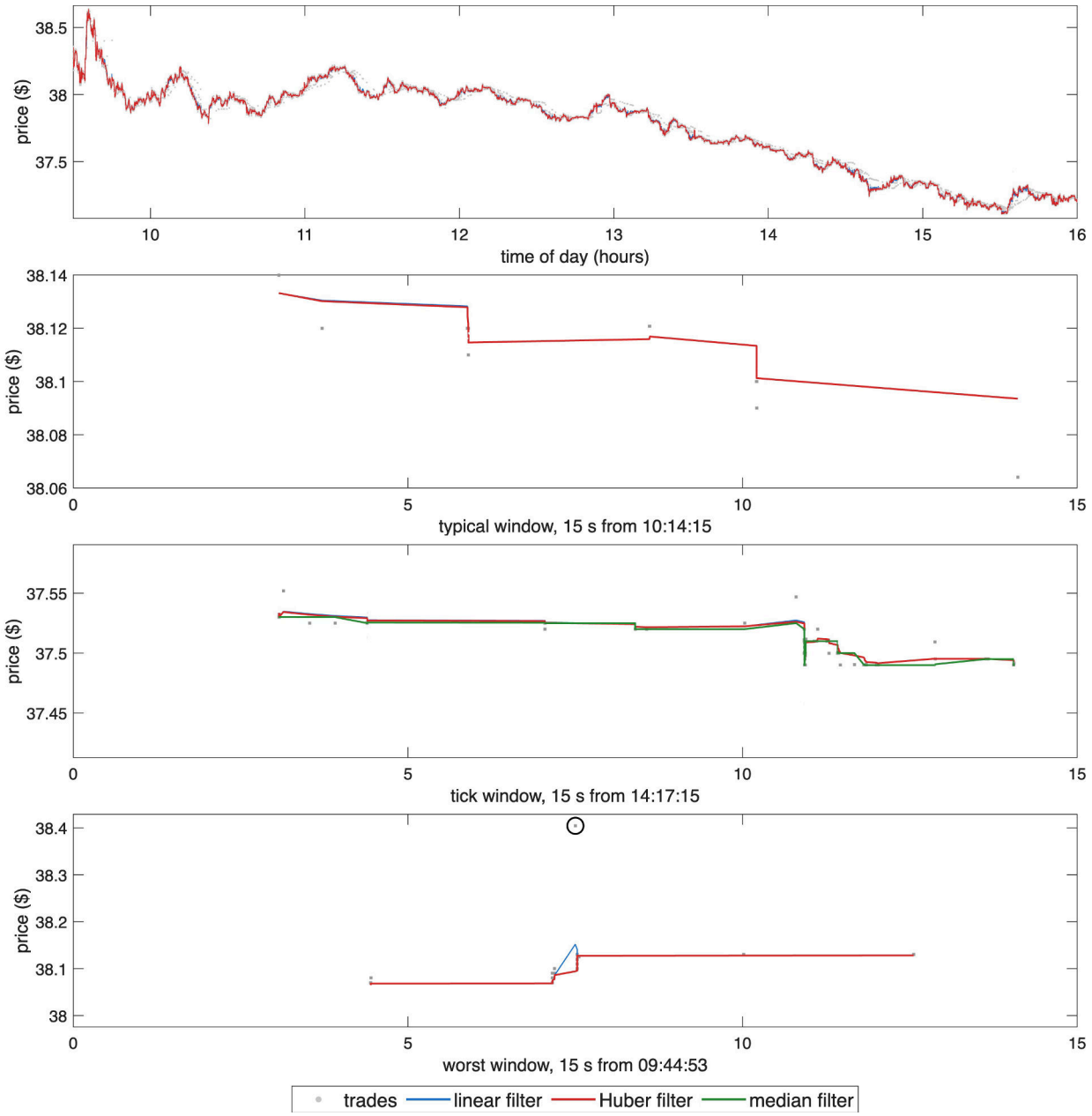

Figure 1: DKNG trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:14:15. 3rd: 15-second window at 14:17:15 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

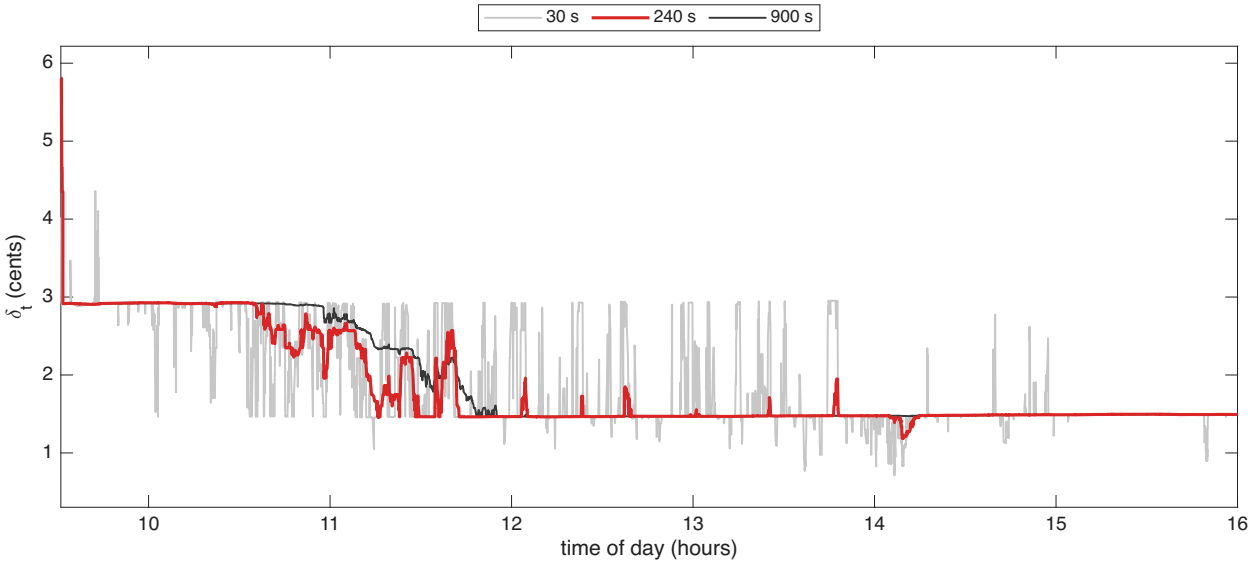

Figure 2: The Huber threshold δ_t in cents for `DKNG` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

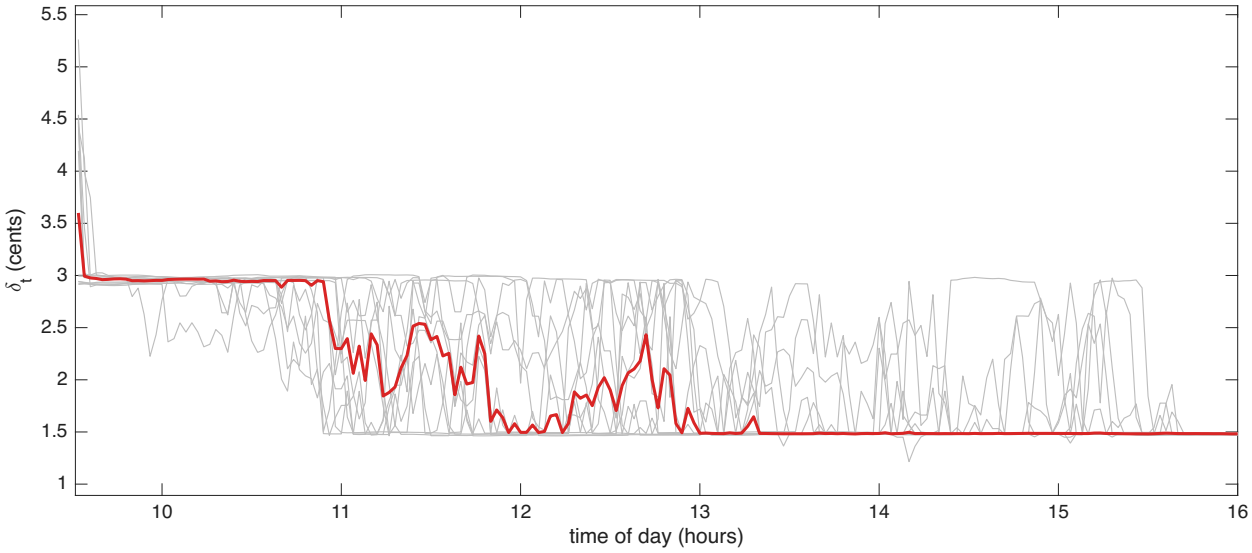

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `DKNG` trading days in Oct 2024 (grey lines) and their median (red line).

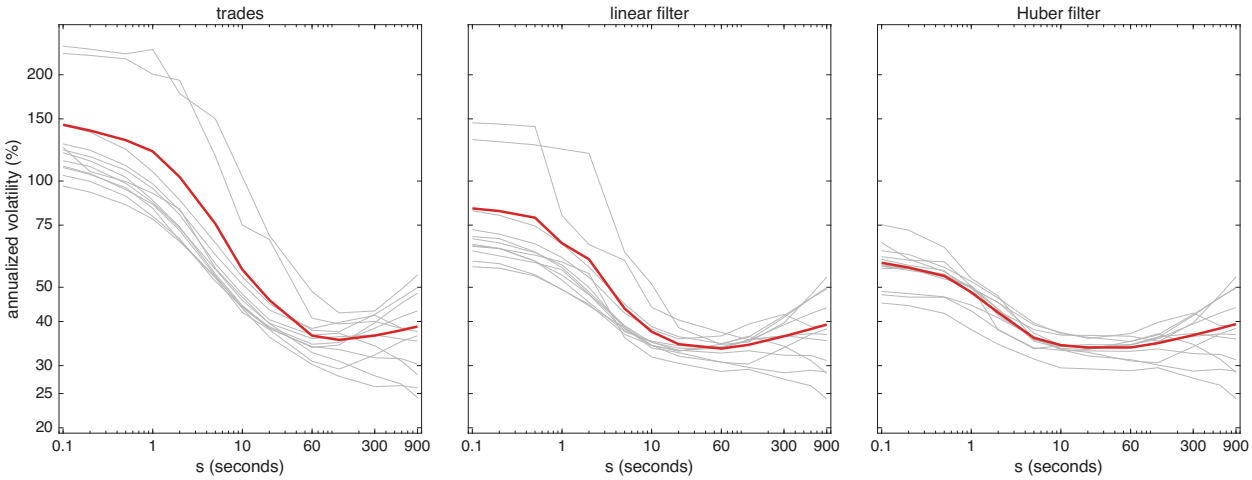

Figure 4: Volatility signature curves for `DKNG` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

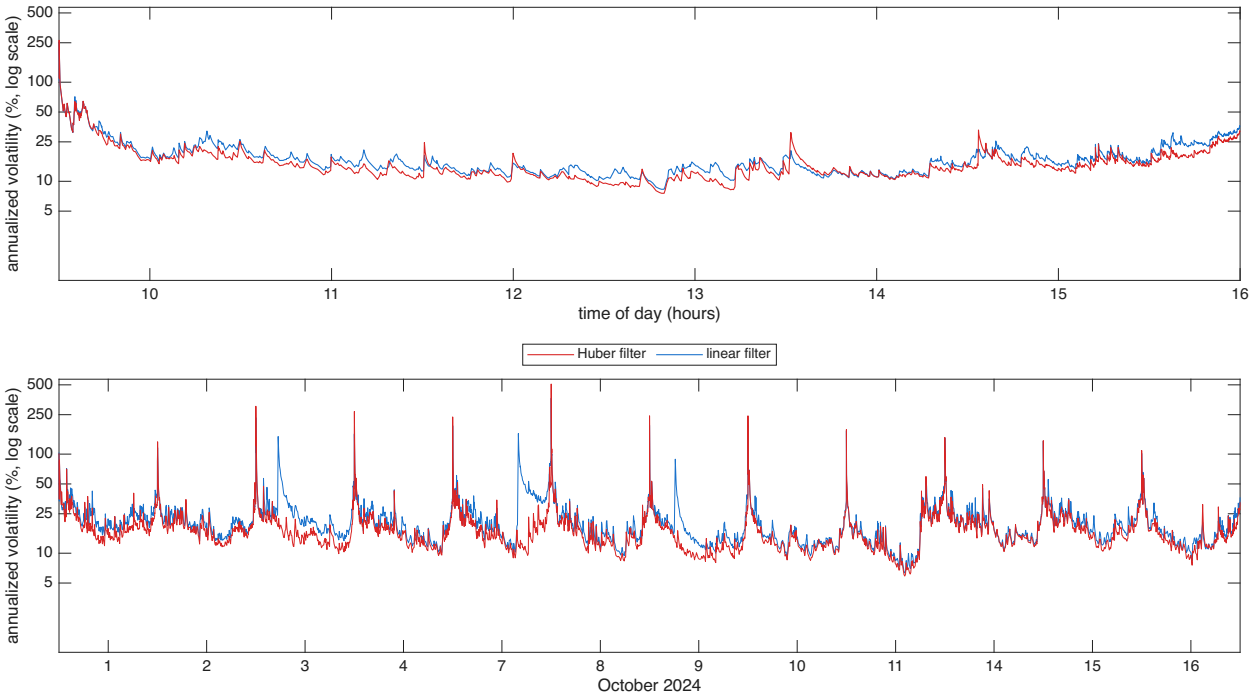

Figure 5: Filtered spot volatility for `DKNG` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 46$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

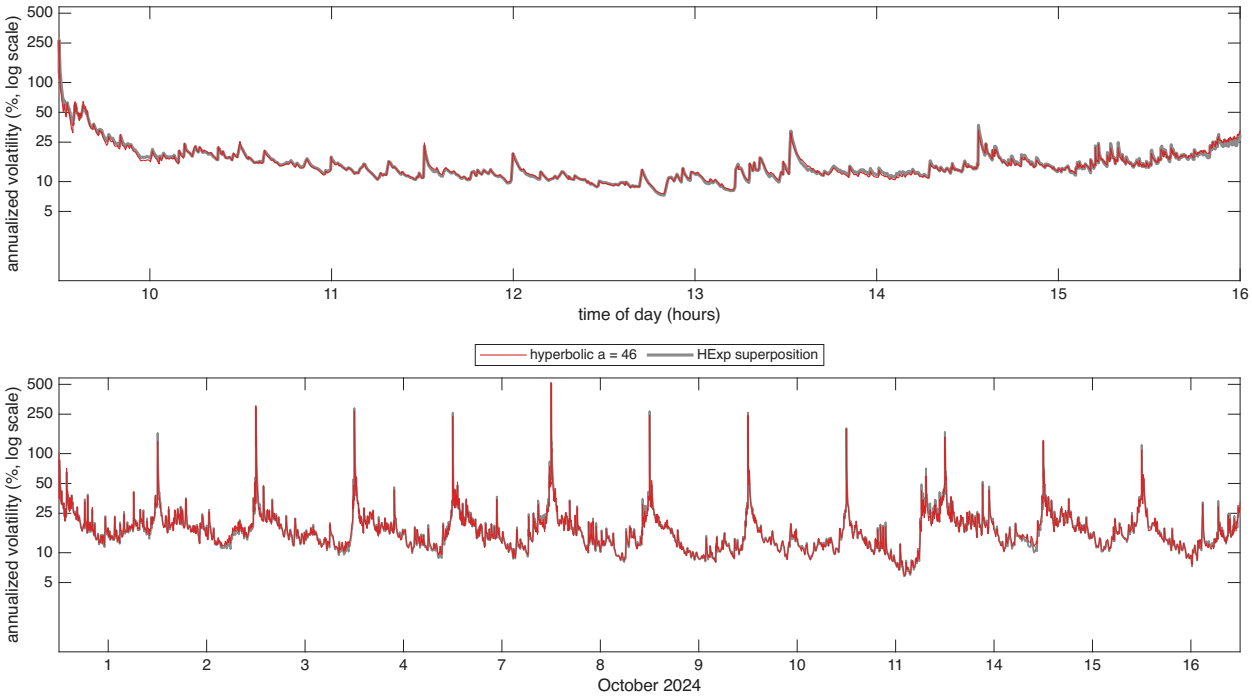

Figure 6: Hyperbolic against superposition weights for `DKNG` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 46$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 30, 150, 750, 3750 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

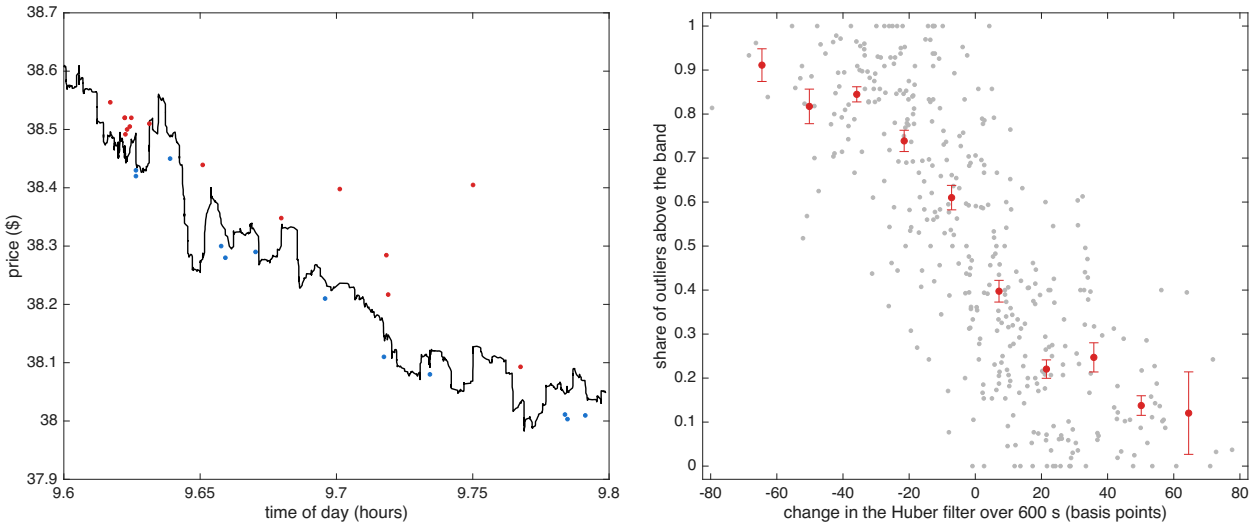

Figure 7: Outliers & robust price for `DKNG` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

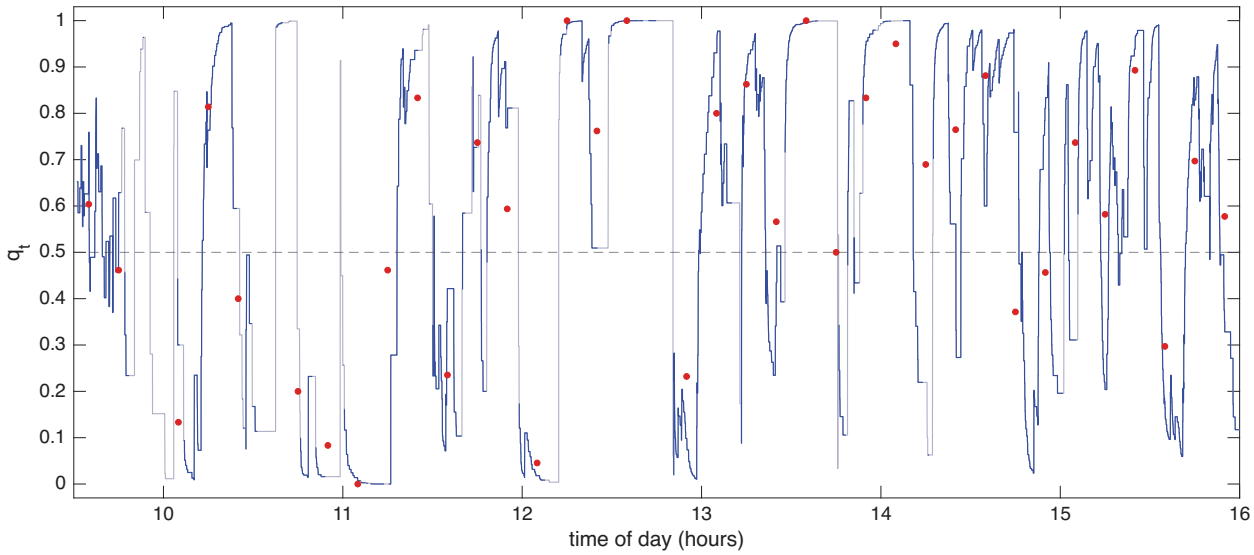

Figure 8: The directional conditional probability for DKNG, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 68 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 8 to 26112 trades, against 46 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 28 trades for this stock.

Diagnostics.

TXN 1 to 16 October 2024

trades on 16 October 2024	57,830
average trading rate	2.7 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,314 (4.00%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.37%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	51 s [43, 60]
Mandelbrot a minimizing QLIKE	36 trades
a within 0.0005 of the minimum	26 to 51
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	26% $\to$ 39%
Huber $\hat{\sigma}$, before $\to$ peak	28% $\to$ 40%
median spot volatility, Huber / linear input[5]	13.9 / 15.8 %
hyperbolic against HExp, correlation of logs	0.9983
HExp QLIKE advantage, leave-one-day-out[6]	-0.00348
15-minute signature level	14 to 36 %
Huber signature, 1 min over 1 s, 12 day average	1.47

TXN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:50; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.578	0.468	0.478	2.036	1.800	1.065	0.868	4.877
2	0.577	0.528	0.553	1.698	1.597	1.069	0.943	3.709
5	0.679	0.609	0.638	1.320	1.486	1.130	1.038	2.598
10	0.740	0.666	0.684	1.118	1.366	1.131	1.072	1.999
25	0.777	0.719	0.720	0.948	1.196	1.077	1.050	1.492
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.01	0.18	−0.42	1.91	1.65	2.83	2.59
2	0.10	0.15	0.08	−0.03	1.98	1.75	2.77	2.58
5	0.02	0.00	−0.01	−0.02	2.03	1.89	2.81	2.57
10	0.01	0.02	−0.01	0.01	2.17	2.17	2.98	2.82
25	0.01	−0.02	−0.01	−0.01	2.59	2.76	3.26	2.81

Table 1: Trade-lag differences, `TXN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.85	3.46	3.46	6.39	21.2	14.3	14.4	75.2
1 s	2.79	2.57	2.54	4.17	19.7	15.6	15.6	56.4
5 s	2.14	2.05	2.06	2.77	21.0	19.2	19.3	37.2
10 s	1.98	1.93	1.93	2.43	21.9	20.9	20.8	30.7
1 min	1.77	1.79	1.77	1.96	24.5	24.3	24.1	26.2
15 min	2.13	2.17	2.16	2.24	33.0	32.9	32.9	33.3
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.05	0.00	0.03	−0.17	4.30	4.21	4.19	3.04
1 s	−0.08	0.00	0.08	−0.22	4.06	3.28	3.36	4.57
5 s	0.03	0.07	0.09	−0.17	4.23	5.09	5.17	3.76
10 s	−0.03	0.00	−0.01	−0.18	4.34	4.36	4.26	4.55
1 min	0.07	0.08	0.08	0.03				

Table 2: Clock-lag differences, `TXN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.42	1.75	1.83
s.d. of $\hat{\epsilon}_t$ in cents	3.65	3.45	3.34
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.84
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.22	0.28	0.18
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.15	0.08
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.03	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−1.19,1.18	−1.29,1.29
95% C.I. for $\theta_t - X_t$	—	−2.15,2.13	−2.42,2.36
99% C.I. for $\theta_t - X_t$	—	−3.91,3.50	−4.84,4.12
s.e. of $\theta_t - X_t$ in cents	—	1.09	1.35

Table 3: Residual summaries, `TXN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.84, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.10	2.54	2.10 to 3.12	2.10 to 3.18
Residual divided by δ_t	2.80	2.79	2.15 to 3.20	2.29 to 3.07
Repricing windows removed	2.63	2.70	2.10 to 3.24	2.10 to 3.23
Calm windows only	2.65	2.66		
Every 10th trade	2.50	2.13	1.85 to 2.64	1.73 to 2.22
Six-trade half-life	2.09	2.86	2.09 to 3.51	2.23 to 3.26

Table 4: Hill estimator of the tail index of the Huber filter residual, `TXN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 30.2 to 40.3 cents over the days and removes 0% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 17.8 cents on 16 October 2024, 92% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

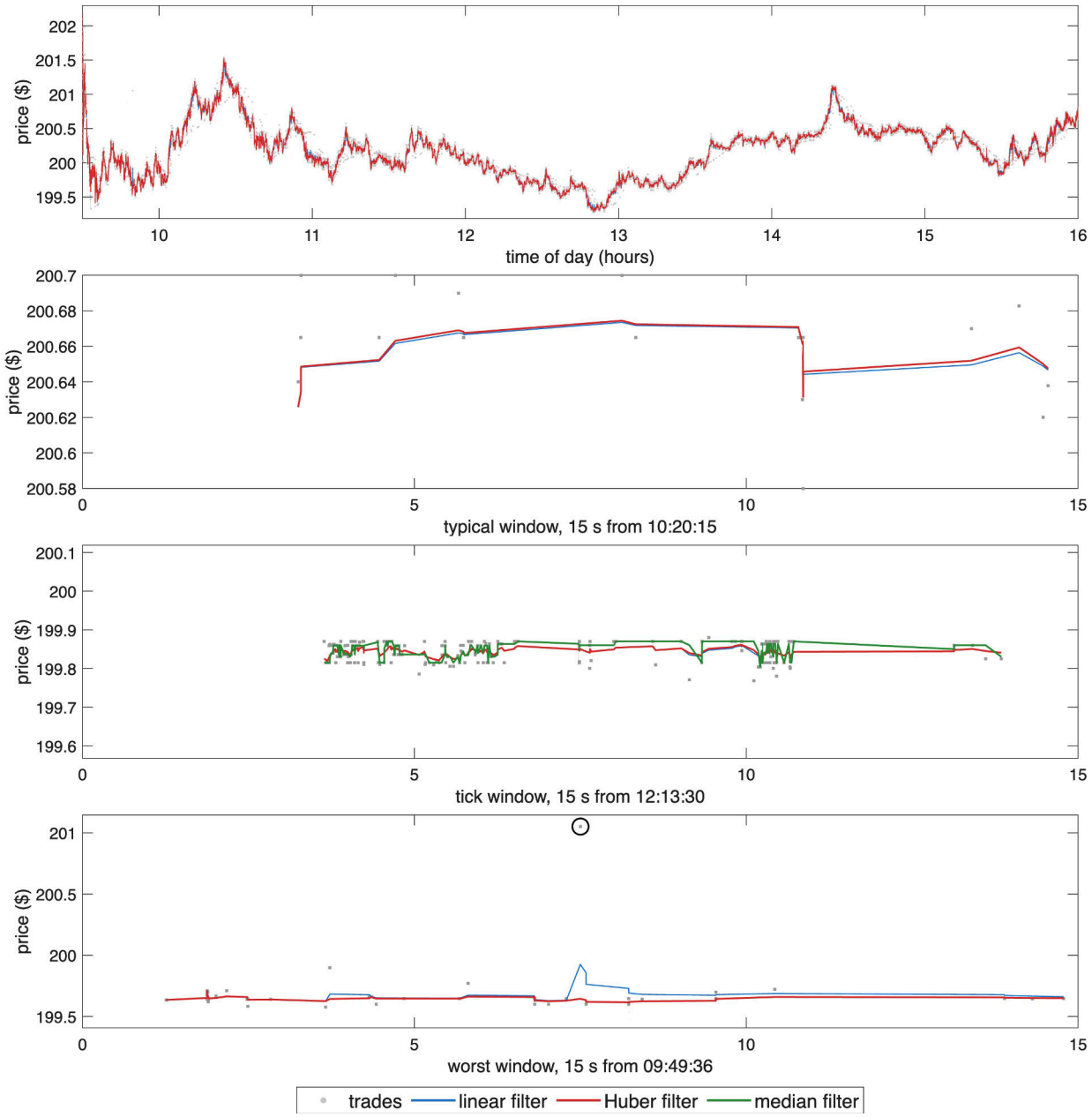

Figure 1: `TXN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:20:15. 3rd: 15-second window at 12:13:30 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

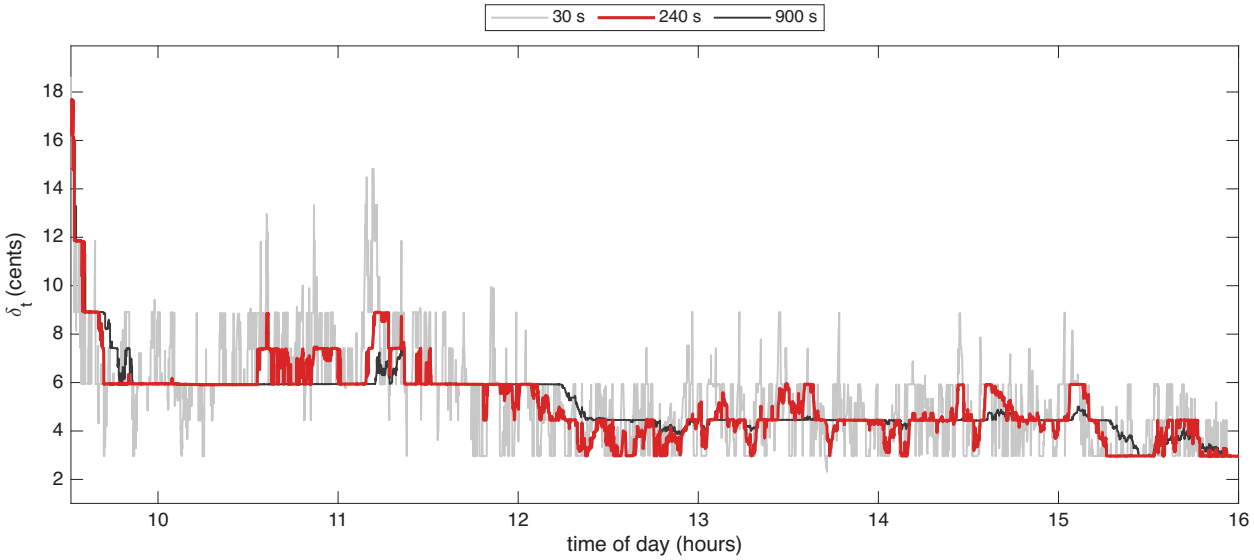

Figure 2: The Huber threshold δ_t in cents for `TXN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

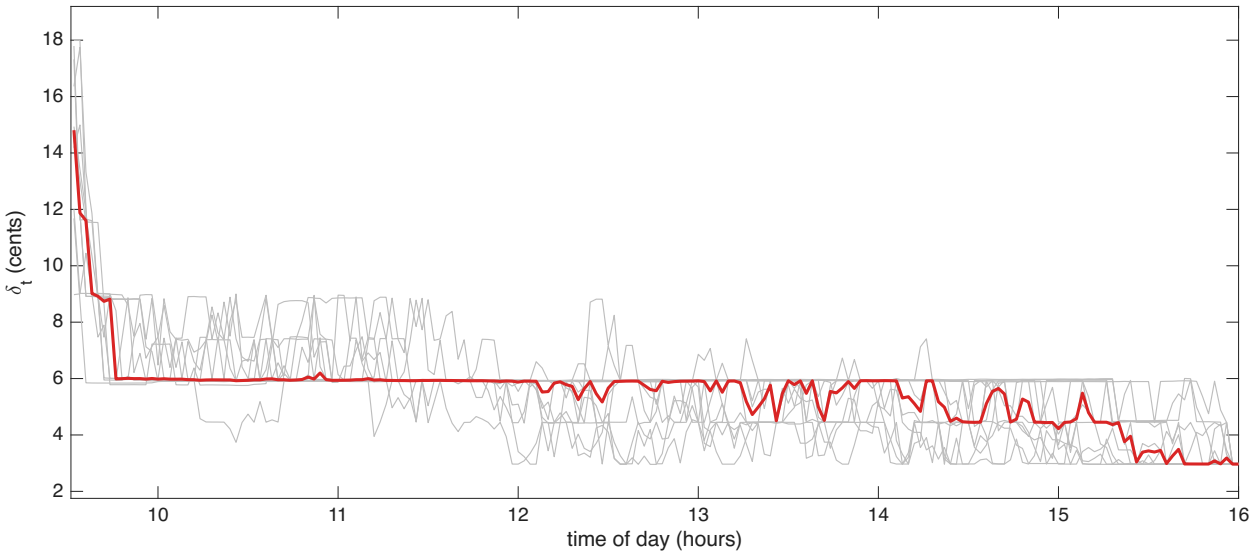

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `TXN` trading days in Oct 2024 (grey lines) and their median (red line).

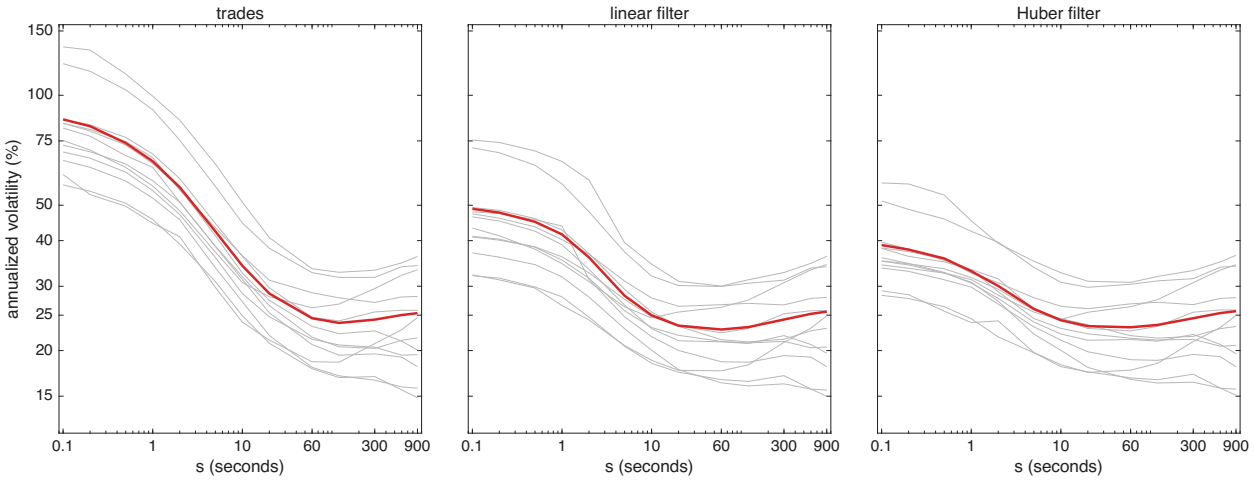

Figure 4: Volatility signature curves for `TXN` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

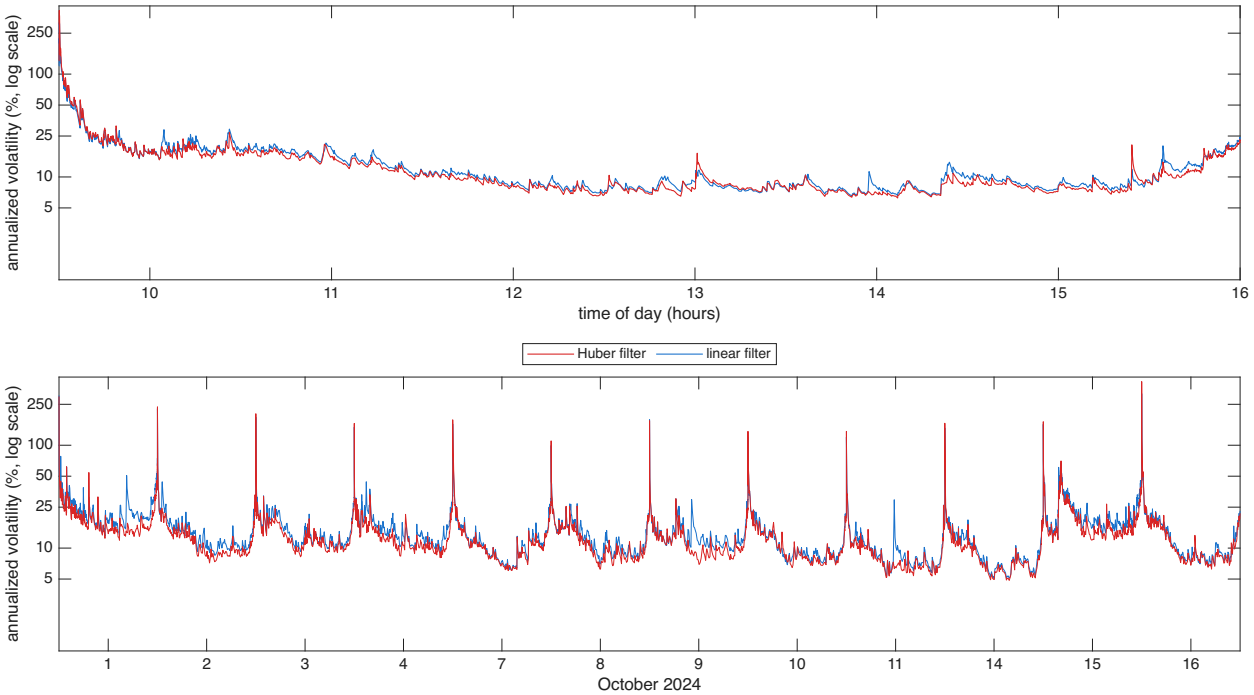

Figure 5: Filtered spot volatility for `TXN` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 36$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

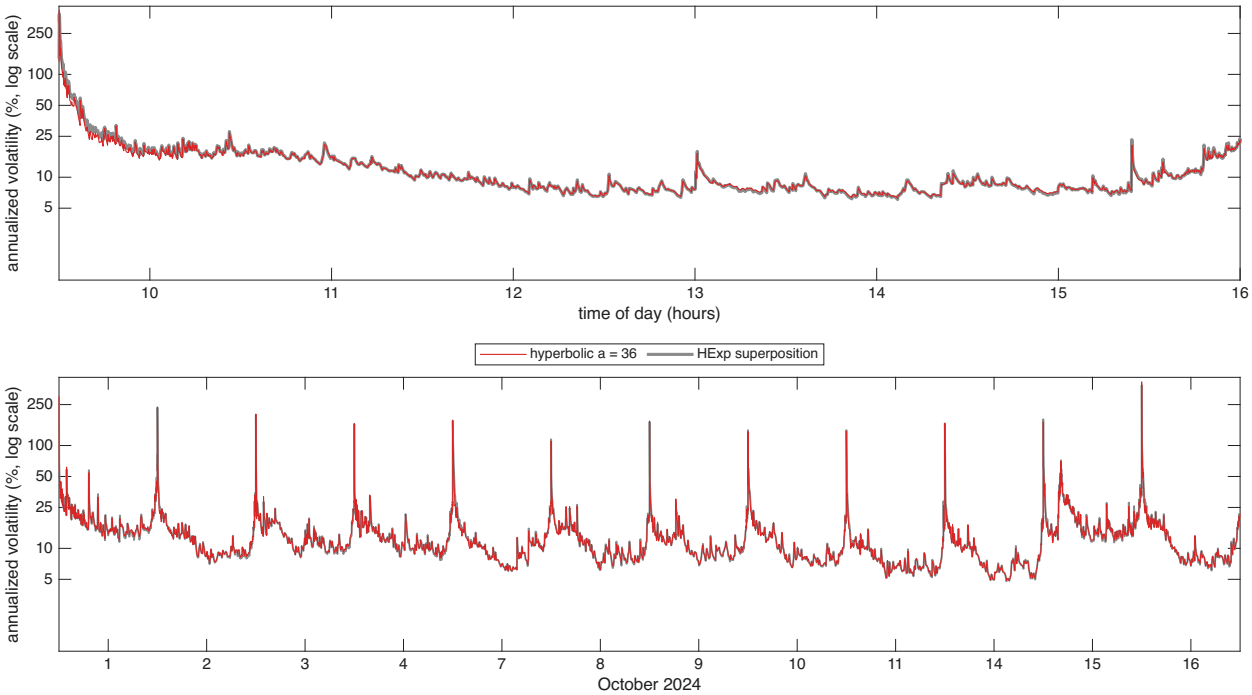

Figure 6: Hyperbolic against superposition weights for `TXN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 36$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

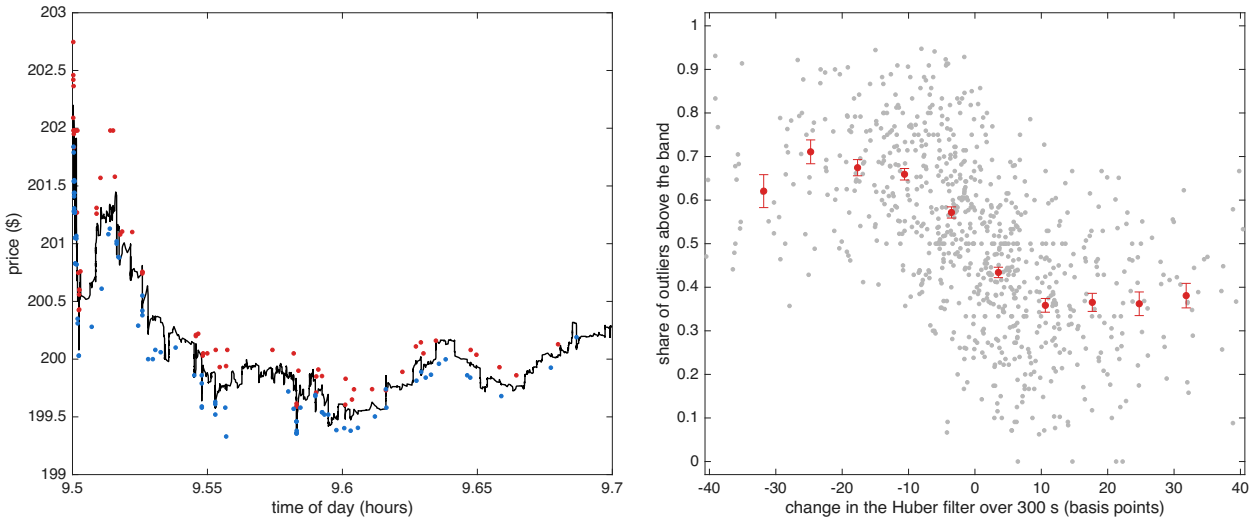

Figure 7: Outliers & robust price for `TXN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

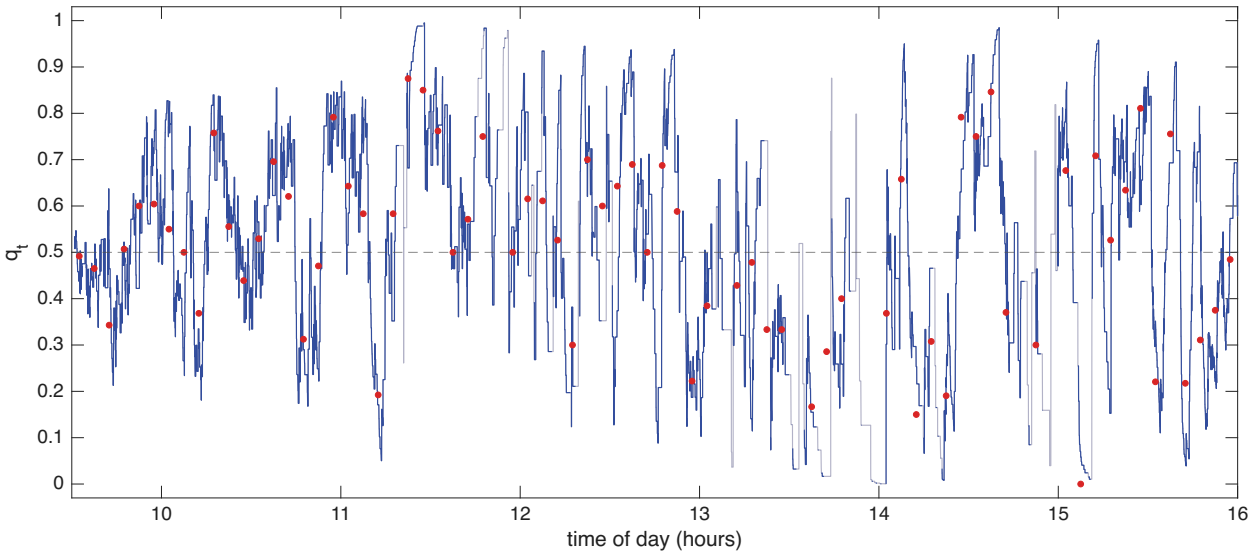

Figure 8: The directional conditional probability for `TXN`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 51 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

No diagnostic flags.

Diagnostics.

PANW 1 to 16 October 2024

trades on 16 October 2024	40,492
average trading rate	2.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,004 (4.95%)
price half-life: run / minimizer[1]	3.0 / 2.7 trades
cost of running 3 trades	+0.23%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	60 s [50, 72]
Mandelbrot a minimizing QLIKE	59 trades
a within 0.0005 of the minimum	40 to 95
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	17% $\rightarrow$ 27%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	16% $\rightarrow$ 20%
median spot volatility, Huber / linear input[5]	15.3 / 17.3 %
hyperbolic against HExp, correlation of logs	0.9986
HExp QLIKE advantage, leave-one-day-out[6]	+0.00077
15-minute signature level	18 to 40 %
Huber signature, 1 min over 1 s, 12 day average	1.31

PANW over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:31; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.703	1.439	1.406	7.067	4.218	2.477	2.298	13.540
2	1.548	1.536	1.547	5.601	3.652	2.564	2.403	10.147
5	1.689	1.617	1.646	4.031	3.263	2.568	2.457	6.820
10	1.687	1.600	1.629	3.162	2.792	2.381	2.360	5.042
25	1.635	1.551	1.551	2.399	2.373	2.170	2.152	3.543
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.25	0.07	0.09	−0.44	2.63	3.20	3.66	3.95
2	0.07	0.02	0.01	−0.03	2.71	2.72	3.13	3.59
5	−0.03	−0.03	−0.03	−0.01	2.88	2.65	3.15	3.42
10	0.00	−0.01	−0.01	0.00	3.25	3.12	3.41	3.52
25	−0.01	−0.01	−0.01	−0.01	3.18	3.42	4.17	3.61

Table 1: Trade-lag differences, `PANW`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	8.78	8.01	7.96	17.18	29.0	18.5	17.9	105.7
1 s	6.56	6.12	6.13	11.25	24.5	18.4	17.9	80.4
5 s	5.14	4.87	4.87	7.30	21.9	19.6	19.9	47.6
10 s	4.85	4.58	4.59	6.21	21.1	19.7	20.0	37.8
1 min	4.23	4.04	4.01	4.66	21.0	20.5	20.3	24.0
15 min	4.86	4.74	4.74	5.04	26.5	26.4	26.3	26.7
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.04	0.01	0.02	−0.12	4.71	4.82	5.87	4.57
1 s	−0.07	−0.04	−0.04	−0.39	3.85	3.91	4.88	3.17
5 s	−0.02	0.01	0.00	−0.38	5.30	5.31	6.03	3.67
10 s	0.04	0.08	0.05	−0.25	8.96	8.67	11.17	4.11
1 min	0.12	0.13	0.14	0.01				

Table 2: Clock-lag differences, `PANW`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	4.47	5.24	5.40
s.d. of $\hat{\epsilon}_t$ in cents	9.58	9.00	8.83
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.65
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.15	0.09	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.03	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	−0.01	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−3.23,3.08	−3.44,3.21
95% C.I. for $\theta_t - X_t$	—	−5.78,5.35	−6.37,5.61
99% C.I. for $\theta_t - X_t$	—	−9.21,8.75	−10.43,9.32
s.e. of $\theta_t - X_t$ in cents	—	2.69	3.09

Table 3: Residual summaries, `PANW`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.65, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.86	3.53	2.32 to 3.53	2.06 to 3.53
Residual divided by δ_t	3.14	3.47	2.10 to 3.14	2.18 to 3.47
Repricing windows removed	3.06	3.65	2.30 to 3.51	2.24 to 3.65
Calm windows only	3.03	3.45		
Every 10th trade	2.28	3.02	2.18 to 3.31	1.73 to 3.39
Six-trade half-life	2.75	3.90	2.48 to 3.81	2.20 to 3.90

Table 4: Hill estimator of the tail index of the Huber filter residual, `PANW`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 70.6 to 93.4 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 46.7 cents on 16 October 2024, 93% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

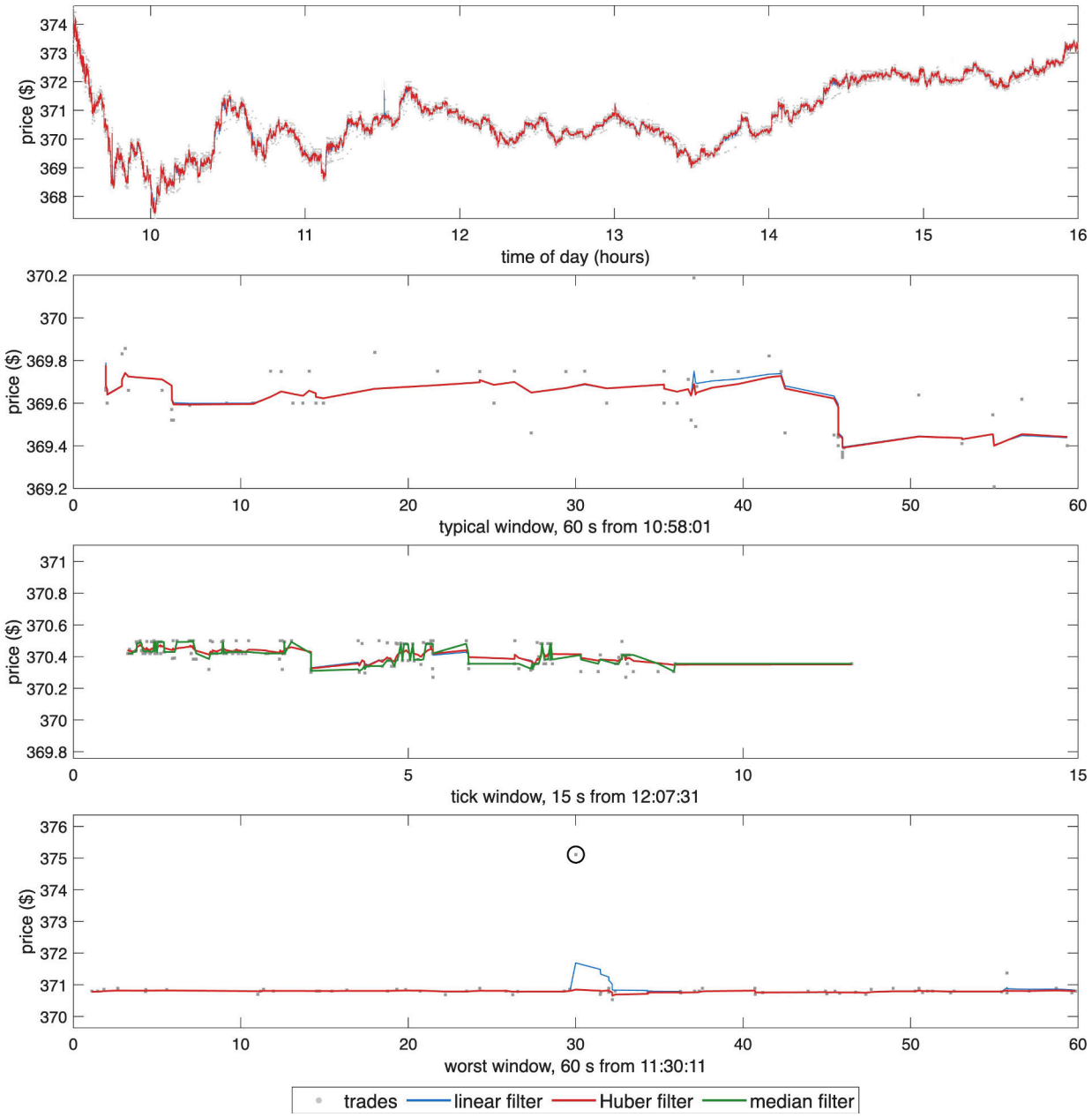

Figure 1: `PANW` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:58:01. 3rd: 15-second window at 12:07:31 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

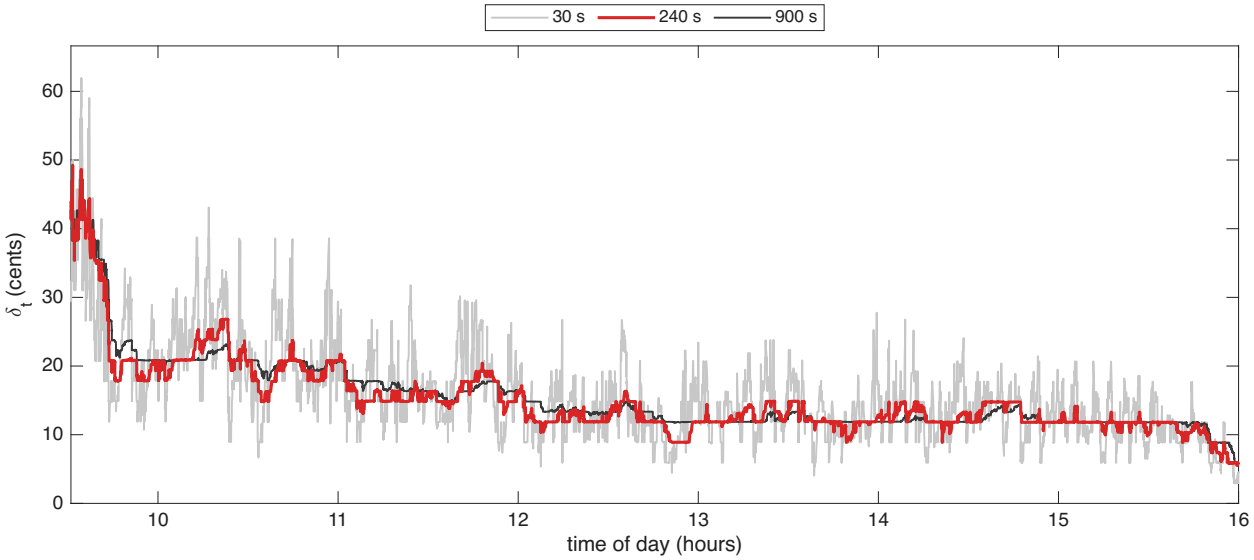

Figure 2: The Huber threshold δ_t in cents for `PANW` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

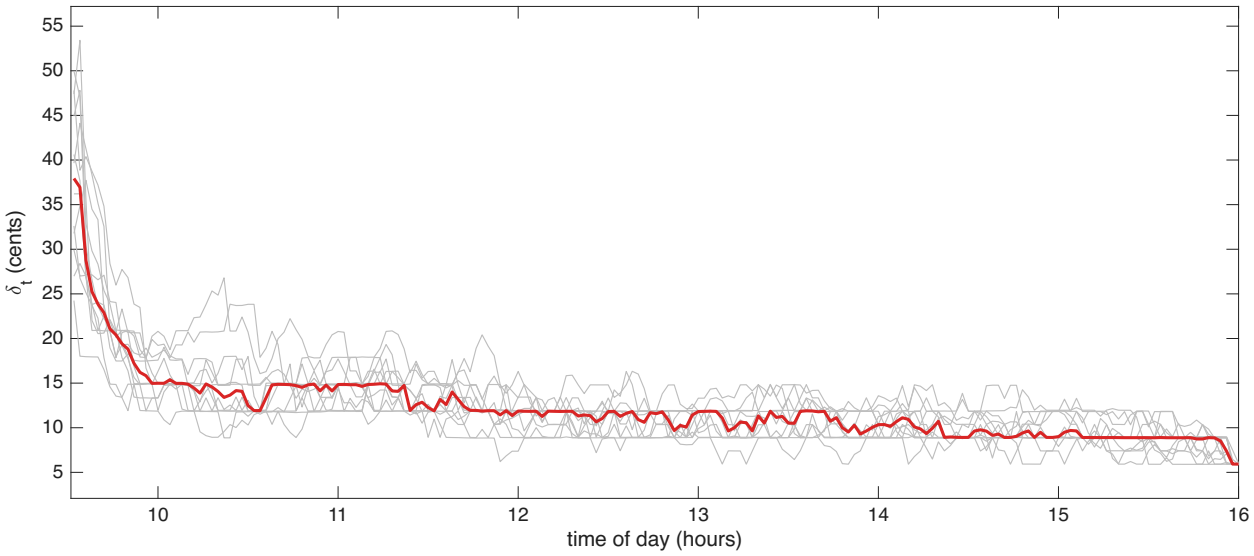

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `PANW` trading days in Oct 2024 (grey lines) and their median (red line).

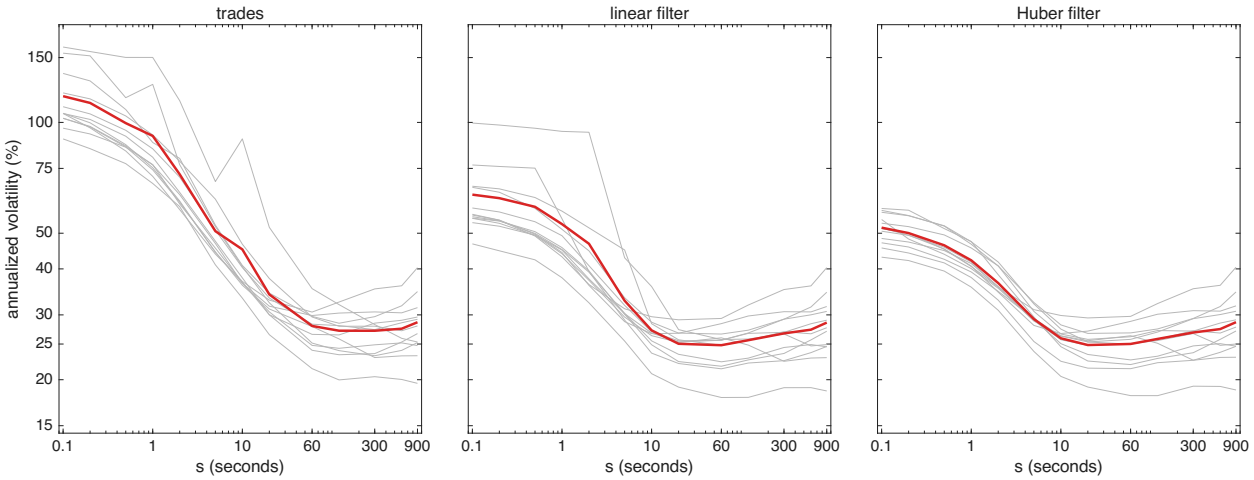

Figure 4: Volatility signature curves for `PANW` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

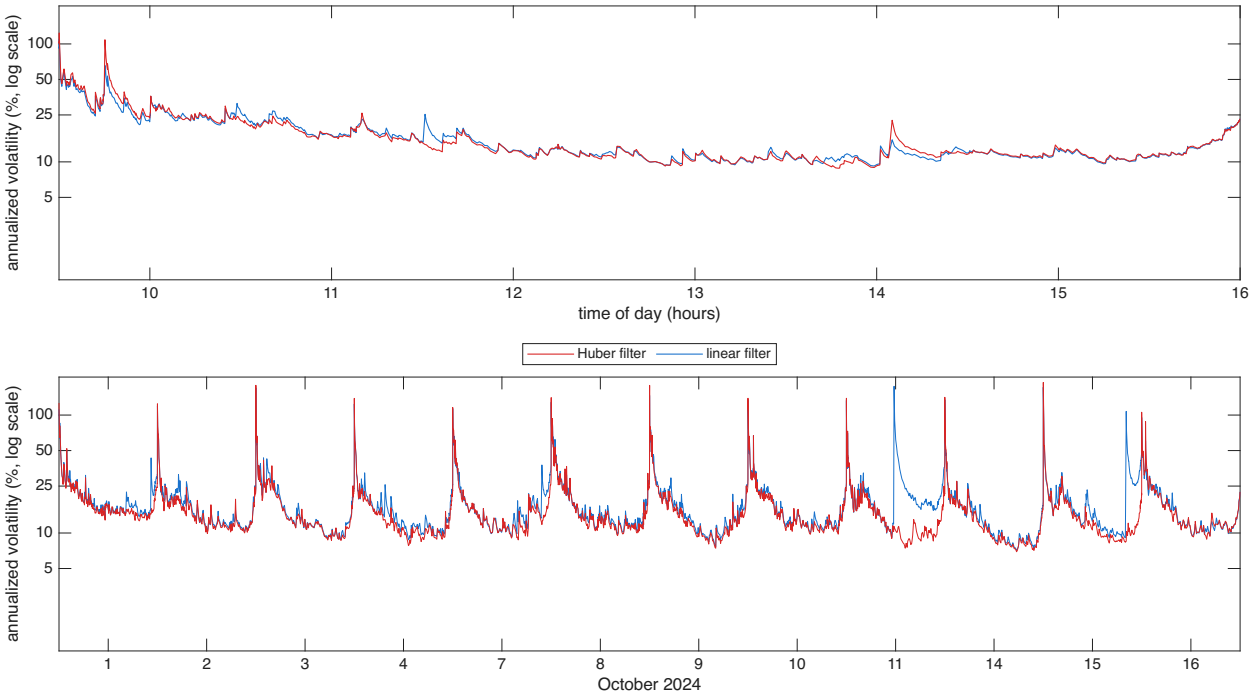

Figure 5: Filtered spot volatility for `PANW` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 59$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

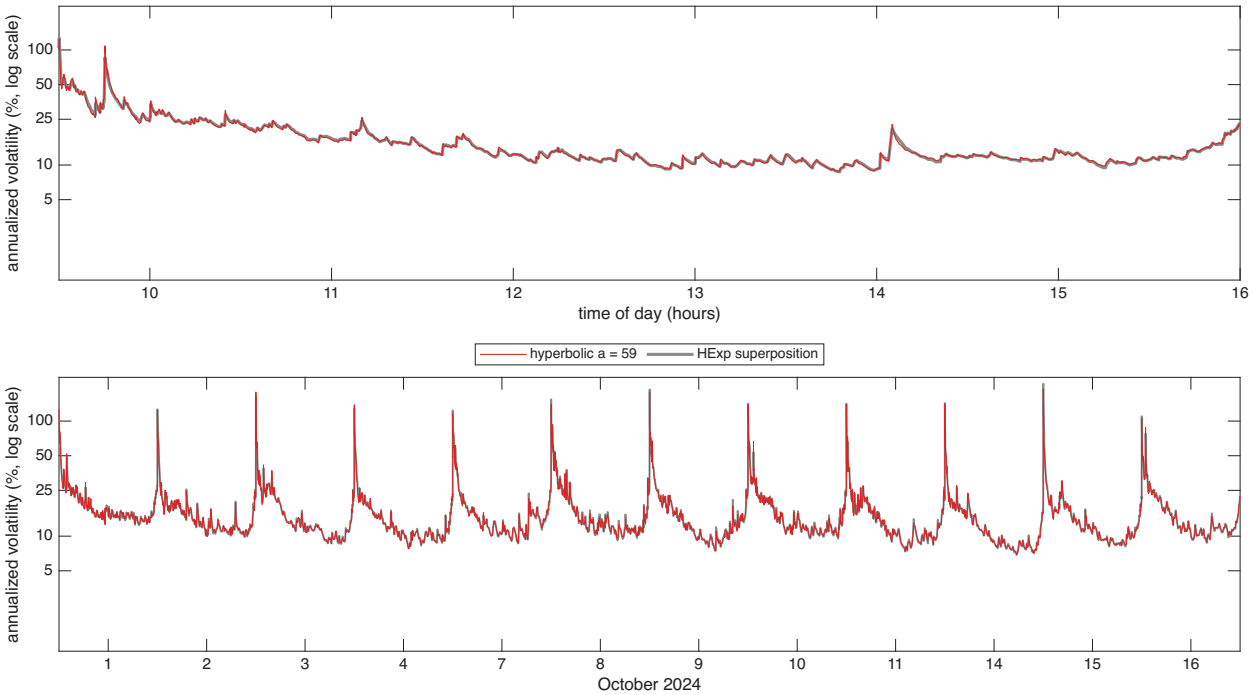

Figure 6: Hyperbolic against superposition weights for PANW over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 59$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

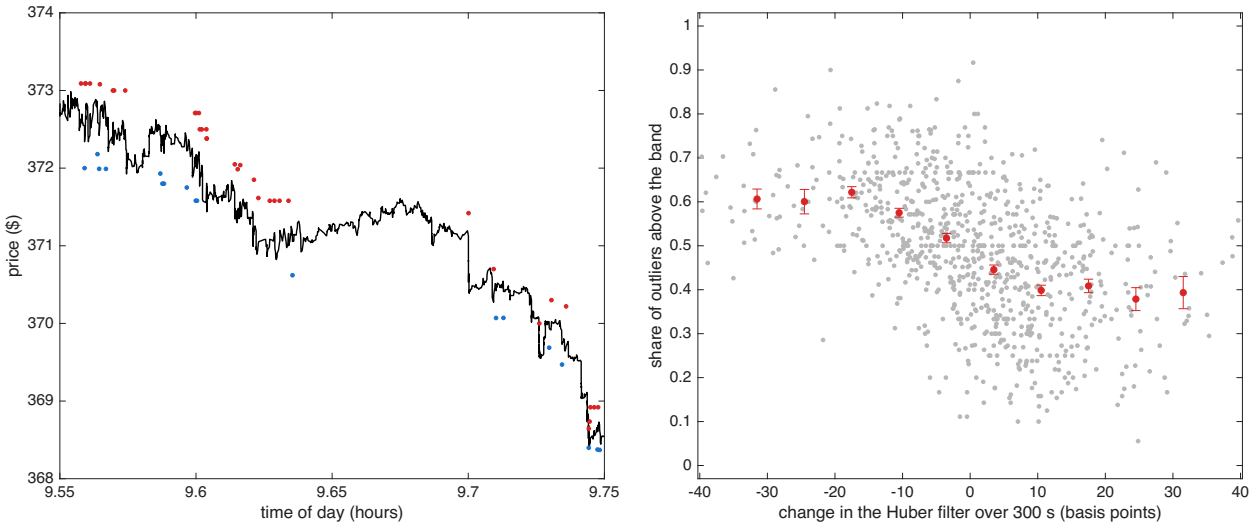

Figure 7: Outliers & robust price for PANW on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

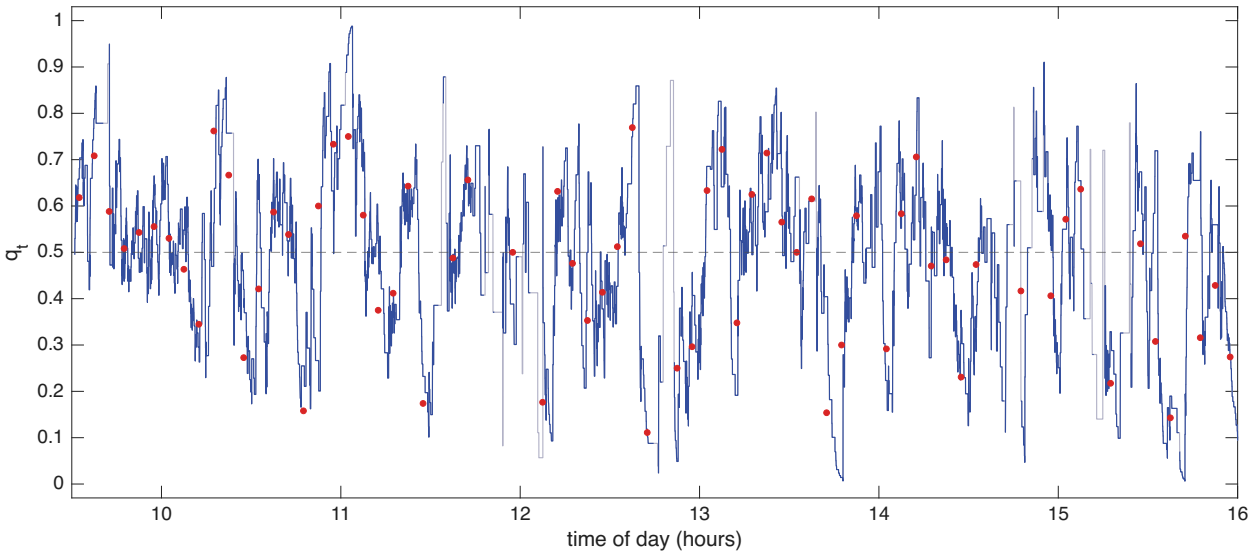

Figure 8: The directional conditional probability for PANW, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 60 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Diagnostics.

ABNB 1 to 16 October 2024

trades on 16 October 2024	60,646
average trading rate	2.3 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	2,110 (3.48%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.26%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	49 s [41, 58]
Mandelbrot a minimizing QLIKE	39 trades
a within 0.0005 of the minimum	27 to 55
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	14% $\rightarrow$ 115%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	12% $\rightarrow$ 20%
median spot volatility, Huber / linear input[5]	13.3 / 15.3 %
hyperbolic against HExp, correlation of logs	0.9958
HExp QLIKE advantage, leave-one-day-out[6]	+0.00213
15-minute signature level	13 to 50 %
Huber signature, 1 min over 1 s, 12 day average	1.54

ABNB over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:14; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.464	0.371	0.391	1.743	1.496	0.987	1.029	6.215
2	0.440	0.407	0.448	1.429	1.207	0.869	1.042	4.565
5	0.506	0.462	0.510	1.080	1.083	0.861	1.016	2.982
10	0.542	0.498	0.537	0.894	0.962	0.830	0.963	2.186
25	0.584	0.550	0.565	0.757	0.908	0.846	0.902	1.537
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.35	−0.23	0.02	−0.46	2.15	1.76	2.18	2.37
2	0.14	0.16	−0.02	−0.02	2.16	1.87	2.23	2.36
5	−0.02	−0.01	−0.02	0.00	2.12	1.96	2.13	2.43
10	0.01	0.00	0.00	0.01	2.26	2.31	2.28	2.46
25	−0.07	−0.09	−0.01	−0.02	2.94	3.15	2.87	2.62

Table 1: Trade-lag differences, `ABNB`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.47	3.19	3.26	5.91	33.0	25.1	28.8	172.3
1 s	2.28	2.14	2.22	3.51	23.3	18.7	22.8	129.6
5 s	1.75	1.67	1.70	2.35	27.0	25.4	28.2	124.5
10 s	1.60	1.54	1.57	1.99	28.9	28.0	32.4	75.4
1 min	1.70	1.67	1.68	1.81	35.5	35.0	35.5	46.3
15 min	2.38	2.37	2.38	2.44	50.5	50.3	50.1	52.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.13	−0.09	−0.02	−0.32	2.72	3.27	4.20	2.89
1 s	0.09	0.09	0.05	−0.23	4.42	5.22	6.52	3.62
5 s	0.09	0.13	0.11	−0.38	5.17	5.94	5.70	4.26
10 s	0.17	0.18	0.17	−0.13	6.19	6.11	5.44	5.05
1 min	0.19	0.20	0.19	0.11				

Table 2: Clock-lag differences, `ABNB`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.17	1.42	1.50
s.d. of $\hat{\epsilon}_t$ in cents	4.47	4.38	3.91
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.18
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.21	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.11	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.04	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.03	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.95,0.92	−1.04,1.02
95% C.I. for $\theta_t - X_t$	—	−1.81,1.80	−2.05,2.07
99% C.I. for $\theta_t - X_t$	—	−2.81,3.18	−3.50,4.11
s.e. of $\theta_t - X_t$ in cents	—	0.84	1.28

Table 3: Residual summaries, `ABNB`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.18, is above two, while the Hill index of the Huber residual itself on this day is 1.77, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.20	1.77	2.20 to 3.17	1.77 to 3.39
Residual divided by δ_t	2.48	2.19	2.29 to 4.08	2.19 to 3.90
Repricing windows removed	2.19	2.00	2.19 to 3.19	2.00 to 3.44
Calm windows only	2.28	1.93		
Every 10th trade	2.31	1.79	1.80 to 3.45	1.79 to 2.93
Six-trade half-life	2.40	1.81	2.40 to 3.33	1.81 to 3.67

Table 4: Hill estimator of the tail index of the Huber filter residual, `ABNB`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 20.2 to 28.2 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 13.9 cents on 16 October 2024, 87% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

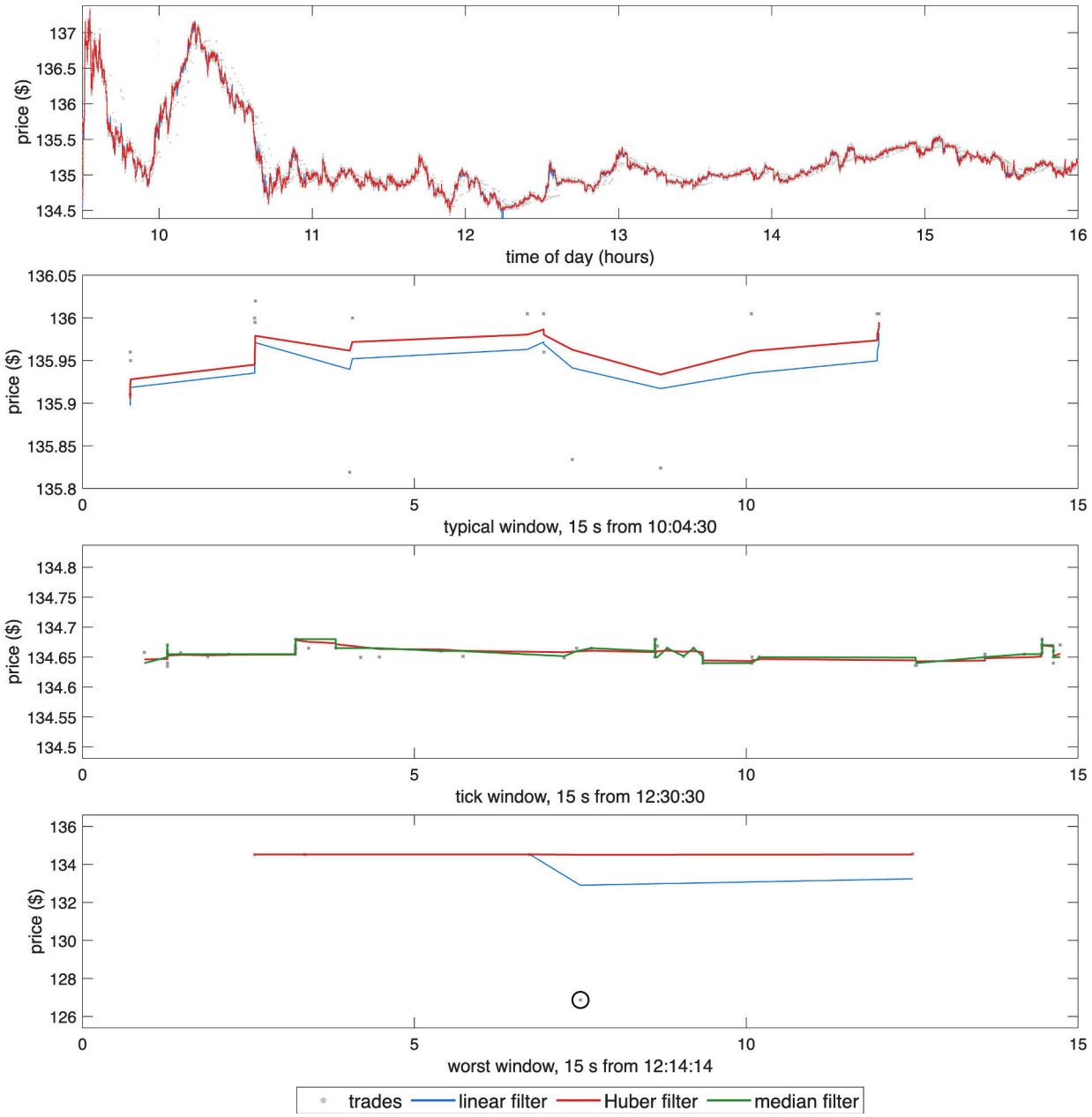

Figure 1: ABNB trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 15-second window at the day's median trade count, from 10:04:30. 3rd: 15-second window at 12:30:30 in which the price reverses direction most often, with the median filter m_t added. Bottom: 15-second window around the largest Huber residual after 09:45, circled.

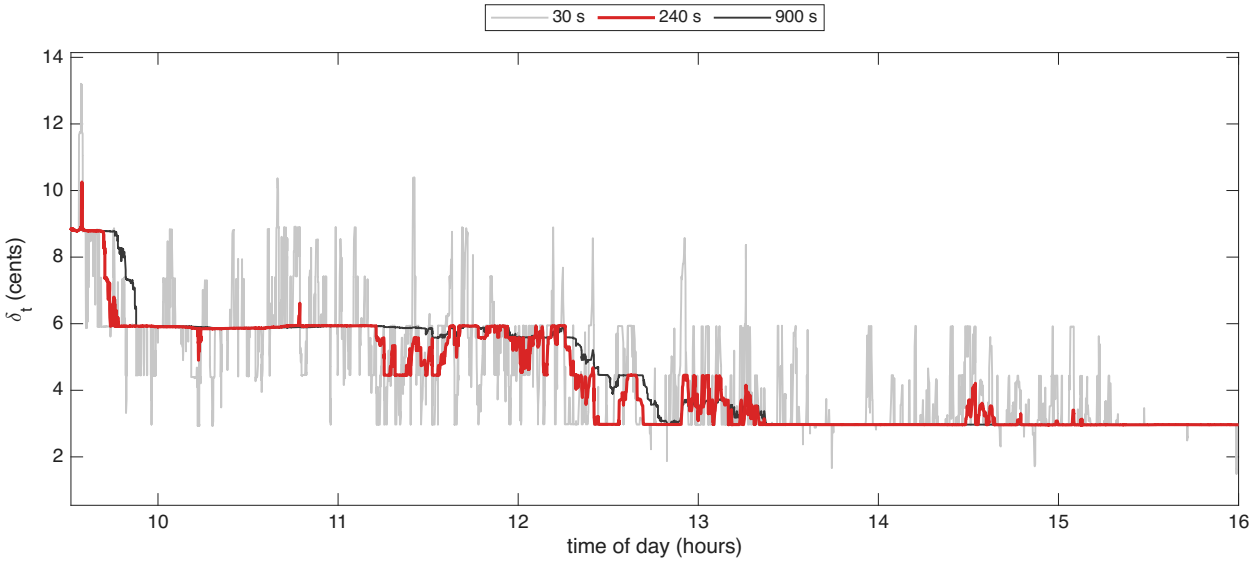

Figure 2: The Huber threshold δ_t in cents for `ABNB` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

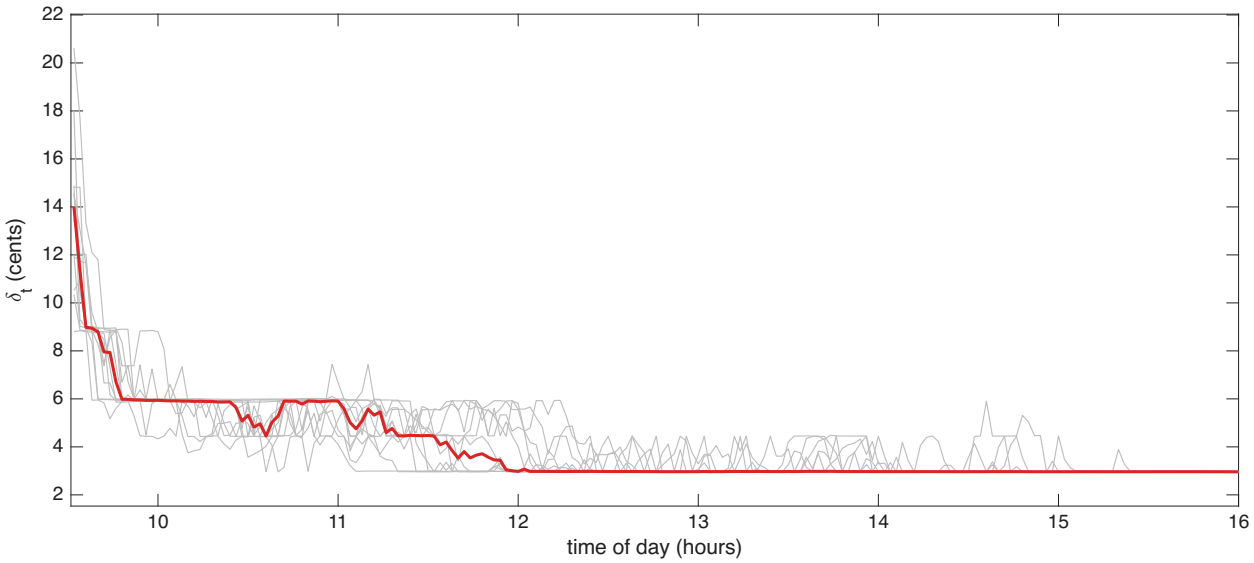

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `ABNB` trading days in Oct 2024 (grey lines) and their median (red line).

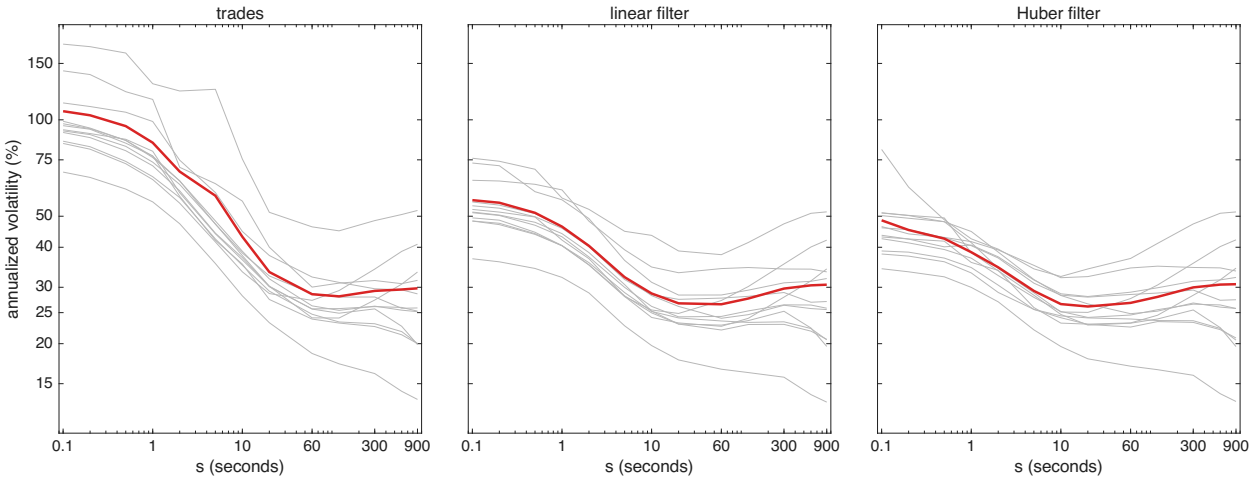

Figure 4: Volatility signature curves for ABNB over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

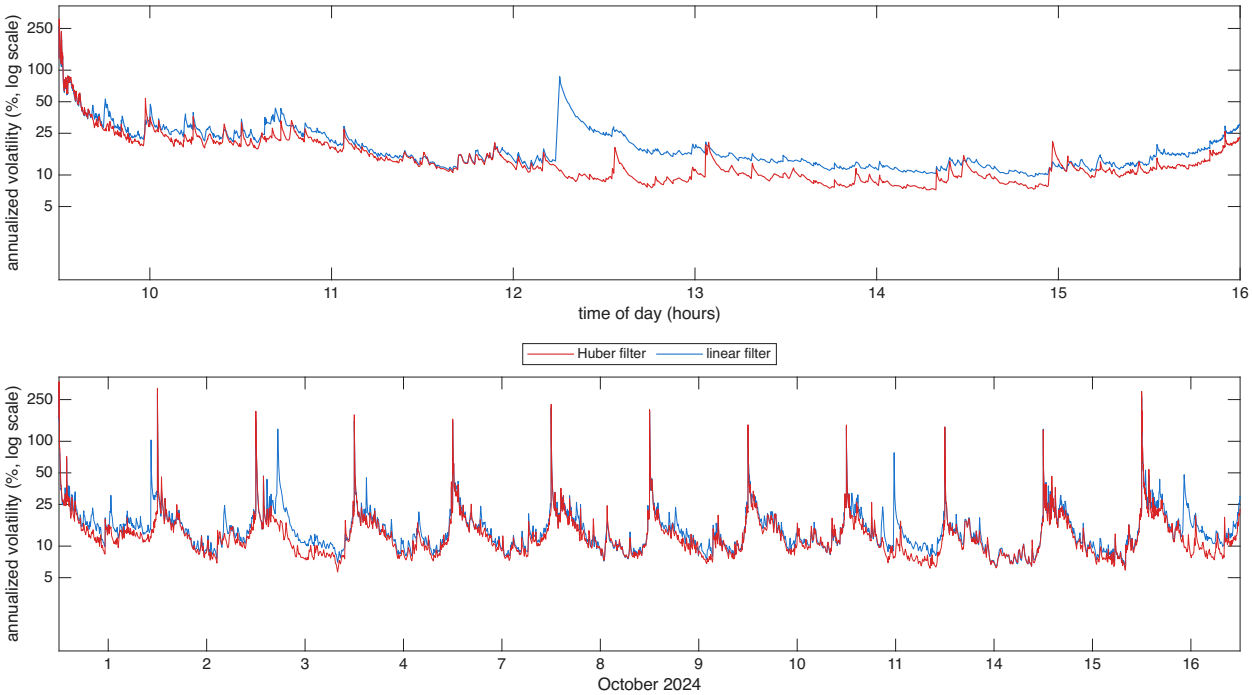

Figure 5: Filtered spot volatility for ABNB over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 39$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

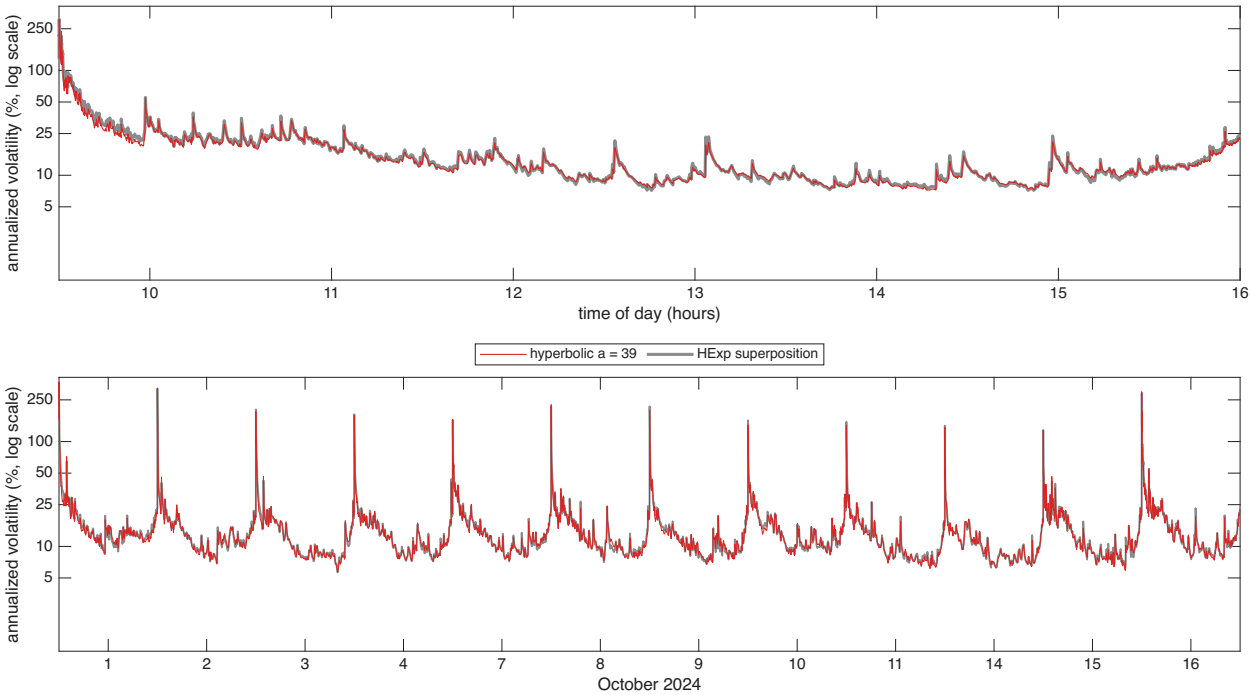

Figure 6: Hyperbolic against superposition weights for ABNB over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 39$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

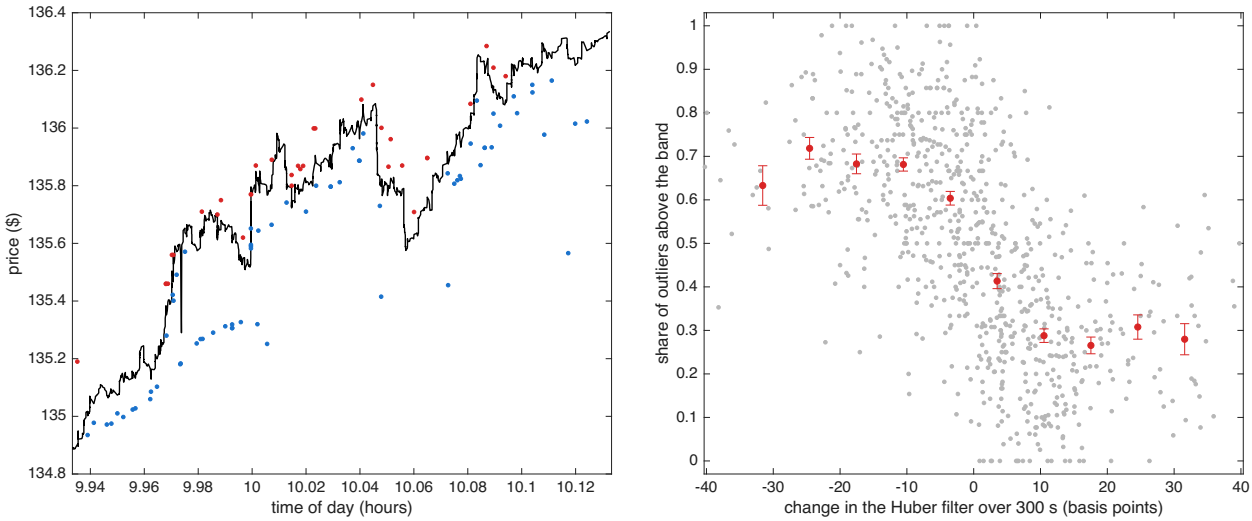

Figure 7: Outliers & robust price for ABNB on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

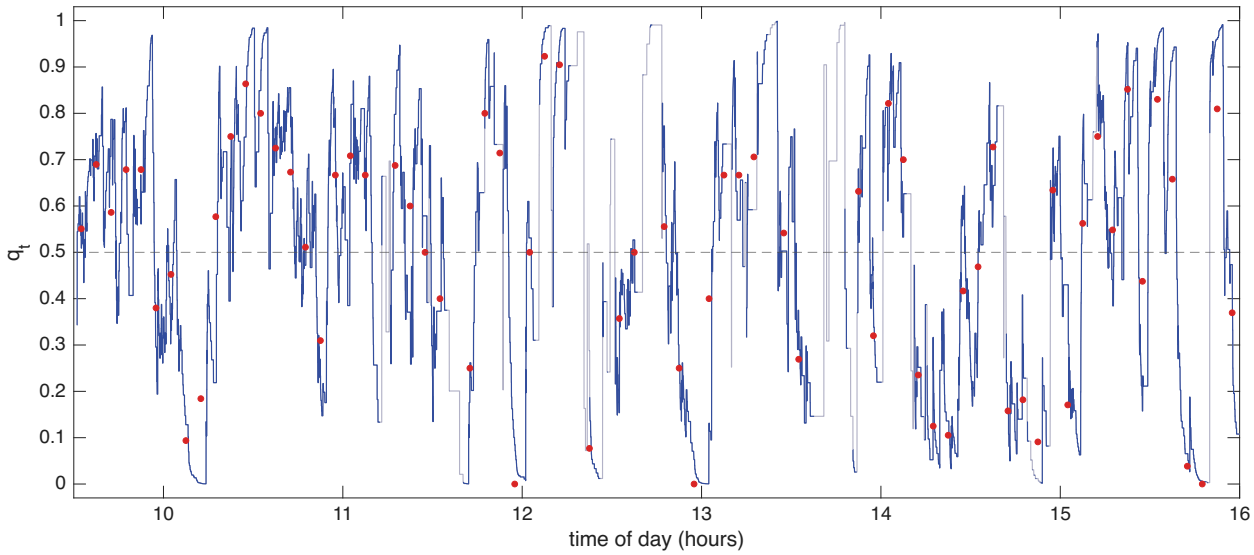

Figure 8: The directional conditional probability for ABNB, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 49 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

Selecting a on single days gives values from 13 to 139 trades, against 39 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 23 trades for this stock.

Diagnostics.

GILD 1 to 16 October 2024

trades on 16 October 2024	44,282
average trading rate	2.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	738 (1.67%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+3.01%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	116 s [96, 144]
Mandelbrot a minimizing QLIKE	25 trades
a within 0.0005 of the minimum	18 to 37
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	8% $\to$ 20%
Huber $\hat{\sigma}$, before $\to$ peak	7% $\to$ 20%
median spot volatility, Huber / linear input[5]	8.4 / 10.1 %
hyperbolic against HExp, correlation of logs	0.9976
HExp QLIKE advantage, leave-one-day-out[6]	-0.00129
15-minute signature level	13 to 25 %
Huber signature, 1 min over 1 s, 12 day average	1.67

GILD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 15:54; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.129	0.110	0.114	0.456	0.410	0.210	0.208	1.137
2	0.133	0.127	0.135	0.385	0.364	0.225	0.227	0.865
5	0.161	0.153	0.162	0.309	0.329	0.249	0.257	0.600
10	0.184	0.174	0.181	0.274	0.326	0.280	0.285	0.486
25	0.209	0.199	0.202	0.249	0.327	0.306	0.308	0.397
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.15	0.19	−0.42	2.28	2.15	3.18	3.46
2	−0.04	0.05	0.12	−0.02	2.34	2.21	2.93	3.14
5	−0.02	0.08	0.06	−0.01	2.66	2.63	3.12	3.28
10	−0.04	0.00	0.01	−0.03	2.70	2.68	2.99	3.08
25	−0.02	−0.02	−0.02	−0.02	4.19	4.74	4.87	3.78

Table 1: Trade-lag differences, `GILD`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.78	0.70	0.73	1.32	11.1	7.0	7.3	38.6
1 s	0.53	0.49	0.50	0.84	10.0	7.8	8.1	31.4
5 s	0.41	0.39	0.40	0.58	10.7	9.8	10.0	22.1
10 s	0.40	0.38	0.39	0.51	11.8	11.1	11.2	18.3
1 min	0.37	0.36	0.36	0.40	13.1	12.8	12.6	14.2
15 min	0.32	0.32	0.32	0.33	12.6	12.6	12.4	13.0
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.06	0.01	0.02	−0.14	3.26	3.26	3.09	4.40
1 s	−0.09	−0.01	−0.01	−0.25	3.26	4.03	3.59	4.08
5 s	0.00	0.03	0.01	−0.26	3.73	2.85	2.71	4.54
10 s	−0.02	0.01	−0.01	−0.26	4.02	3.59	3.52	4.61
1 min	0.00	0.03	0.03	−0.05				

Table 2: Clock-lag differences, `GILD`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.33	0.42	0.44
s.d. of $\hat{\epsilon}_t$ in cents	0.85	0.80	0.79
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.38
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.28	0.21
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.17	0.11
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.09	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.03	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.01	−0.02
80% C.I. for $\theta_t - X_t$	—	−0.32,0.33	−0.34,0.34
95% C.I. for $\theta_t - X_t$	—	−0.62,0.62	−0.66,0.66
99% C.I. for $\theta_t - X_t$	—	−1.12,1.10	−1.25,1.25
s.e. of $\theta_t - X_t$ in cents	—	0.32	0.35

Table 3: Residual summaries, `GILD`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.38, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.54	2.74	1.99 to 3.18	1.95 to 3.68
Residual divided by δ_t	2.28	2.77	1.85 to 3.15	2.02 to 3.26
Repricing windows removed	2.48	2.81	1.99 to 3.18	1.94 to 3.68
Calm windows only	2.28	2.65		
Every 10th trade	1.90	2.10	1.39 to 2.88	1.71 to 2.37
Six-trade half-life	2.71	2.89	2.17 to 3.42	2.11 to 4.07

Table 4: Hill estimator of the tail index of the Huber filter residual, `GILD`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 10.1 to 20.2 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 5.0 cents on 16 October 2024, 96% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

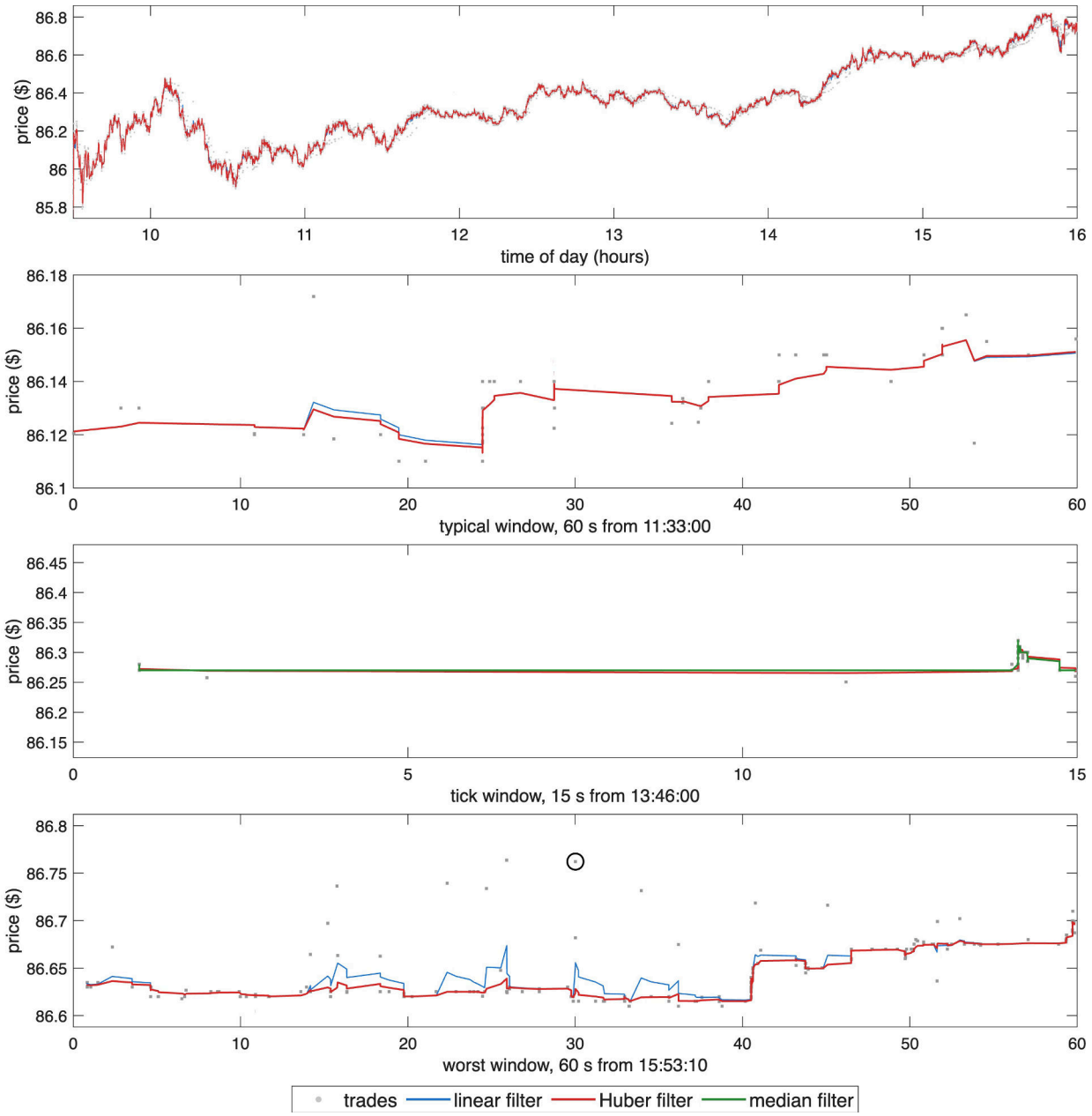

Figure 1: `GILD` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 11:33:00. 3rd: 15-second window at 13:46:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

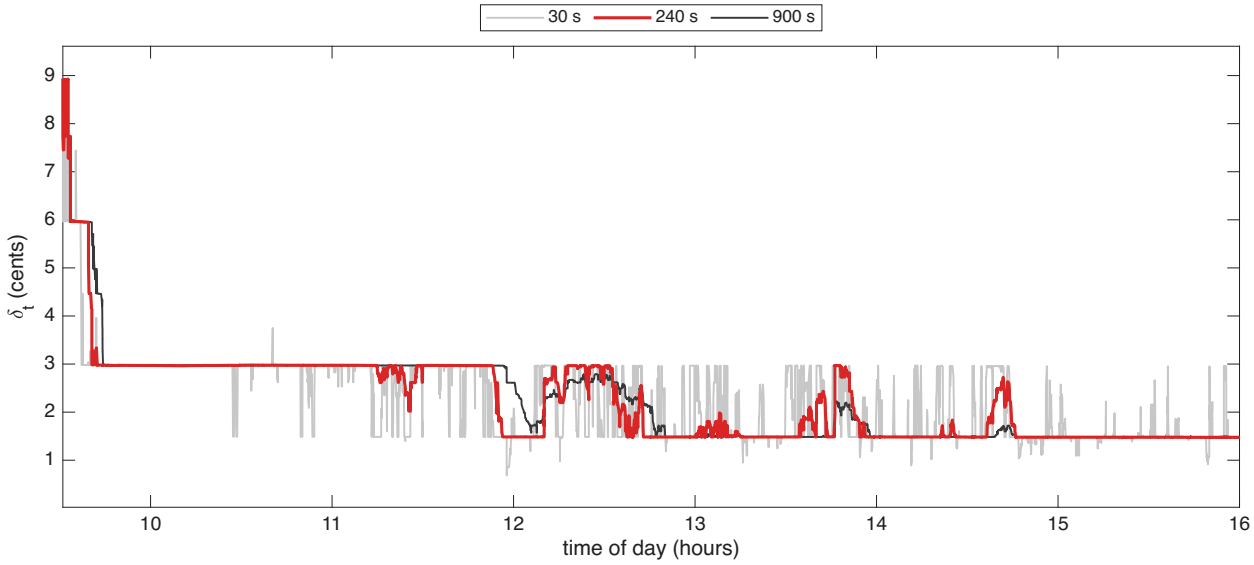

Figure 2: The Huber threshold δ_t in cents for `GILD` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

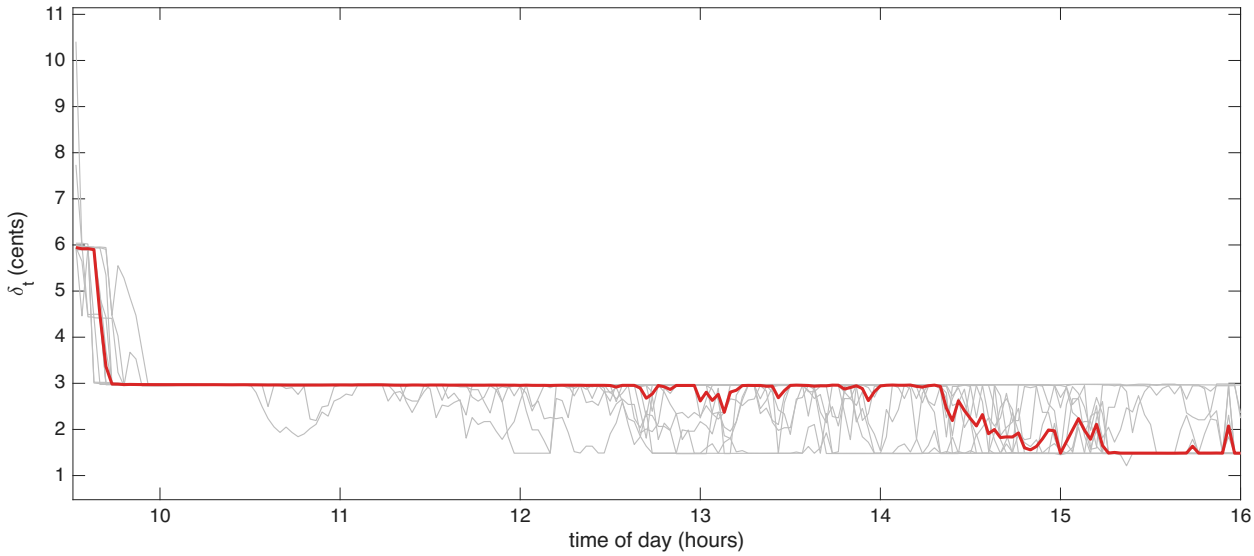

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `GILD` trading days in Oct 2024 (grey lines) and their median (red line).

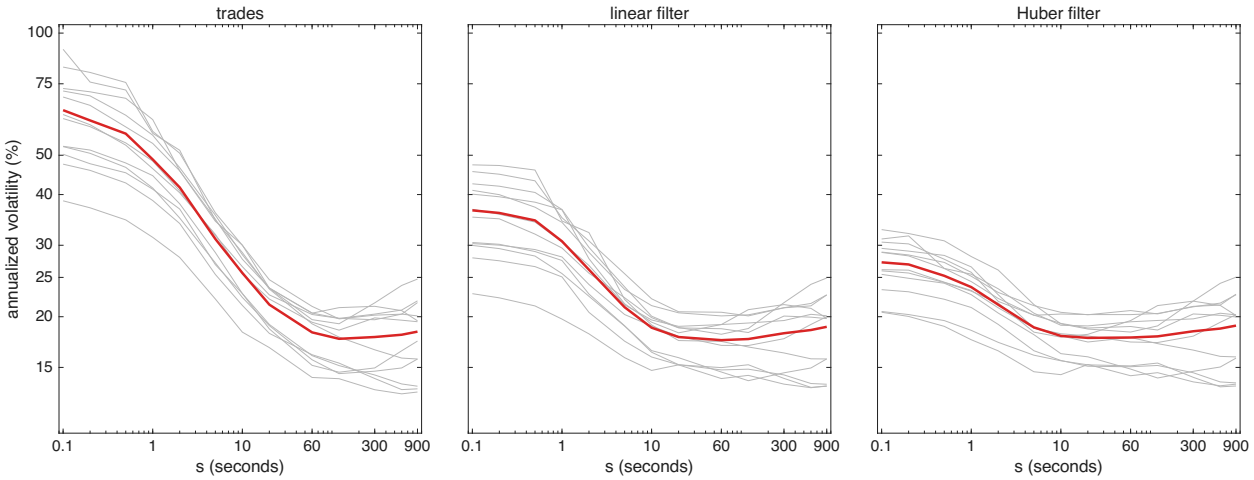

Figure 4: Volatility signature curves for `GILD` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

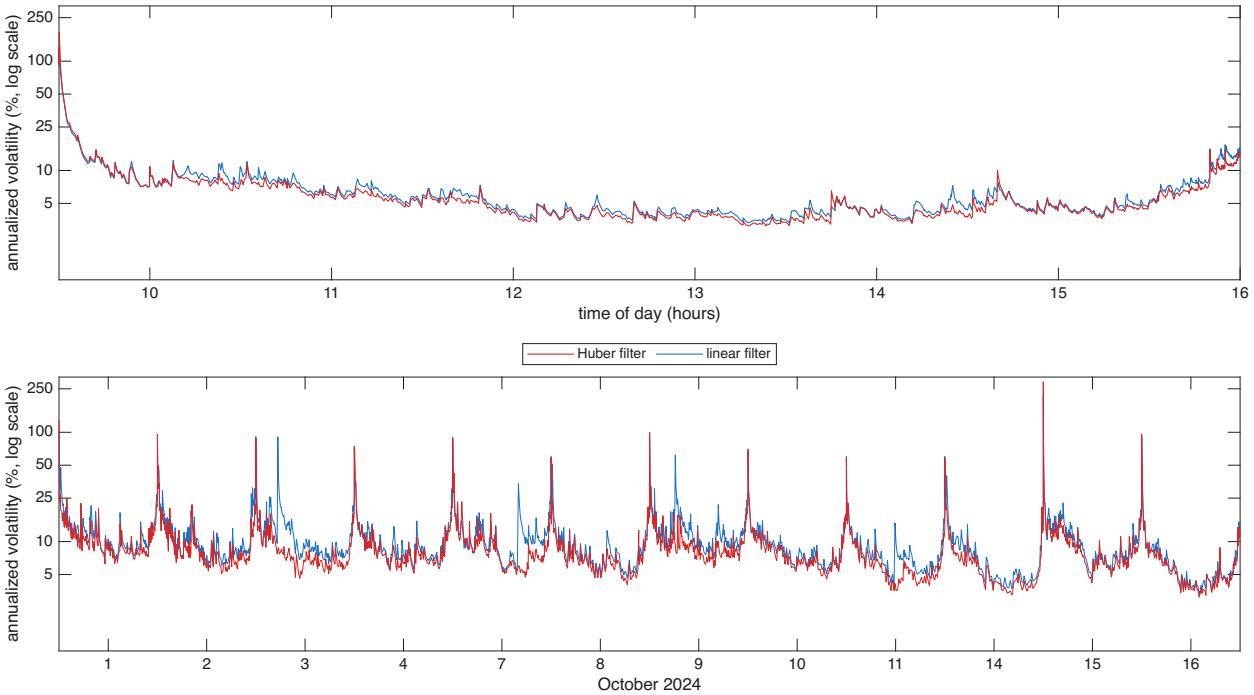

Figure 5: Filtered spot volatility for `GILD` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 25$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

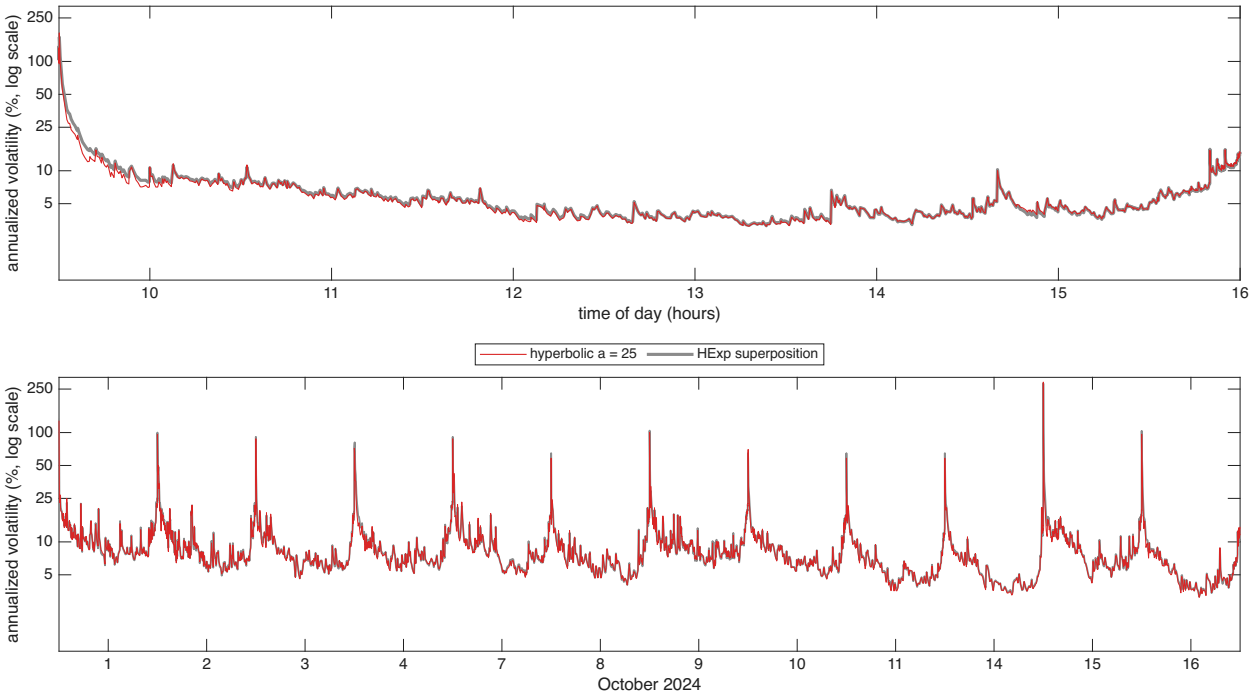

Figure 6: Hyperbolic against superposition weights for `GILD` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 25$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

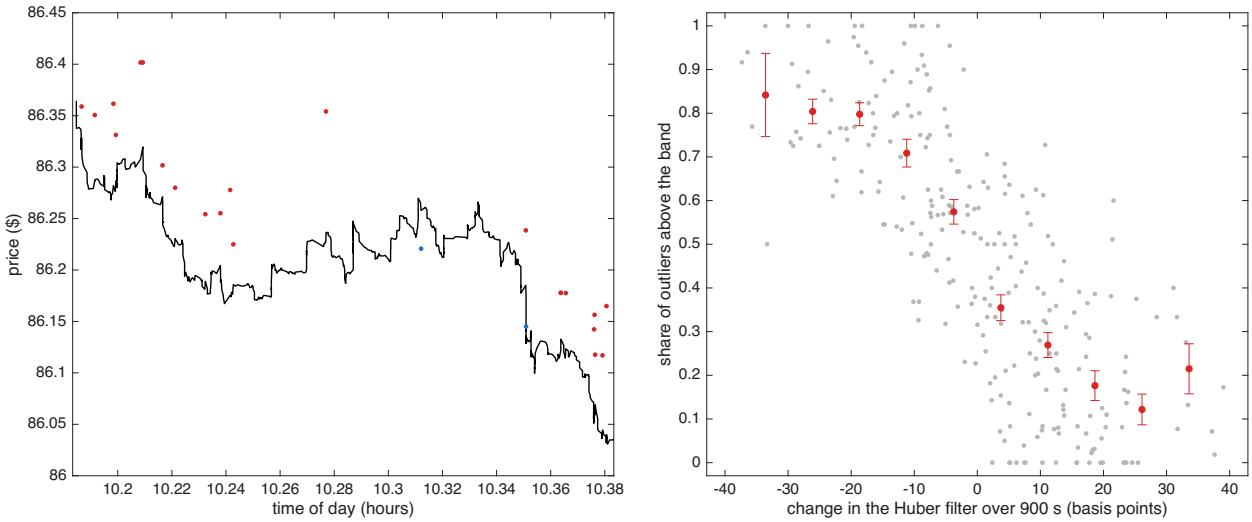

Figure 7: Outliers & robust price for `GILD` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 900 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 900 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

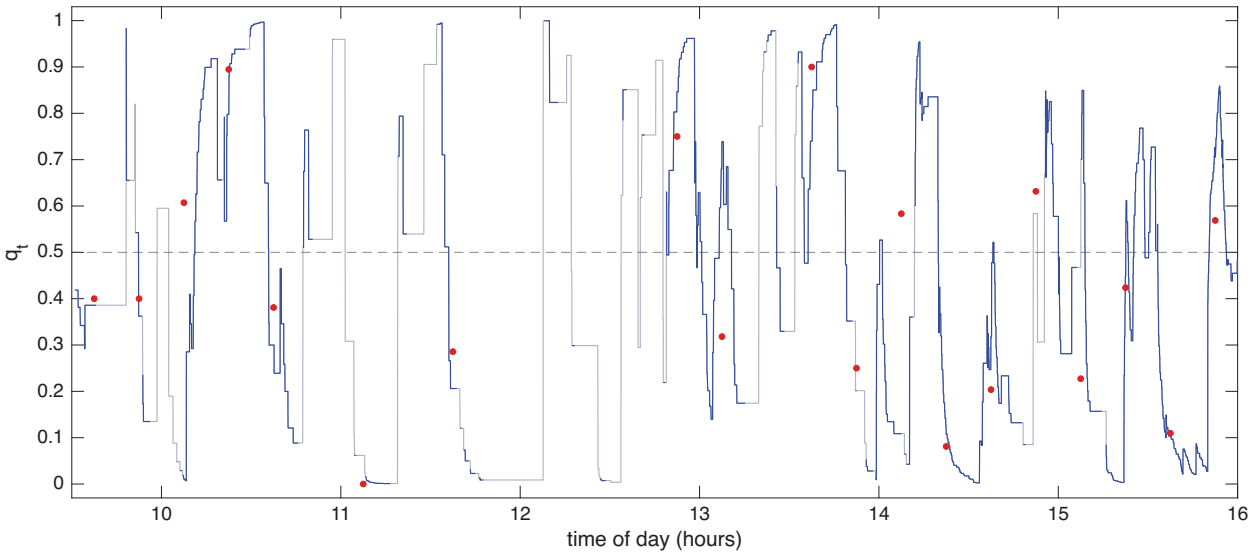

Figure 8: The directional conditional probability for GILD, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 116 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 900 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 6 to 1632 trades, against 25 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 22 trades for this stock.
The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 15:54.

Diagnostics.

DDOG 1 to 16 October 2024

trades on 16 October 2024	43,337
average trading rate	2.1 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,240 (2.86%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.21%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.01%
direction half-life[3]	88 s [72, 111]
Mandelbrot a minimizing QLIKE	78 trades
a within 0.0005 of the minimum	53 to 120
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	31% $\rightarrow$ 42%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	31% $\rightarrow$ 33%
median spot volatility, Huber / linear input[5]	13.3 / 15.6 %
hyperbolic against HExp, correlation of logs	0.9979
HExp QLIKE advantage, leave-one-day-out[6]	+0.00645
15-minute signature level	17 to 41 %
Huber signature, 1 min over 1 s, 12 day average	1.60

DDOG over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:45; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.398	0.320	0.338	1.439	1.070	0.555	0.598	3.339
2	0.401	0.364	0.393	1.205	0.958	0.597	0.650	2.535
5	0.475	0.423	0.455	0.930	0.892	0.659	0.703	1.734
10	0.523	0.469	0.491	0.796	0.860	0.713	0.738	1.362
25	0.562	0.525	0.529	0.684	0.816	0.752	0.756	1.039
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.20	0.16	0.18	−0.42	2.84	2.57	3.31	3.85
2	0.05	0.11	0.07	−0.03	3.10	2.71	3.27	3.66
5	−0.02	0.04	0.02	0.01	3.88	3.38	3.50	3.85
10	−0.02	−0.02	−0.02	−0.01	3.71	3.55	3.68	3.73
25	0.01	0.01	0.01	−0.01	5.10	5.15	5.23	4.05

Table 1: Trade-lag differences, DDOG, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.35	2.10	2.15	3.94	17.9	11.7	13.5	76.6
1 s	1.61	1.49	1.53	2.59	17.1	13.2	14.9	62.9
5 s	1.23	1.16	1.20	1.73	19.1	17.4	18.5	42.4
10 s	1.09	1.04	1.08	1.48	20.8	19.6	20.2	35.2
1 min	0.99	0.98	0.99	1.12	24.3	23.7	23.4	27.1
15 min	1.23	1.24	1.25	1.31	29.2	28.9	28.9	29.9
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.03	0.03	−0.11	5.69	6.54	7.01	3.86
1 s	−0.02	0.03	−0.01	−0.32	9.35	12.73	13.08	4.36
5 s	−0.07	−0.04	−0.06	−0.32	7.14	7.39	7.48	5.27
10 s	0.05	0.10	0.07	−0.29	7.16	6.88	6.77	6.36
1 min	−0.07	−0.07	−0.06	−0.11				

Table 2: Clock-lag differences, DDOG, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.01	1.25	1.30
s.d. of $\hat{\epsilon}_t$ in cents	2.47	2.35	2.27
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.50
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.25	0.17
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.14	0.08
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.06	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.87,0.86	−0.94,0.93
95% C.I. for $\theta_t - X_t$	—	−1.71,1.60	−1.91,1.74
99% C.I. for $\theta_t - X_t$	—	−3.08,2.69	−3.50,2.96
s.e. of $\theta_t - X_t$ in cents	—	0.82	0.93

Table 3: Residual summaries, DDOG, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.50, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.58	3.02	1.95 to 3.29	2.41 to 3.84
Residual divided by δ_t	2.25	2.98	1.83 to 3.33	2.39 to 3.02
Repricing windows removed	2.57	2.99	1.99 to 3.24	2.34 to 3.83
Calm windows only	2.28	3.08		
Every 10th trade	1.98	2.10	1.68 to 2.71	1.77 to 2.59
Six-trade half-life	2.78	3.29	2.04 to 3.53	2.61 to 4.40

Table 4: Hill estimator of the tail index of the Huber filter residual, DDOG, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 20.2 to 30.2 cents over the days and removes 0% to 4% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 10.1 cents on 16 October 2024, 86% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

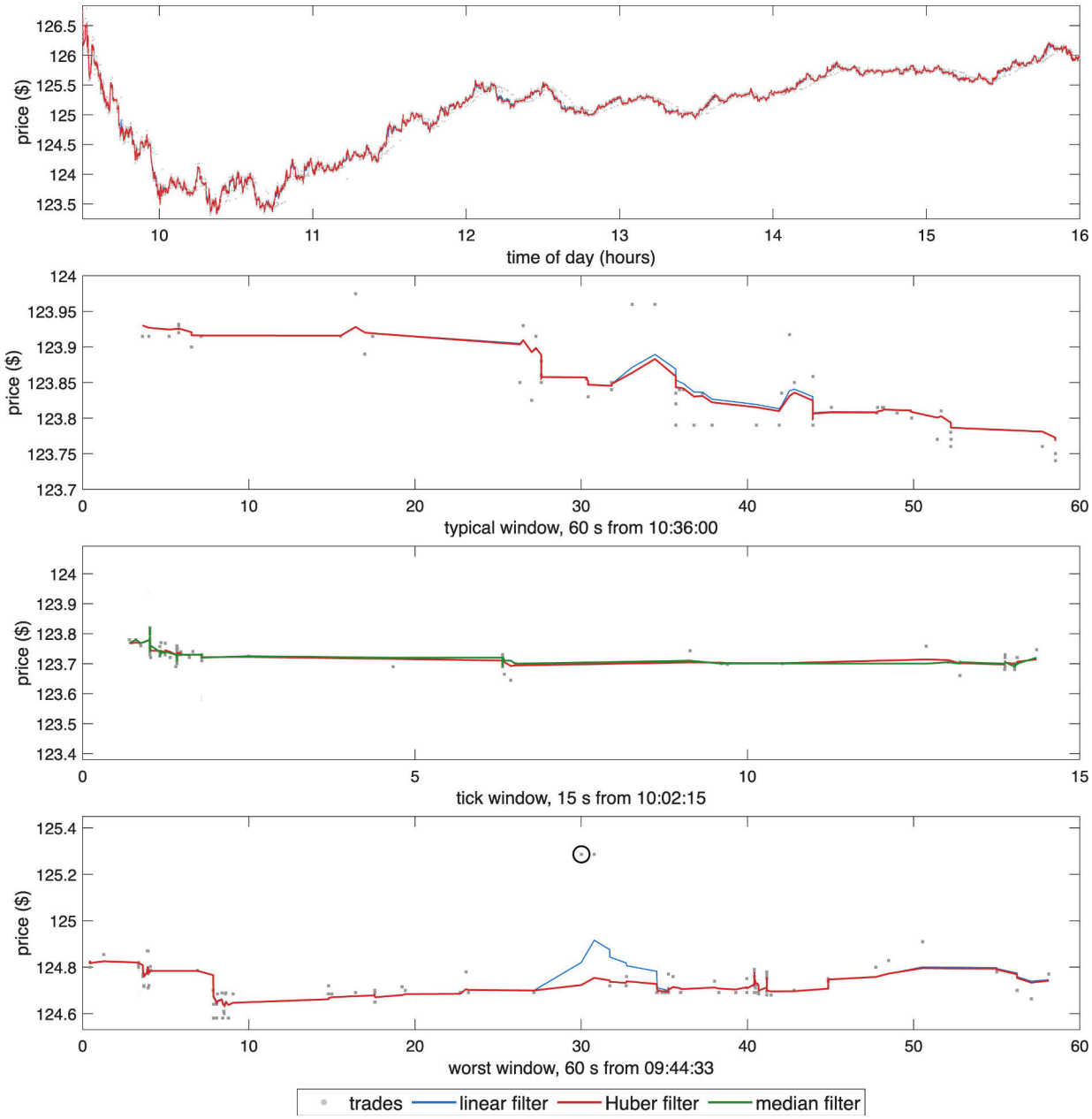

Figure 1: DDOG trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:36:00. 3rd: 15-second window at 10:02:15 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

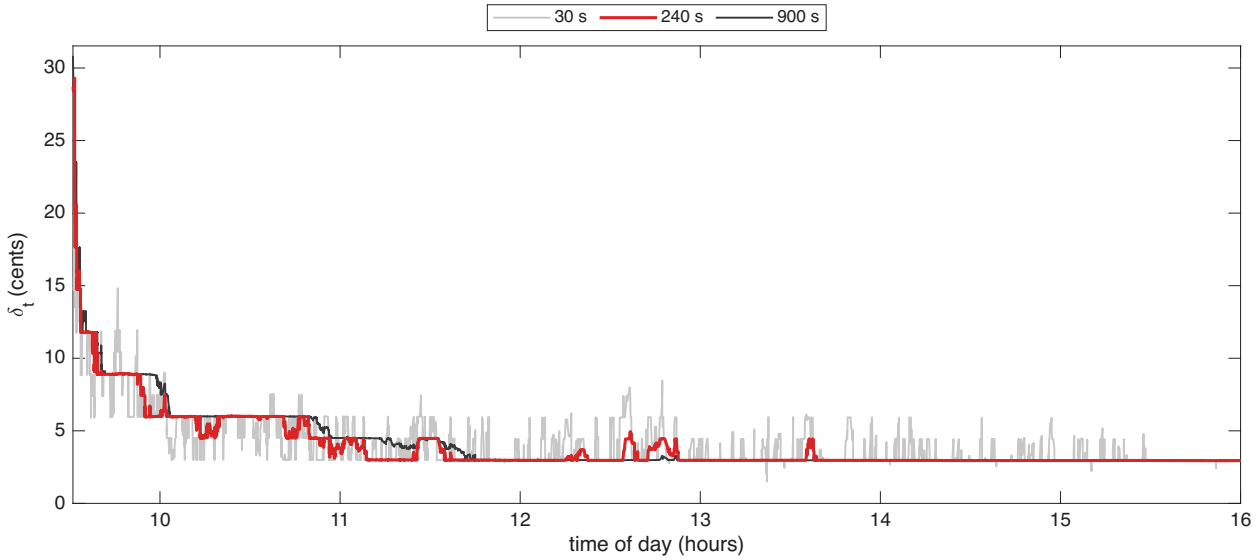

Figure 2: The Huber threshold δ_t in cents for DDOG on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

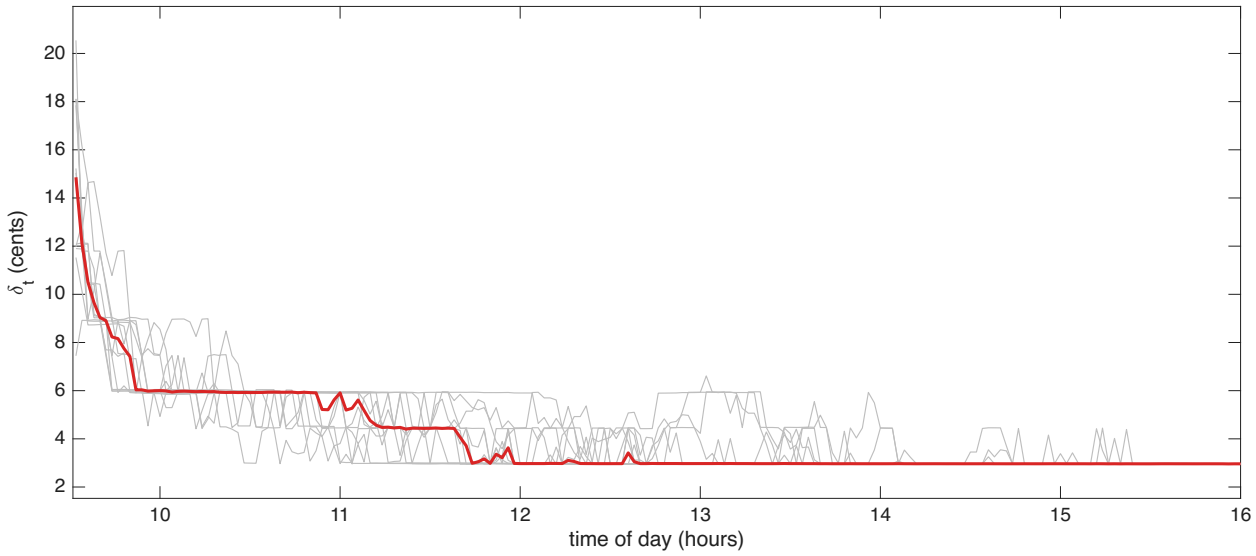

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 DDOG trading days in Oct 2024 (grey lines) and their median (red line).

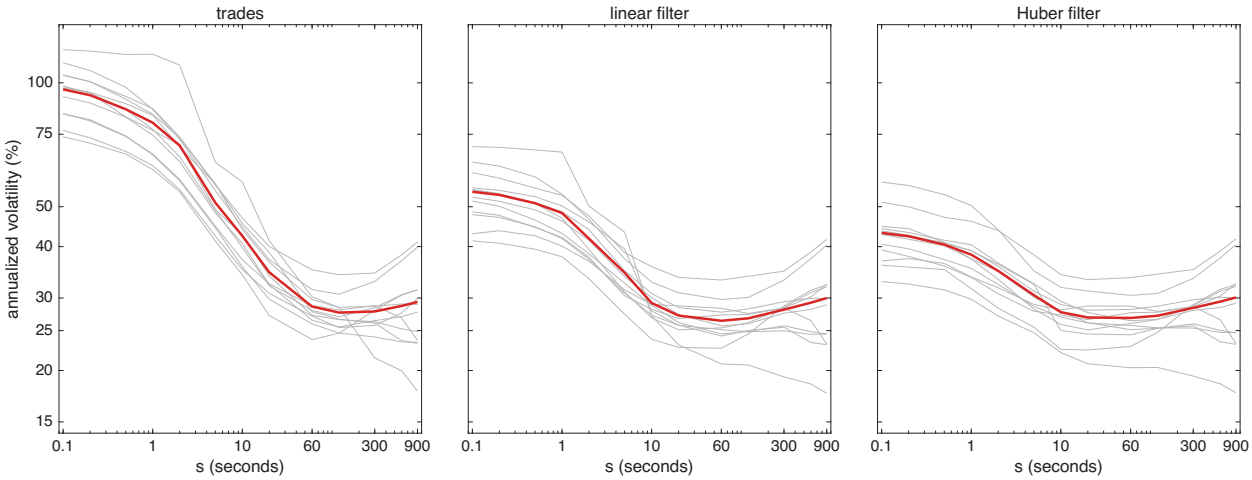

Figure 4: Volatility signature curves for DDOG over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

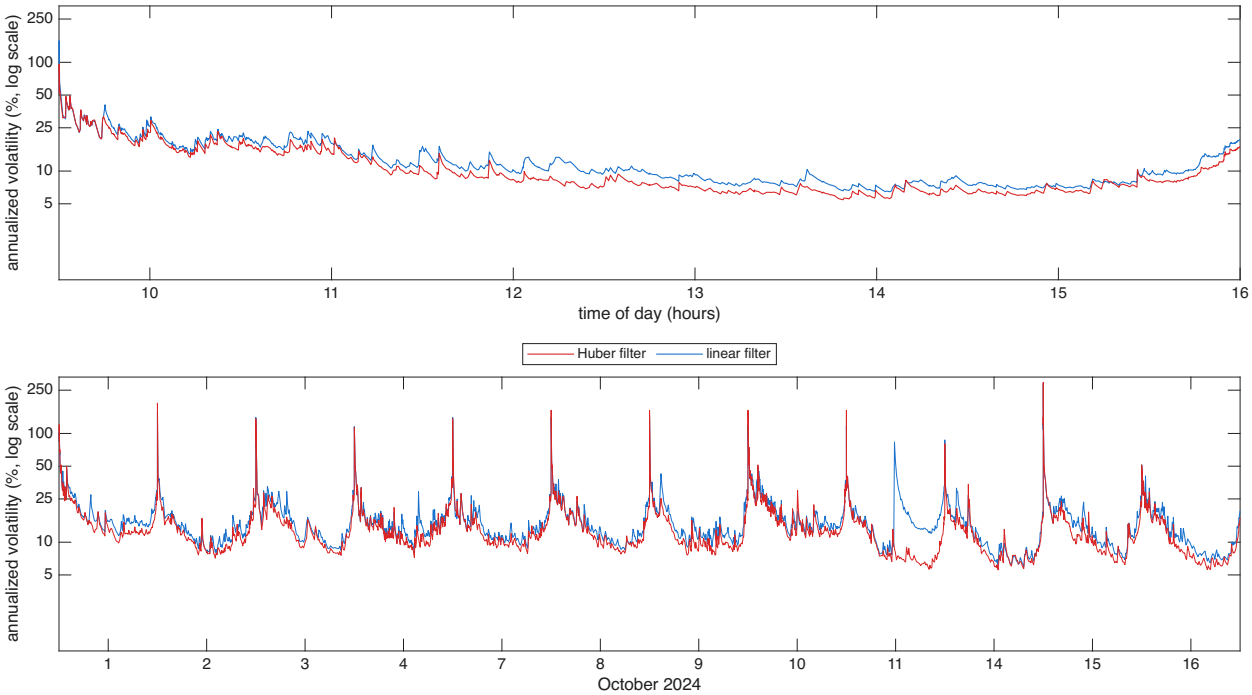

Figure 5: Filtered spot volatility for DDOG over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 78$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

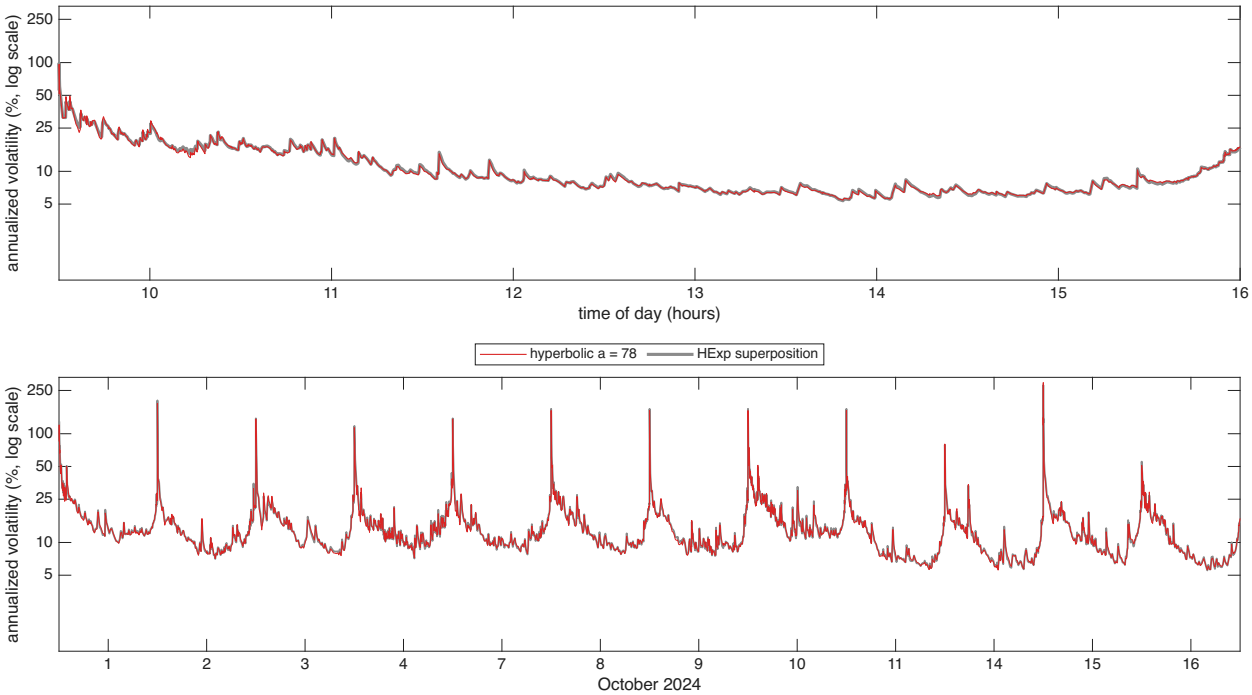

Figure 6: Hyperbolic against superposition weights for `DDOG` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 78$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

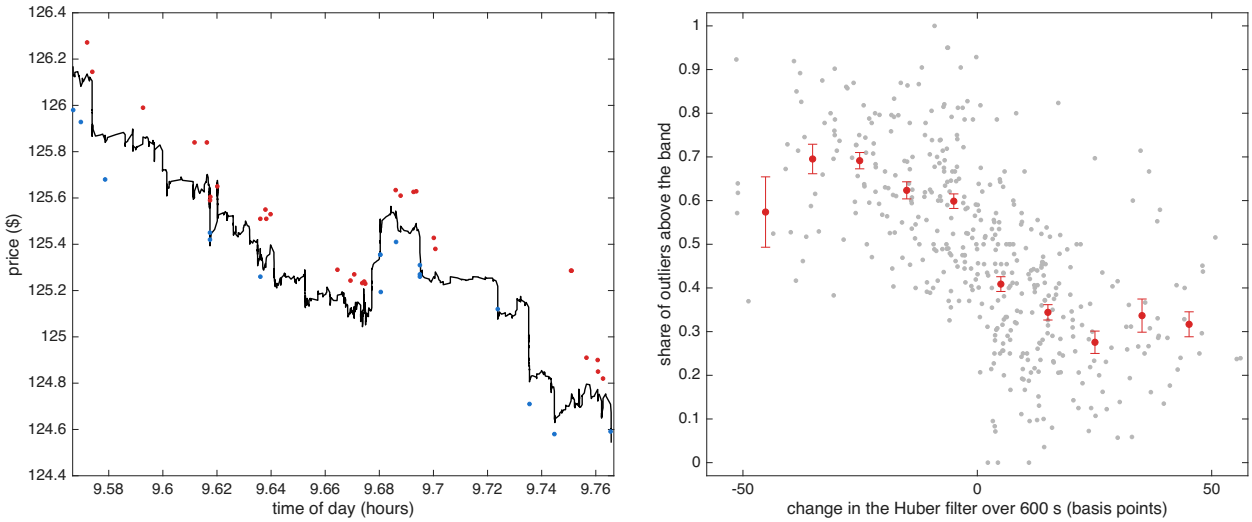

Figure 7: Outliers & robust price for `DDOG` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

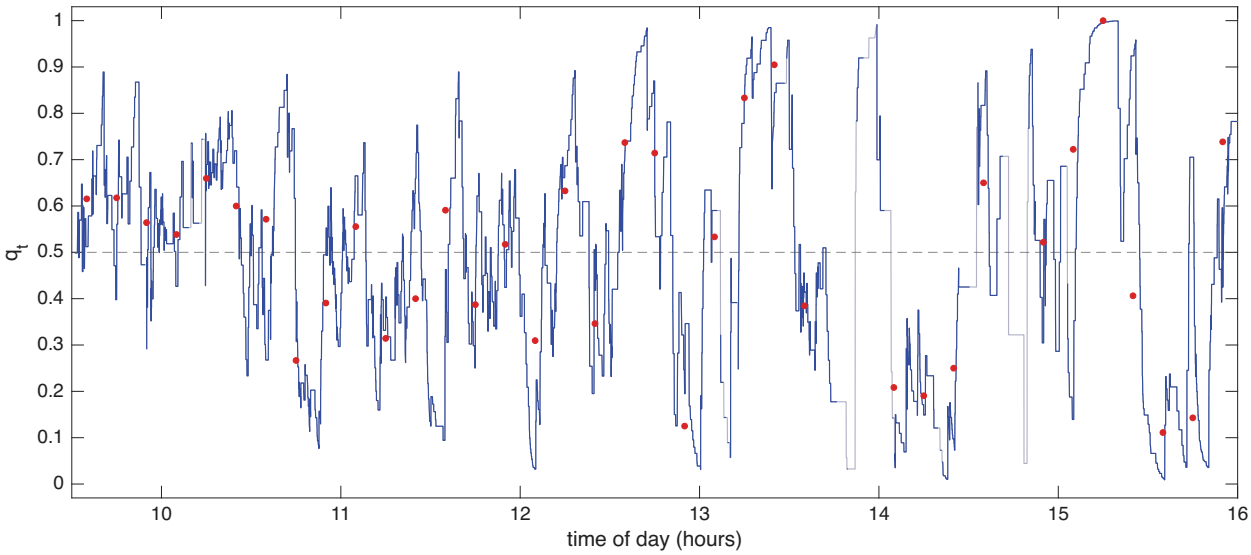

Figure 8: The directional conditional probability for DDOG, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 88 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 10 to 26112 trades, against 78 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 21 trades for this stock.

On 15 October 2024, 27 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 11 per cent on any other day. Every number that day contributes to carries that burst.

Diagnostics.

ETSY 1 to 16 October 2024

trades on 16 October 2024	43,591
average trading rate	1.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,155 (2.65%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.25%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.02%
direction half-life[3]	62 s [54, 77]
Mandelbrot a minimizing QLIKE	27 trades
a within 0.0005 of the minimum	20 to 37
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	23% $\rightarrow$ 73%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	19% $\rightarrow$ 79%
median spot volatility, Huber / linear input[5]	17.6 / 19.4 %
hyperbolic against HExp, correlation of logs	0.9947
HExp QLIKE advantage, leave-one-day-out[6]	+0.01917
15-minute signature level	16 to 68 %
Huber signature, 1 min over 1 s, 12 day average	1.57

`ETSY` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:23; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.233	0.193	0.213	0.910	0.648	0.345	0.411	2.410
2	0.228	0.217	0.247	0.757	0.566	0.367	0.431	1.813
5	0.265	0.249	0.280	0.576	0.516	0.392	0.444	1.210
10	0.289	0.271	0.295	0.484	0.477	0.405	0.441	0.907
25	0.318	0.302	0.311	0.415	0.470	0.437	0.448	0.670
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.24	0.13	0.10	−0.43	2.89	2.28	3.29	3.93
2	0.06	0.06	0.01	−0.05	3.00	2.33	3.65	3.96
5	−0.02	0.01	−0.02	−0.01	2.96	2.62	3.75	4.17
10	0.01	0.02	−0.01	−0.01	3.41	3.36	4.41	4.20
25	0.01	0.01	0.00	0.00	4.15	4.92	5.66	4.29

Table 1: Trade-lag differences, ETSY, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.32	1.20	1.28	2.53	29.9	18.6	23.3	139.8
1 s	0.92	0.84	0.89	1.64	28.9	20.8	24.3	114.4
5 s	0.72	0.68	0.70	1.07	29.5	25.5	27.5	72.7
10 s	0.65	0.62	0.64	0.90	31.0	28.4	29.2	59.9
1 min	0.64	0.63	0.63	0.71	35.4	34.1	33.1	39.5
15 min	0.74	0.74	0.73	0.74	43.2	42.8	42.2	43.0
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.02	0.04	−0.05	3.83	4.31	4.75	5.74
1 s	−0.02	0.04	−0.02	−0.35	5.00	5.47	5.15	5.93
5 s	−0.08	−0.01	−0.05	−0.38	4.69	5.03	5.19	6.35
10 s	0.01	0.07	0.04	−0.34	5.17	5.32	5.45	6.77
1 min	0.25	0.27	0.28	0.12				

Table 2: Clock-lag differences, ETSY, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.62	0.77	0.82
s.d. of $\hat{\epsilon}_t$ in cents	1.78	1.69	1.57
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.24
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.23	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.05	0.11	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.04	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.03	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.51,0.52	−0.57,0.57
95% C.I. for $\theta_t - X_t$	—	−0.96,0.98	−1.14,1.13
99% C.I. for $\theta_t - X_t$	—	−1.54,1.68	−1.89,2.18
s.e. of $\theta_t - X_t$ in cents	—	0.46	0.57

Table 3: Residual summaries, `ETSY`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.24, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.52	2.07	2.52 to 3.36	2.07 to 3.21
Residual divided by δ_t	2.54	2.26	2.51 to 3.18	2.26 to 3.47
Repricing windows removed	2.48	2.08	2.48 to 3.36	2.08 to 3.21
Calm windows only	2.37	2.10		
Every 10th trade	1.37	2.04	1.37 to 3.44	1.72 to 2.98
Six-trade half-life	2.73	2.19	2.73 to 3.64	2.19 to 3.50

Table 4: Hill estimator of the tail index of the Huber filter residual, `ETSY`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.1 to 20.2 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 10.1 cents on 16 October 2024, 97% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

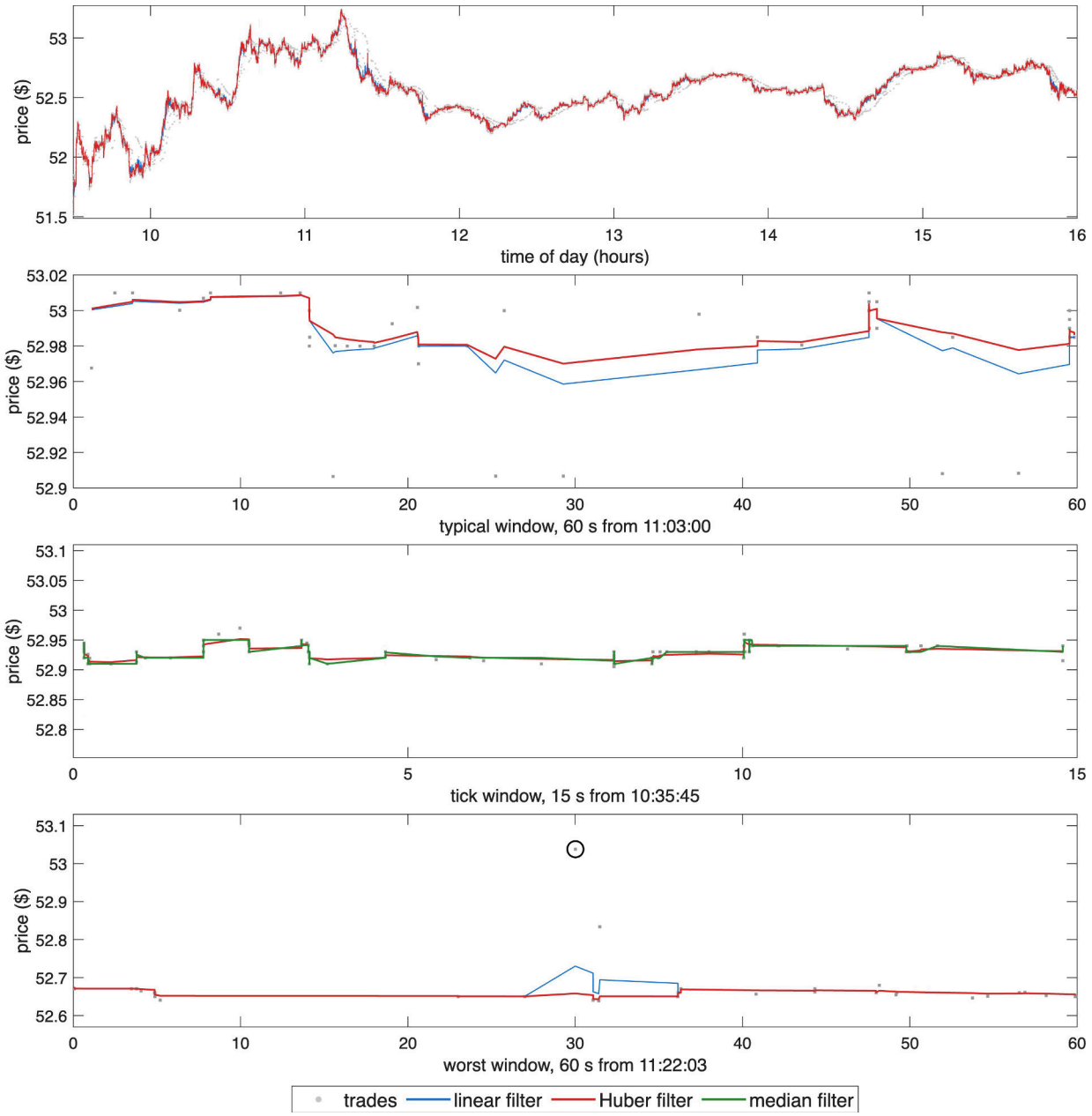

Figure 1: ETSY trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 11:03:00. 3rd: 15-second window at 10:35:45 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

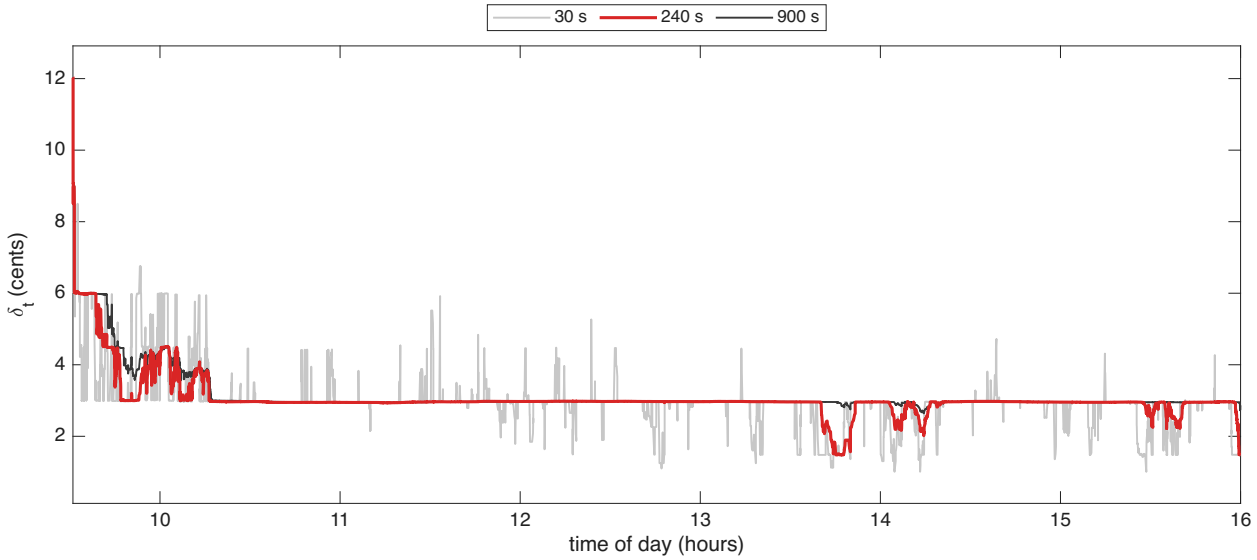

Figure 2: The Huber threshold δ_t in cents for ETSY on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

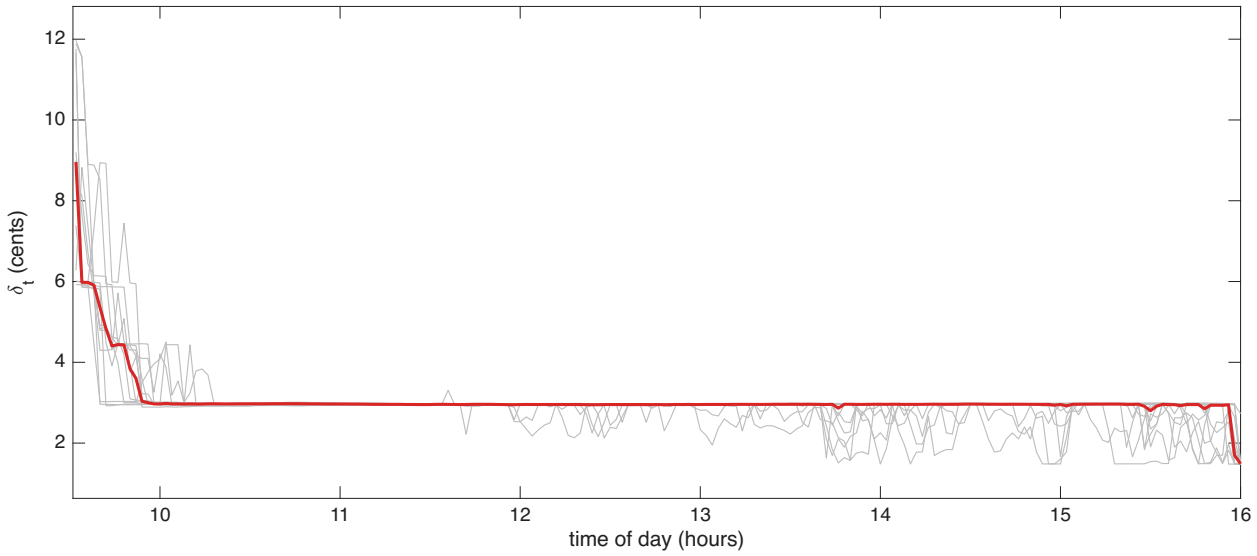

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 ETSY trading days in Oct 2024 (grey lines) and their median (red line).

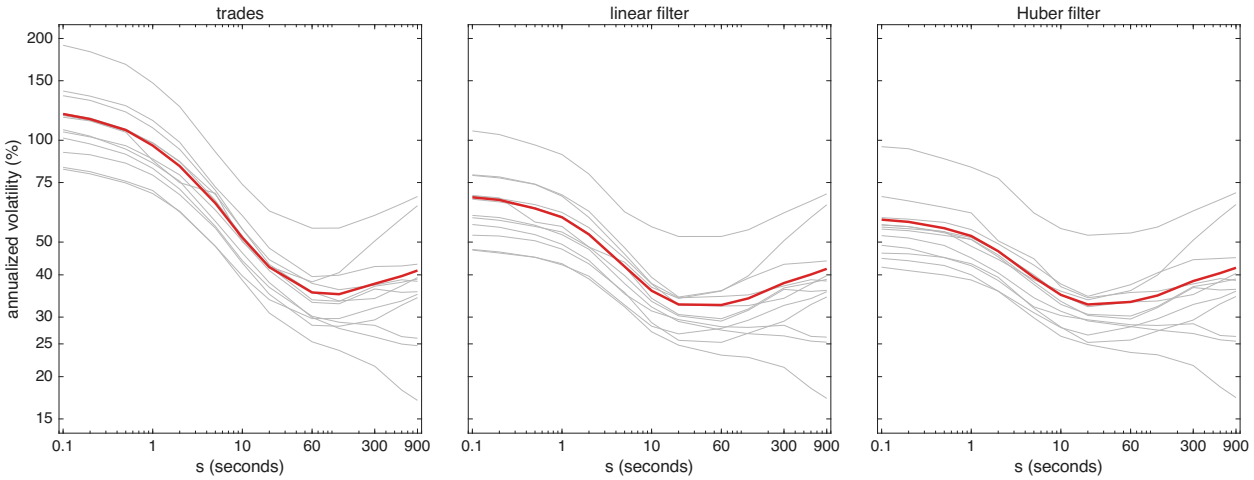

Figure 4: Volatility signature curves for `ETSY` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

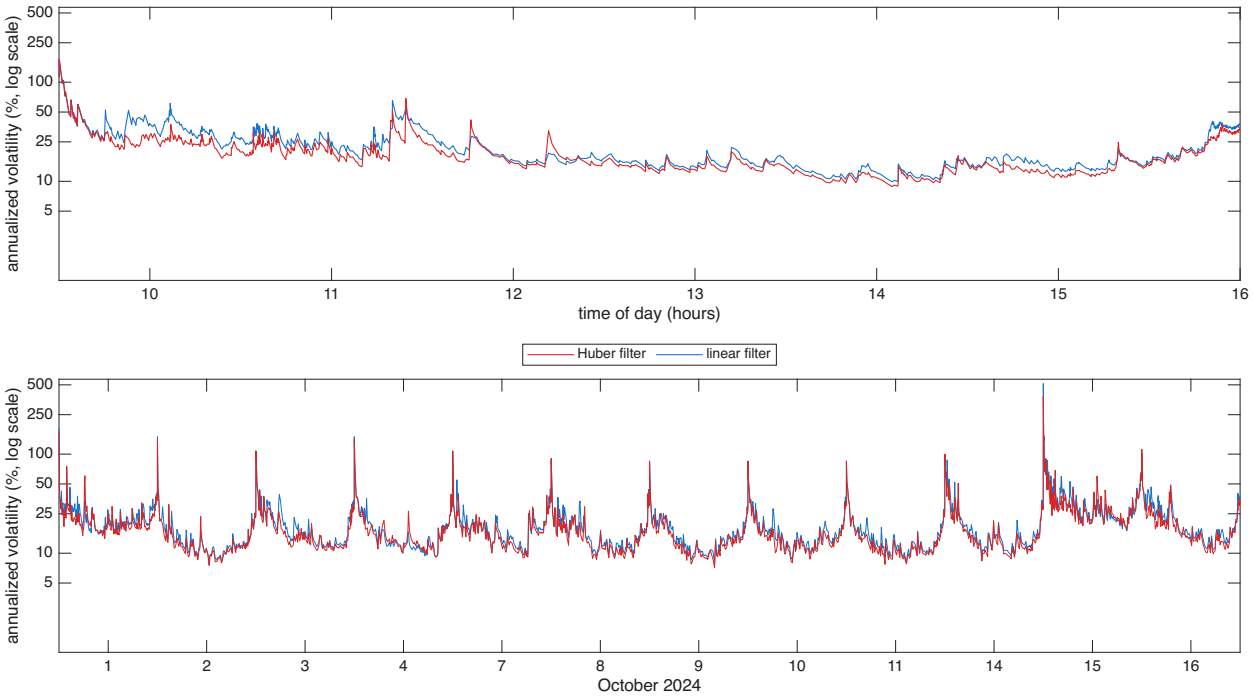

Figure 5: Filtered spot volatility for `ETSY` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 27$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

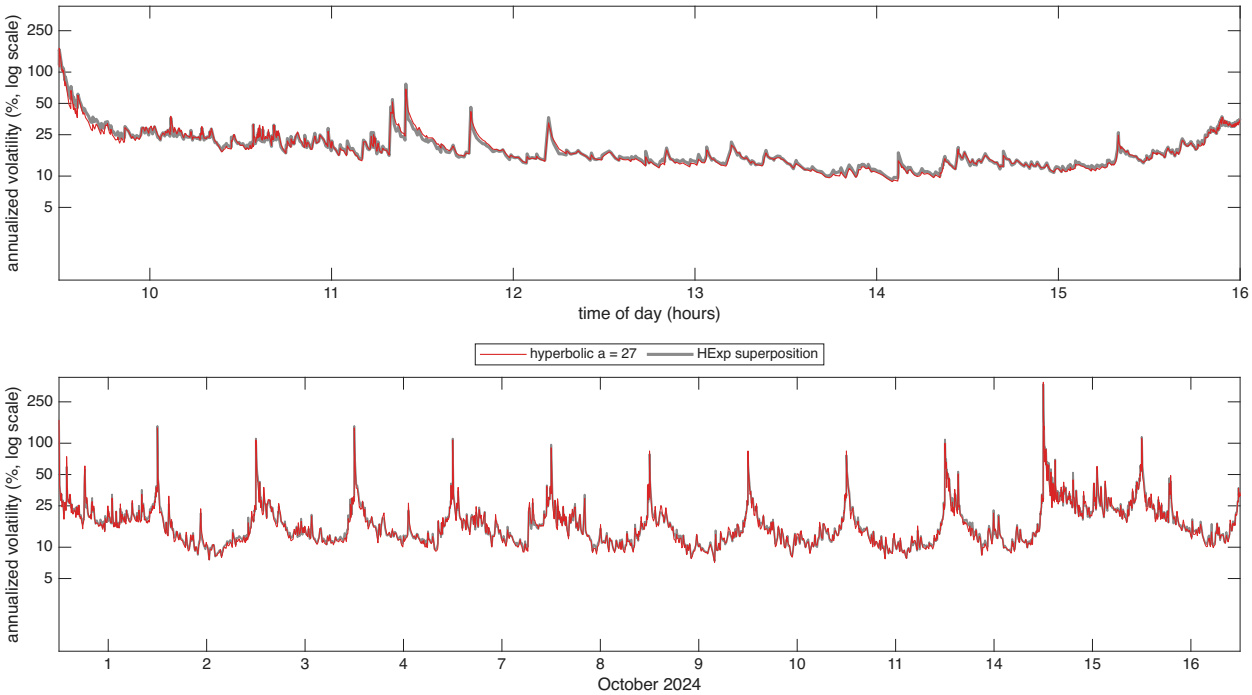

Figure 6: Hyperbolic against superposition weights for `ETSY` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 27$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

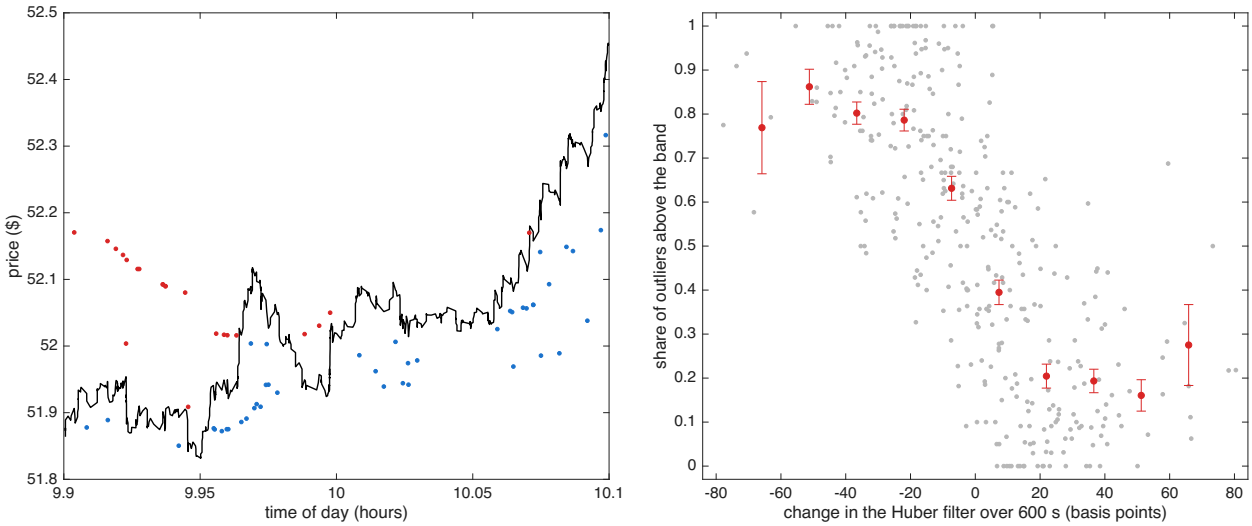

Figure 7: Outliers & robust price for `ETSY` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

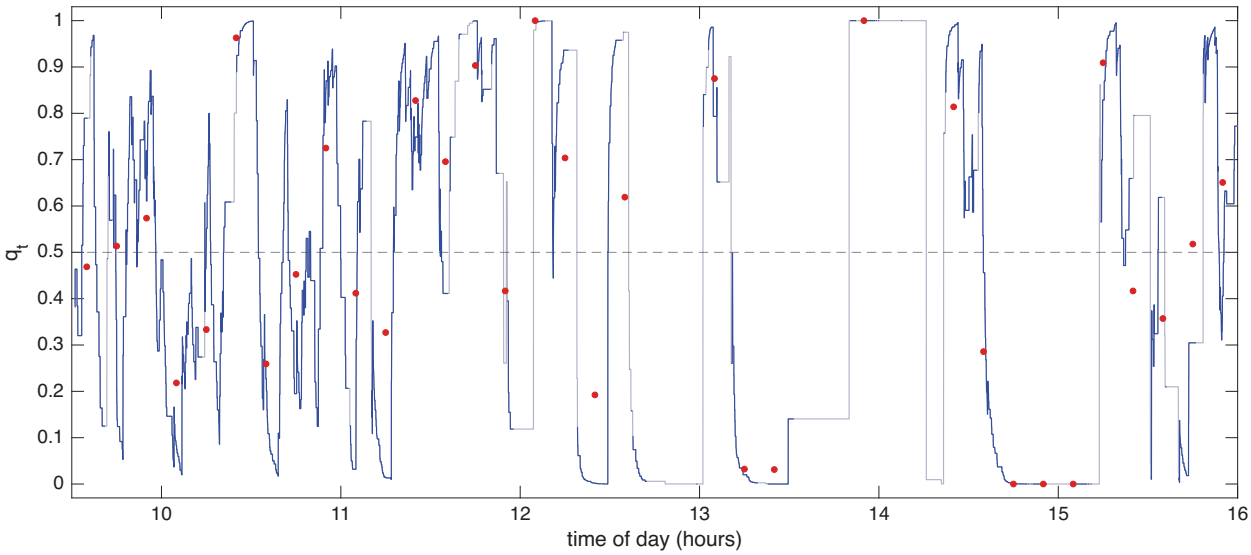

Figure 8: The directional conditional probability for ETSY, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 62 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 10 to 204 trades, against 27 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 19 trades for this stock.

Diagnostics.

LCID 1 to 16 October 2024

trades on 16 October 2024	35,435
average trading rate	1.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	9 (0.03%)
price half-life: run / minimizer[1]	3.0 / 5.0 trades
cost of running 3 trades	+2.44%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.01%
direction half-life[3]	415 s [173, 1500]
Mandelbrot a minimizing QLIKE	74 trades
a within 0.0005 of the minimum	45 to 102
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	36% $\rightarrow$ 38%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	36% $\rightarrow$ 36%
median spot volatility, Huber / linear input[5]	49.5 / 49.7 %
hyperbolic against HExp, correlation of logs	0.9976
HExp QLIKE advantage, leave-one-day-out[6]	+0.00237
15-minute signature level	32 to 70 %
Huber signature, 1 min over 1 s, 12 day average	0.93

LCID over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 13:43; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.054	0.038	0.039	0.174	0.141	0.055	0.056	0.318
2	0.049	0.042	0.042	0.142	0.119	0.059	0.060	0.244
5	0.052	0.045	0.045	0.103	0.103	0.062	0.062	0.165
10	0.051	0.044	0.044	0.083	0.088	0.061	0.061	0.126
25	0.041	0.037	0.037	0.058	0.065	0.050	0.050	0.084
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.29	0.15	0.15	−0.41	8.22	6.23	5.48	5.46
2	0.06	0.03	0.03	−0.05	12.40	7.24	6.85	4.90
5	−0.02	−0.02	−0.02	0.01	19.41	8.35	8.31	5.49
10	−0.02	−0.05	−0.05	−0.01	14.70	8.27	8.27	4.90
25	−0.01	0.00	0.00	−0.01	13.75	7.14	7.15	4.76

Table 1: Trade-lag differences, `LCID`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.21	0.18	0.18	0.40	100.4	43.6	43.8	255.9
1 s	0.15	0.13	0.13	0.25	86.1	43.6	43.7	203.5
5 s	0.11	0.10	0.10	0.16	73.3	49.1	49.2	128.1
10 s	0.09	0.08	0.08	0.13	65.1	47.5	47.6	98.7
1 min	0.06	0.06	0.06	0.08	46.4	40.9	40.9	54.7
15 min	0.05	0.05	0.05	0.06	38.7	37.6	37.6	38.7
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.08	−0.01	−0.01	−0.13	19546414805016.32	4.86	4.86	6.12
1 s	−0.16	−0.01	−0.01	−0.33	11.99	4.72	4.72	6.81
5 s	−0.26	−0.15	−0.15	−0.40	7.64	8.46	8.46	13.96
10 s	−0.29	−0.16	−0.16	−0.40	12.92	8.17	8.18	7.39
1 min	−0.12	−0.06	−0.06	−0.22				

Table 2: Clock-lag differences, `LCID`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.11	0.15	0.15
s.d. of $\hat{\epsilon}_t$ in cents	0.24	0.21	0.21
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	5.83
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.12	0.15	0.15
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.03	0.03	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	−0.01	−0.05	−0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	−0.10,0.10	−0.10,0.10
95% C.I. for $\theta_t - X_t$	—	−0.16,0.16	−0.16,0.16
99% C.I. for $\theta_t - X_t$	—	−0.22,0.22	−0.22,0.22
s.e. of $\theta_t - X_t$ in cents	—	0.08	0.08

Table 3: Residual summaries, `LCID`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 5.83, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	7.36	5.54	2.38 to 10.79	3.12 to 8.56
Residual divided by δ_t	6.60	6.03	3.34 to 8.56	3.06 to 8.15
Repricing windows removed	7.36	5.54	2.38 to 10.79	3.12 to 8.56
Calm windows only	7.36	5.54		
Every 10th trade	5.99	6.54	3.53 to 6.86	4.41 to 8.65
Six-trade half-life	7.71	5.39	2.14 to 10.84	2.81 to 8.45

Table 4: Hill estimator of the tail index of the Huber filter residual, `LCID`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 7.1 to 8.0 cents over the days and removes 0% to 0% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 3.8 cents on 16 October 2024, 100% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

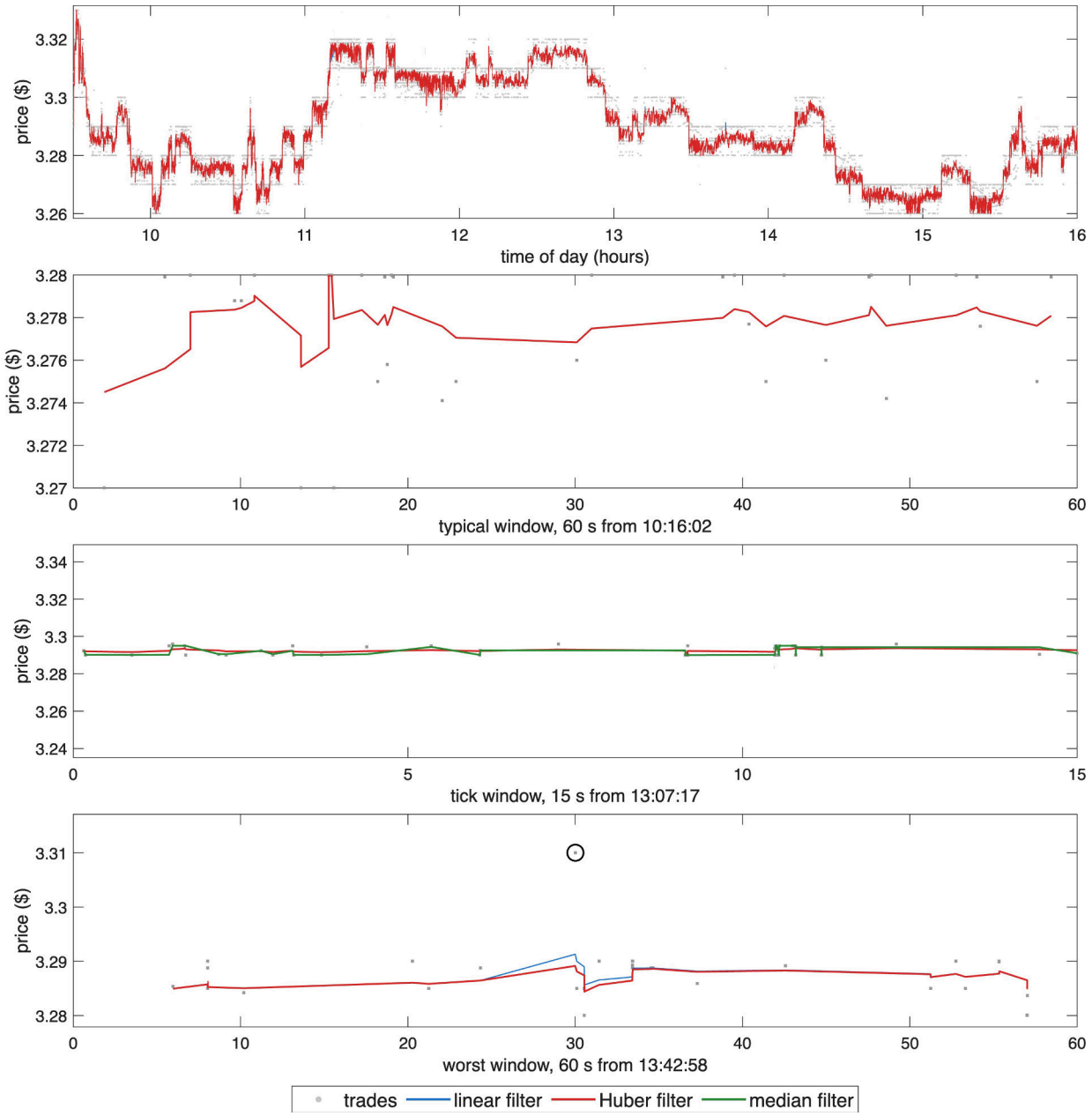

Figure 1: `LCID` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:16:02. 3rd: 15-second window at 13:07:17 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

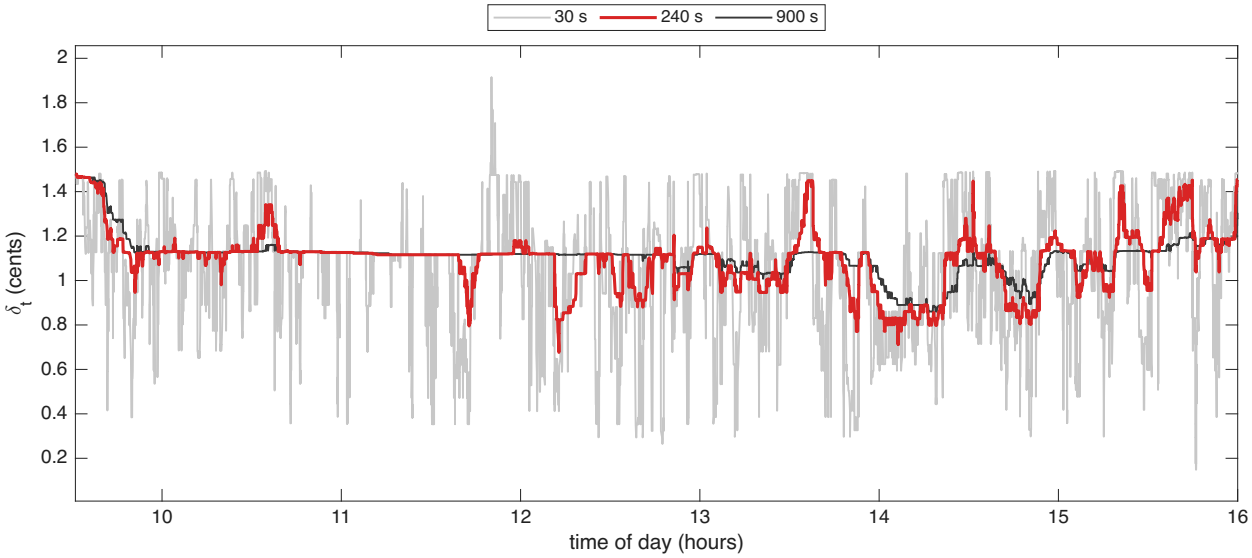

Figure 2: The Huber threshold δ_t in cents for `LCID` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

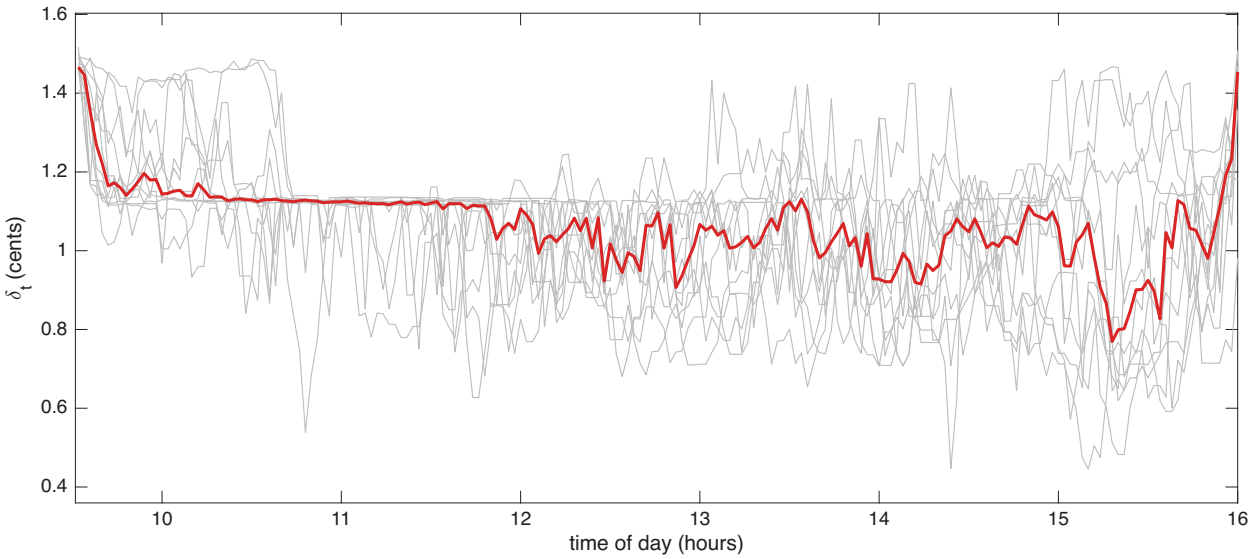

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `LCID` trading days in Oct 2024 (grey lines) and their median (red line).

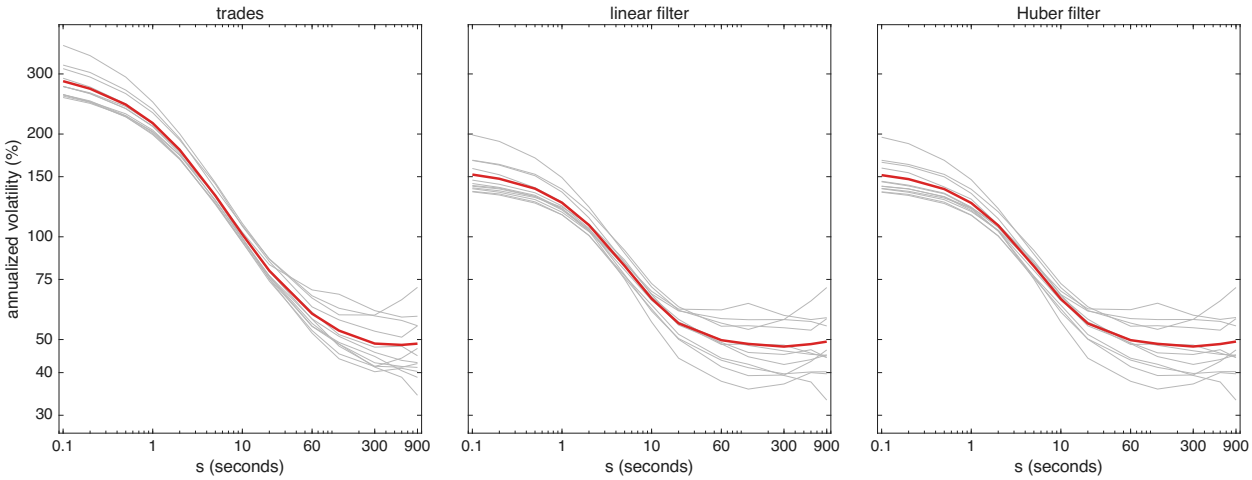

Figure 4: Volatility signature curves for `LCID` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

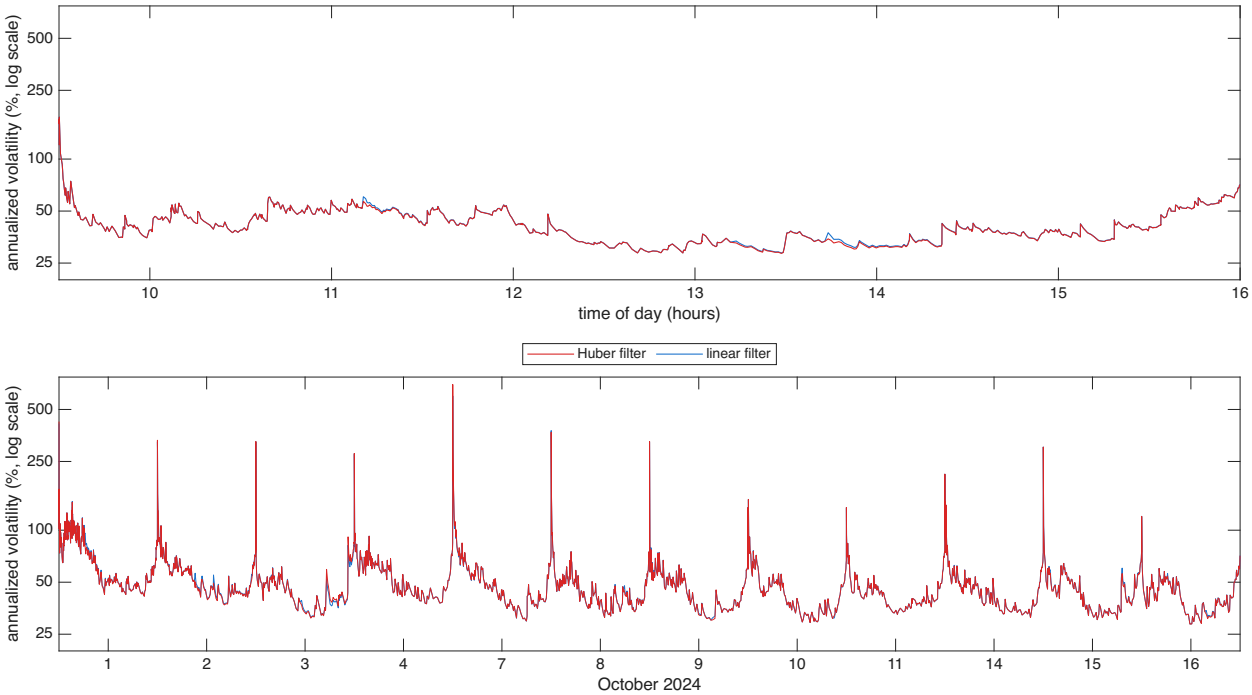

Figure 5: Filtered spot volatility for `LCID` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 74$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

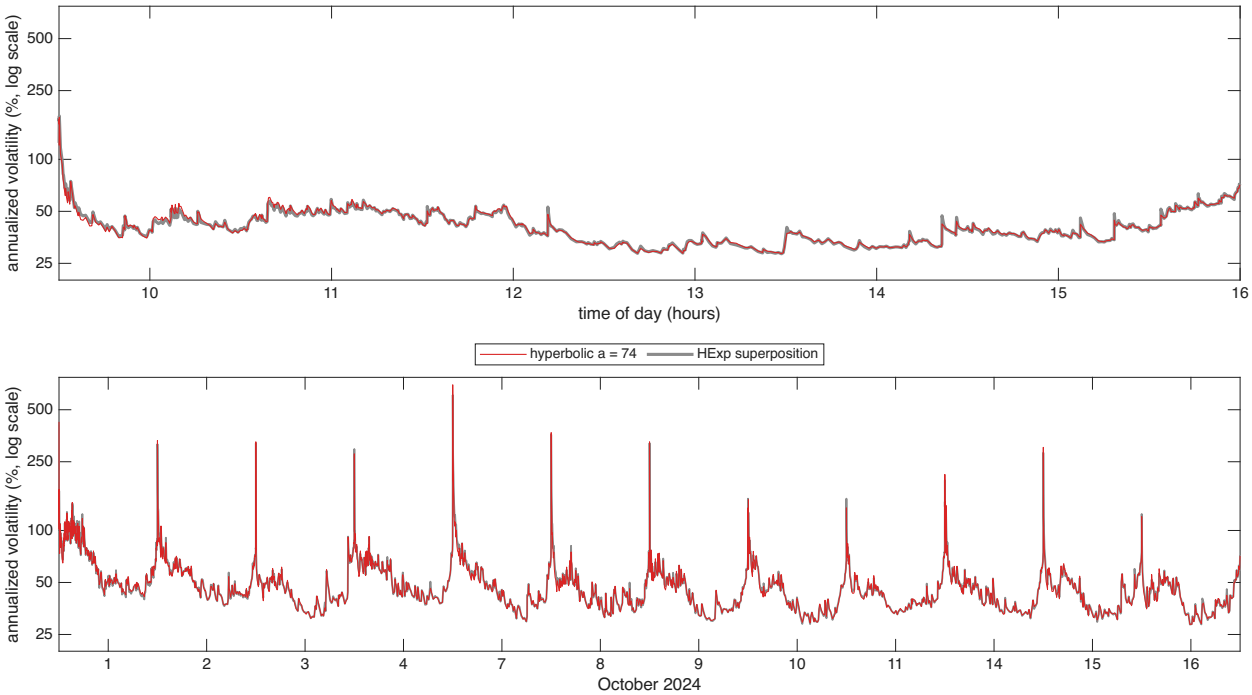

Figure 6: Hyperbolic against superposition weights for `LCID` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 74$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

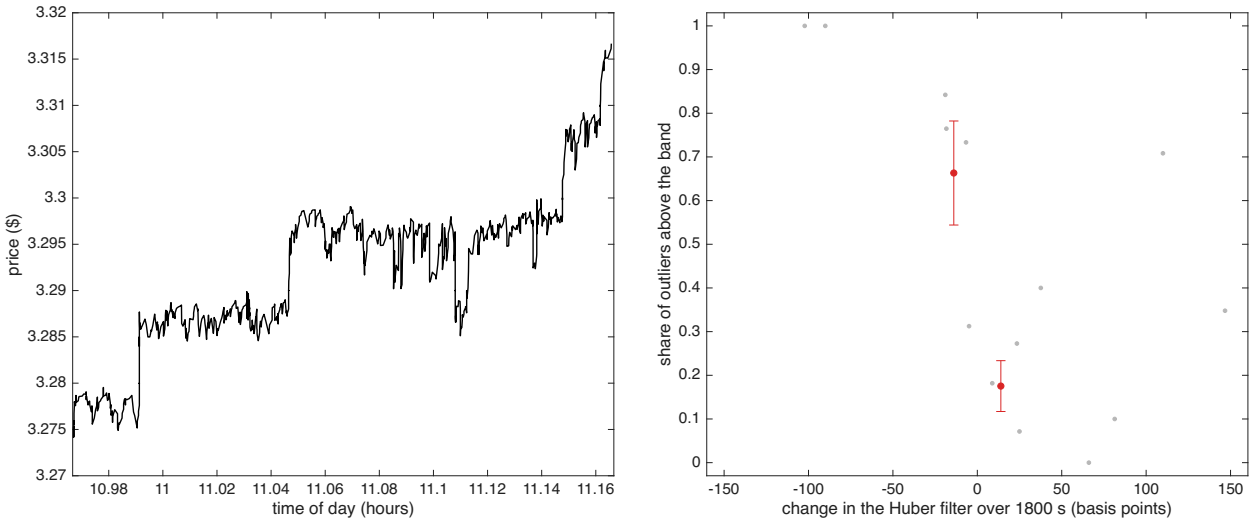

Figure 7: Outliers & robust price for `LCID` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

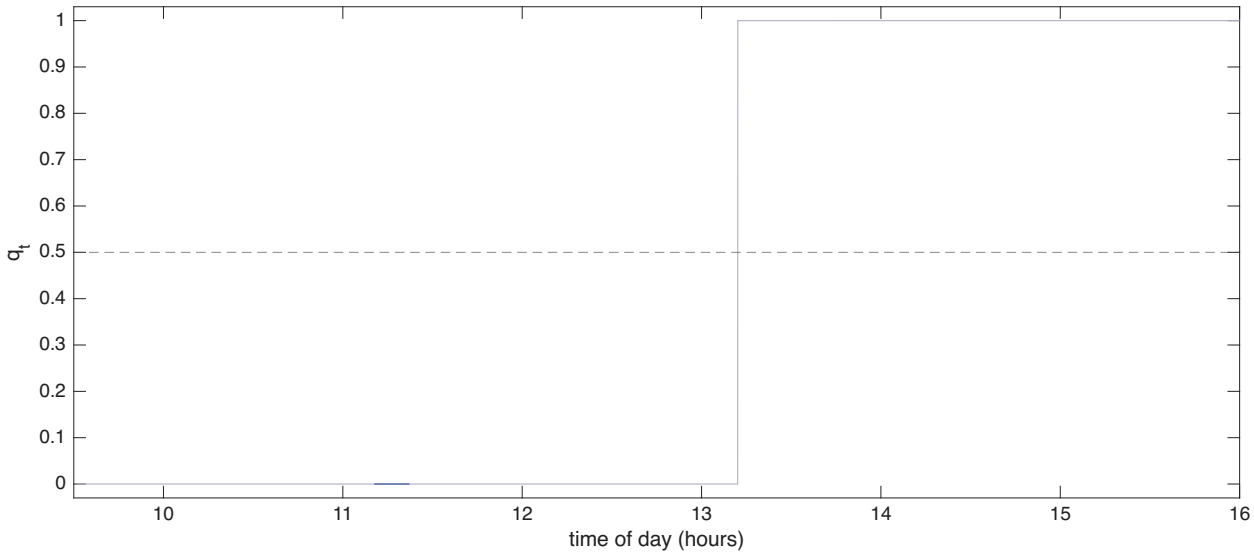

Figure 8: The directional conditional probability for `LCID`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 415 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

At the horizons k = 2, 3, 5, 10, 30 and 100 the half-life minimizing the k-trade-ahead Huber forecast loss is 4, 5, 5, 5, 5 and 5 trades, within one grid point of the end of the 1 to 5 trade grid searched.
The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
Only 9 trades lie beyond the Huber threshold on 16 October 2024.

Diagnostics.

TTD 1 to 16 October 2024

trades on 16 October 2024	27,736
average trading rate	1.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	921 (3.32%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.55%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	100 s [86, 117]
Mandelbrot a minimizing QLIKE	23 trades
a within 0.0005 of the minimum	17 to 31
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	9% $\to$ 29%
Huber $\hat{\sigma}$, before $\to$ peak	9% $\to$ 15%
median spot volatility, Huber / linear input[5]	12.6 / 14.2 %
hyperbolic against HExp, correlation of logs	0.9929
HExp QLIKE advantage, leave-one-day-out[6]	+0.01577
15-minute signature level	16 to 39 %
Huber signature, 1 min over 1 s, 12 day average	1.64

TTD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 13:01; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.440	0.347	0.363	1.571	1.077	0.563	0.618	3.538
2	0.428	0.390	0.421	1.316	0.943	0.605	0.661	2.679
5	0.501	0.453	0.486	1.007	0.874	0.665	0.707	1.804
10	0.548	0.498	0.520	0.853	0.840	0.713	0.735	1.388
25	0.592	0.552	0.558	0.733	0.834	0.772	0.777	1.078
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.15	0.14	−0.43	2.75	2.48	2.99	3.37
2	0.07	0.09	0.05	−0.04	2.91	2.82	3.92	3.40
5	0.01	0.04	0.02	0.02	4.16	3.65	4.23	3.56
10	0.01	0.02	0.02	0.01	5.19	4.48	4.70	3.57
25	0.00	0.00	0.00	0.00	7.77	6.67	6.60	4.31

Table 1: Trade-lag differences, `TTD`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	1.62	1.45	1.51	3.42	17.3	10.7	12.5	73.8
1 s	1.26	1.15	1.21	2.43	16.4	11.6	13.1	63.4
5 s	0.97	0.91	0.94	1.51	16.6	14.2	15.1	44.3
10 s	0.89	0.83	0.86	1.23	17.5	15.9	16.4	35.1
1 min	0.74	0.73	0.72	0.85	18.6	18.0	17.8	22.2
15 min	0.97	0.98	0.98	1.00	22.9	22.9	22.8	23.5
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.01	0.05	0.06	−0.06	8.12	6.51	7.13	6.16
1 s	−0.05	−0.01	−0.09	−0.44	10.26	7.38	7.86	6.83
5 s	−0.01	0.02	0.00	−0.41	4.75	4.62	4.66	5.77
10 s	−0.09	−0.02	−0.04	−0.31	6.01	5.63	5.17	5.32
1 min	0.08	0.12	0.12	−0.04				

Table 2: Clock-lag differences, `TTD`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.09	1.34	1.39
s.d. of $\hat{\epsilon}_t$ in cents	2.59	2.47	2.37
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.22	0.23	0.14
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.13	0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.06	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.03	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.01
80% C.I. for $\theta_t - X_t$	—	−0.98,0.94	−1.06,1.00
95% C.I. for $\theta_t - X_t$	—	−1.76,1.60	−1.96,1.76
99% C.I. for $\theta_t - X_t$	—	−2.65,2.61	−3.10,3.13
s.e. of $\theta_t - X_t$ in cents	—	0.82	0.94

Table 3: Residual summaries, `TTD`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.00, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.47	2.76	1.95 to 3.11	2.13 to 3.32
Residual divided by δ_t	3.49	2.42	2.23 to 3.49	2.14 to 4.51
Repricing windows removed	2.47	2.76	2.04 to 3.12	2.15 to 3.30
Calm windows only	2.39	2.69		
Every 10th trade	—	—	2.00 to 3.19	1.41 to 3.46
Six-trade half-life	2.65	2.99	2.07 to 3.37	2.25 to 3.48

Table 4: Hill estimator of the tail index of the Huber filter residual, `TTD`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.2 to 30.4 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 15.1 cents on 16 October 2024, 98% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

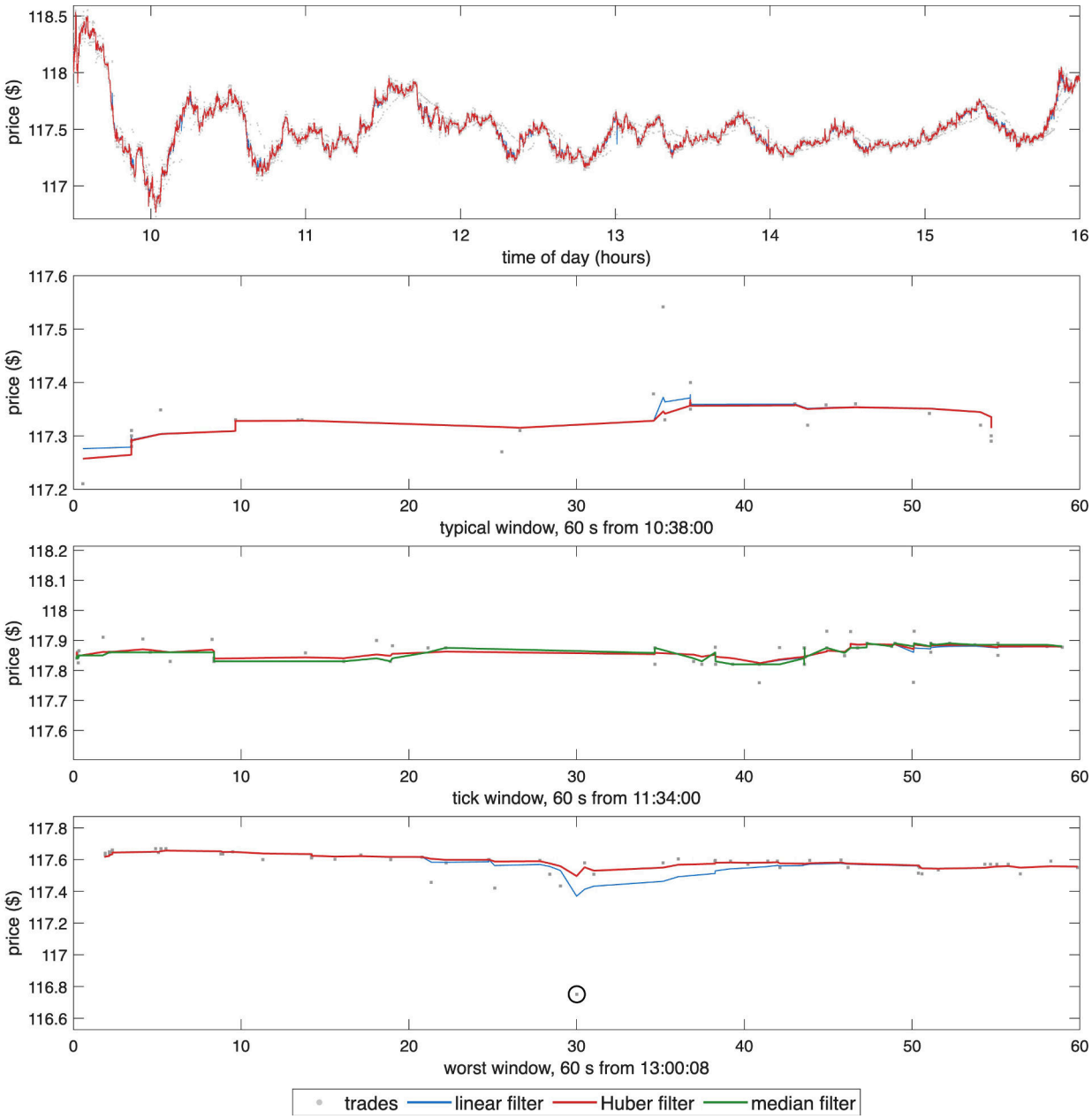

Figure 1: TTD trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:38:00. 3rd: 60-second window at 11:34:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

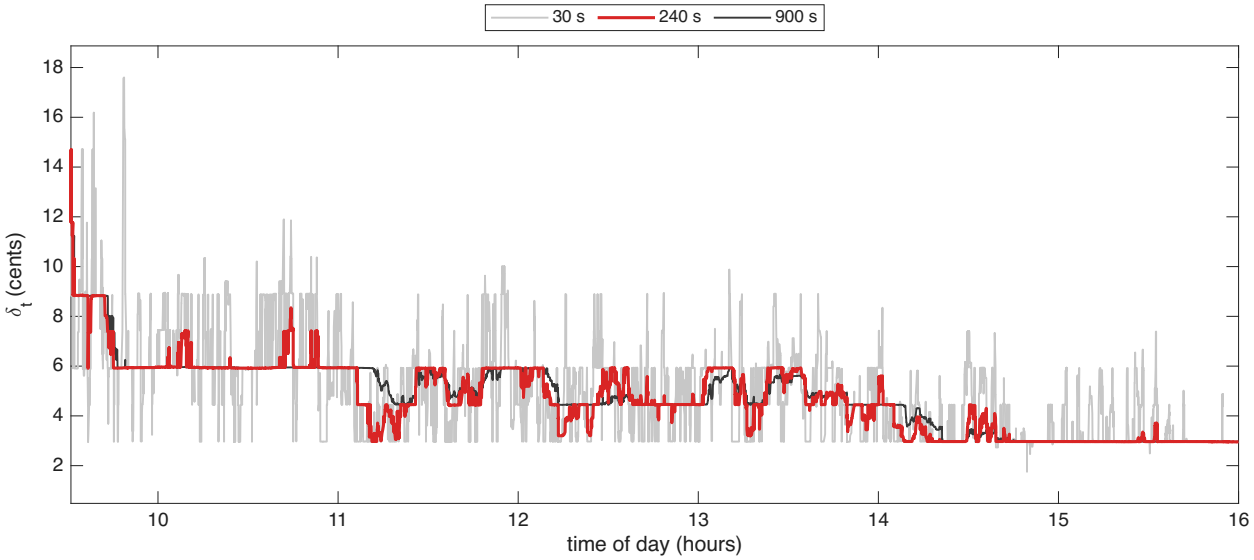

Figure 2: The Huber threshold δ_t in cents for `TTD` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

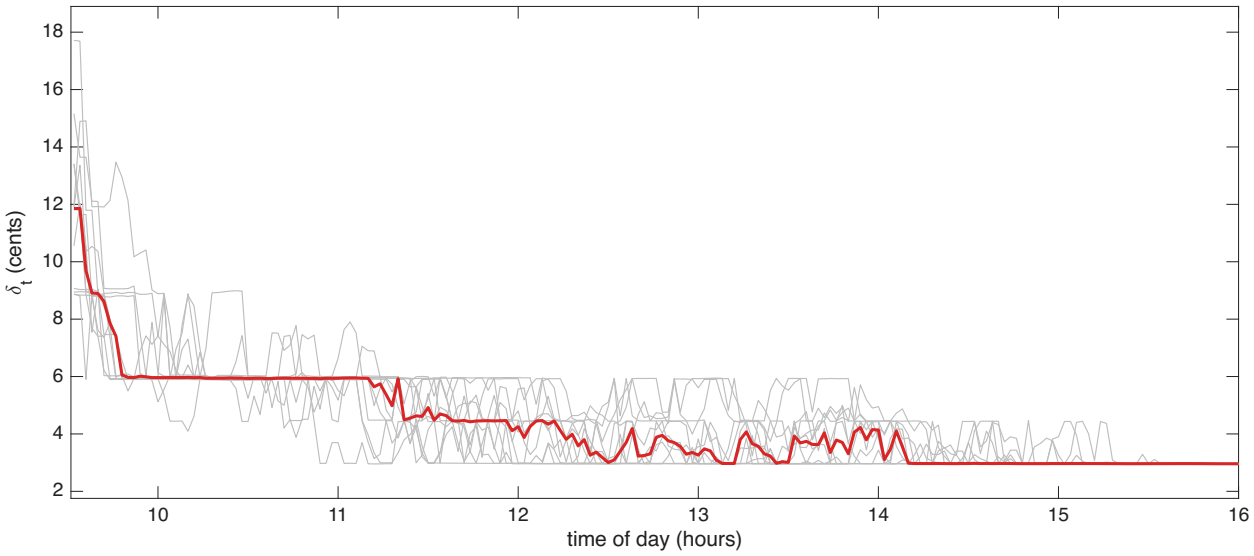

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `TTD` trading days in Oct 2024 (grey lines) and their median (red line).

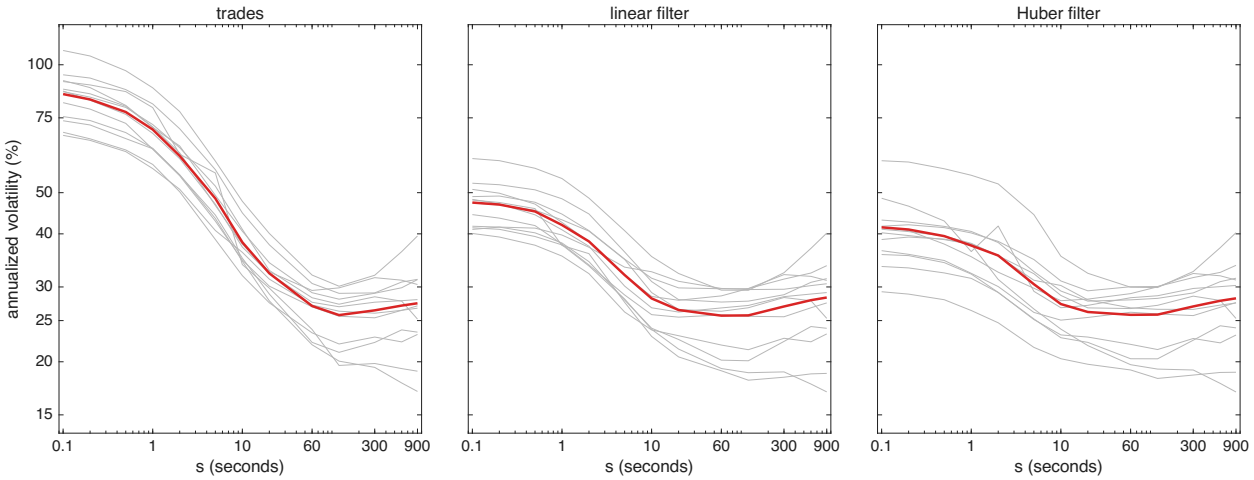

Figure 4: Volatility signature curves for `TTD` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

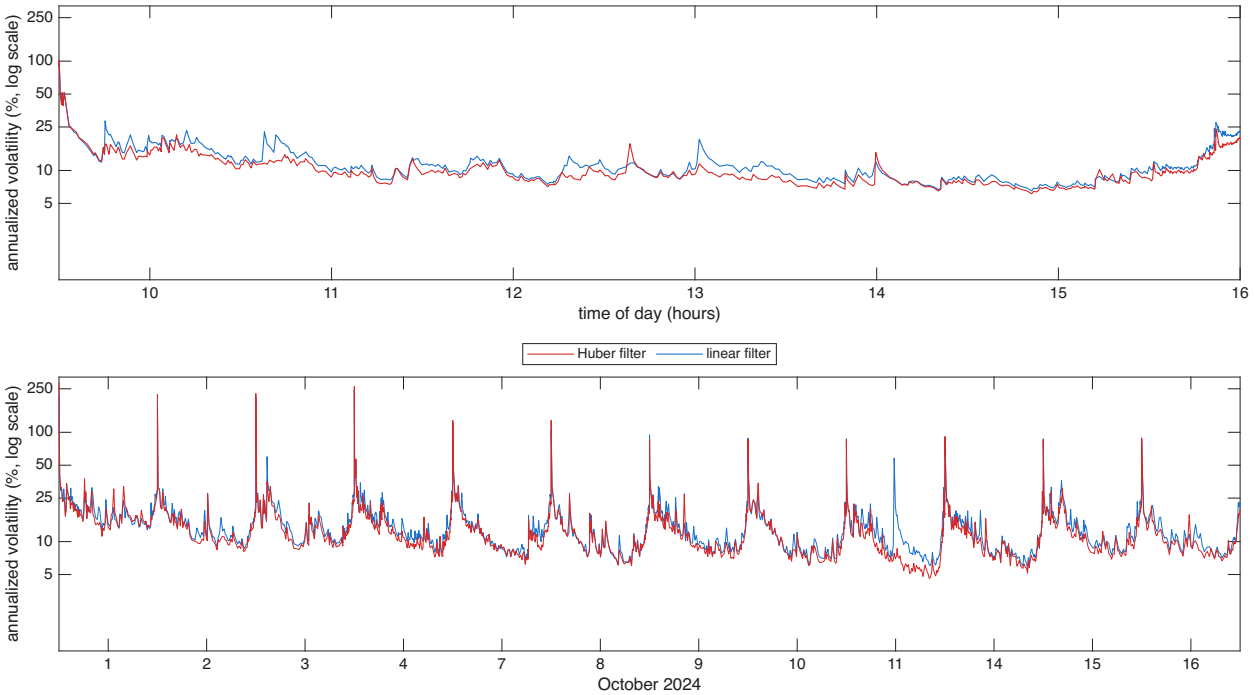

Figure 5: Filtered spot volatility for `TTD` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 23$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

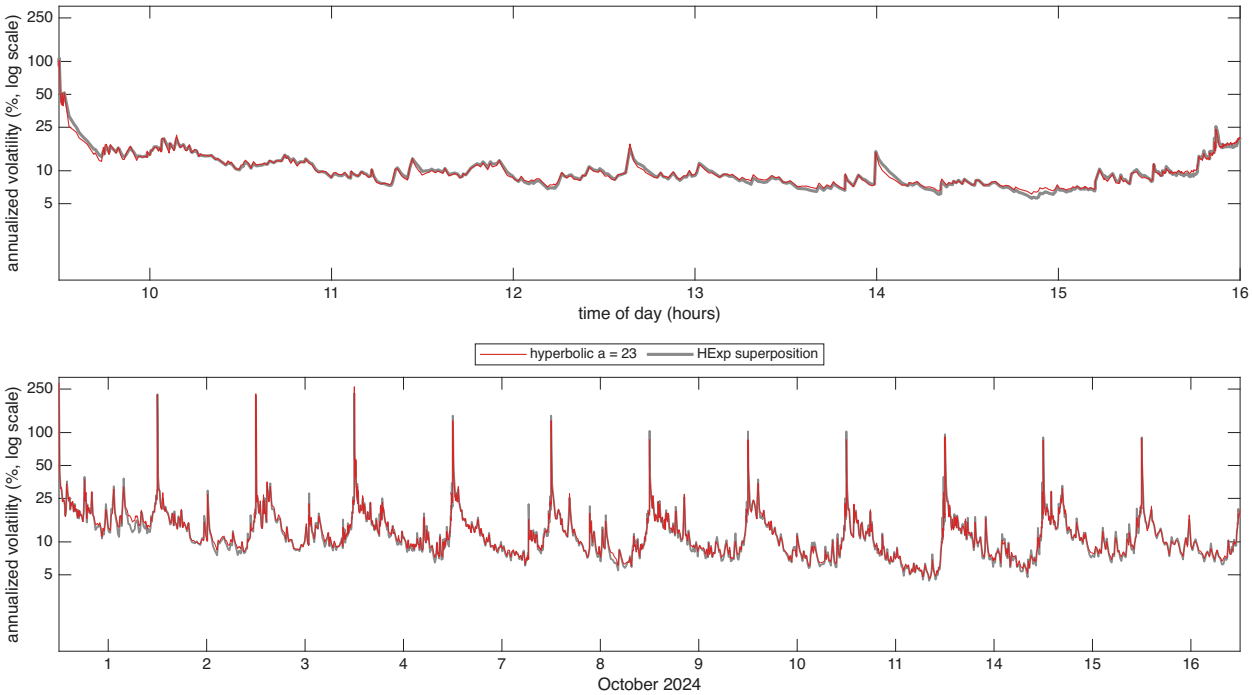

Figure 6: Hyperbolic against superposition weights for `TTD` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 23$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

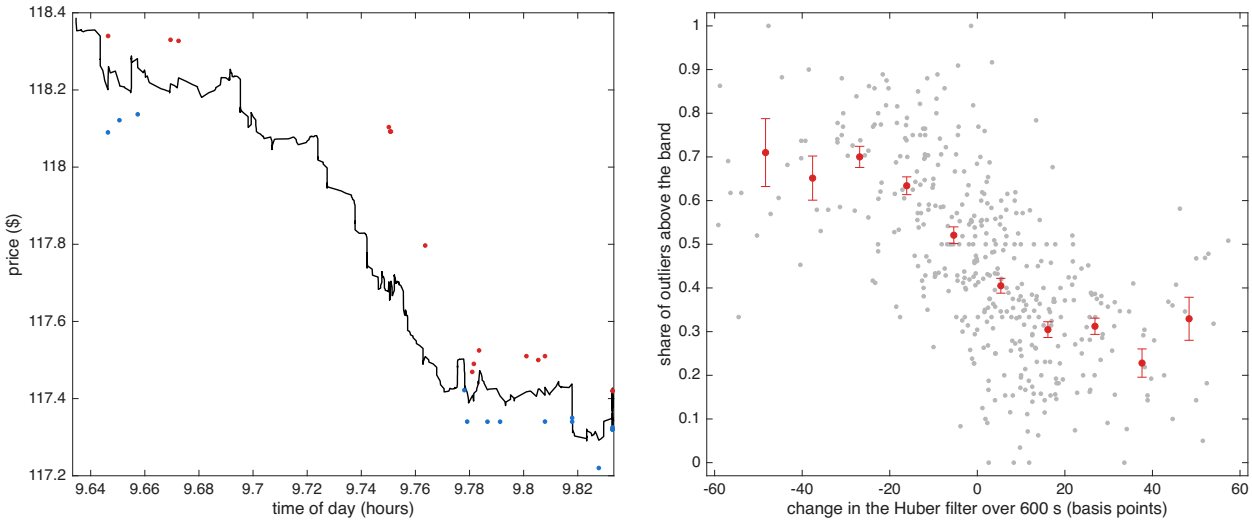

Figure 7: Outliers & robust price for `TTD` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

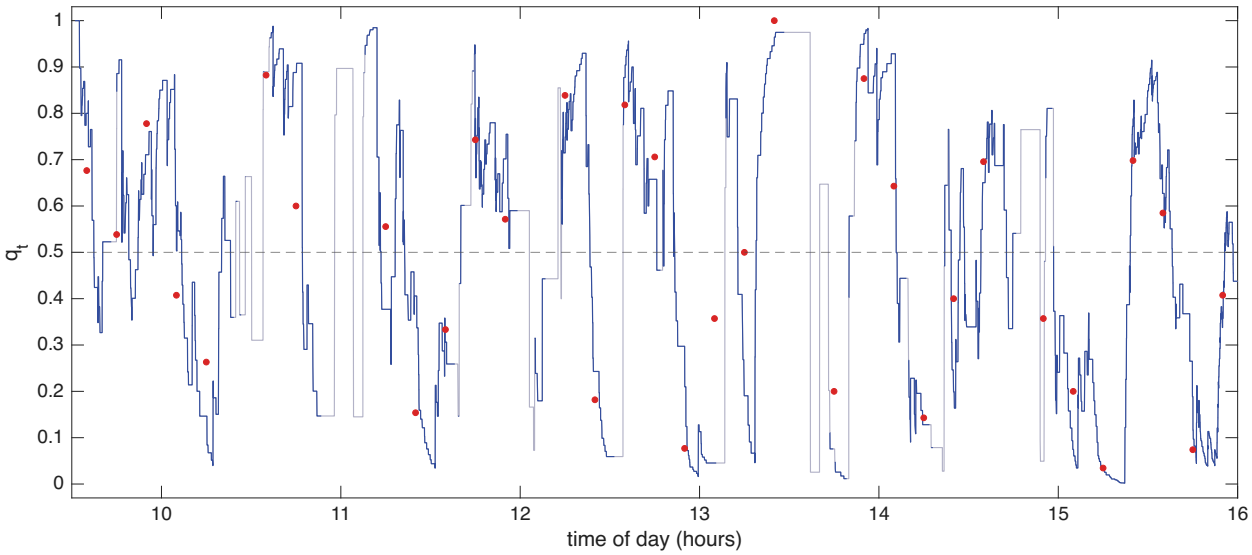

Figure 8: The directional conditional probability for TTD, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 100 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 9 to 90 trades, against 23 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 16 trades for this stock.
The tail index of the residual cannot be estimated for every 10th trade, on 2 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

ROKU 1 to 16 October 2024

trades on 16 October 2024	43,596
average trading rate	1.5 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,489 (3.42%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+1.84%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.01%
direction half-life[3]	77 s [64, 97]
Mandelbrot a minimizing QLIKE	31 trades
a within 0.0005 of the minimum	23 to 43
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	61% $\to$ 64%
Huber $\hat{\sigma}$, before $\to$ peak	54% $\to$ 56%
median spot volatility, Huber / linear input[5]	18.7 / 20.8 %
hyperbolic against HExp, correlation of logs	0.9975
HExp QLIKE advantage, leave-one-day-out[6]	+0.01334
15-minute signature level	25 to 55 %
Huber signature, 1 min over 1 s, 12 day average	1.63

ROKU over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:52; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.388	0.322	0.336	1.434	0.959	0.537	0.584	3.267
2	0.388	0.366	0.392	1.196	0.872	0.581	0.636	2.500
5	0.463	0.429	0.459	0.927	0.840	0.655	0.698	1.716
10	0.514	0.479	0.500	0.796	0.822	0.715	0.738	1.337
25	0.584	0.555	0.557	0.711	0.850	0.804	0.808	1.076
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.17	0.17	0.19	−0.41	3.65	3.67	2.96	3.39
2	0.06	0.11	0.08	−0.04	3.67	3.66	3.48	3.54
5	−0.01	0.05	0.03	0.01	4.50	4.04	4.58	3.68
10	0.00	0.02	0.02	0.00	4.27	4.17	4.74	3.88
25	−0.01	0.01	0.00	0.00	5.52	5.59	5.90	4.16

Table 1: Trade-lag differences, ROKU, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.45	2.29	2.32	4.15	28.3	19.1	21.9	122.2
1 s	1.77	1.71	1.74	2.88	29.1	23.7	25.7	96.1
5 s	1.69	1.62	1.63	2.10	36.0	33.9	34.6	64.6
10 s	1.64	1.59	1.61	1.94	39.3	38.1	38.3	56.6
1 min	1.66	1.65	1.64	1.73	50.9	50.3	49.8	52.8
15 min	1.73	1.72	1.72	1.74	55.1	55.0	54.9	55.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.07	0.08	0.08	−0.10	4.71	4.34	4.27	7.16
1 s	0.32	0.35	0.31	−0.15	10.65	10.68	10.71	6.89
5 s	0.43	0.51	0.52	0.24	8.34	11.23	11.19	6.19
10 s	0.25	0.27	0.25	0.13	11.48	19.04	24.35	16.56
1 min	0.16	0.16	0.16	0.12				

Table 2: Clock-lag differences, ROKU, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.01	1.24	1.29
s.d. of $\hat{\epsilon}_t$ in cents	2.44	2.34	2.24
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.98
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.24	0.27	0.18
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.16	0.08
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.05	0.08	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.04	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−0.90,0.88	−0.98,0.95
95% C.I. for $\theta_t - X_t$	—	−1.79,1.68	−1.99,1.86
99% C.I. for $\theta_t - X_t$	—	−2.91,2.83	−3.20,3.24
s.e. of $\theta_t - X_t$ in cents	—	0.81	0.93

Table 3: Residual summaries, `ROKU`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.98, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.17	2.90	2.17 to 3.25	2.34 to 3.62
Residual divided by δ_t	2.85	3.29	2.50 to 4.28	2.78 to 4.10
Repricing windows removed	2.17	2.92	2.17 to 3.38	2.34 to 3.62
Calm windows only	2.22	3.18		
Every 10th trade	1.93	2.33	1.55 to 2.83	1.56 to 3.37
Six-trade half-life	2.37	3.24	2.37 to 3.46	2.48 to 3.93

Table 4: Hill estimator of the tail index of the Huber filter residual, `ROKU`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.2 to 30.4 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 10.1 cents on 16 October 2024, 82% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

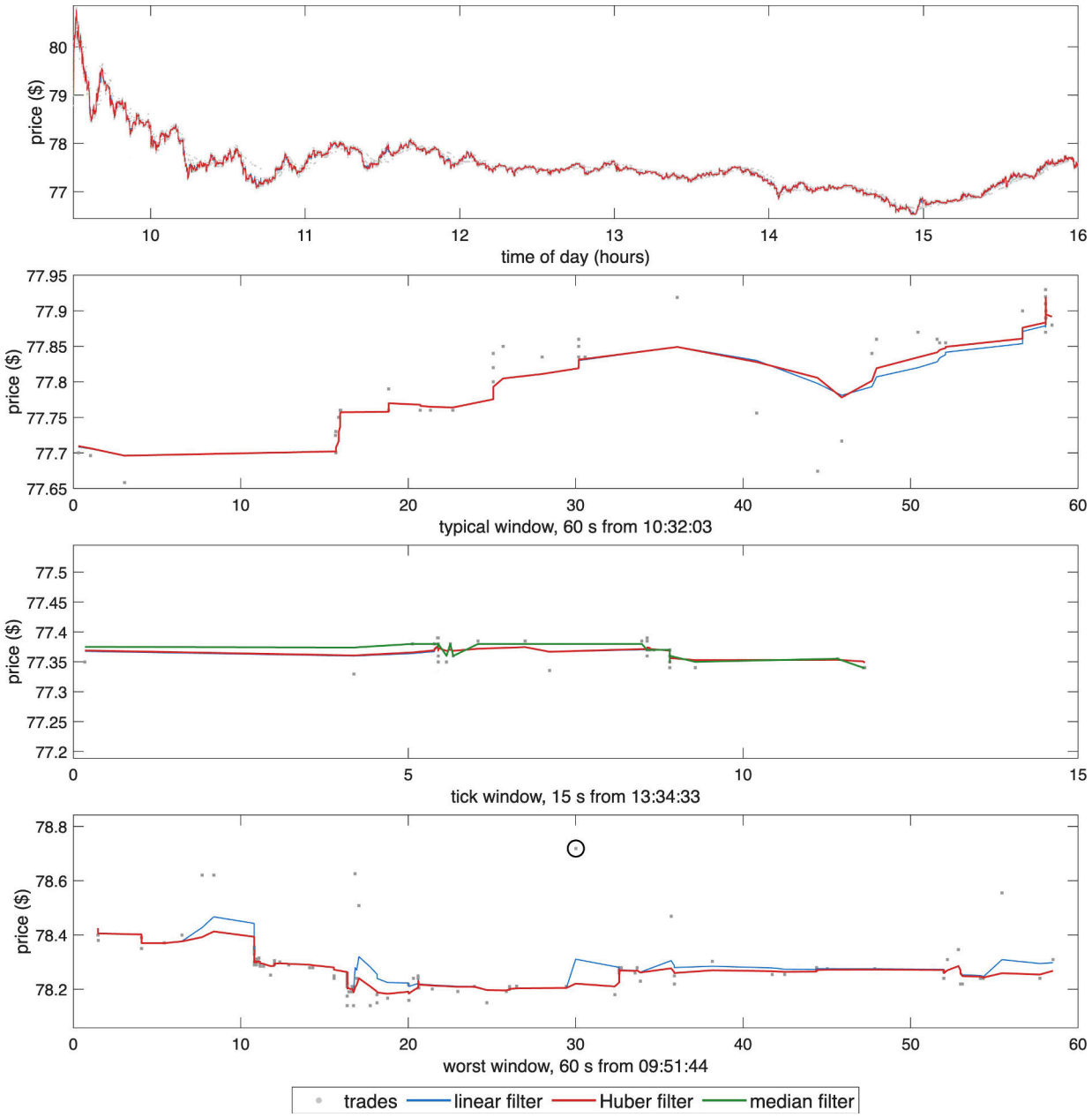

Figure 1: ROKU trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:32:03. 3rd: 15-second window at 13:34:33 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

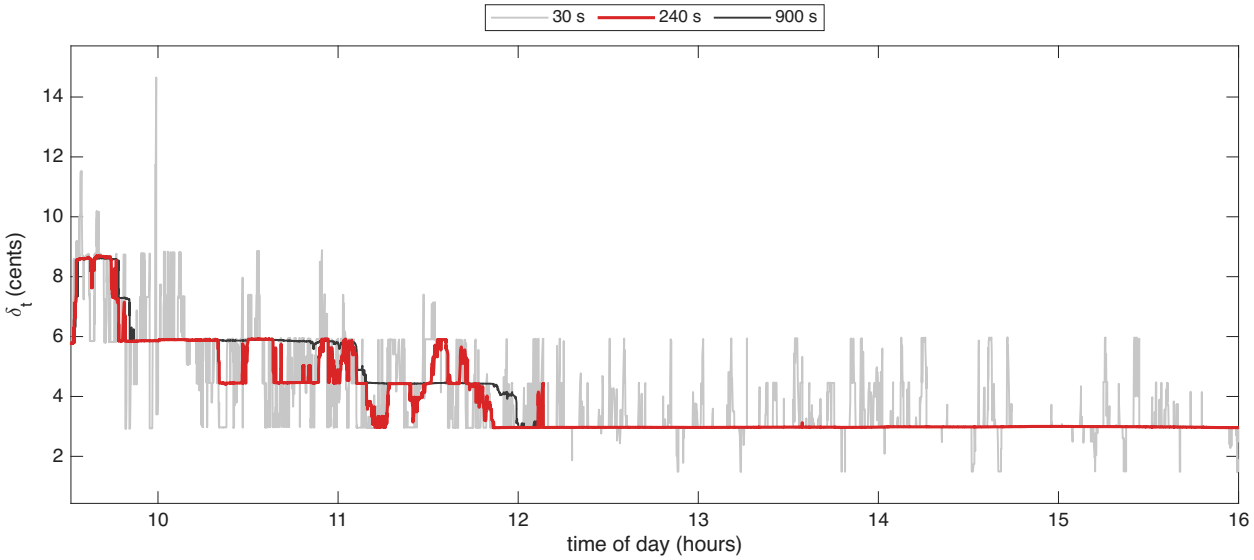

Figure 2: The Huber threshold δ_t in cents for ROKU on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

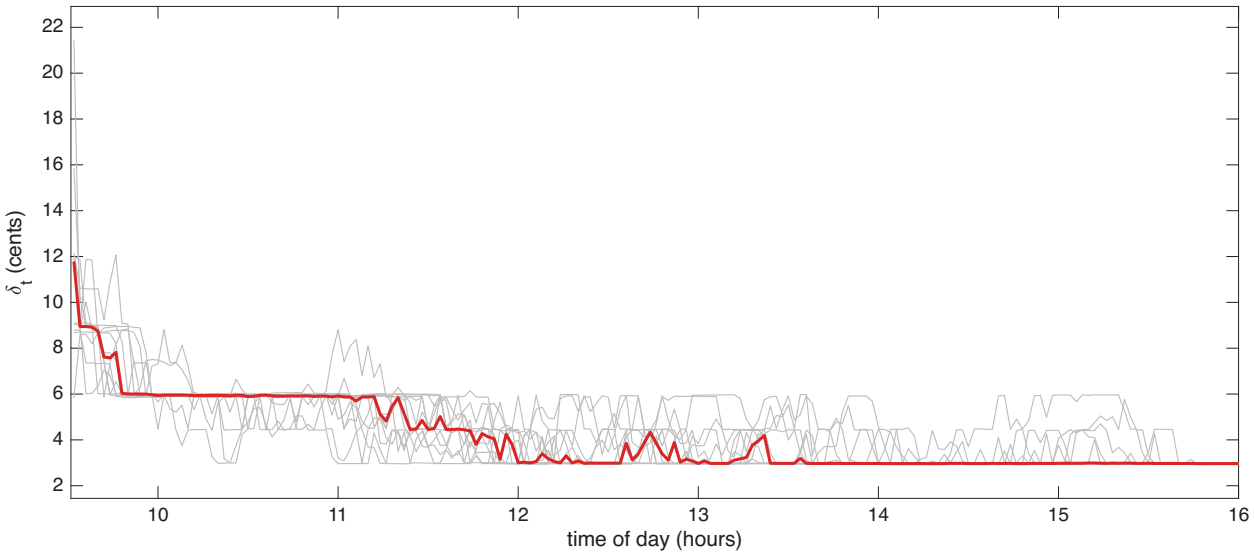

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 ROKU trading days in Oct 2024 (grey lines) and their median (red line).

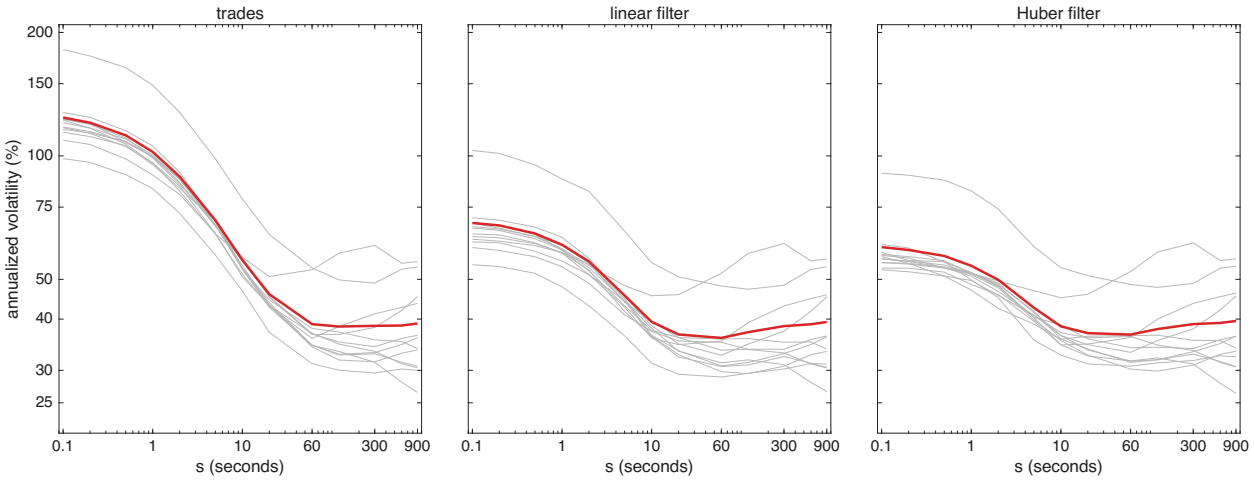

Figure 4: Volatility signature curves for ROKU over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

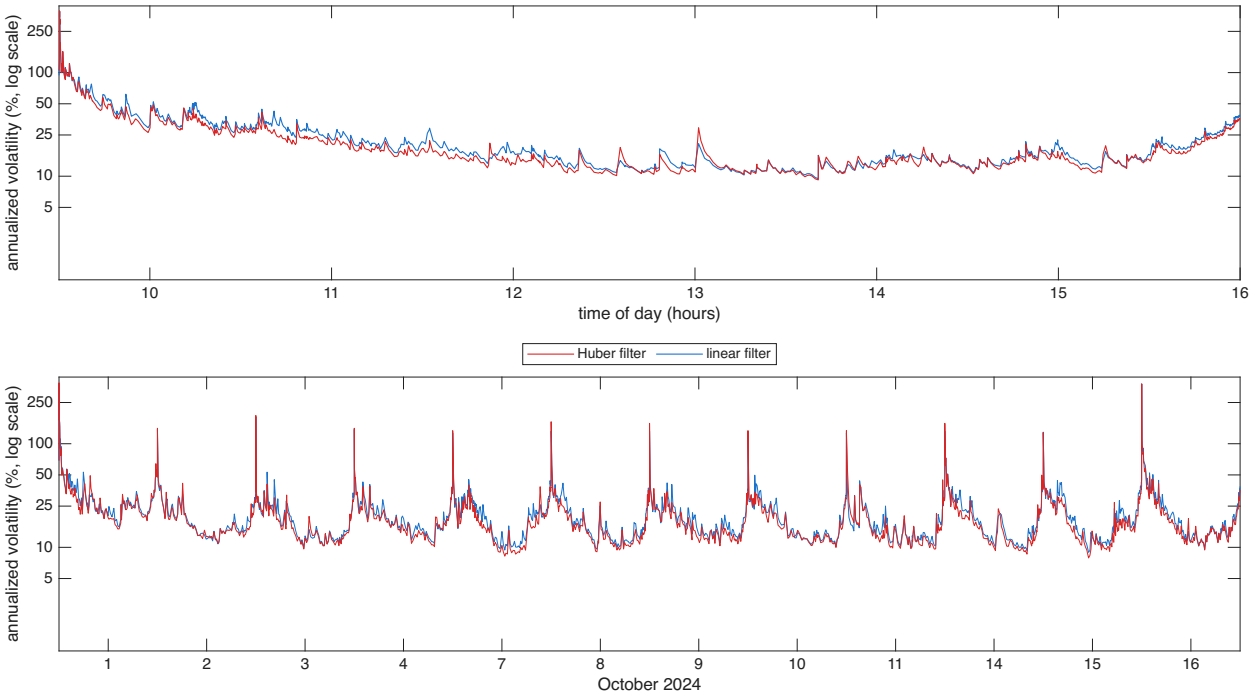

Figure 5: Filtered spot volatility for ROKU over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 31$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

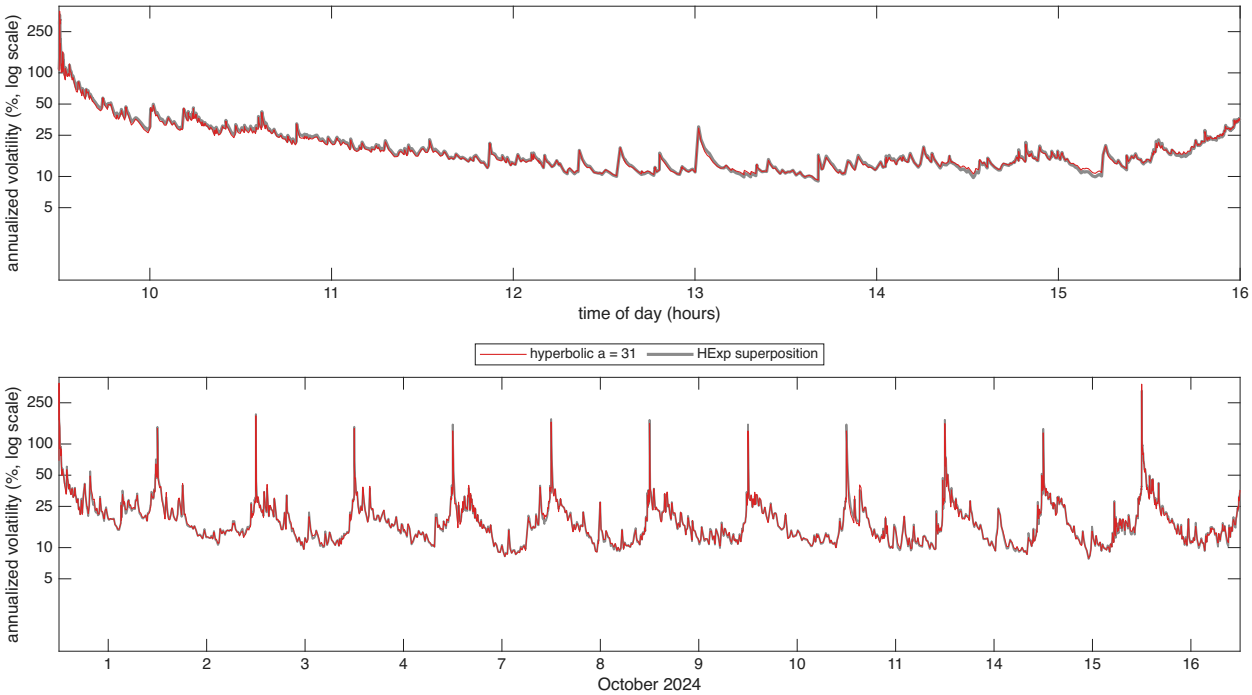

Figure 6: Hyperbolic against superposition weights for ROKU over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 31$ in red, drawn on top, and the HExp superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

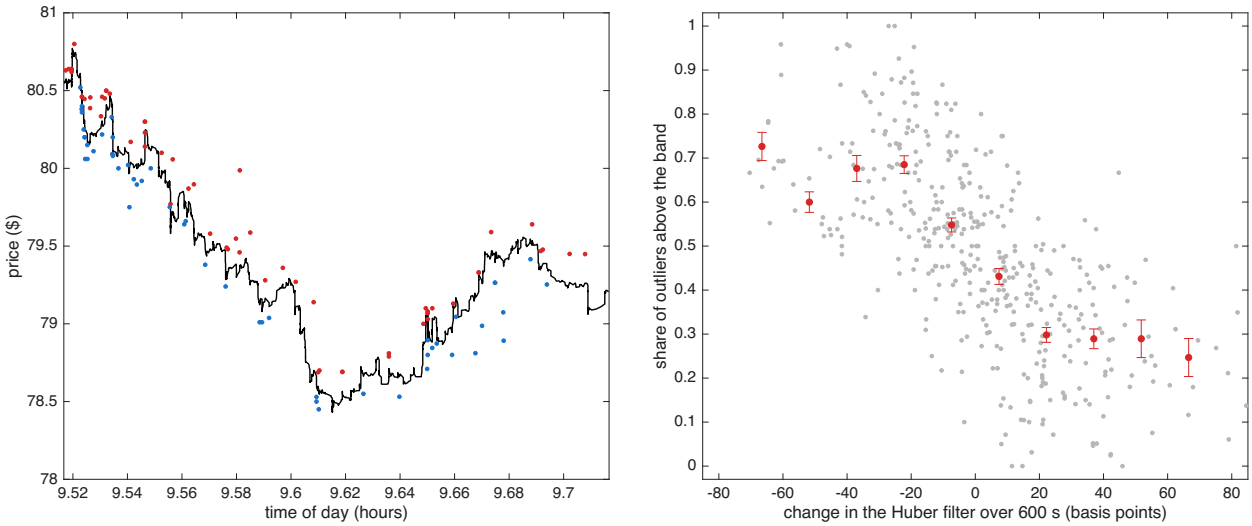

Figure 7: Outliers & robust price for ROKU on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

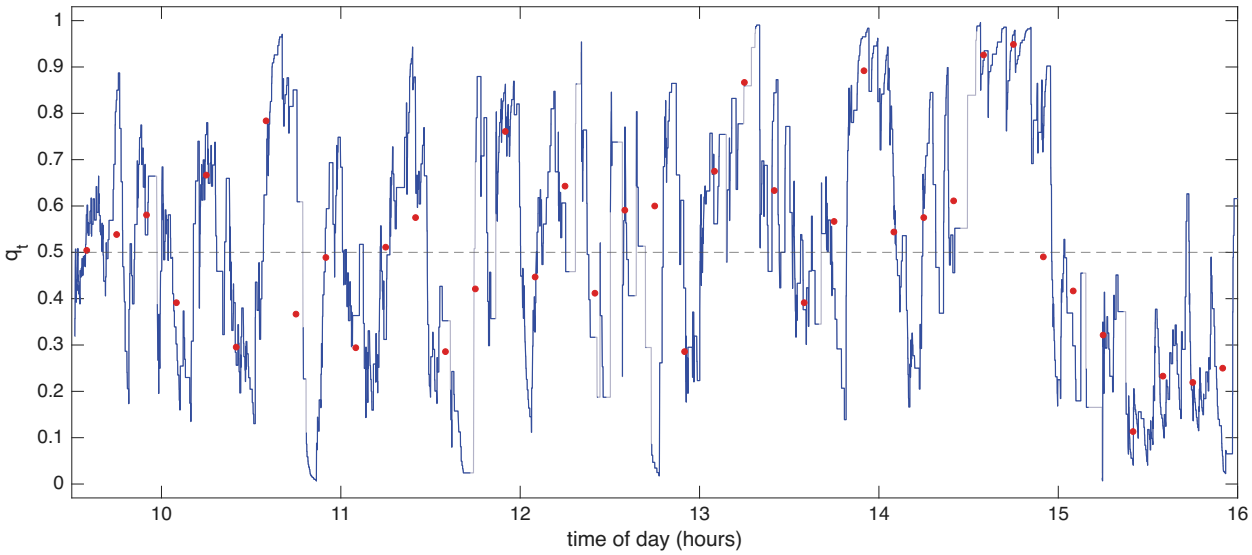

Figure 8: The directional conditional probability for ROKU, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 77 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 10 to 204 trades, against 31 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 15 trades for this stock.
The tail index of the residual cannot be estimated for every 10th trade, on 2 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
The largest Huber residual on 16 October 2024 occurs at 09:35, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:52.

Diagnostics.

TEAM 1 to 16 October 2024

trades on 16 October 2024	25,426
average trading rate	1.5 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,235 (4.86%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.82%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	95 s [81, 112]
Mandelbrot a minimizing QLIKE	22 trades
a within 0.0005 of the minimum	17 to 29
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	25% $\rightarrow$ 56%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	23% $\rightarrow$ 57%
median spot volatility, Huber / linear input[5]	17.0 / 19.6 %
hyperbolic against HExp, correlation of logs	0.9966
HExp QLIKE advantage, leave-one-day-out[6]	+0.03943
15-minute signature level	27 to 73 %
Huber signature, 1 min over 1 s, 12 day average	1.68

`TEAM` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:50; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.056	0.845	0.868	3.824	2.645	1.482	1.558	9.064
2	1.035	0.948	1.001	3.190	2.391	1.604	1.644	6.676
5	1.189	1.079	1.137	2.438	2.249	1.733	1.767	4.619
10	1.265	1.157	1.194	2.013	2.091	1.777	1.814	3.500
25	1.328	1.251	1.258	1.687	1.992	1.865	1.877	2.667
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.18	0.17	0.11	−0.46	2.71	2.34	3.17	3.44
2	0.03	0.04	0.07	0.00	2.88	2.41	3.86	4.32
5	−0.06	−0.02	−0.01	−0.02	3.12	2.79	3.62	4.10
10	−0.01	0.00	0.00	0.01	3.46	3.63	4.89	3.89
25	0.00	0.00	−0.01	−0.02	4.96	5.21	6.40	4.15

Table 1: Trade-lag differences, `TEAM`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	5.22	4.73	4.76	8.58	24.8	16.3	17.9	106.3
1 s	3.62	3.33	3.35	5.56	23.2	16.4	17.9	93.5
5 s	2.74	2.55	2.58	4.03	24.1	20.8	21.5	68.9
10 s	2.38	2.28	2.31	3.31	24.3	22.5	23.0	56.0
1 min	1.87	1.82	1.81	2.19	26.7	26.4	25.9	35.8
15 min	2.25	2.24	2.25	2.38	35.0	35.0	34.9	37.6
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.02	0.05	0.06	−0.10	5.16	6.72	6.46	5.13
1 s	−0.01	0.05	0.07	−0.19	6.32	8.56	7.72	5.53
5 s	−0.04	0.02	0.00	−0.29	6.45	10.11	8.50	5.84
10 s	−0.05	−0.02	−0.03	−0.35	7.91	6.21	7.99	8.14
1 min	0.02	0.03	0.06	−0.09				

Table 2: Clock-lag differences, `TEAM`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.61	3.18	3.33
s.d. of $\hat{\epsilon}_t$ in cents	6.50	6.23	5.98
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.16
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.24	0.24	0.12
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.13	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.02	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.02	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−2.25,2.09	−2.50,2.27
95% C.I. for $\theta_t - X_t$	—	−4.35,3.83	−4.85,4.39
99% C.I. for $\theta_t - X_t$	—	−7.10,6.24	−8.20,7.78
s.e. of $\theta_t - X_t$ in cents	—	1.95	2.31

Table 3: Residual summaries, `TEAM`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.16, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.44	2.80	2.05 to 3.21	2.05 to 3.33
Residual divided by δ_t	2.44	3.34	1.90 to 3.14	2.48 to 4.38
Repricing windows removed	2.42	2.88	2.12 to 3.19	2.03 to 3.33
Calm windows only	2.45	2.78		
Every 10th trade	—	—	1.73 to 2.66	1.88 to 3.03
Six-trade half-life	2.53	3.08	2.14 to 3.76	2.21 to 3.66

Table 4: Hill estimator of the tail index of the Huber filter residual, `TEAM`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 11 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 40.4 to 60.4 cents over the days and removes 0% to 4% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 30.2 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

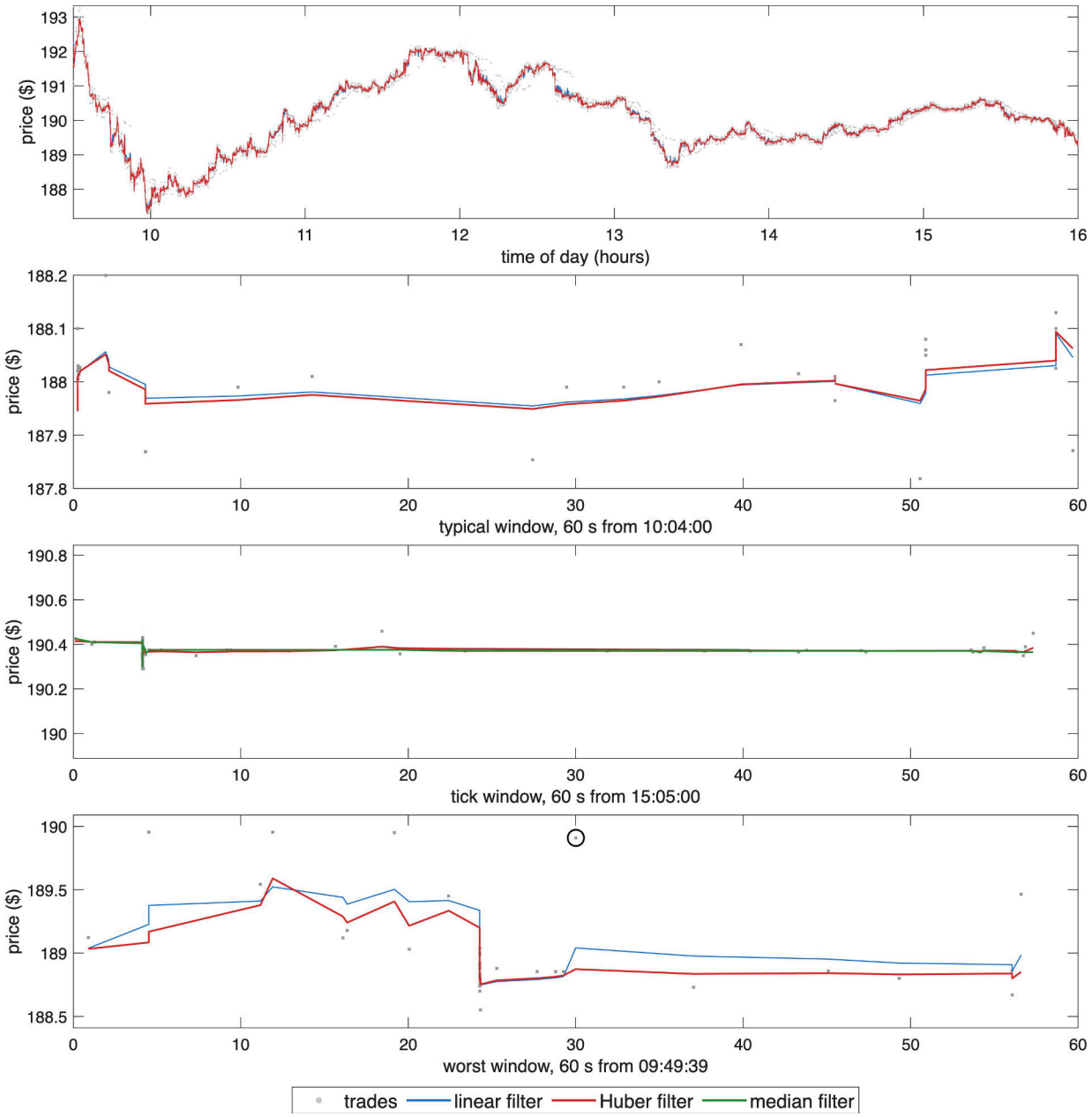

Figure 1: TEAM trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:04:00. 3rd: 60-second window at 15:05:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

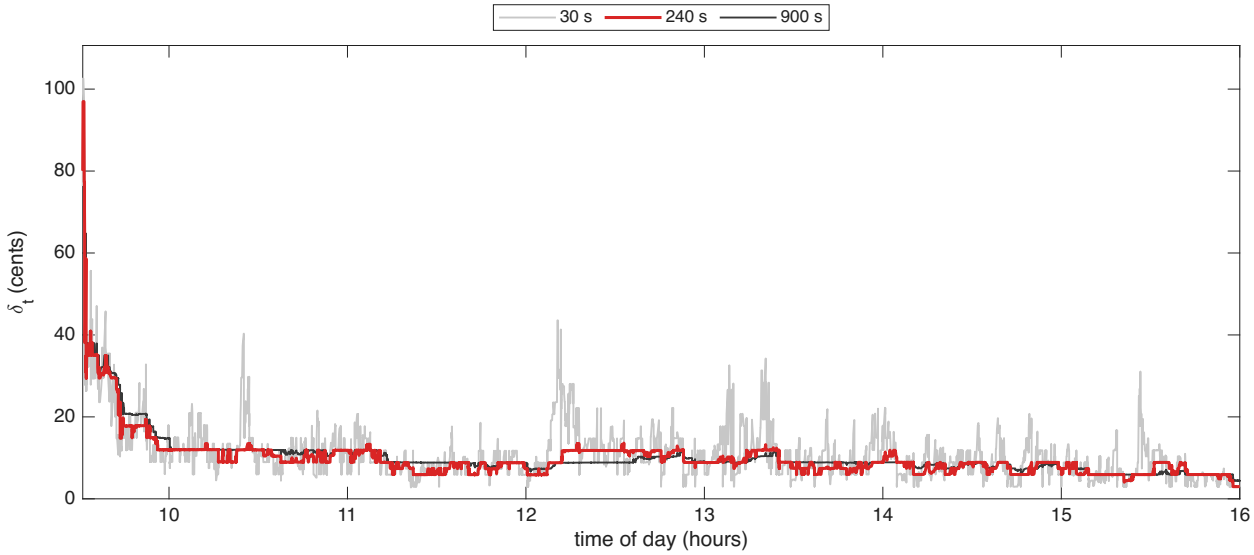

Figure 2: The Huber threshold δ_t in cents for `TEAM` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

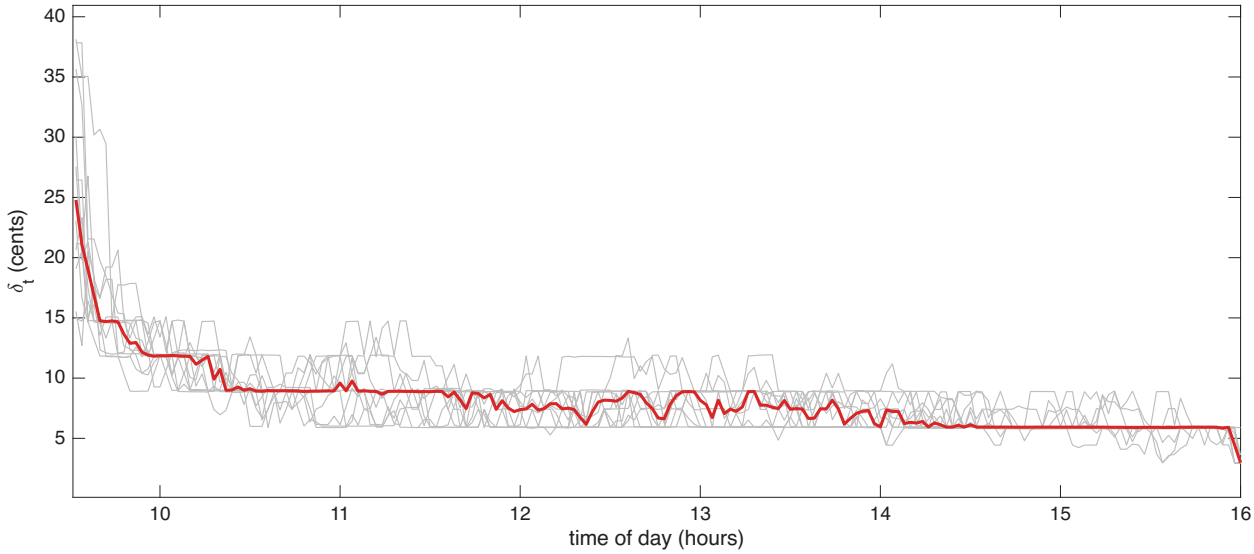

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `TEAM` trading days in Oct 2024 (grey lines) and their median (red line).

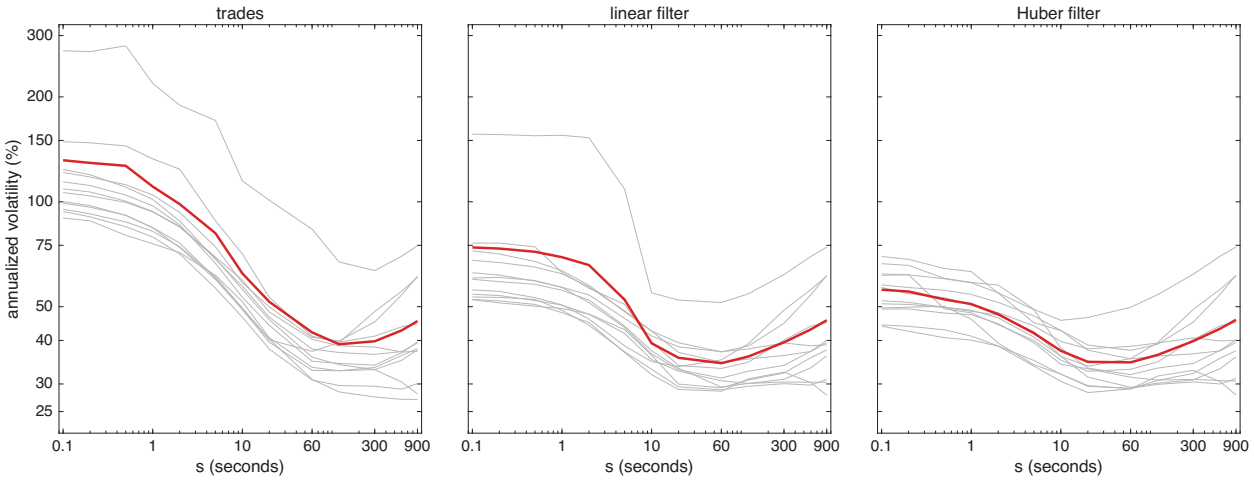

Figure 4: Volatility signature curves for `TEAM` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

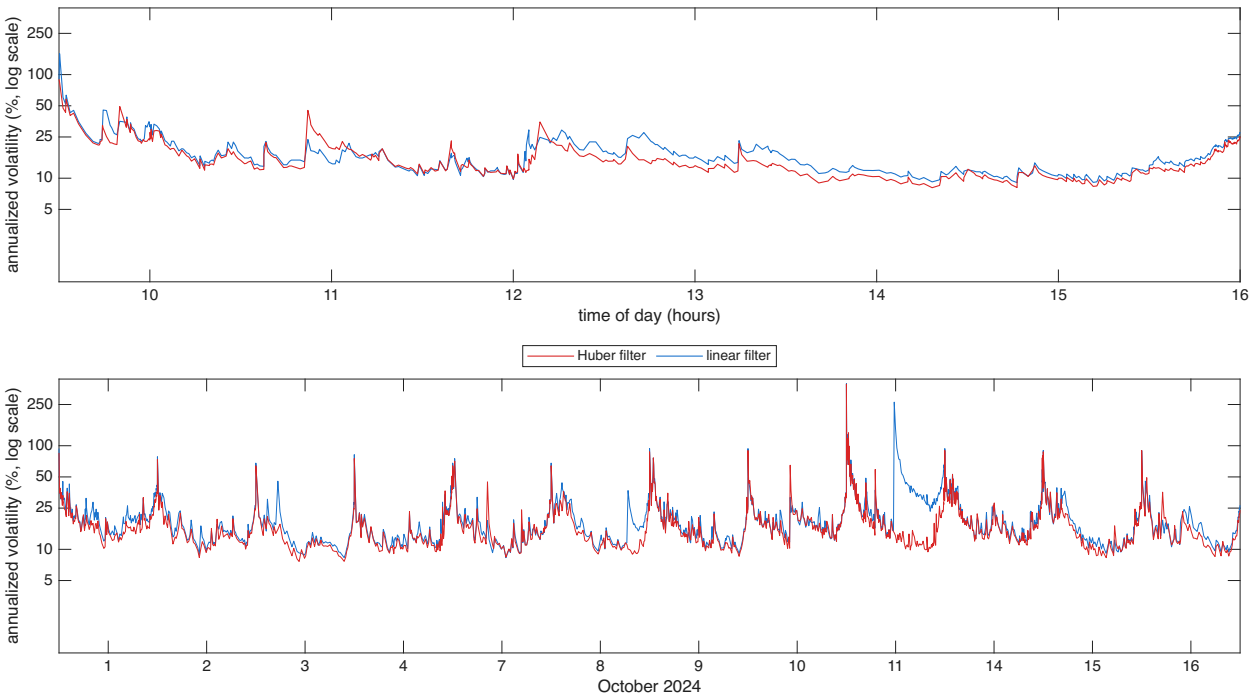

Figure 5: Filtered spot volatility for `TEAM` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 22$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

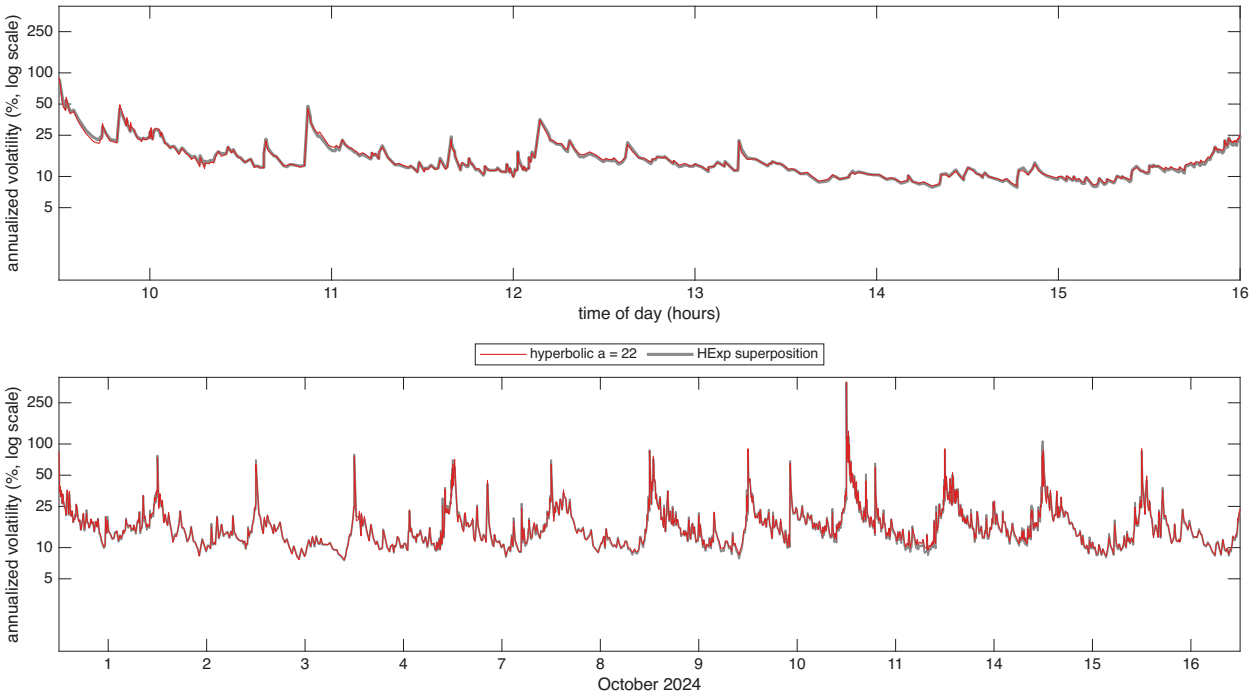

Figure 6: Hyperbolic against superposition weights for `TEAM` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 22$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 30, 150, 750, 3750 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

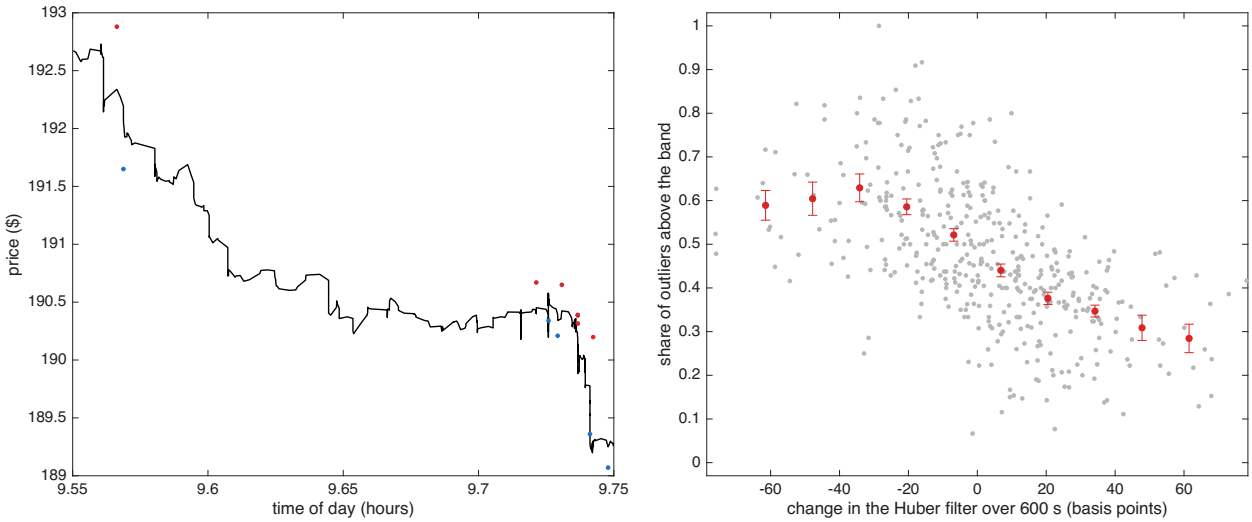

Figure 7: Outliers & robust price for `TEAM` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

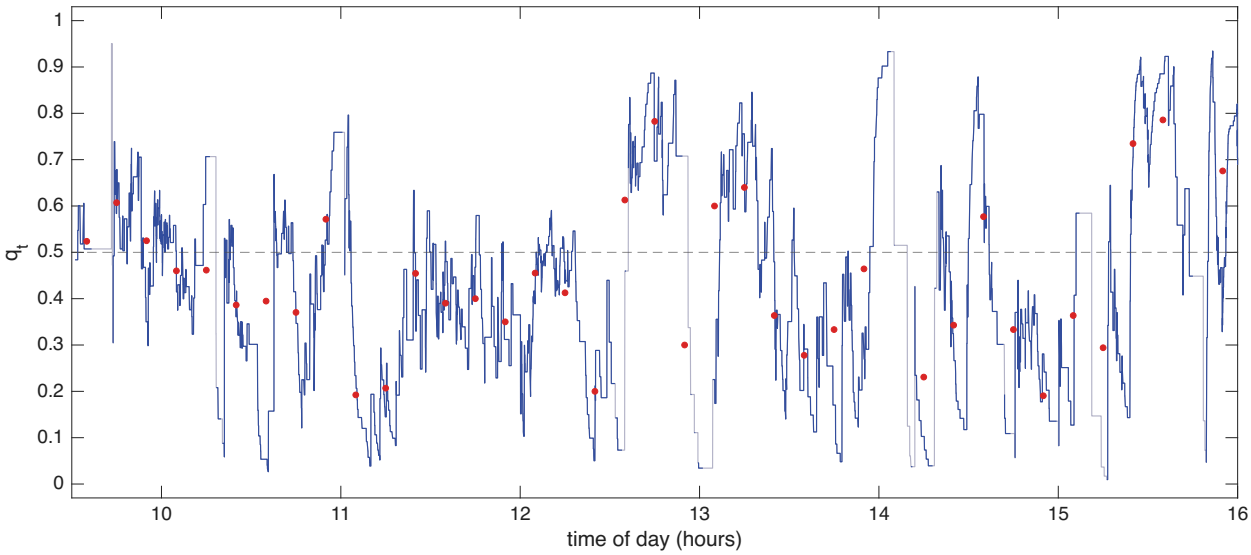

Figure 8: The directional conditional probability for `TEAM`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 95 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 9 to 86 trades, against 22 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 15 trades for this stock.
The tail index of the residual cannot be estimated for every 10th trade, on 4 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:50.

Diagnostics.

INTU 1 to 16 October 2024

trades on 16 October 2024	31,146
average trading rate	1.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,599 (5.13%)
price half-life: run / minimizer[1]	3.0 / 2.8 trades
cost of running 3 trades	+0.29%
scale half-life: run / minimizer[2]	240 / 180 s
cost of running 240 s	+0.04%
direction half-life[3]	87 s [70, 112]
Mandelbrot a minimizing QLIKE	31 trades
a within 0.0005 of the minimum	21 to 45
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	8% $\rightarrow$ 71%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	7% $\rightarrow$ 13%
median spot volatility, Huber / linear input[5]	11.6 / 12.9 %
hyperbolic against HExp, correlation of logs	0.9990
HExp QLIKE advantage, leave-one-day-out[6]	+0.01647
15-minute signature level	14 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.33

INTU over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:14; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.215	1.861	1.856	8.667	6.098	3.263	3.467	20.470
2	2.140	2.050	2.102	7.072	5.514	3.487	3.623	15.119
5	2.425	2.251	2.318	5.276	5.084	3.715	3.771	10.284
10	2.514	2.321	2.378	4.260	4.454	3.624	3.684	7.683
25	2.512	2.356	2.372	3.376	3.893	3.485	3.523	5.511
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.18	0.14	0.09	−0.45	2.46	2.83	2.67	2.55
2	0.05	0.06	0.04	0.00	2.42	2.90	2.95	2.83
5	−0.04	−0.04	−0.02	0.00	2.62	2.87	2.58	2.78
10	0.00	−0.01	−0.02	−0.01	3.28	3.77	3.32	2.85
25	−0.01	−0.02	0.00	−0.01	3.76	4.75	4.42	2.96

Table 1: Trade-lag differences, `INTU`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	11.56	10.71	10.66	20.59	19.0	12.3	14.3	84.5
1 s	8.38	8.02	8.02	13.76	18.1	12.7	14.4	73.7
5 s	6.46	6.29	6.35	9.83	18.1	14.6	15.4	55.8
10 s	5.80	5.68	5.75	8.59	17.8	15.9	16.7	49.8
1 min	5.08	5.06	5.05	6.31	18.0	17.7	17.8	24.6
15 min	6.07	6.12	6.12	6.82	29.0	29.0	28.9	29.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.02	0.02	0.02	−0.06	5.74	5.55	5.00	3.19
1 s	−0.04	0.02	0.04	−0.17	4.51	5.40	5.13	2.70
5 s	−0.11	−0.06	0.01	−0.20	4.06	5.29	5.35	2.38
10 s	0.03	0.12	−0.02	−0.66	4.47	5.08	4.01	4.14
1 min	0.12	0.14	0.14	−0.13				

Table 2: Clock-lag differences, `INTU`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	5.72	6.89	7.13
s.d. of $\hat{\epsilon}_t$ in cents	14.71	13.94	13.33
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.67
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.19	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.08	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	−0.02	0.00
80% C.I. for $\theta_t - X_t$	—	−4.45,4.28	−4.76,4.55
95% C.I. for $\theta_t - X_t$	—	−8.43,7.90	−9.18,8.64
99% C.I. for $\theta_t - X_t$	—	−16.16,13.53	−17.87,15.62
s.e. of $\theta_t - X_t$ in cents	—	4.12	4.79

Table 3: Residual summaries, `INTU`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.67, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.35	2.90	2.35 to 3.56	2.35 to 3.63
Residual divided by δ_t	2.79	2.96	2.79 to 4.35	2.28 to 4.45
Repricing windows removed	2.40	2.88	2.40 to 3.68	2.61 to 3.65
Calm windows only	2.31	3.09		
Every 10th trade	1.73	2.77	1.73 to 3.78	2.09 to 3.39
Six-trade half-life	2.41	2.92	2.41 to 3.57	2.29 to 3.92

Table 4: Hill estimator of the tail index of the Huber filter residual, `INTU`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 105.5 to 171.4 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 60.7 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

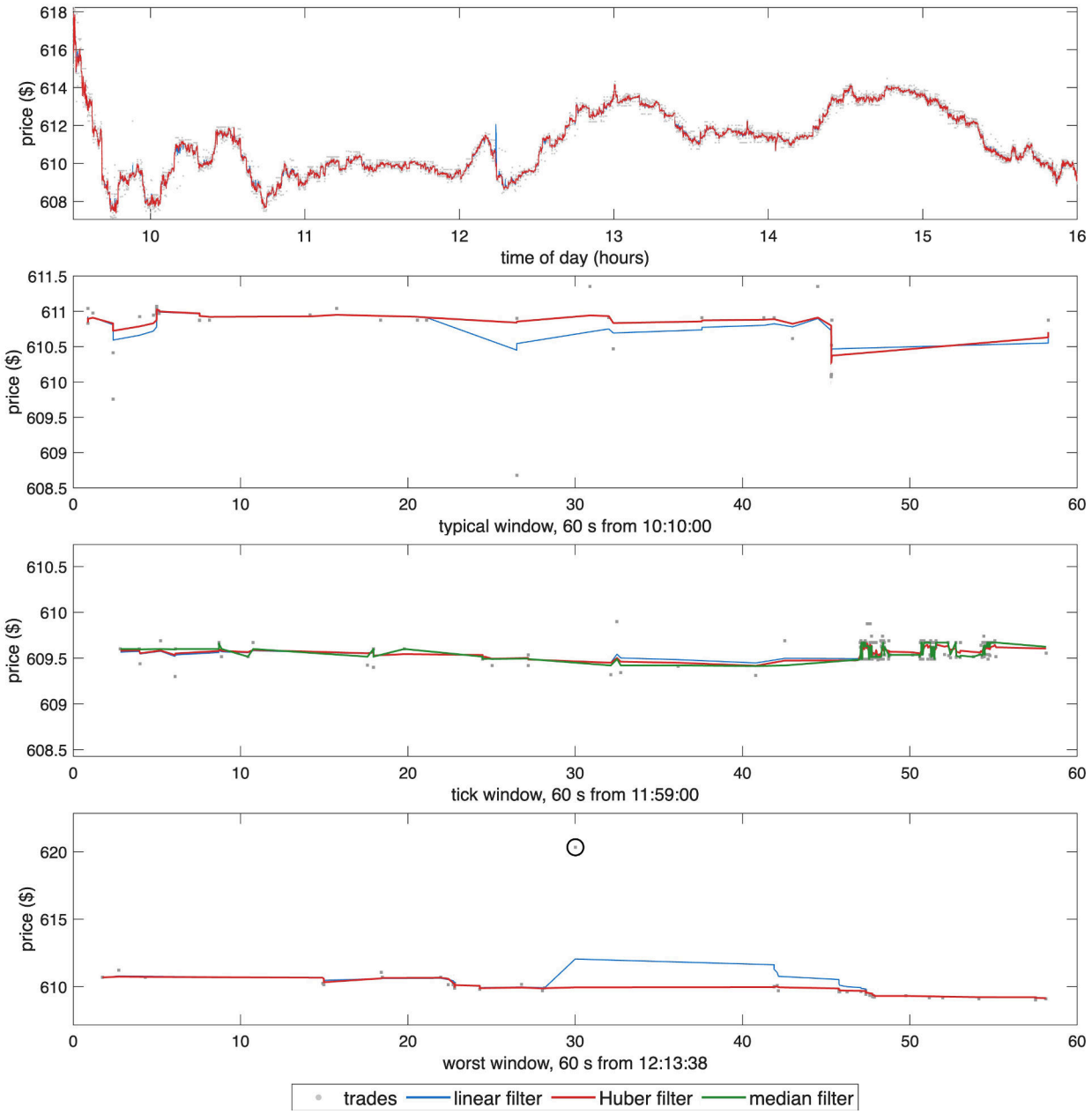

Figure 1: `INTU` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:10:00. 3rd: 60-second window at 11:59:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

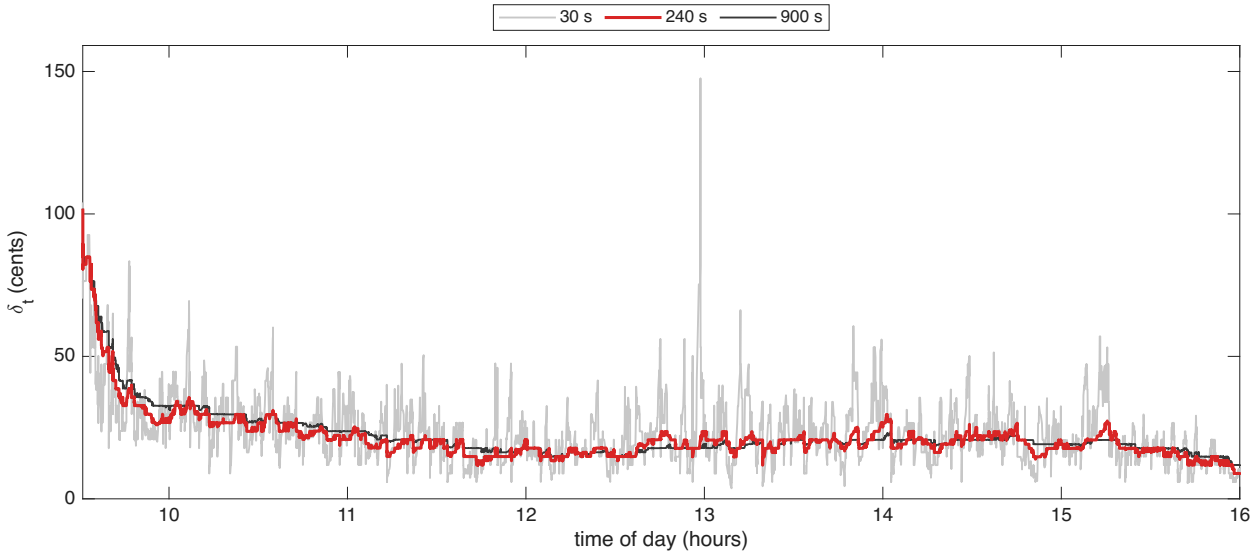

Figure 2: The Huber threshold δ_t in cents for `INTU` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

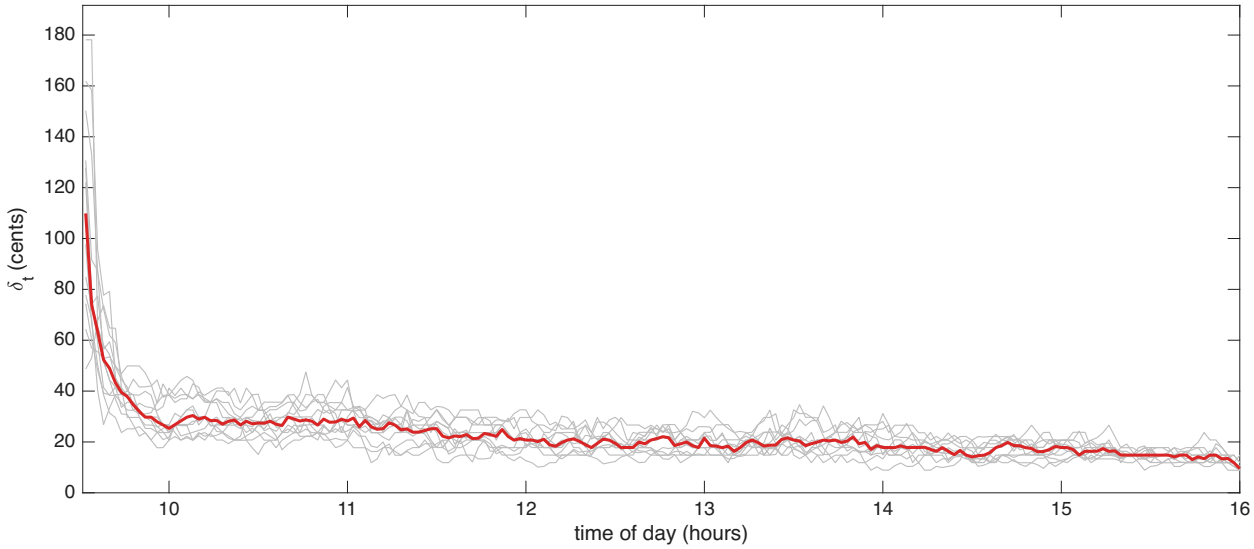

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `INTU` trading days in Oct 2024 (grey lines) and their median (red line).

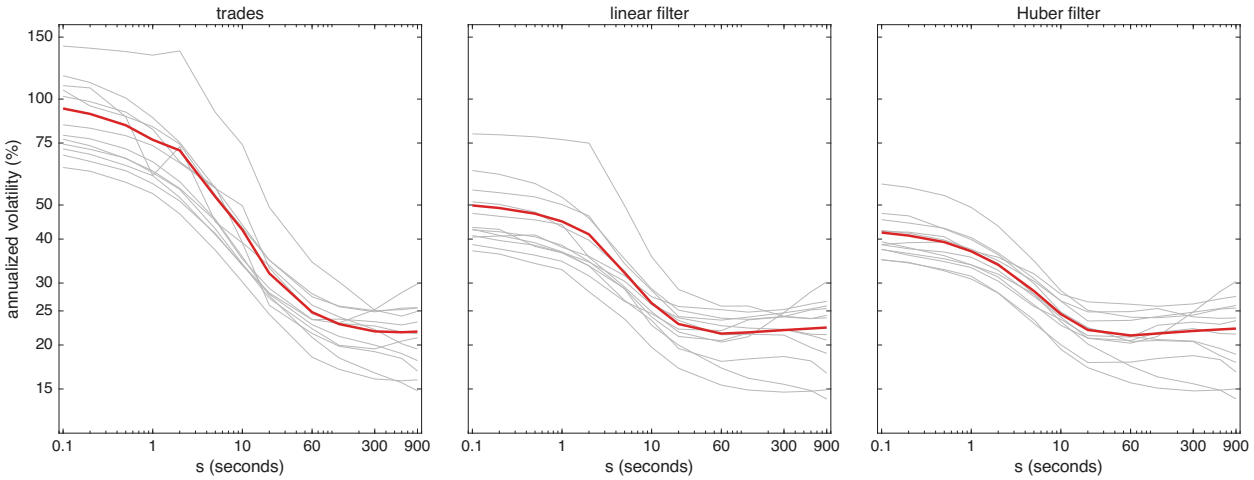

Figure 4: Volatility signature curves for `INTU` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

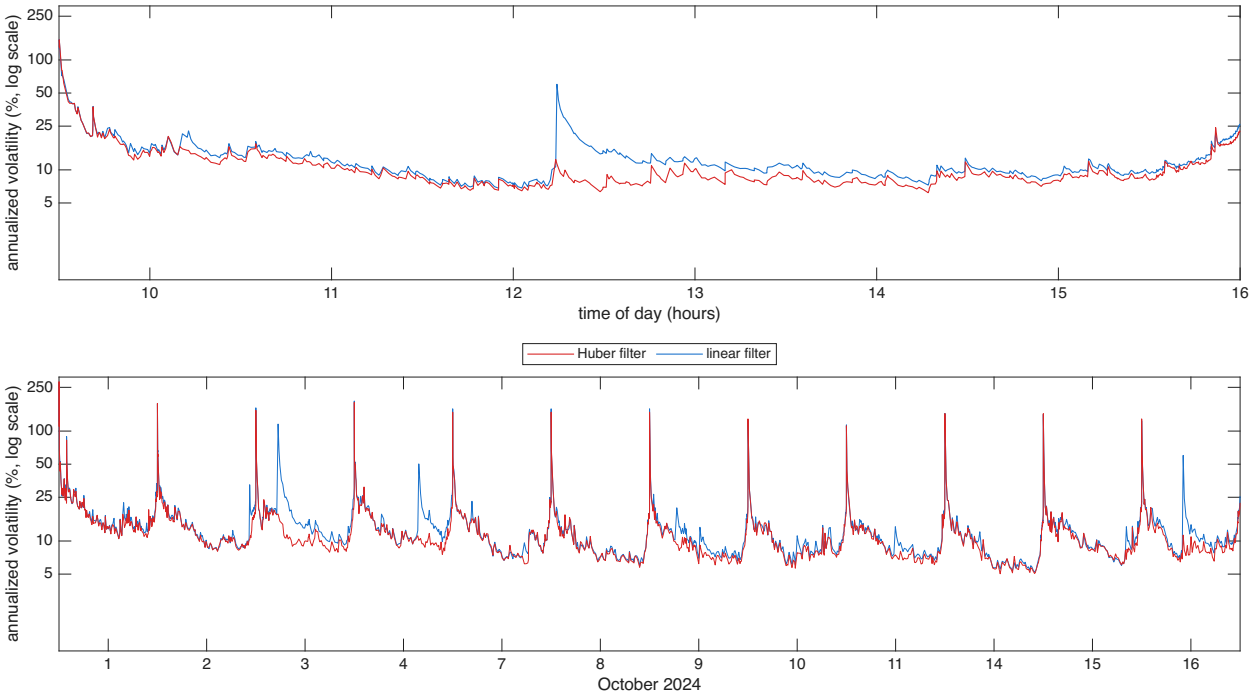

Figure 5: Filtered spot volatility for `INTU` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 31$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

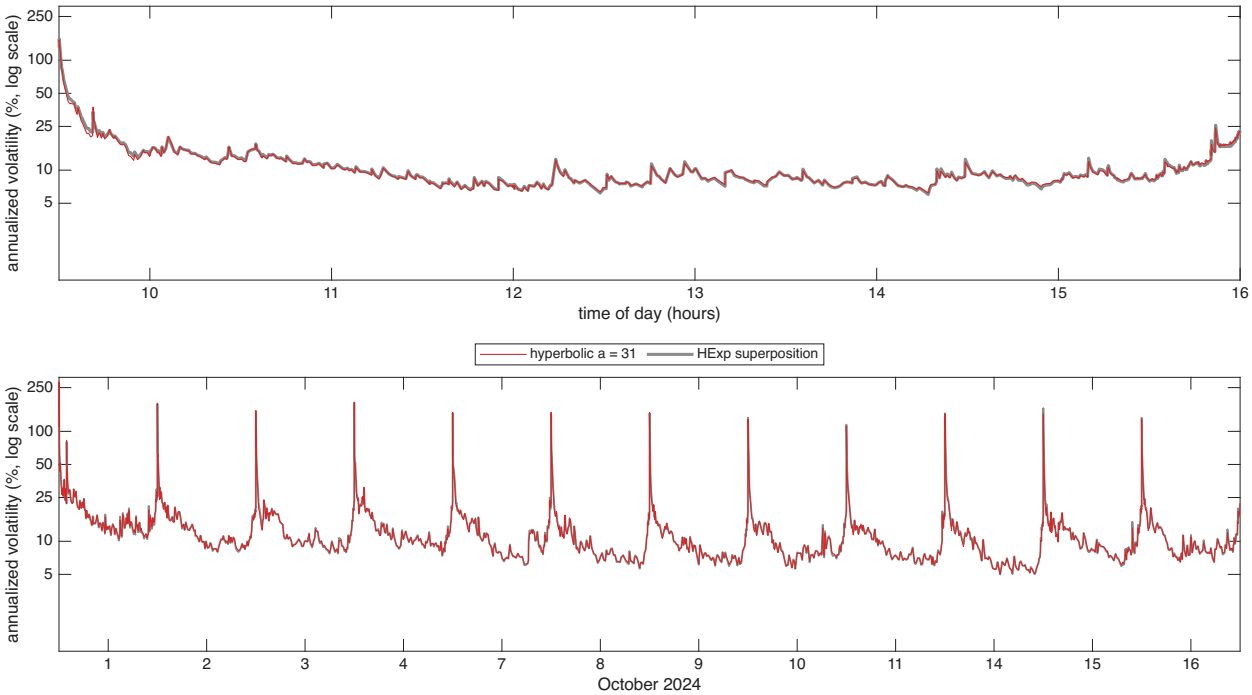

Figure 6: Hyperbolic against superposition weights for `INTU` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 31$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

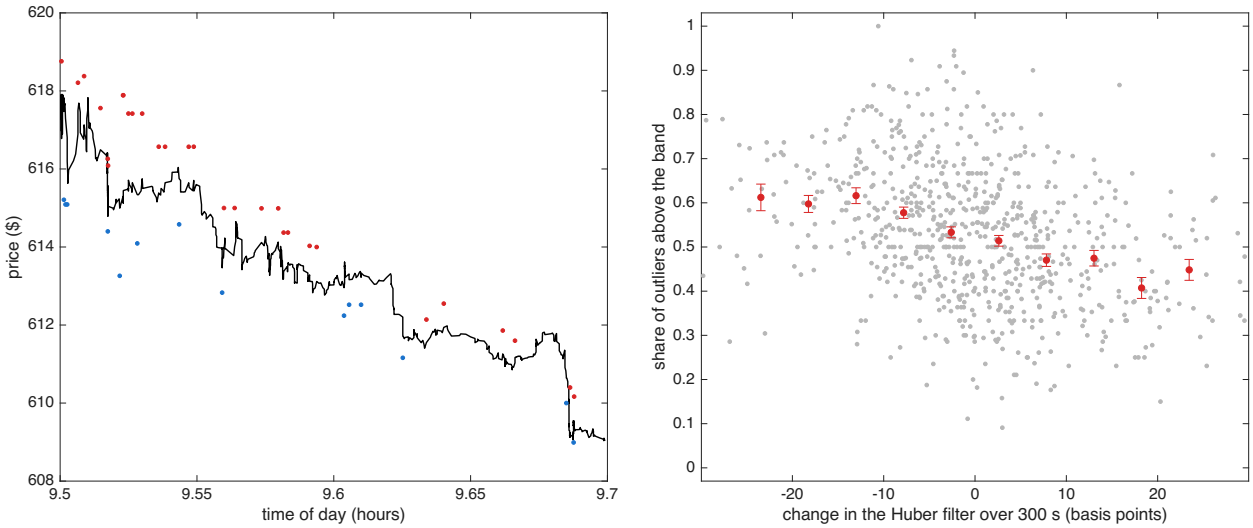

Figure 7: Outliers & robust price for `INTU` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 300 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 300 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

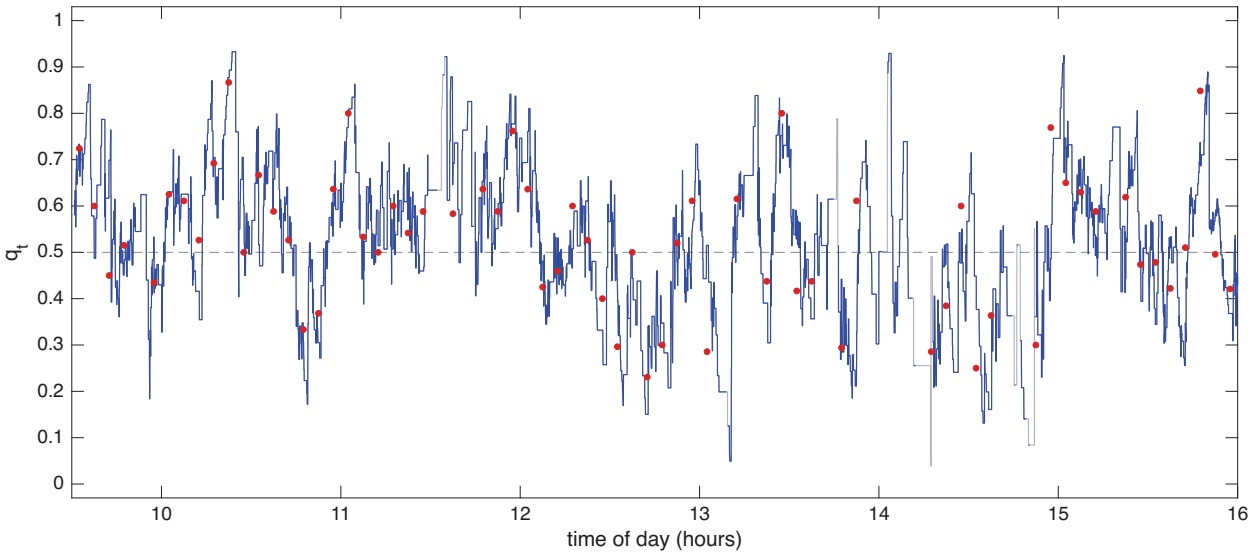

Figure 8: The directional conditional probability for INTU, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 87 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 300 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

On 14 October 2024, 43 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 35 per cent on any other day. Every number that day contributes to carries that burst.

The tail index of the residual cannot be estimated for every 10th trade, on 3 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

AMGN 1 to 16 October 2024

trades on 16 October 2024	34,042
average trading rate	1.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,415 (4.16%)
price half-life: run / minimizer[1]	3.0 / 2.8 trades
cost of running 3 trades	+0.33%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	88 s [73, 111]
Mandelbrot a minimizing QLIKE	26 trades
a within 0.0005 of the minimum	18 to 39
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	9% $\to$ 37%
Huber $\hat{\sigma}$, before $\to$ peak	9% $\to$ 11%
median spot volatility, Huber / linear input[5]	11.1 / 11.3 %
hyperbolic against HExp, correlation of logs	0.9990
HExp QLIKE advantage, leave-one-day-out[6]	+0.01326
15-minute signature level	14 to 23 %
Huber signature, 1 min over 1 s, 12 day average	1.25

AMGN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 12:56; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.074	0.870	0.865	4.074	2.677	1.424	1.471	8.572
2	1.015	0.949	0.967	3.318	2.351	1.511	1.551	6.417
5	1.127	1.024	1.048	2.450	2.156	1.606	1.623	4.353
10	1.146	1.041	1.056	1.950	1.924	1.588	1.602	3.273
25	1.106	1.025	1.031	1.511	1.657	1.493	1.502	2.330
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.13	0.11	−0.44	3.05	2.47	2.82	2.98
2	0.07	0.05	0.03	−0.03	3.10	2.71	2.60	2.81
5	0.00	−0.01	−0.02	−0.01	2.86	2.77	2.62	2.75
10	−0.01	−0.01	−0.02	0.00	3.00	3.08	3.00	2.68
25	0.01	0.01	0.01	0.02	3.55	3.37	3.48	3.05

Table 1: Trade-lag differences, `AMGN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	4.90	4.37	4.36	8.95	17.0	10.5	11.6	68.4
1 s	3.45	3.21	3.24	5.89	15.0	10.4	11.4	54.3
5 s	2.53	2.37	2.40	3.79	15.0	12.3	12.8	36.1
10 s	2.29	2.16	2.18	3.15	14.4	12.9	13.3	28.8
1 min	1.73	1.67	1.68	2.06	13.7	13.3	13.4	17.9
15 min	2.05	2.04	2.06	2.14	15.5	15.5	15.6	17.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.07	0.01	0.03	−0.09	3.30	4.04	4.32	3.54
1 s	−0.06	−0.01	−0.02	−0.31	3.69	4.21	4.35	3.08
5 s	−0.11	−0.05	−0.06	−0.37	3.80	4.40	4.65	3.32
10 s	−0.08	−0.03	−0.05	−0.38	4.10	5.01	5.16	2.67
1 min	0.01	0.04	0.04	−0.12				

Table 2: Clock-lag differences, `AMGN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.67	3.22	3.32
s.d. of $\hat{\epsilon}_t$ in cents	6.14	5.79	5.64
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.84
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.17	0.11
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.05	0.06	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.00	0.00	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.00	0.01
80% C.I. for $\theta_t - X_t$	—	−2.06,2.00	−2.19,2.10
95% C.I. for $\theta_t - X_t$	—	−3.64,3.68	−3.91,3.92
99% C.I. for $\theta_t - X_t$	—	−6.19,5.91	−6.66,6.52
s.e. of $\theta_t - X_t$ in cents	—	1.82	2.07

Table 3: Residual summaries, `AMGN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.84, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.68	2.44	2.18 to 3.81	2.44 to 4.10
Residual divided by δ_t	2.92	2.63	2.87 to 4.29	2.15 to 3.49
Repricing windows removed	3.80	2.79	2.60 to 3.91	2.39 to 4.05
Calm windows only	4.01	2.98		
Every 10th trade	3.63	3.48	2.17 to 3.63	1.75 to 3.64
Six-trade half-life	4.10	2.37	2.15 to 4.10	2.37 to 4.38

Table 4: Hill estimator of the tail index of the Huber filter residual, `AMGN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 60.5 to 81.3 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 30.2 cents on 16 October 2024, 97% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

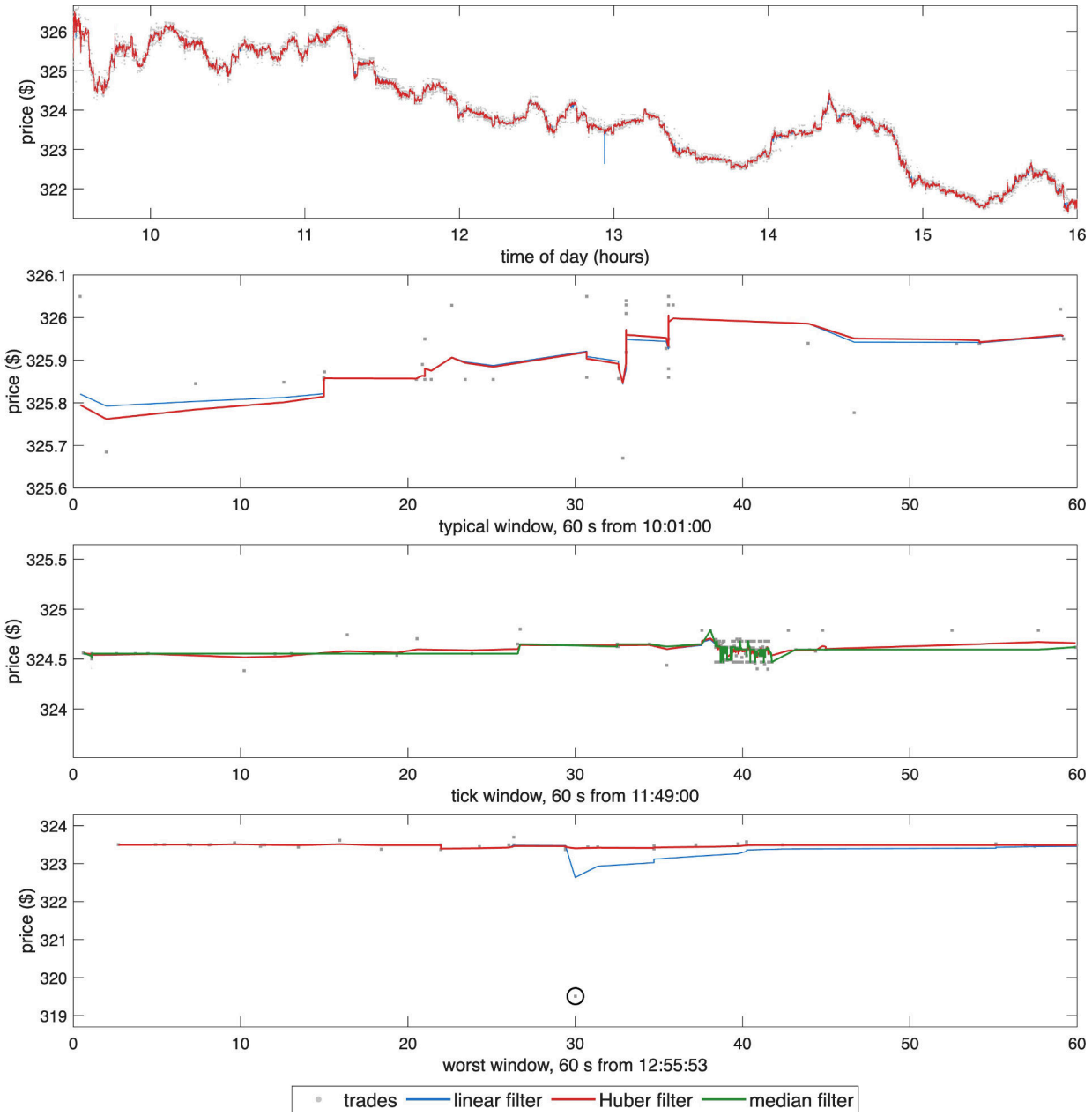

Figure 1: `AMGN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:01:00. 3rd: 60-second window at 11:49:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

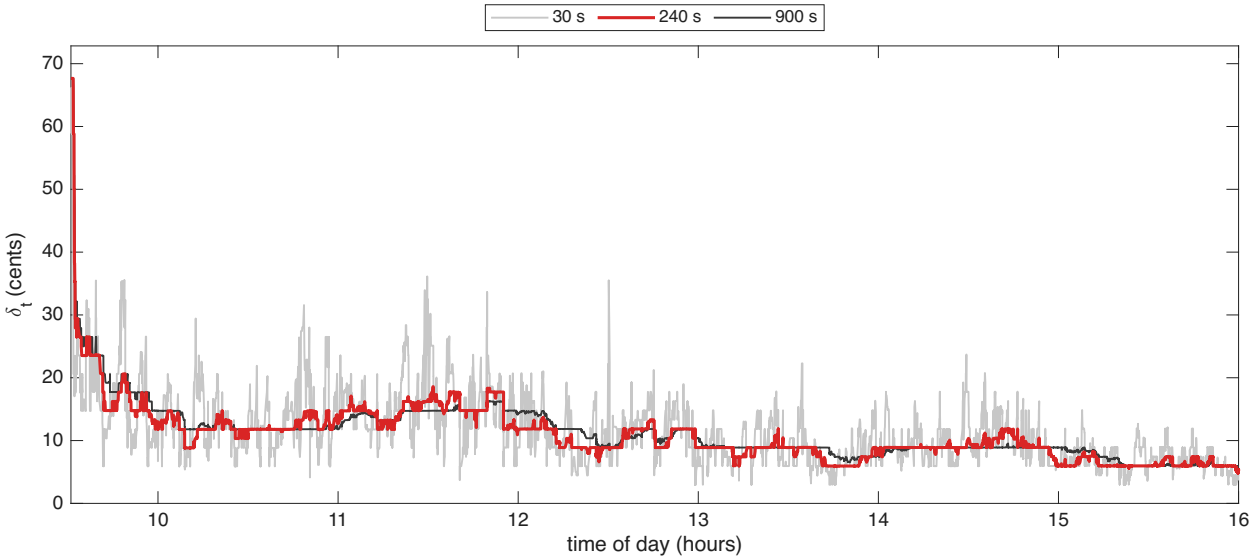

Figure 2: The Huber threshold δ_t in cents for `AMGN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

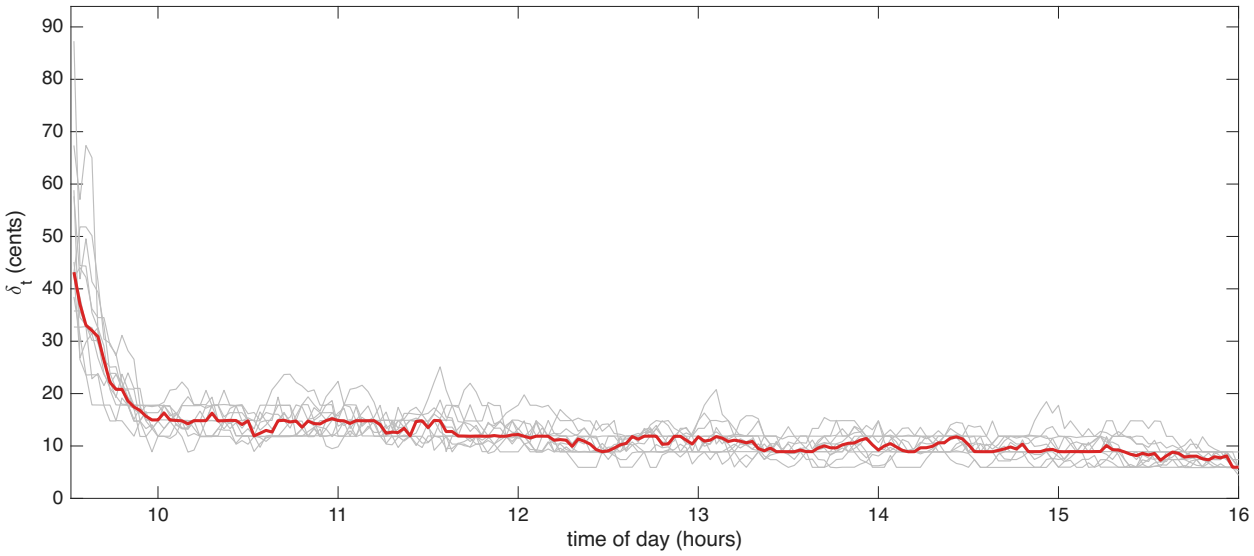

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `AMGN` trading days in Oct 2024 (grey lines) and their median (red line).

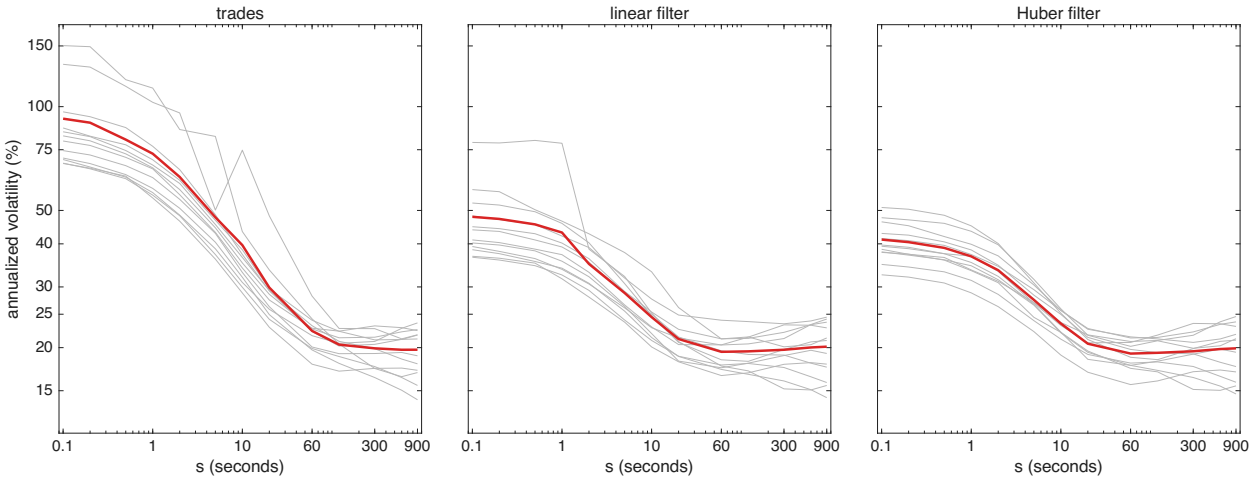

Figure 4: Volatility signature curves for `AMGN` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

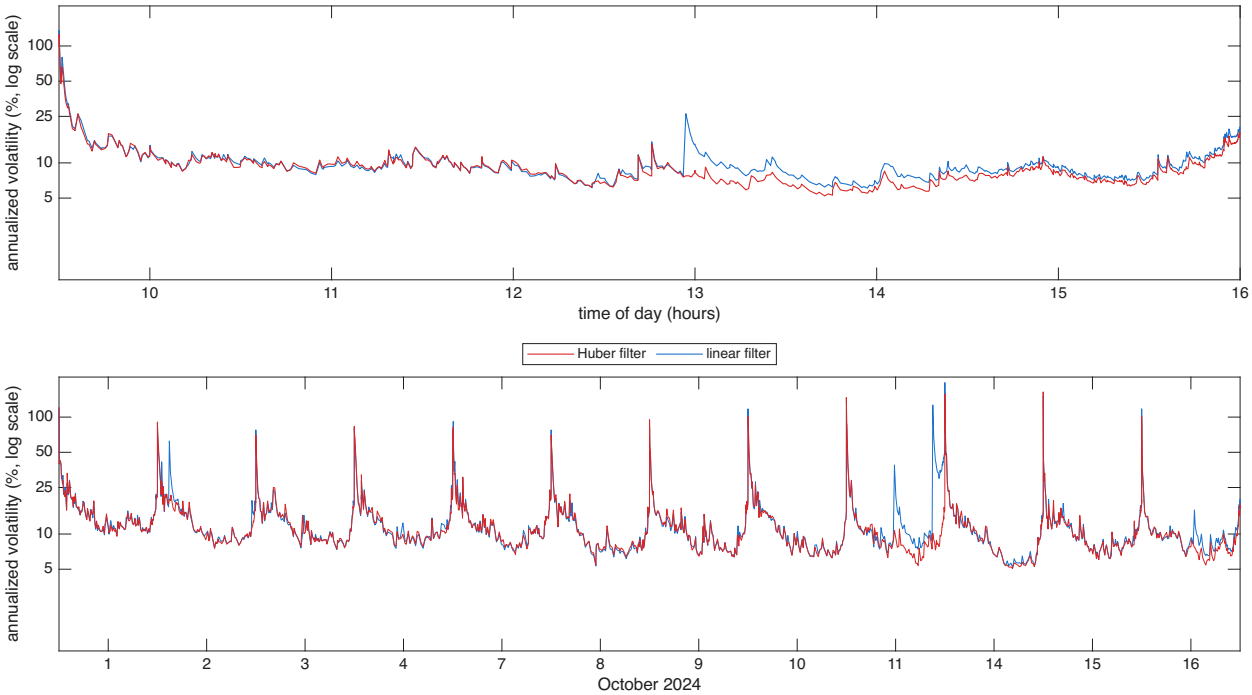

Figure 5: Filtered spot volatility for `AMGN` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 26$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

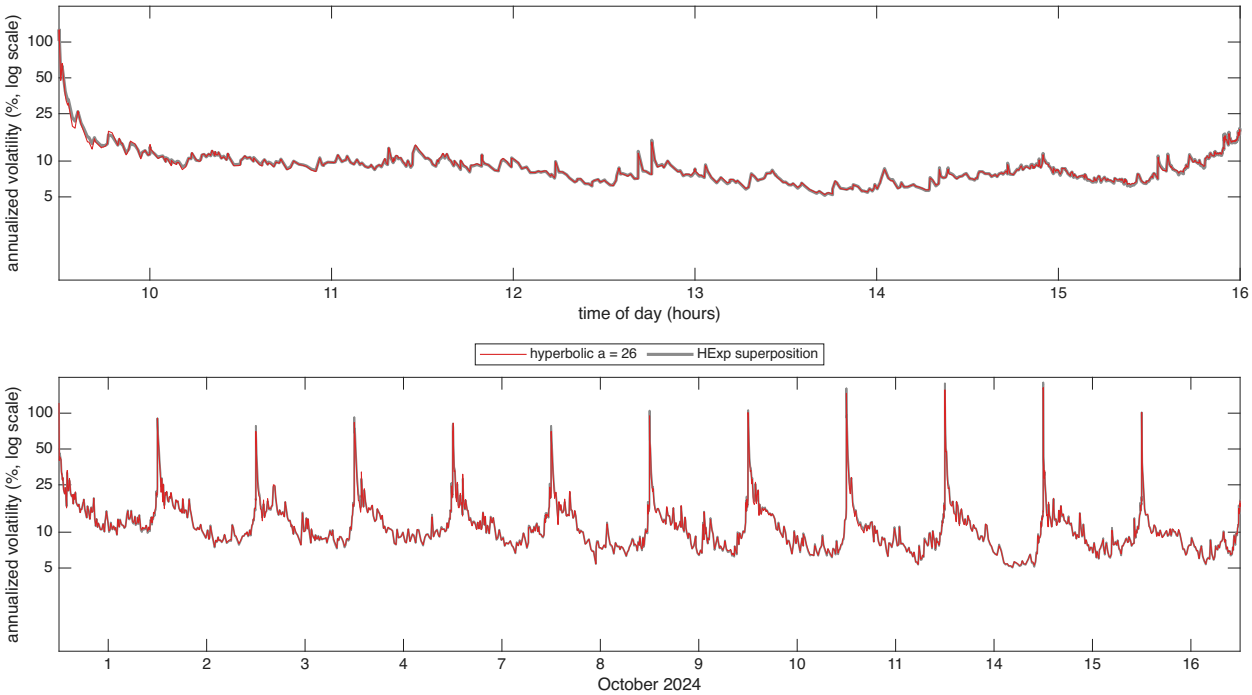

Figure 6: Hyperbolic against superposition weights for `AMGN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 26$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

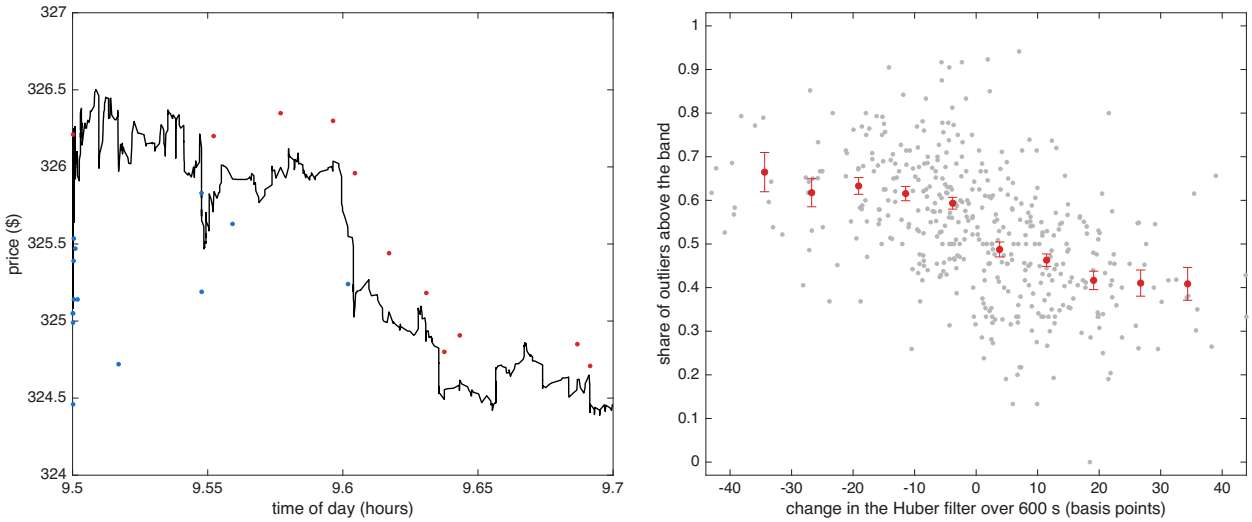

Figure 7: Outliers & robust price for `AMGN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

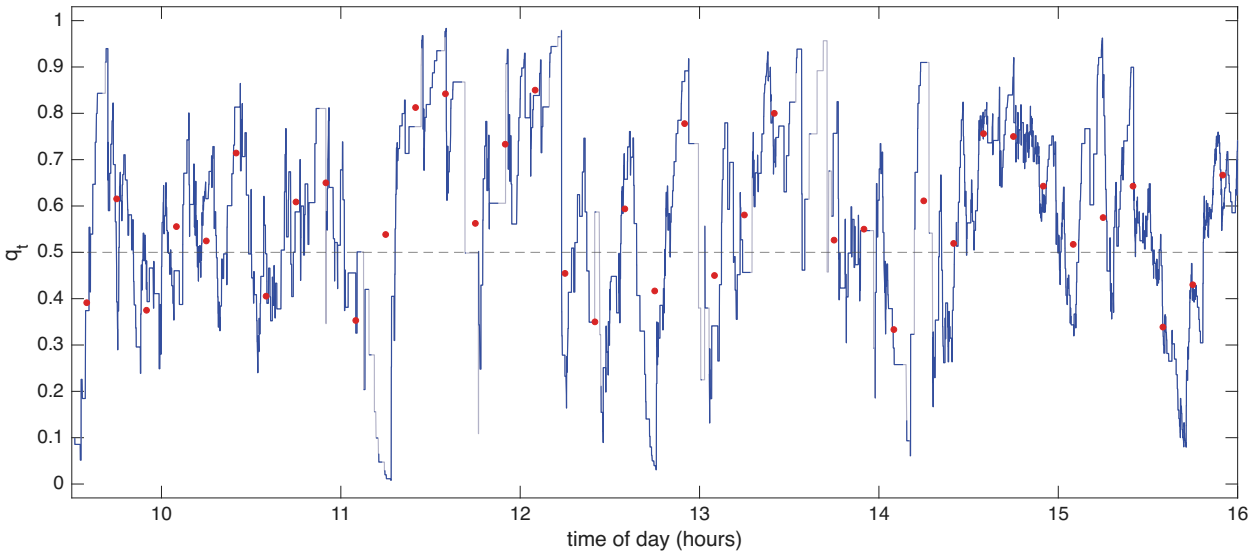

Figure 8: The directional conditional probability for `AMGN`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 88 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

On 14 October 2024, 28 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 22 per cent on any other day. Every number that day contributes to carries that burst.

The tail index of the residual cannot be estimated for every 10th trade, on 4 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

ISRG 1 to 16 October 2024

trades on 16 October 2024	28,106
average trading rate	1.3 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,346 (4.79%)
price half-life: run / minimizer[1]	3.0 / 3.1 trades
cost of running 3 trades	+0.07%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	84 s [71, 101]
Mandelbrot a minimizing QLIKE	28 trades
a within 0.0005 of the minimum	19 to 41
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	15% $\rightarrow$ 38%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	15% $\rightarrow$ 21%
median spot volatility, Huber / linear input[5]	12.0 / 12.6 %
hyperbolic against HExp, correlation of logs	0.9980
HExp QLIKE advantage, leave-one-day-out[6]	+0.02372
15-minute signature level	11 to 26 %
Huber signature, 1 min over 1 s, 12 day average	1.19

ISRG over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:05; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.020	1.554	1.533	7.237	4.550	2.407	2.546	14.793
2	1.877	1.682	1.705	5.862	3.991	2.563	2.693	11.136
5	2.072	1.817	1.851	4.345	3.653	2.723	2.818	7.541
10	2.054	1.832	1.858	3.433	3.258	2.711	2.788	5.661
25	1.921	1.780	1.791	2.630	2.828	2.602	2.643	4.097
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.13	0.12	−0.43	3.05	4.03	2.66	2.65
2	0.06	0.05	0.03	−0.03	3.18	3.99	3.15	2.69
5	−0.03	−0.02	−0.02	0.01	3.55	4.21	4.23	2.73
10	−0.02	−0.01	−0.01	0.01	3.91	4.22	4.29	2.88
25	−0.01	−0.02	−0.02	0.00	3.70	3.88	4.56	3.44

Table 1: Trade-lag differences, `ISRG`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	8.52	7.56	7.60	15.32	19.2	11.6	12.6	73.3
1 s	5.85	5.28	5.28	9.73	17.0	11.3	12.1	60.1
5 s	4.34	3.99	3.97	6.28	16.0	12.8	13.2	41.4
10 s	3.85	3.59	3.60	5.50	16.1	14.0	14.4	32.8
1 min	2.90	2.77	2.77	3.49	14.3	13.8	13.9	19.1
15 min	3.03	2.97	2.96	3.26	19.4	19.3	19.2	20.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.06	−0.03	−0.07	−0.23	3.95	4.34	3.94	3.15
1 s	−0.05	0.00	0.00	−0.24	4.97	4.15	3.76	3.37
5 s	−0.09	−0.03	−0.02	−0.28	4.28	5.32	5.28	3.65
10 s	−0.22	−0.14	−0.14	−0.45	4.38	5.11	5.42	4.27
1 min	0.15	0.18	0.18	−0.05				

Table 2: Clock-lag differences, `ISRG`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	4.70	5.69	5.89
s.d. of $\hat{\epsilon}_t$ in cents	10.53	9.98	9.78
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.66
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.16	0.12
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.07	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.00	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	−0.02	−0.02
80% C.I. for $\theta_t - X_t$	—	−3.50,3.48	−3.69,3.65
95% C.I. for $\theta_t - X_t$	—	−6.36,6.31	−6.85,6.94
99% C.I. for $\theta_t - X_t$	—	−11.19,10.70	−12.97,11.43
s.e. of $\theta_t - X_t$ in cents	—	3.19	3.60

Table 3: Residual summaries, `ISRG`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.66, is above two, while the Hill index of the Huber residual itself on this day is 1.95, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.38	1.95	2.61 to 3.62	1.95 to 4.07
Residual divided by δ_t	4.47	2.35	2.66 to 5.42	2.13 to 4.40
Repricing windows removed	3.35	1.90	2.87 to 3.87	1.90 to 4.30
Calm windows only	3.67	1.90		
Every 10th trade	—	—	2.03 to 3.85	1.74 to 3.46
Six-trade half-life	3.44	2.09	2.55 to 3.80	2.09 to 4.43

Table 4: Hill estimator of the tail index of the Huber filter residual, `ISRG`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 13 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 100.8 to 141.2 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 50.4 cents on 16 October 2024, 95% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

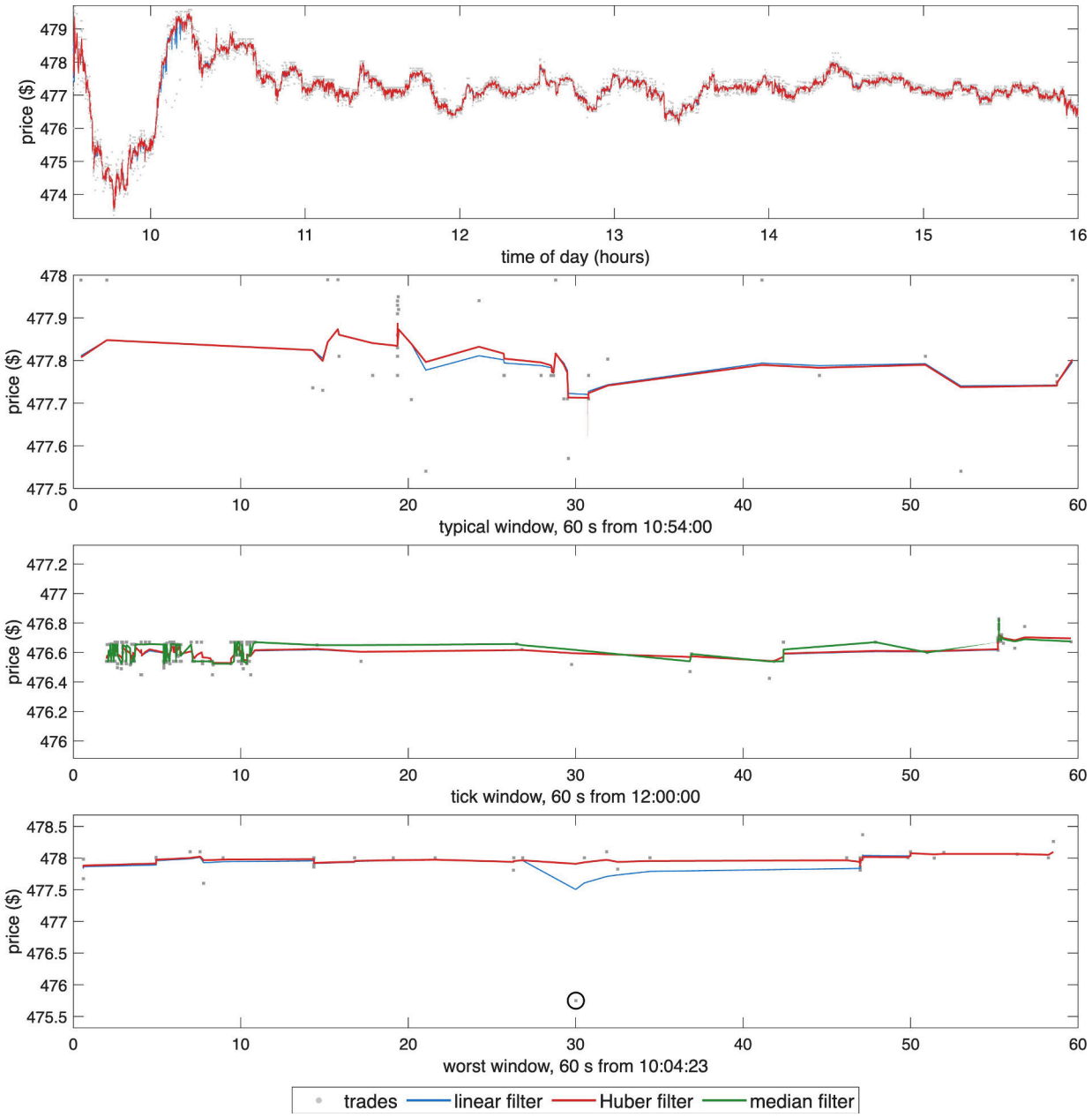

Figure 1: `ISRG` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:54:00. 3rd: 60-second window at 12:00:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

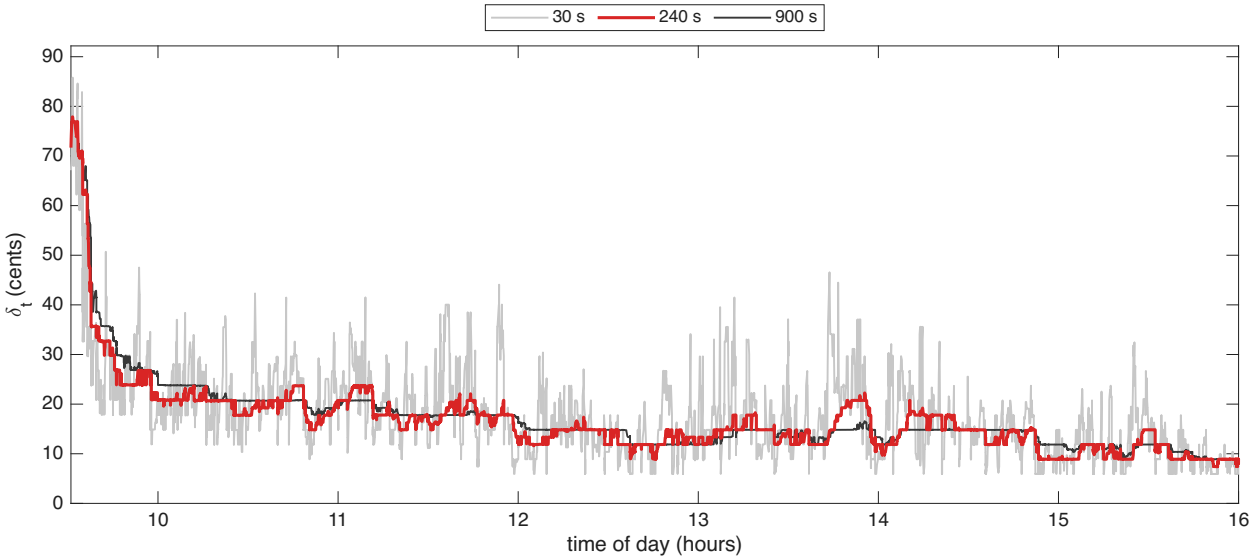

Figure 2: The Huber threshold δ_t in cents for `ISRG` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

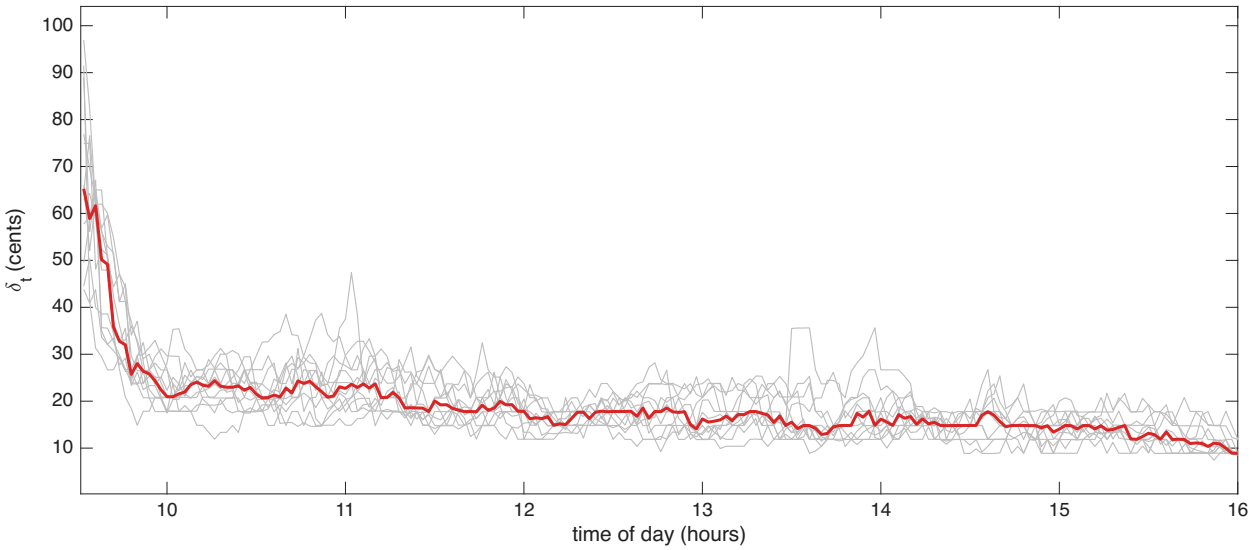

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `ISRG` trading days in Oct 2024 (grey lines) and their median (red line).

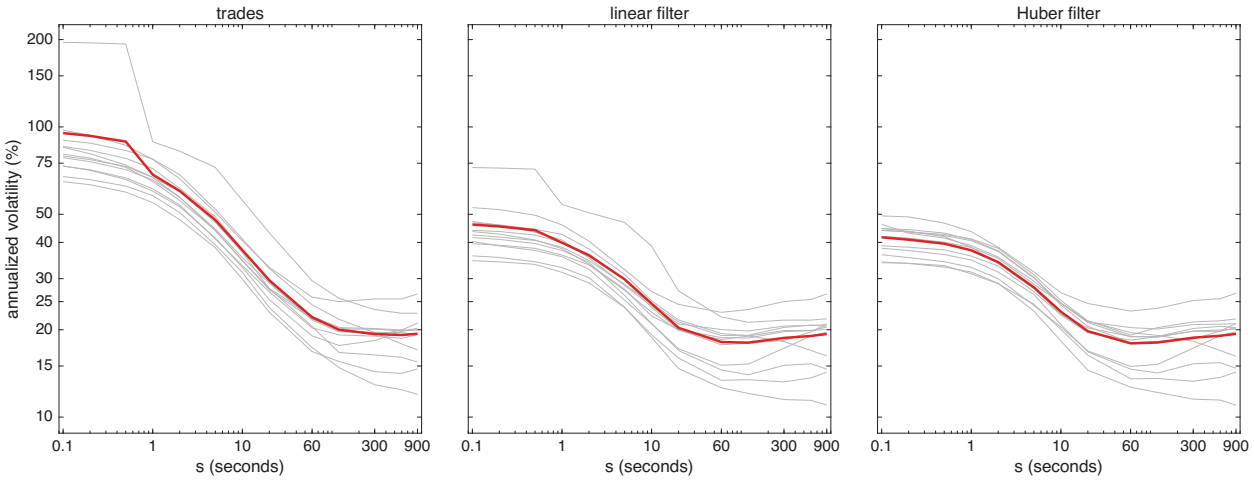

Figure 4: Volatility signature curves for `ISRG` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

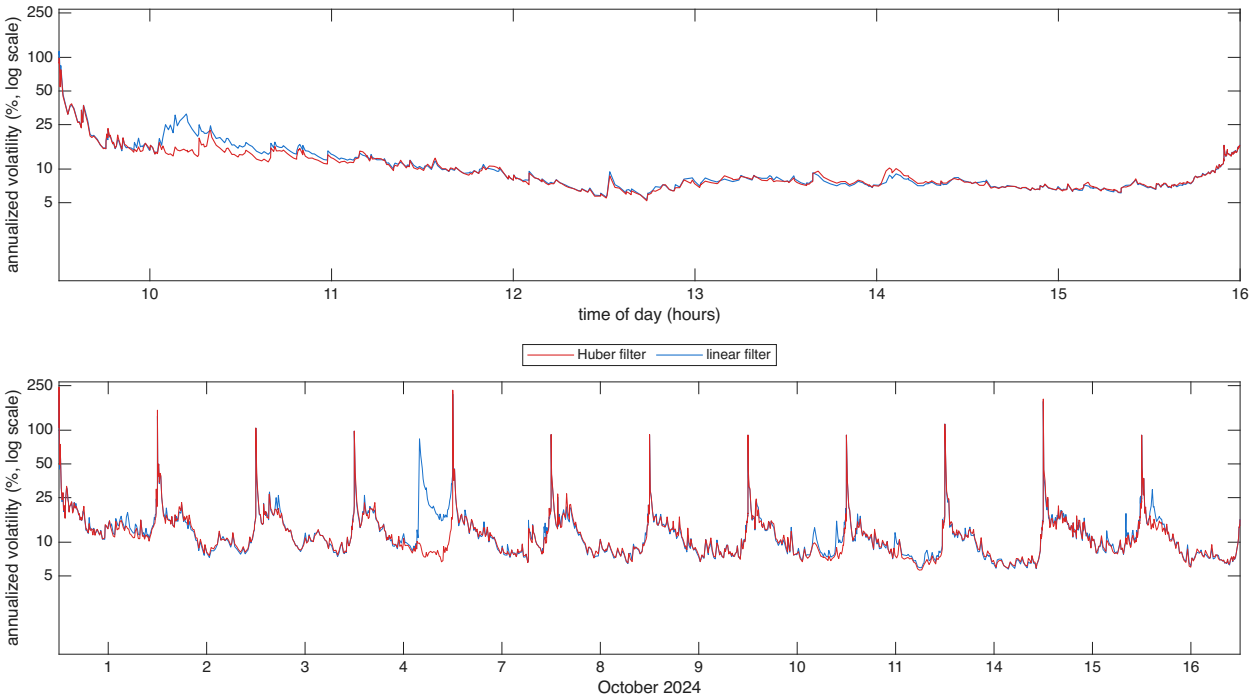

Figure 5: Filtered spot volatility for `ISRG` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 28$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

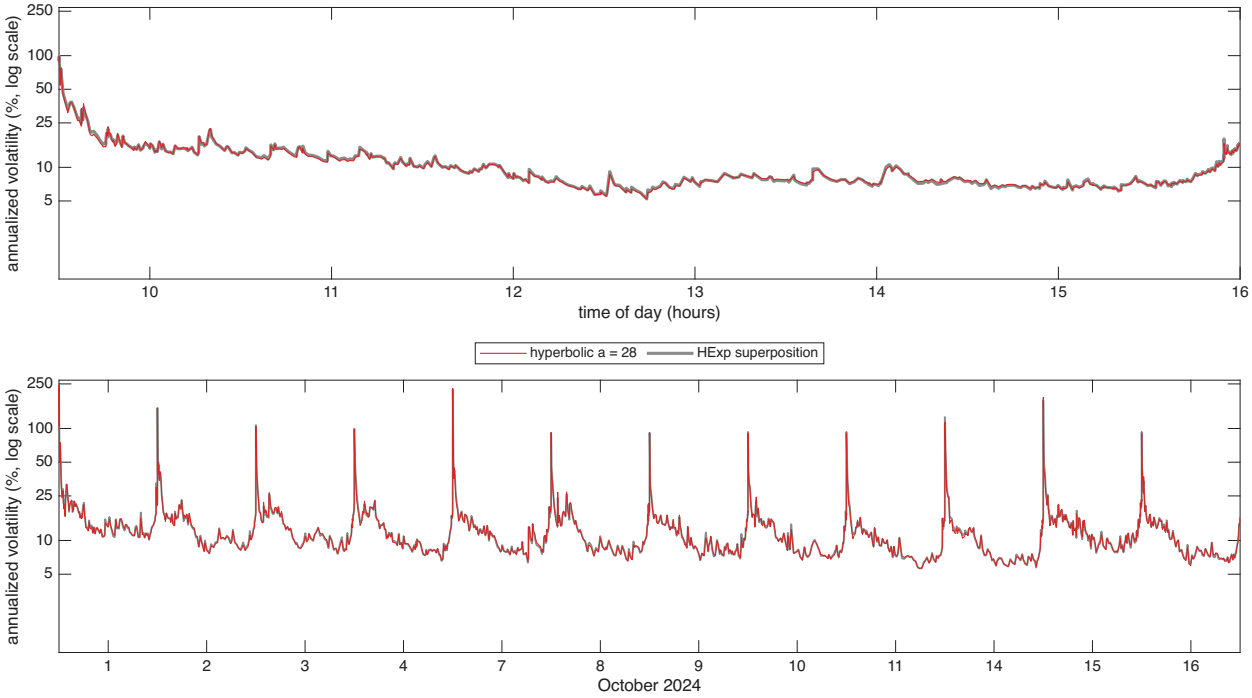

Figure 6: Hyperbolic against superposition weights for `ISRG` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 28$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

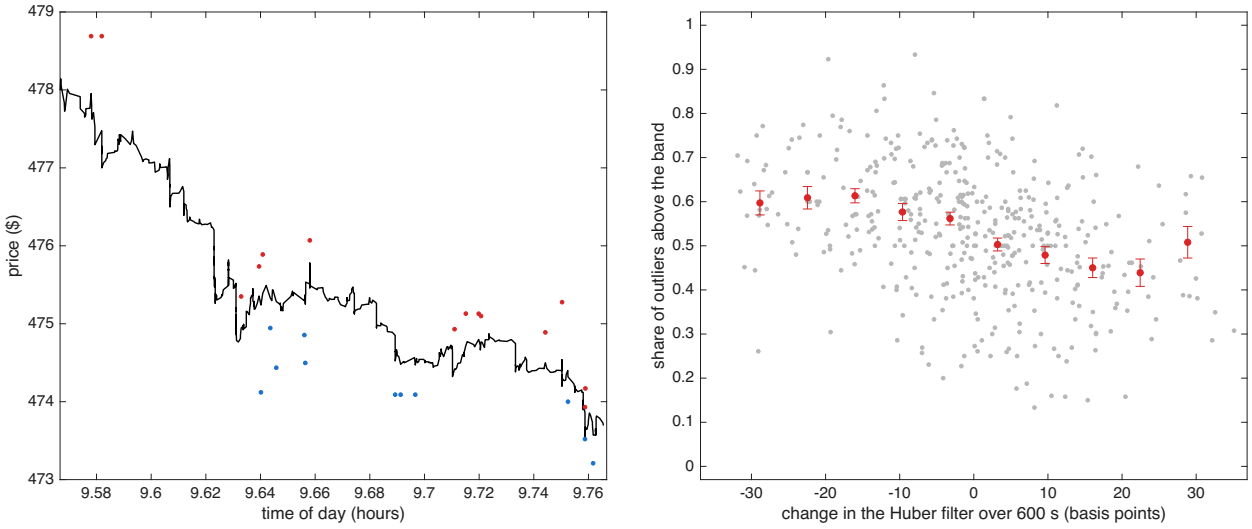

Figure 7: Outliers & robust price for `ISRG` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

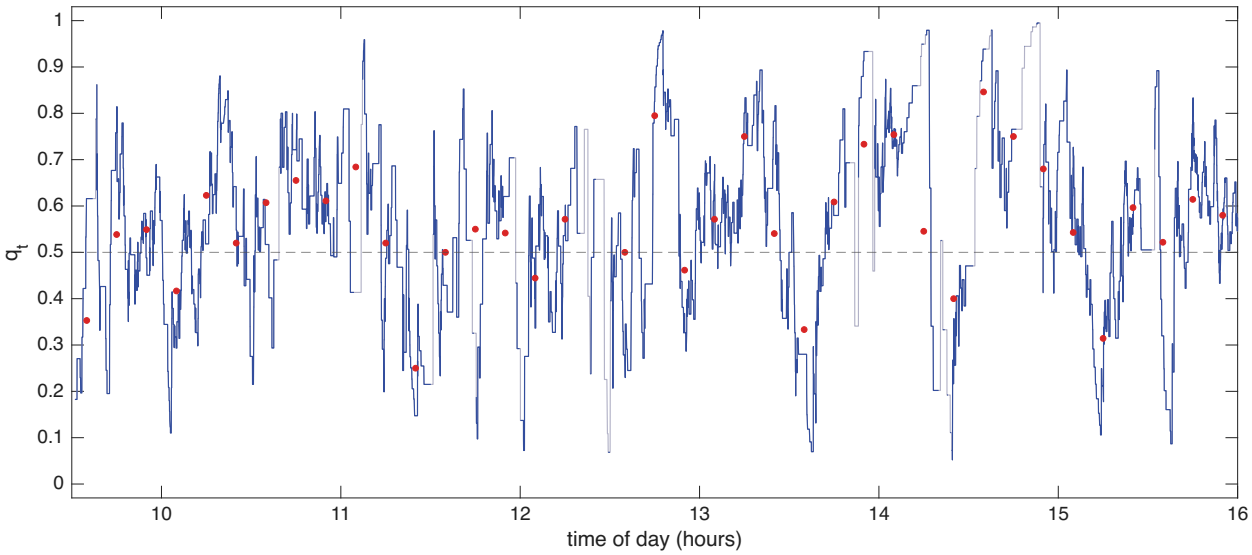

Figure 8: The directional conditional probability for `ISRG`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 84 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

The tail index of the residual cannot be estimated for every 10th trade, on 8 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:05.

Diagnostics.

VRTX 1 to 16 October 2024

trades on 16 October 2024	22,614
average trading rate	1.1 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,156 (5.11%)
price half-life: run / minimizer[1]	3.0 / 3.1 trades
cost of running 3 trades	+0.27%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.02%
direction half-life[3]	101 s [86, 120]
Mandelbrot a minimizing QLIKE	50 trades
a within 0.0005 of the minimum	34 to 74
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	14% $\to$ 26%
Huber $\hat{\sigma}$, before $\to$ peak	14% $\to$ 24%
median spot volatility, Huber / linear input[5]	10.6 / 11.1 %
hyperbolic against HExp, correlation of logs	0.9990
HExp QLIKE advantage, leave-one-day-out[6]	+0.04333
15-minute signature level	11 to 29 %
Huber signature, 1 min over 1 s, 12 day average	1.28

VRTX over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:12; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.466	1.918	1.871	8.849	6.456	3.393	3.131	17.897
2	2.330	2.064	2.076	7.184	5.605	3.529	3.344	13.464
5	2.559	2.230	2.251	5.255	5.127	3.751	3.565	9.194
10	2.546	2.268	2.283	4.216	4.444	3.622	3.513	6.966
25	2.441	2.240	2.237	3.273	3.785	3.387	3.335	5.048
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.25	0.08	0.14	−0.43	2.34	2.48	4.24	3.93
2	0.10	0.10	0.06	−0.03	2.78	2.47	3.78	4.07
5	0.00	0.00	0.00	0.03	2.38	2.60	3.46	4.29
10	−0.03	−0.02	−0.02	0.00	3.01	3.99	4.40	4.87
25	−0.02	0.00	−0.01	0.00	3.92	5.29	5.49	6.73

Table 1: Trade-lag differences, `VRTX`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	10.18	9.12	9.01	17.78	24.7	14.9	14.0	77.8
1 s	6.92	6.33	6.27	11.81	21.1	14.5	14.4	65.6
5 s	5.39	5.01	5.00	8.36	19.7	15.9	16.1	48.5
10 s	4.77	4.51	4.50	6.84	20.1	17.9	17.8	39.9
1 min	3.86	3.74	3.73	4.87	21.1	20.5	20.4	26.1
15 min	3.50	3.49	3.46	3.83	19.5	19.1	18.9	20.6
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.00	0.02	0.01	−0.09	4.51	7.84	6.84	5.20
1 s	−0.05	0.01	0.02	−0.23	3.36	5.71	5.17	7.43
5 s	−0.07	0.00	0.01	−0.33	6.31	6.36	6.15	8.73
10 s	−0.11	−0.04	−0.06	−0.34	10.29	6.00	5.66	9.85
1 min	0.09	0.09	0.07	−0.30				

Table 2: Clock-lag differences, `VRTX`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	5.77	6.93	7.18
s.d. of $\hat{\epsilon}_t$ in cents	12.83	12.03	11.94
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	4.50
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.18	0.18	0.14
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.06	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.00	0.00
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−4.19,4.36	−4.45,4.55
95% C.I. for $\theta_t - X_t$	—	−8.27,8.35	−9.14,8.78
99% C.I. for $\theta_t - X_t$	—	−14.67,13.65	−16.38,15.20
s.e. of $\theta_t - X_t$ in cents	—	3.95	4.52

Table 3: Residual summaries, `VRTX`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 4.50, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.86	2.86	2.61 to 4.59	2.31 to 4.52
Residual divided by δ_t	3.10	3.46	2.71 to 5.21	2.21 to 4.64
Repricing windows removed	3.15	2.96	2.76 to 4.52	2.31 to 4.51
Calm windows only	2.95	3.04		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.96	3.08	2.64 to 5.05	2.41 to 5.07

Table 4: Hill estimator of the tail index of the Huber filter residual, `VRTX`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 12 to 13 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 80.7 to 181.4 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 60.4 cents on 16 October 2024, 93% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

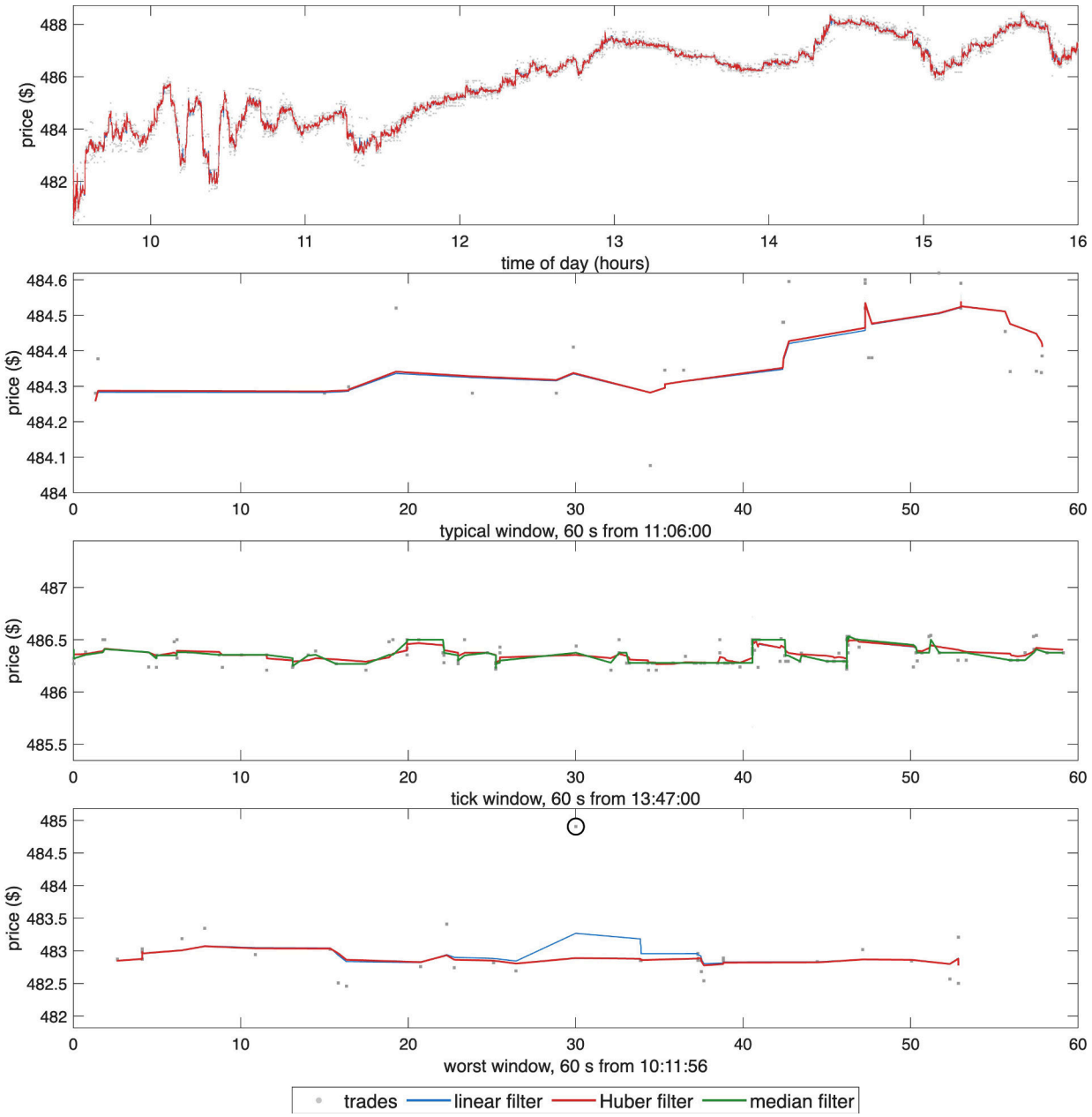

Figure 1: `VRTX` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 11:06:00. 3rd: 60-second window at 13:47:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

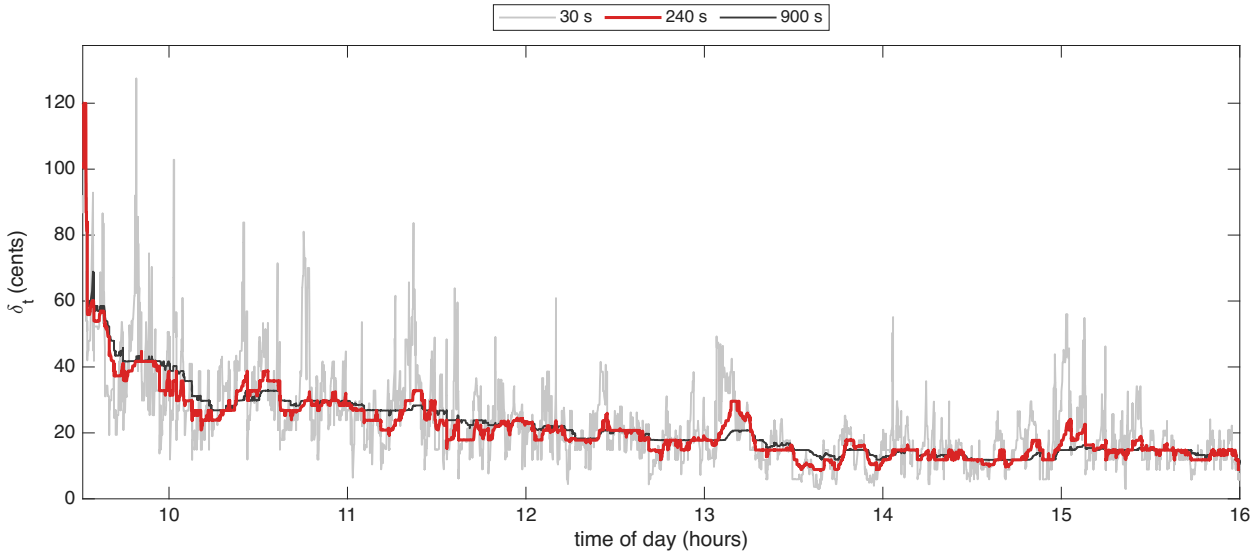

Figure 2: The Huber threshold δ_t in cents for `VRTX` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

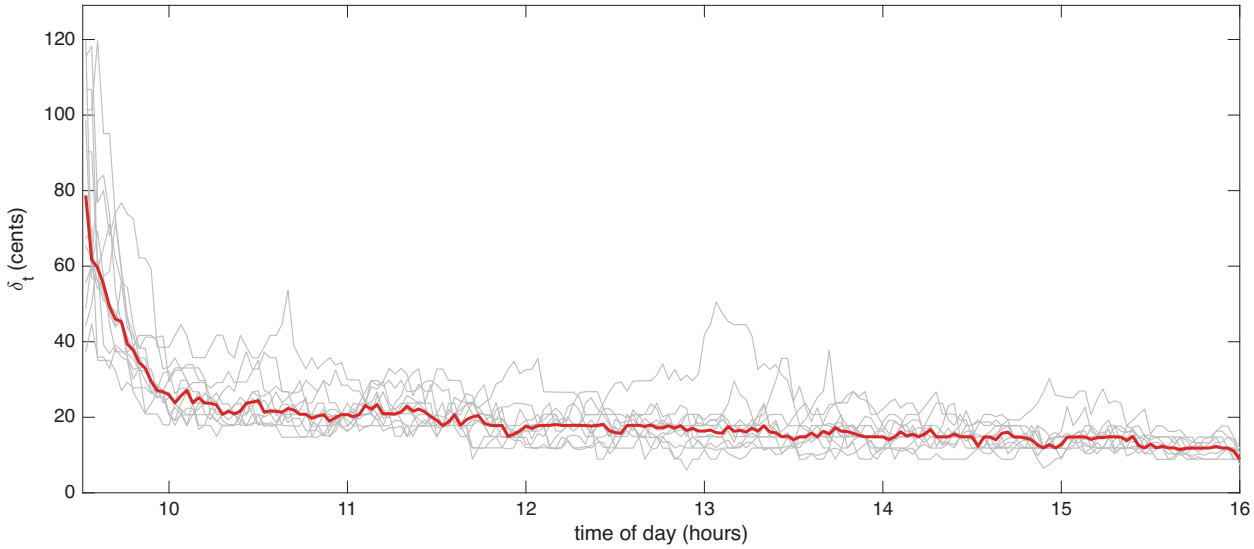

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `VRTX` trading days in Oct 2024 (grey lines) and their median (red line).

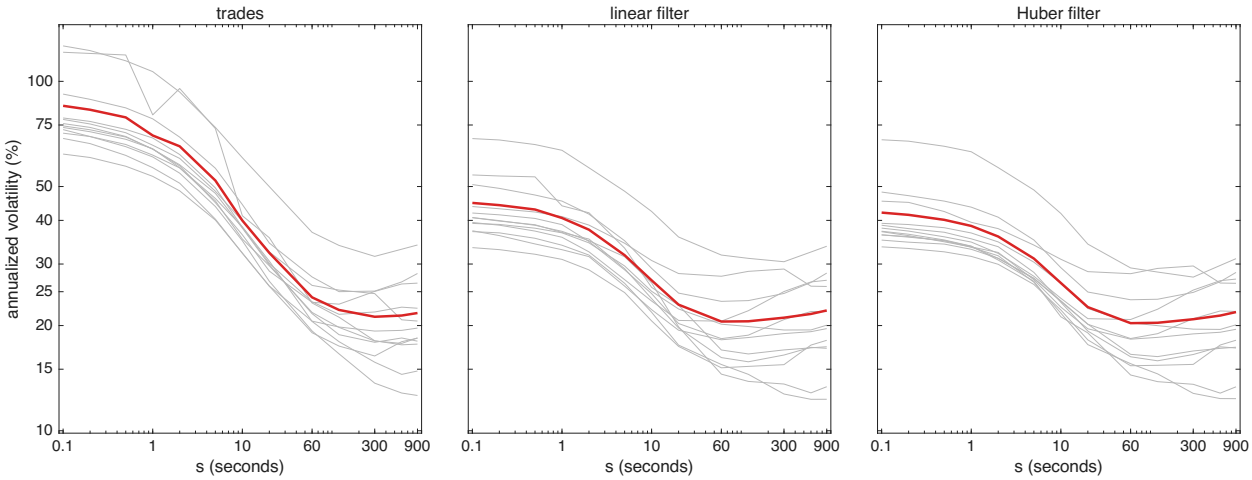

Figure 4: Volatility signature curves for `VRTX` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

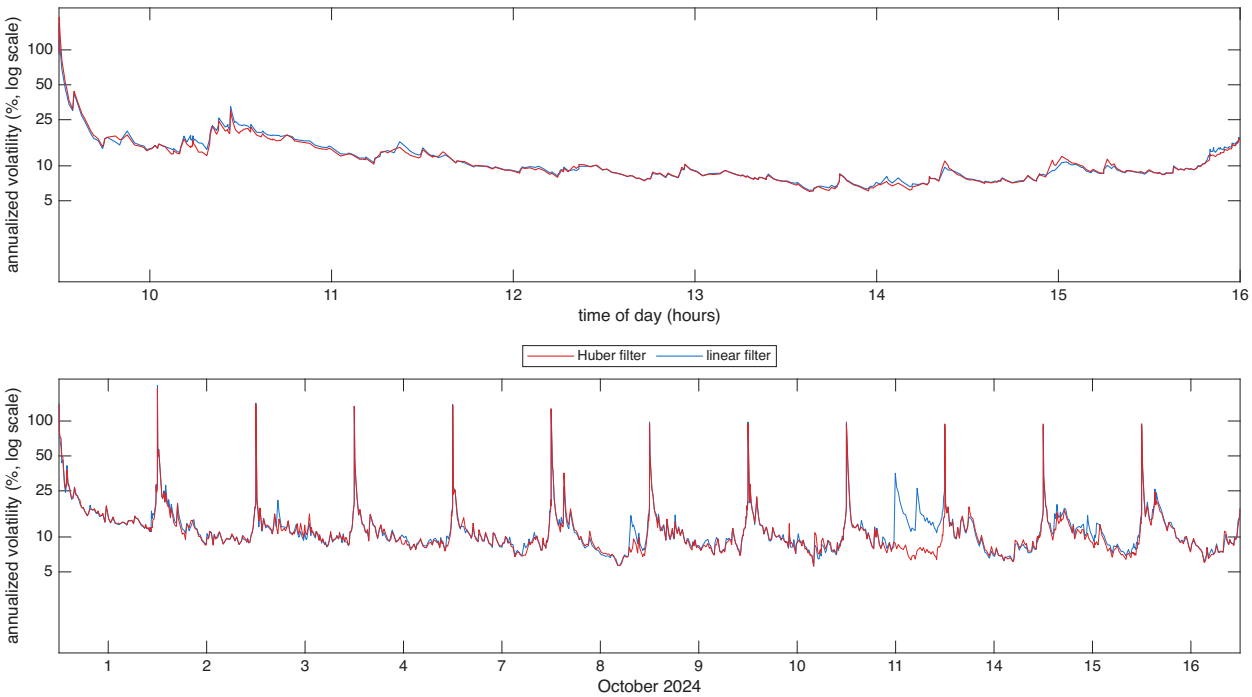

Figure 5: Filtered spot volatility for `VRTX` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 50$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

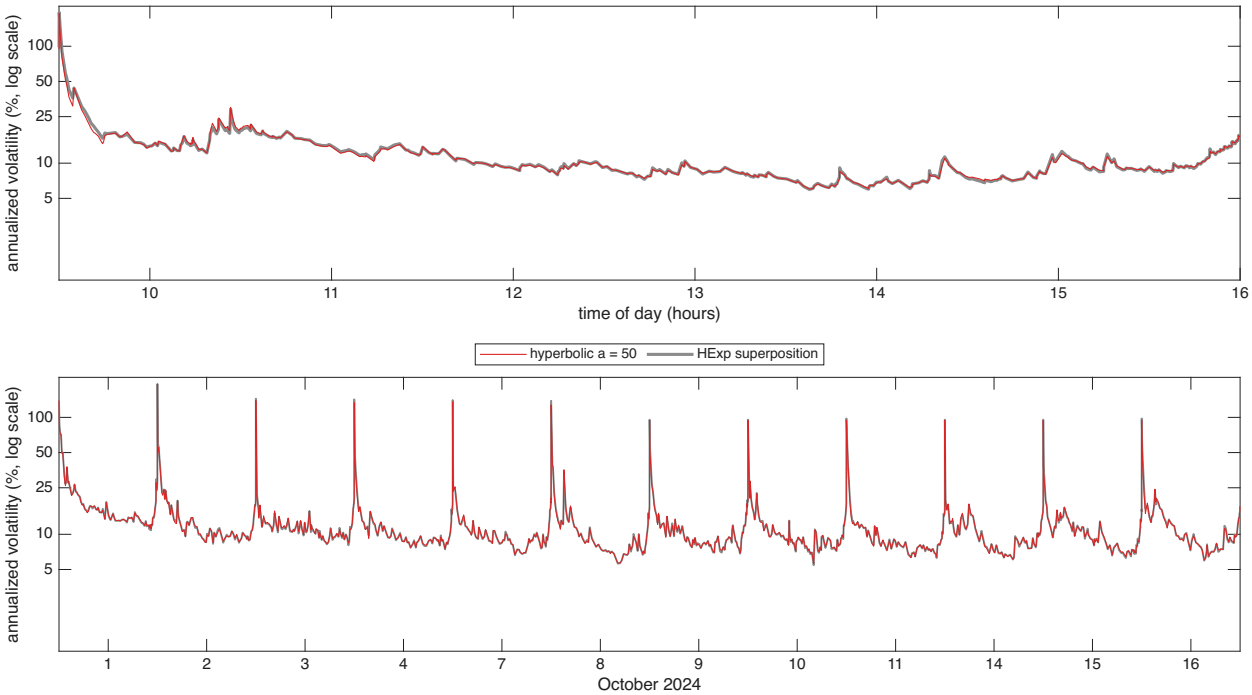

Figure 6: Hyperbolic against superposition weights for `VRTX` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 50$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

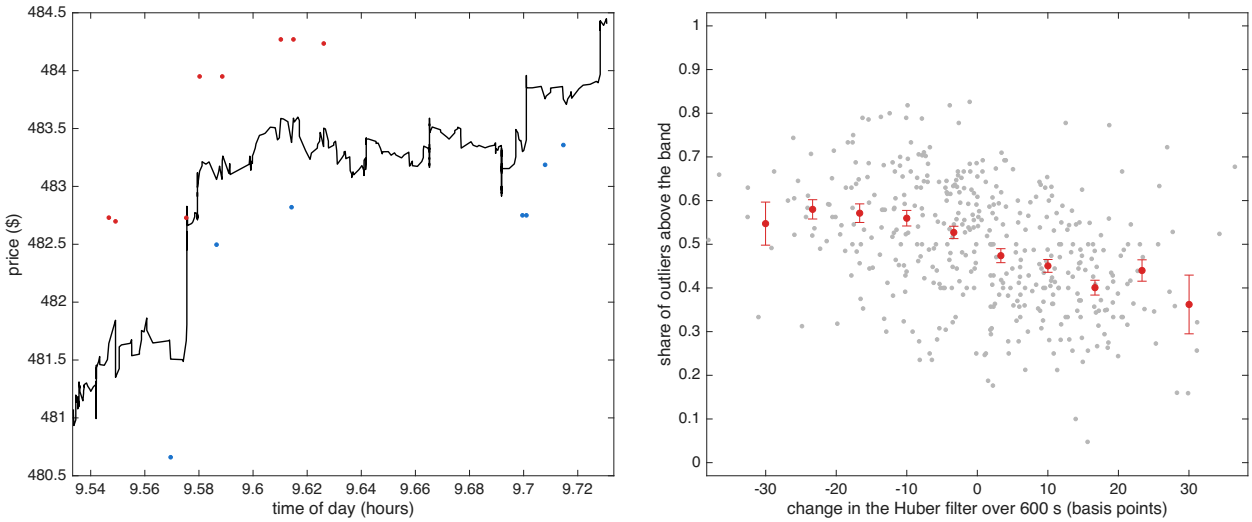

Figure 7: Outliers & robust price for `VRTX` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

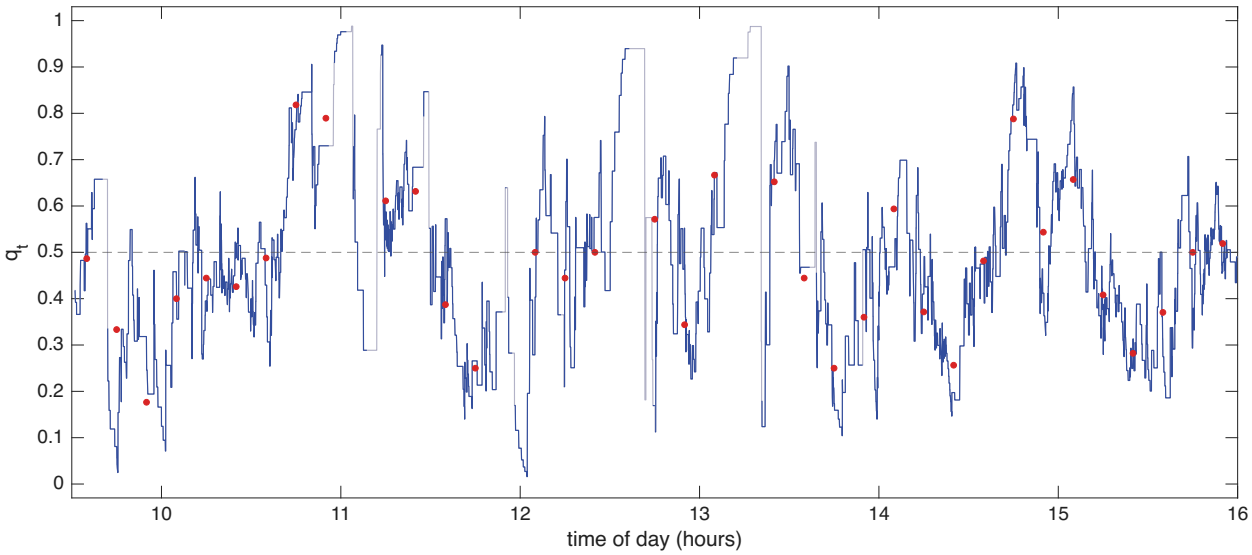

Figure 8: The directional conditional probability for `VRTX`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 101 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 9 to 408 trades, against 50 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 11 trades for this stock.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

REGN 1 to 16 October 2024

trades on 16 October 2024	16,209
average trading rate	0.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	936 (5.77%)
price half-life: run / minimizer[1]	3.0 / 3.0 trades
cost of running 3 trades	+0.11%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.08%
direction half-life[3]	132 s [110, 164]
Mandelbrot a minimizing QLIKE	18 trades
a within 0.0005 of the minimum	14 to 25
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	18% $\to$ 31%
Huber $\hat{\sigma}$, before $\to$ peak	18% $\to$ 41%
median spot volatility, Huber / linear input[5]	13.6 / 14.0 %
hyperbolic against HExp, correlation of logs	0.9965
HExp QLIKE advantage, leave-one-day-out[6]	+0.08232
15-minute signature level	17 to 35 %
Huber signature, 1 min over 1 s, 12 day average	1.28

REGN over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:13; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	6.832	5.428	5.201	25.556	19.517	10.838	9.986	55.712
2	6.249	5.759	5.727	20.334	17.540	11.669	10.560	41.383
5	6.799	6.106	6.118	14.753	15.426	11.760	10.888	28.496
10	6.726	6.051	6.076	11.597	12.902	10.575	10.309	21.372
25	6.192	5.671	5.648	8.744	10.062	8.650	8.441	14.357
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.19	0.16	0.12	−0.45	2.03	2.22	2.24	2.56
2	0.03	0.00	0.06	0.01	1.93	1.90	2.19	2.55
5	−0.01	−0.05	−0.04	−0.01	2.64	1.94	2.26	2.23
10	−0.06	−0.07	−0.04	0.02	2.66	2.26	2.41	2.10
25	−0.08	−0.03	0.01	0.00	2.35	2.25	2.57	2.41

Table 1: Trade-lag differences, `REGN`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	28.20	26.30	26.78	49.51	31.5	19.8	18.8	98.7
1 s	18.86	17.86	18.53	32.58	25.9	18.0	19.0	84.1
5 s	14.06	13.59	14.33	22.19	22.4	17.5	19.0	64.5
10 s	12.54	12.10	12.83	18.84	21.7	18.1	19.7	54.8
1 min	10.02	10.07	10.85	14.21	19.9	18.4	20.0	33.9
15 min	9.32	9.54	10.34	11.68	21.1	20.4	21.9	28.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.09	0.03	0.07	−0.14	3.24	2.91	3.07	2.80
1 s	−0.13	−0.08	−0.03	−0.28	2.66	2.43	2.48	2.28
5 s	−0.14	−0.12	−0.12	−0.35	2.62	2.40	2.46	2.52
10 s	−0.14	−0.09	−0.07	−0.29	2.78	2.57	2.66	2.56
1 min	−0.03	0.02	0.01	−0.28				

Table 2: Clock-lag differences, `REGN`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	16.17	19.26	19.90
s.d. of $\hat{\epsilon}_t$ in cents	39.95	37.57	36.75
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.38
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.14	0.12
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.04	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	0.01	−0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	−0.04	−0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	−0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−11.00,11.36	−11.70,12.05
95% C.I. for $\theta_t - X_t$	—	−20.58,21.80	−22.14,23.13
99% C.I. for $\theta_t - X_t$	—	−38.32,39.15	−41.57,41.31
s.e. of $\theta_t - X_t$ in cents	—	10.69	13.27

Table 3: Residual summaries, `REGN`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.38, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.03	3.01	2.03 to 3.89	2.55 to 4.18
Residual divided by δ_t	3.92	3.71	3.06 to 4.93	2.68 to 5.05
Repricing windows removed	3.41	3.91	2.85 to 4.31	2.55 to 4.18
Calm windows only	3.39	4.10		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.03	3.27	2.03 to 4.11	2.64 to 4.45

Table 4: Hill estimator of the tail index of the Huber filter residual, `REGN`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 10 to 14 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 241.7 to 361.8 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 150.6 cents on 16 October 2024, 97% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

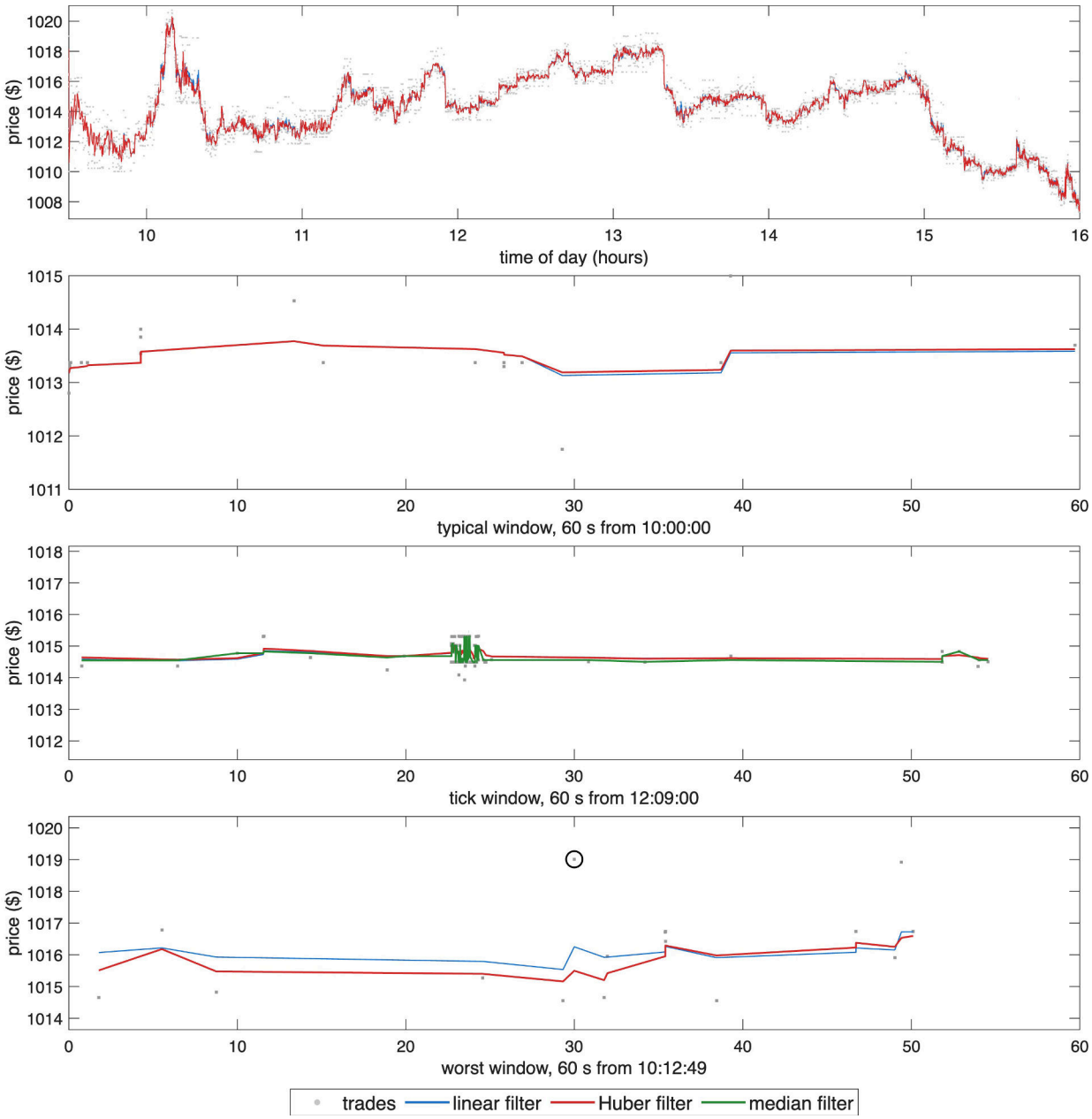

Figure 1: `REGN` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:00:00. 3rd: 60-second window at 12:09:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

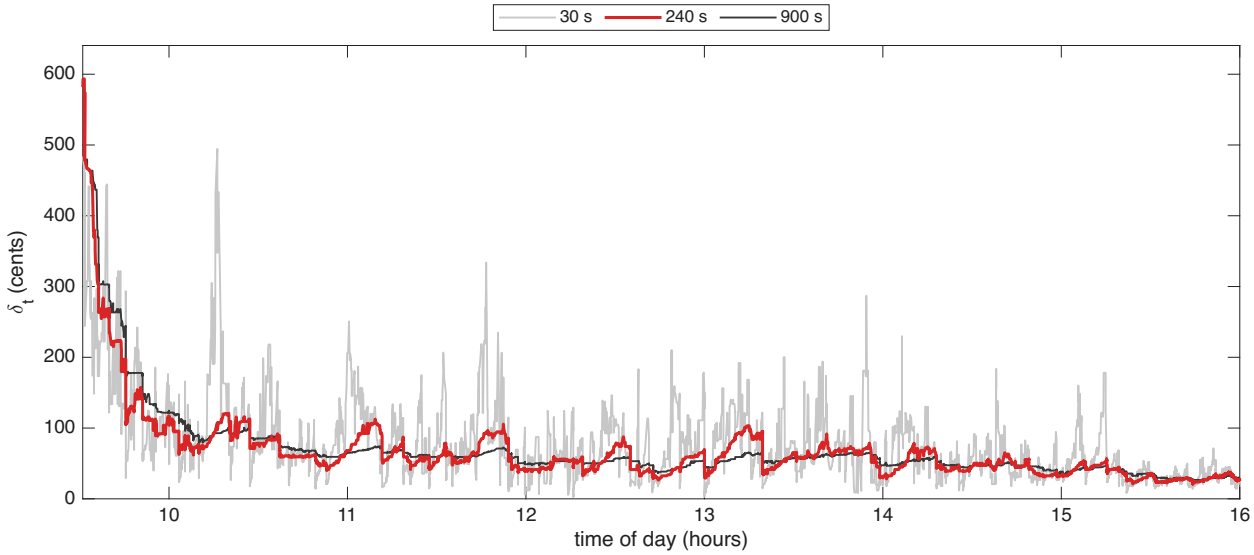

Figure 2: The Huber threshold δ_t in cents for `REGN` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

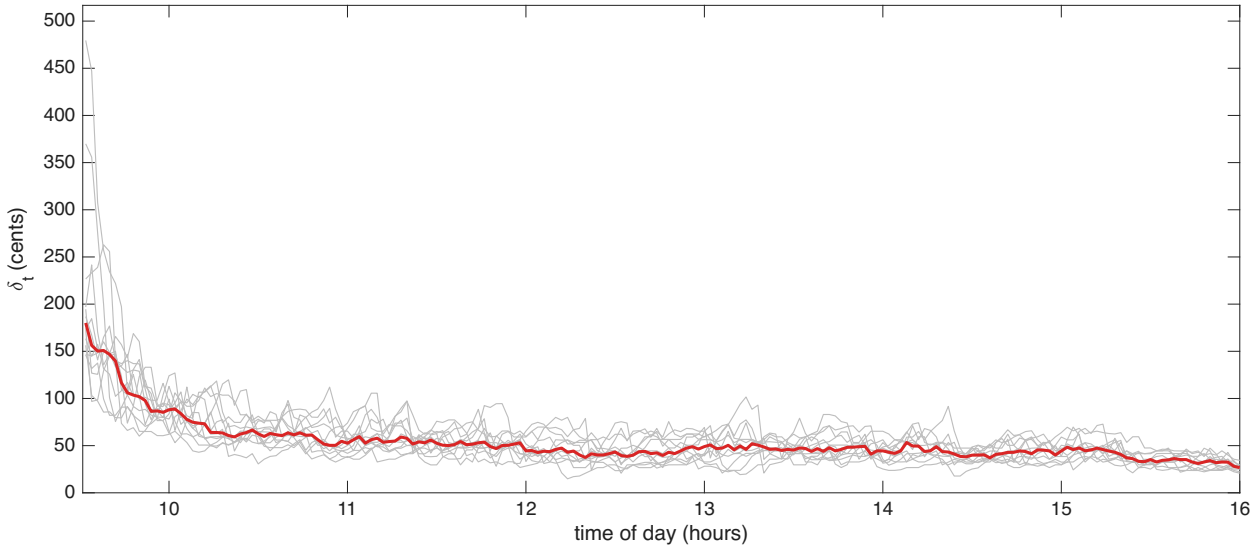

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `REGN` trading days in Oct 2024 (grey lines) and their median (red line).

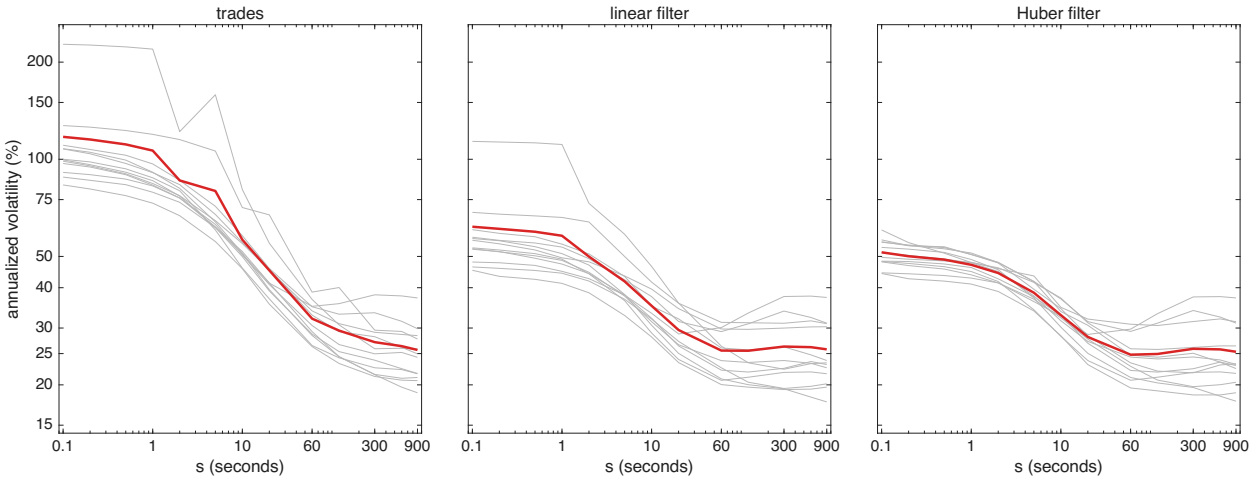

Figure 4: Volatility signature curves for `REGN` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

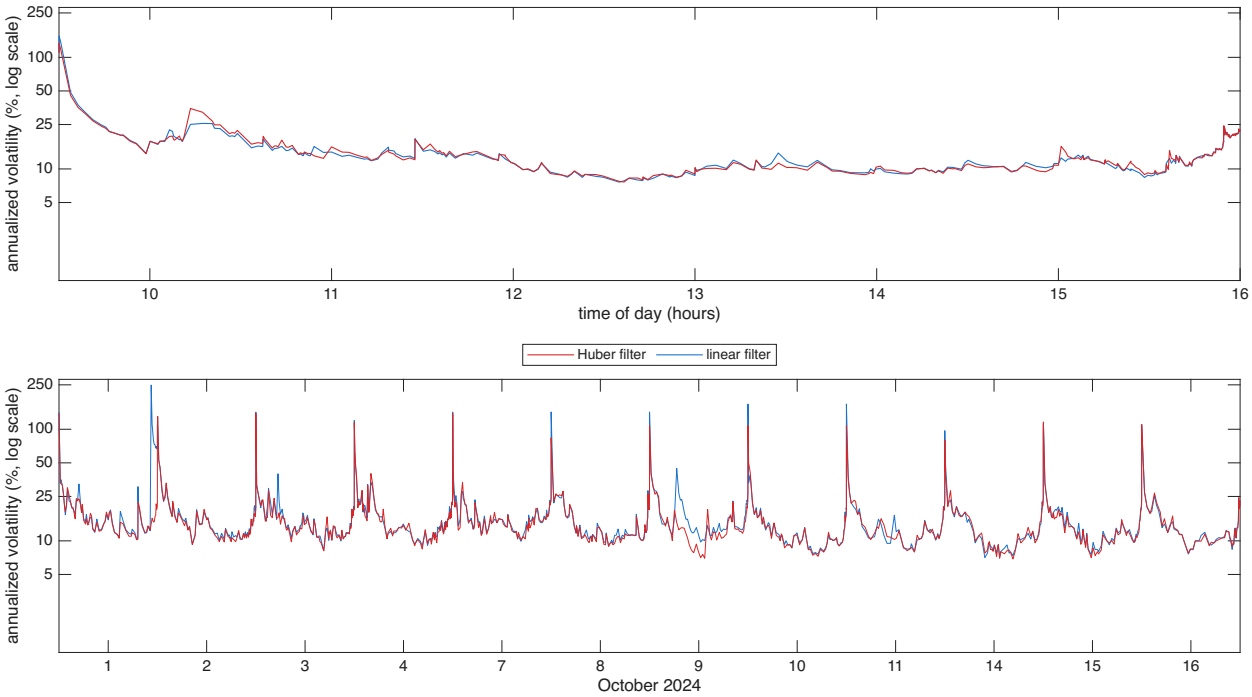

Figure 5: Filtered spot volatility for `REGN` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 18$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

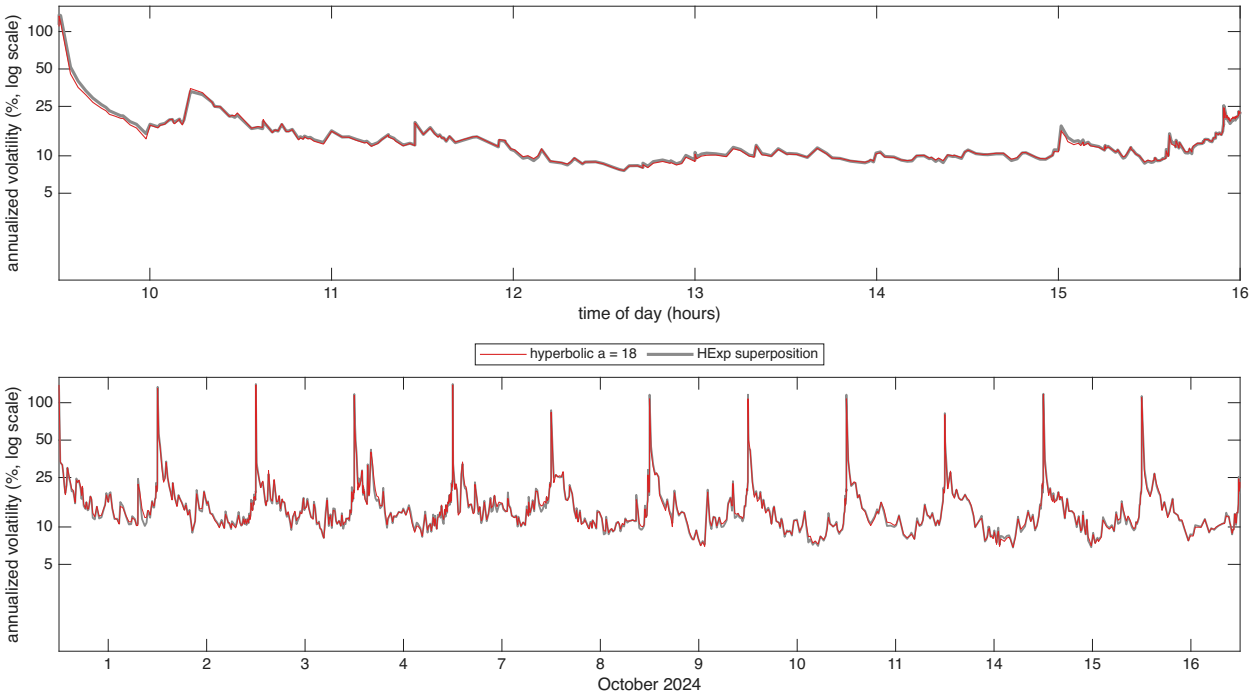

Figure 6: Hyperbolic against superposition weights for `REGN` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 18$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

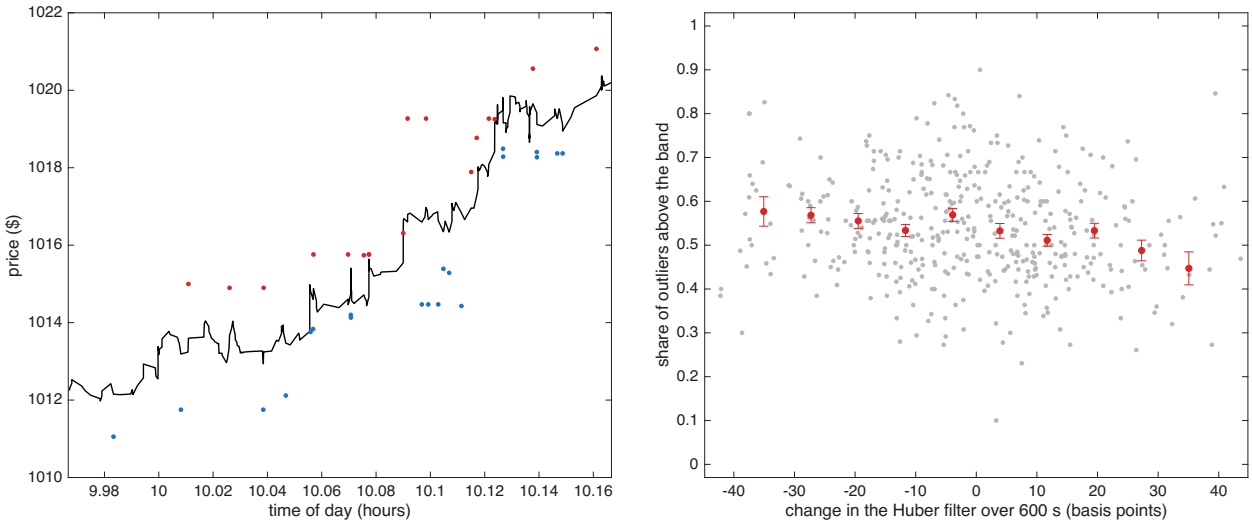

Figure 7: Outliers & robust price for `REGN` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

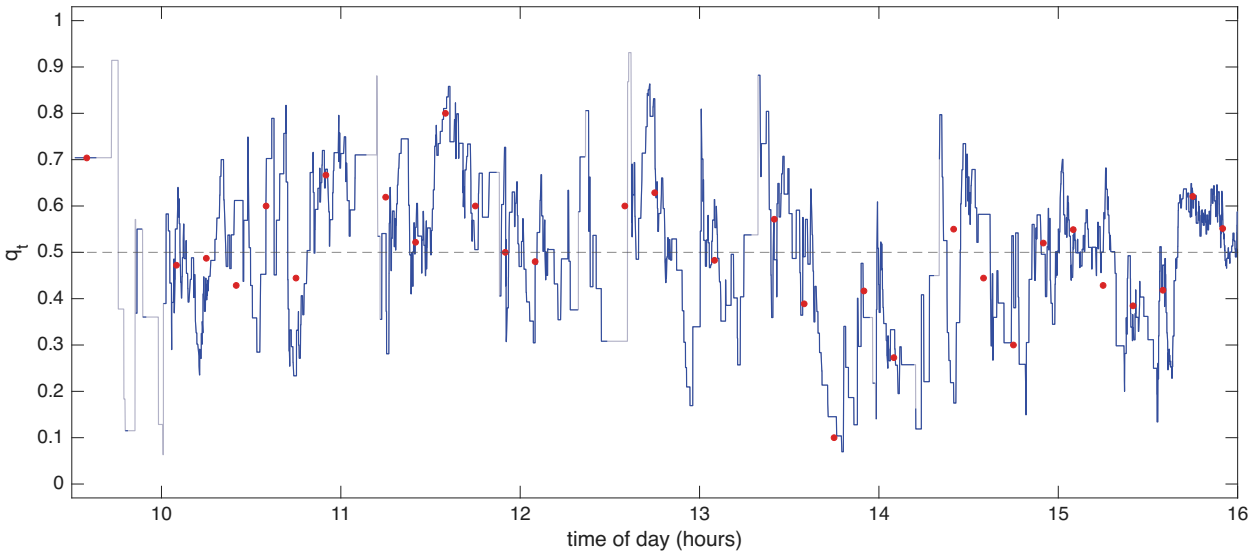

Figure 8: The directional conditional probability for `REGN`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 132 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 7 to 102 trades, against 18 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 9 trades for this stock.

On 15 October 2024, 29 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 27 per cent on any other day. Every number that day contributes to carries that burst.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:13.

Diagnostics.

CROX 1 to 16 October 2024

trades on 16 October 2024	24,819
average trading rate	0.9 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,218 (4.91%)
price half-life: run / minimizer[1]	3.0 / 2.3 trades
cost of running 3 trades	+0.87%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.08%
direction half-life[3]	93 s [78, 112]
Mandelbrot a minimizing QLIKE	24 trades
a within 0.0005 of the minimum	18 to 34
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	21% $\rightarrow$ 29%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	20% $\rightarrow$ 32%
median spot volatility, Huber / linear input[5]	17.2 / 18.1 %
hyperbolic against HExp, correlation of logs	0.9949
HExp QLIKE advantage, leave-one-day-out[6]	+0.09195
15-minute signature level	17 to 42 %
Huber signature, 1 min over 1 s, 12 day average	1.43

CROX over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:41; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.884	0.695	0.685	3.128	2.171	1.231	1.156	6.533
2	0.836	0.754	0.770	2.600	1.888	1.269	1.217	4.954
5	0.908	0.820	0.841	1.923	1.646	1.266	1.246	3.326
10	0.923	0.841	0.857	1.543	1.449	1.217	1.214	2.464
25	0.933	0.865	0.866	1.240	1.300	1.187	1.184	1.810
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.24	0.06	0.11	−0.43	2.70	2.27	3.24	3.48
2	0.02	−0.01	0.01	−0.05	3.04	2.54	3.29	3.80
5	0.00	−0.02	−0.01	0.00	3.25	3.13	3.70	4.20
10	0.02	0.00	0.01	0.02	3.65	3.98	4.76	4.54
25	0.00	−0.01	−0.02	0.00	4.84	4.64	5.60	4.19

Table 1: Trade-lag differences, `CROX`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	4.03	3.68	3.68	6.96	27.5	18.0	18.0	105.9
1 s	2.79	2.57	2.60	4.52	26.8	18.5	18.0	90.6
5 s	1.99	1.87	1.89	3.02	25.3	20.6	20.4	66.8
10 s	1.73	1.66	1.66	2.47	25.5	22.4	22.2	55.4
1 min	1.22	1.20	1.21	1.54	24.8	23.9	23.4	33.3
15 min	1.44	1.46	1.47	1.55	29.3	29.2	29.0	32.0
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.01	0.00	0.01	−0.08	5.55	6.48	6.98	4.29
1 s	0.02	0.10	0.10	−0.16	5.66	8.42	7.42	5.00
5 s	−0.09	0.00	−0.02	−0.32	7.68	7.85	7.53	7.09
10 s	−0.17	−0.13	−0.14	−0.37	6.54	6.83	6.78	6.86
1 min	0.01	0.06	0.07	−0.08				

Table 2: Clock-lag differences, `CROX`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.08	2.51	2.62
s.d. of $\hat{\epsilon}_t$ in cents	4.65	4.41	4.32
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.80
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.16	0.12
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.05	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.01	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	−0.01
80% C.I. for $\theta_t - X_t$	—	−1.65,1.55	−1.81,1.66
95% C.I. for $\theta_t - X_t$	—	−2.89,2.75	−3.21,3.12
99% C.I. for $\theta_t - X_t$	—	−4.61,4.24	−5.23,5.17
s.e. of $\theta_t - X_t$ in cents	—	1.35	1.58

Table 3: Residual summaries, CROX, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.80, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.59	2.97	2.53 to 4.23	2.67 to 3.97
Residual divided by δ_t	4.26	3.20	2.43 to 4.26	2.47 to 3.88
Repricing windows removed	3.73	3.00	2.53 to 4.23	2.66 to 3.97
Calm windows only	3.71	3.08		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	3.75	3.18	2.58 to 4.48	2.69 to 4.42

Table 4: Hill estimator of the tail index of the Huber filter residual, CROX, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 10 to 13 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 40.3 to 60.4 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 21.7 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

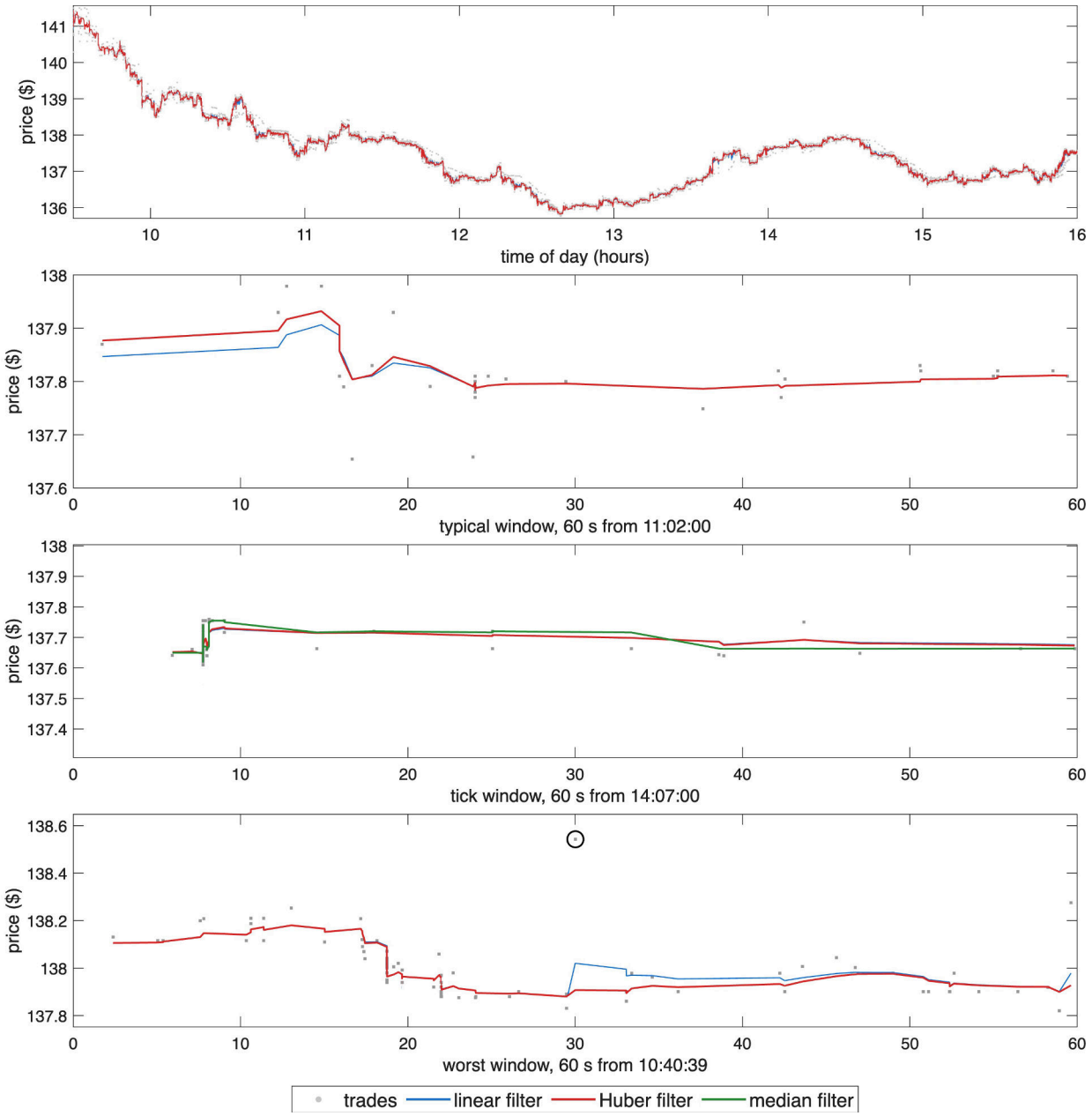

Figure 1: CROX trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 11:02:00. 3rd: 60-second window at 14:07:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

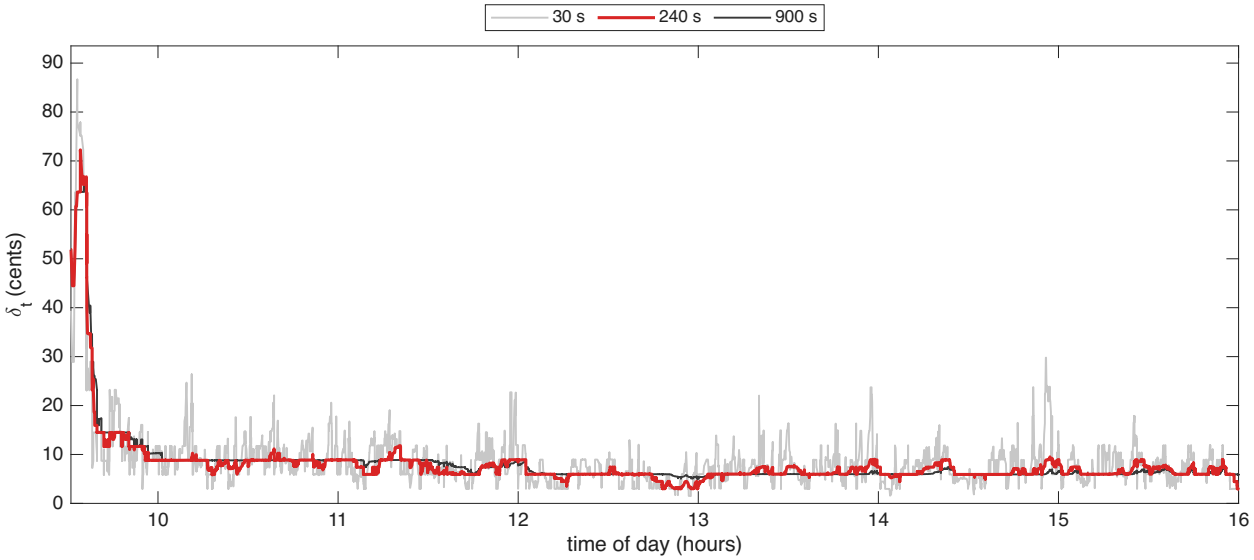

Figure 2: The Huber threshold δ_t in cents for CROX on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

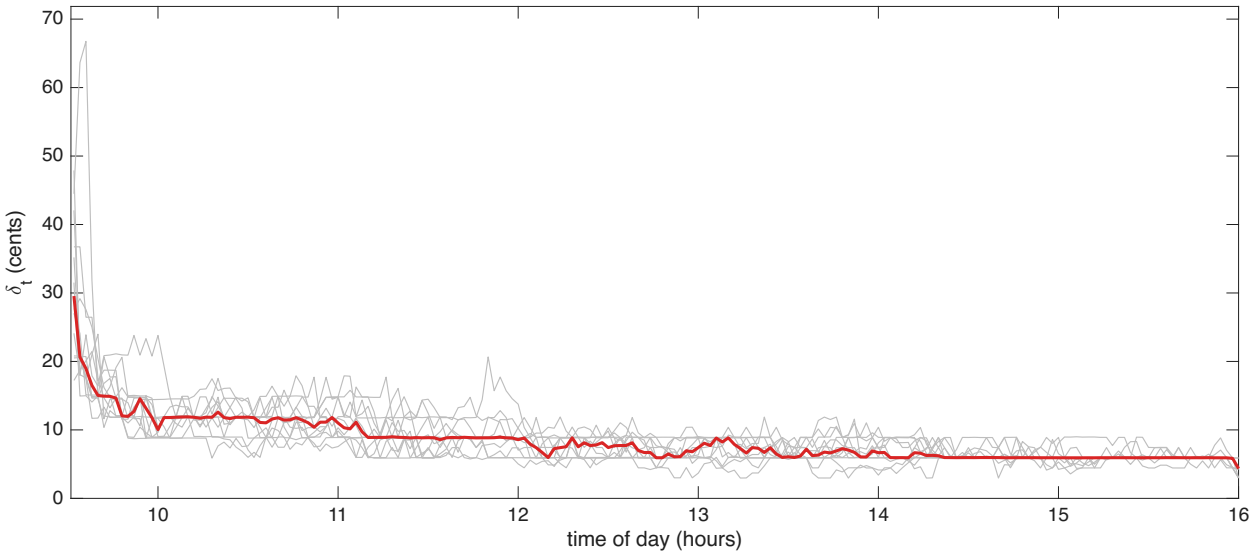

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 CROX trading days in Oct 2024 (grey lines) and their median (red line).

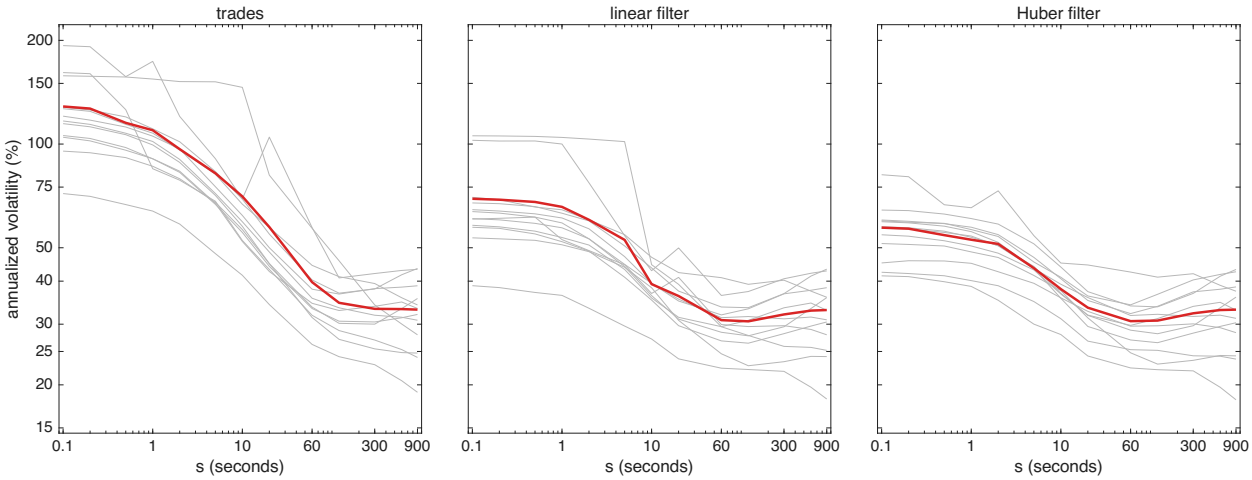

Figure 4: Volatility signature curves for CROX over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

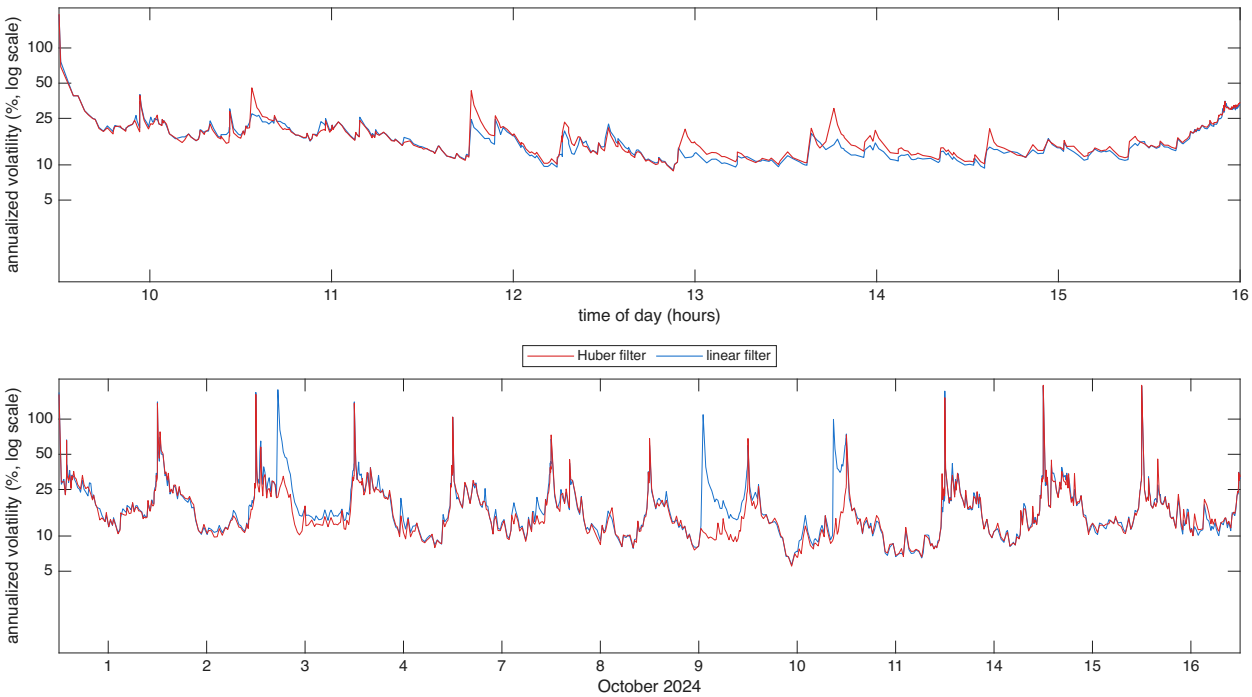

Figure 5: Filtered spot volatility for CROX over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 24$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

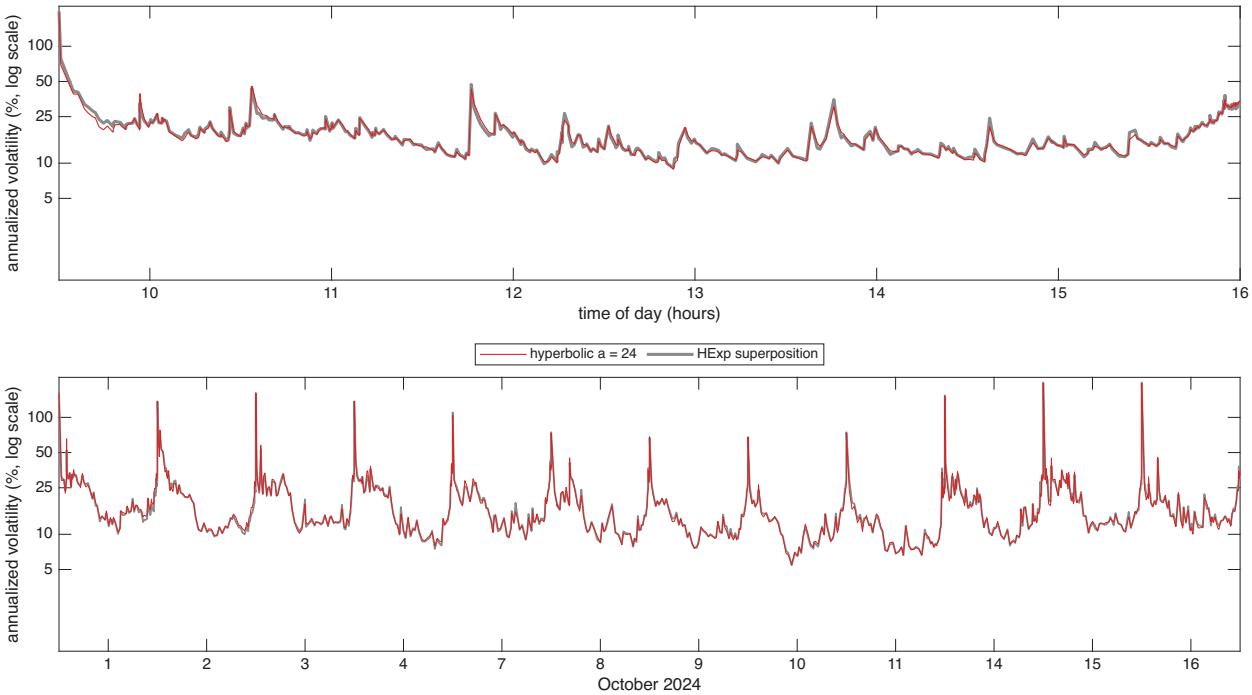

Figure 6: Hyperbolic against superposition weights for `CROX` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 24$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 2, 20, 200, 2000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

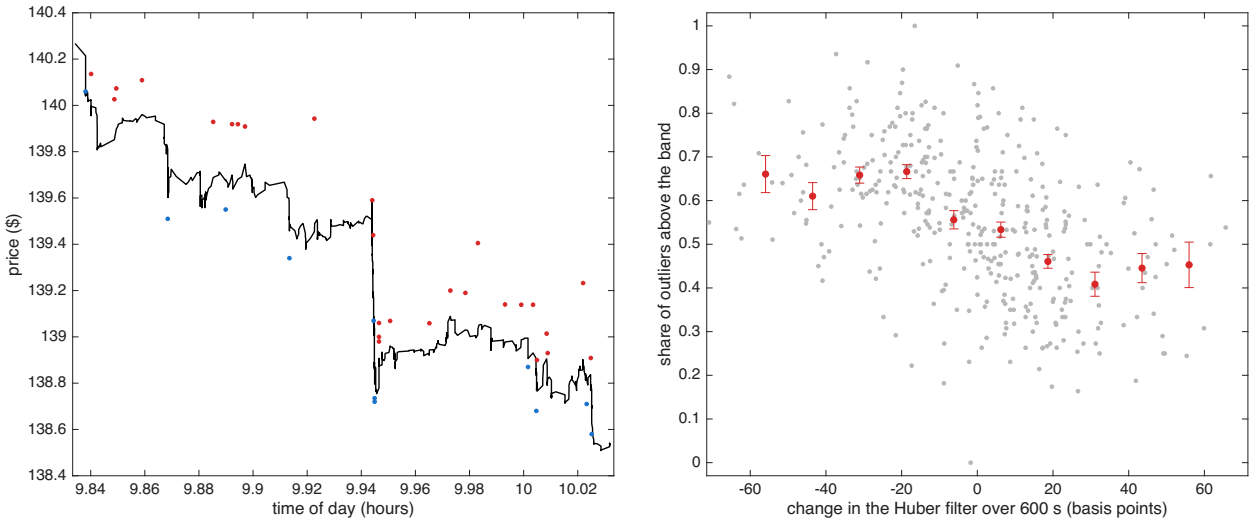

Figure 7: Outliers & robust price for `CROX` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

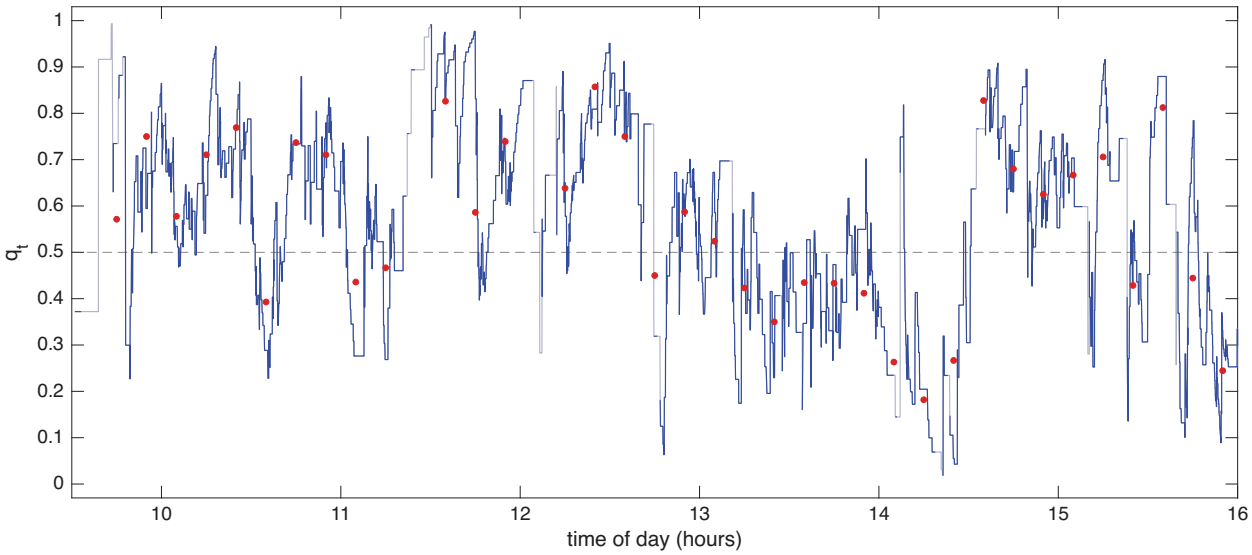

Figure 8: The directional conditional probability for CROX, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 93 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 9 to 102 trades, against 24 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 9 trades for this stock.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:41.

Diagnostics.

MELI 1 to 16 October 2024

trades on 16 October 2024	14,028
average trading rate	0.8 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	968 (6.90%)
price half-life: run / minimizer[1]	3.0 / 5.0 trades
cost of running 3 trades	+1.40%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.12%
direction half-life[3]	160 s [132, 201]
Mandelbrot a minimizing QLIKE	16 trades
a within 0.0005 of the minimum	12 to 23
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	17% $\to$ 38%
Huber $\hat{\sigma}$, before $\to$ peak	18% $\to$ 33%
median spot volatility, Huber / linear input[5]	20.8 / 20.9 %
hyperbolic against HExp, correlation of logs	0.9977
HExp QLIKE advantage, leave-one-day-out[6]	+0.06821
15-minute signature level	18 to 39 %
Huber signature, 1 min over 1 s, 12 day average	0.91

`MELI` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:42; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	22.700	17.833	17.140	88.027	54.264	29.202	28.333	168.250
2	20.153	18.467	18.420	68.760	47.581	29.827	29.349	123.939
5	21.232	18.787	18.948	48.965	40.988	29.271	29.293	84.441
10	20.164	17.933	18.024	37.429	34.261	26.296	26.036	59.683
25	17.416	15.708	15.692	26.832	26.144	22.106	21.940	40.272
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.23	0.04	0.07	−0.46	2.97	4.36	4.09	4.99
2	0.04	0.01	0.02	0.01	3.44	3.87	3.54	4.56
5	−0.05	−0.05	−0.07	−0.02	5.66	3.94	3.35	3.59
10	−0.01	0.00	0.01	0.03	9.85	5.74	5.95	4.16
25	0.00	0.01	0.01	0.01	8.73	6.57	5.41	4.53

Table 1: Trade-lag differences, `MELI`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	69.35	58.72	57.79	140.12	40.9	24.7	24.4	142.7
1 s	44.14	38.61	38.51	89.09	37.5	24.1	23.8	125.6
5 s	32.16	30.05	29.90	59.09	32.3	23.3	23.3	94.3
10 s	27.41	25.87	25.91	46.97	30.7	23.6	23.4	77.3
1 min	19.00	18.20	18.32	27.32	23.1	21.1	21.4	39.9
15 min	18.15	18.47	18.71	20.71	26.6	26.1	26.5	31.3
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.13	−0.10	−0.11	−0.10	7.07	8.77	7.84	5.97
1 s	−0.33	−0.29	−0.25	−0.28	4.22	5.17	6.38	4.55
5 s	−0.25	−0.11	−0.08	−0.43	6.94	7.55	8.51	5.03
10 s	−0.22	−0.12	−0.10	−0.44	7.84	12.41	10.63	4.27
1 min	−0.04	0.03	0.03	−0.30				

Table 2: Clock-lag differences, `MELI`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	54.28	63.96	65.75
s.d. of $\hat{\epsilon}_t$ in cents	115.57	108.61	108.03
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	4.45
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.16	0.07	0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.03	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.00	−0.05	−0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	0.01	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	0.01
80% C.I. for $\theta_t - X_t$	—	−34.32,35.82	−36.03,38.16
95% C.I. for $\theta_t - X_t$	—	−66.58,67.03	−70.01,69.85
99% C.I. for $\theta_t - X_t$	—	−112.52,120.08	−119.54,127.79
s.e. of $\theta_t - X_t$ in cents	—	32.27	35.31

Table 3: Residual summaries, `MELI`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 4.45, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.01	3.36	2.60 to 4.80	2.31 to 5.15
Residual divided by δ_t	5.30	3.86	3.99 to 7.59	2.58 to 6.43
Repricing windows removed	3.01	3.36	2.95 to 4.80	2.31 to 5.15
Calm windows only	4.08	3.32		
Every 10th trade	—	—	3.06 to 3.06	2.60 to 2.60
Six-trade half-life	2.92	3.44	2.71 to 4.90	2.48 to 5.89

Table 4: Hill estimator of the tail index of the Huber filter residual, `MELI`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 9 to 13 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 727.2 to 1130.5 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 514.7 cents on 16 October 2024, 95% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

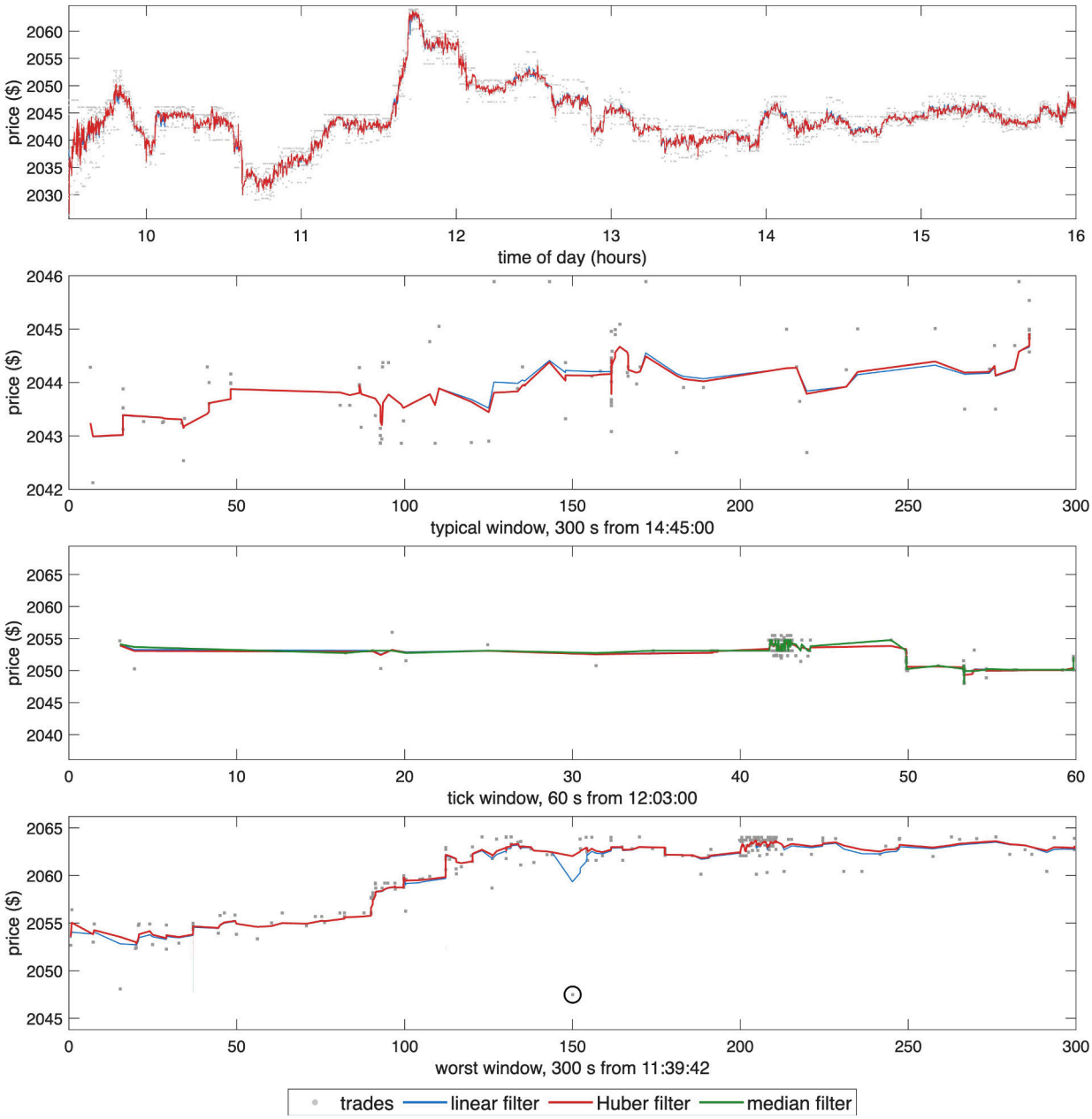

Figure 1: `MELI` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 14:45:00. 3rd: 60-second window at 12:03:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

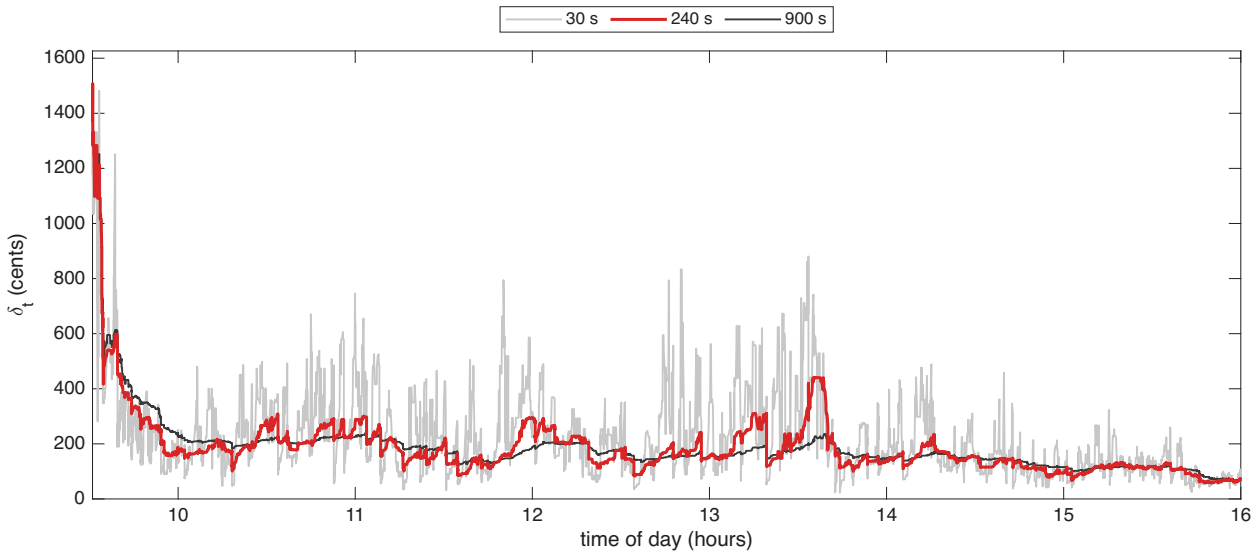

Figure 2: The Huber threshold δ_t in cents for `MELI` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

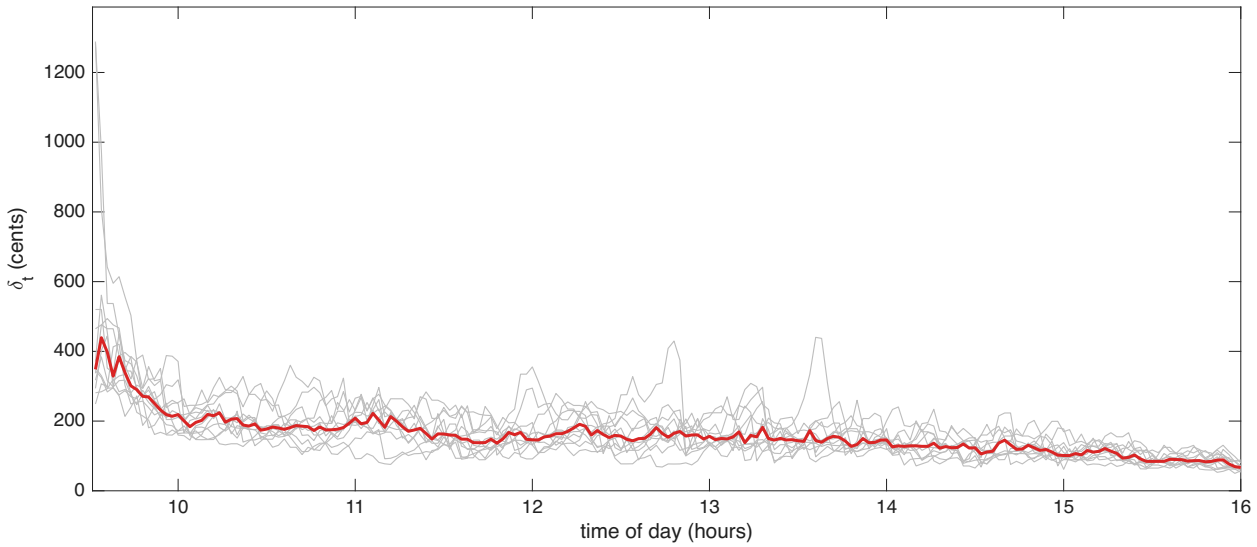

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `MELI` trading days in Oct 2024 (grey lines) and their median (red line).

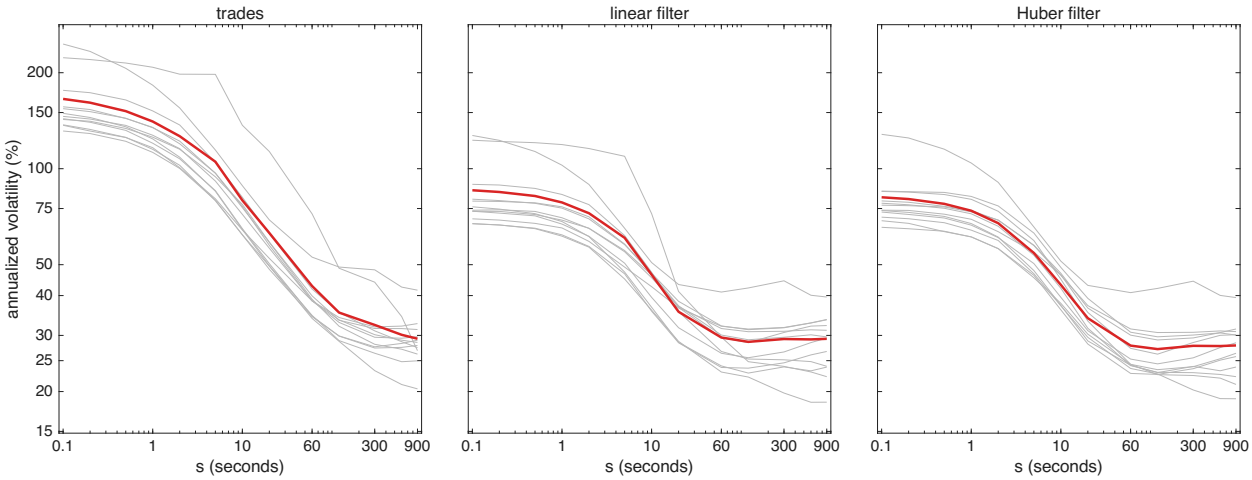

Figure 4: Volatility signature curves for `MELI` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

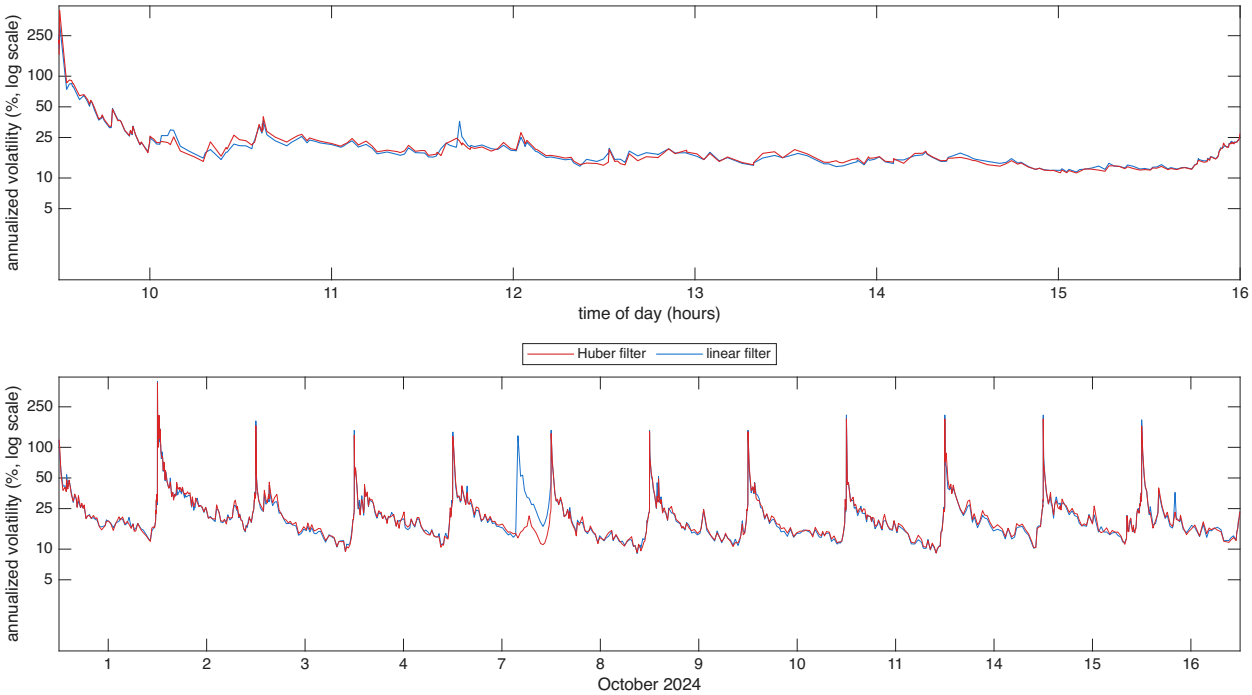

Figure 5: Filtered spot volatility for `MELI` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 16$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

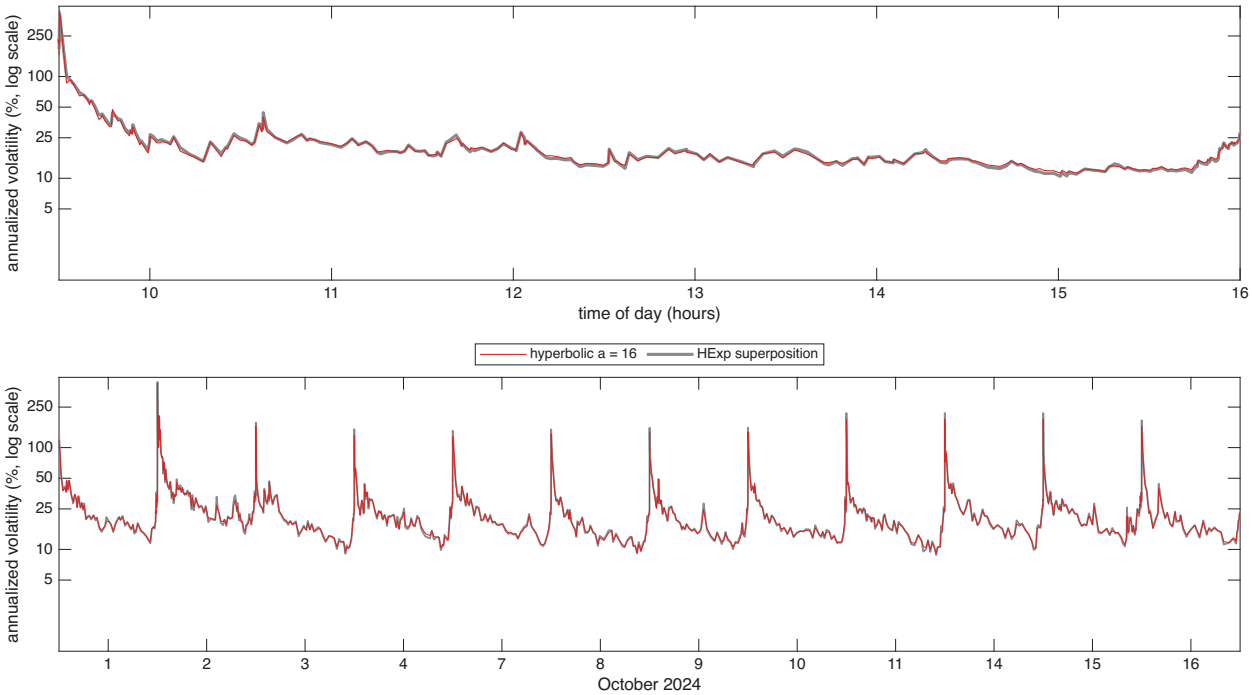

Figure 6: Hyperbolic against superposition weights for `MELI` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 16$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

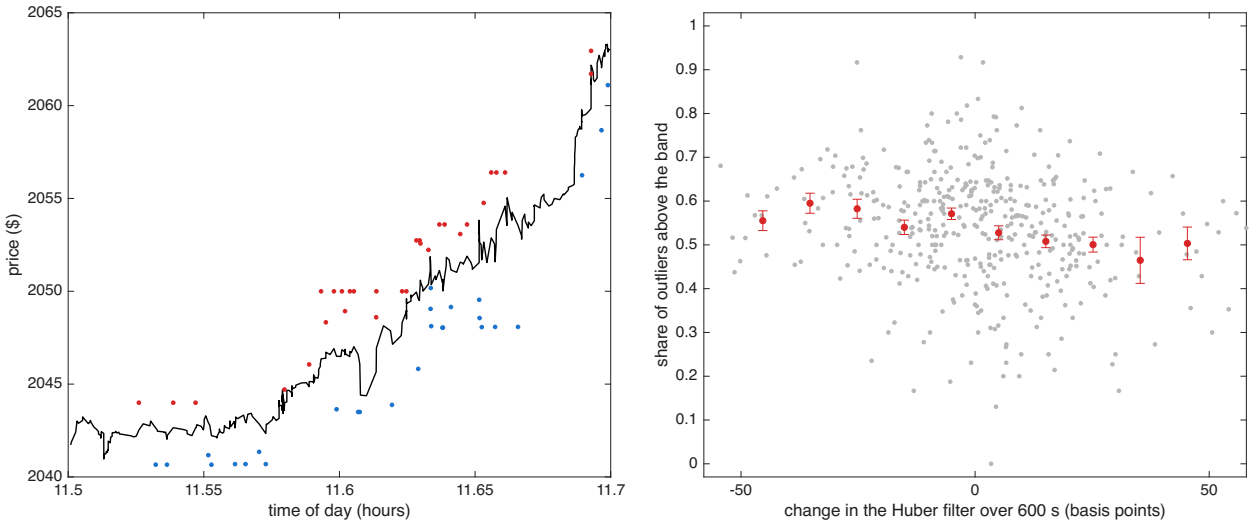

Figure 7: Outliers & robust price for `MELI` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

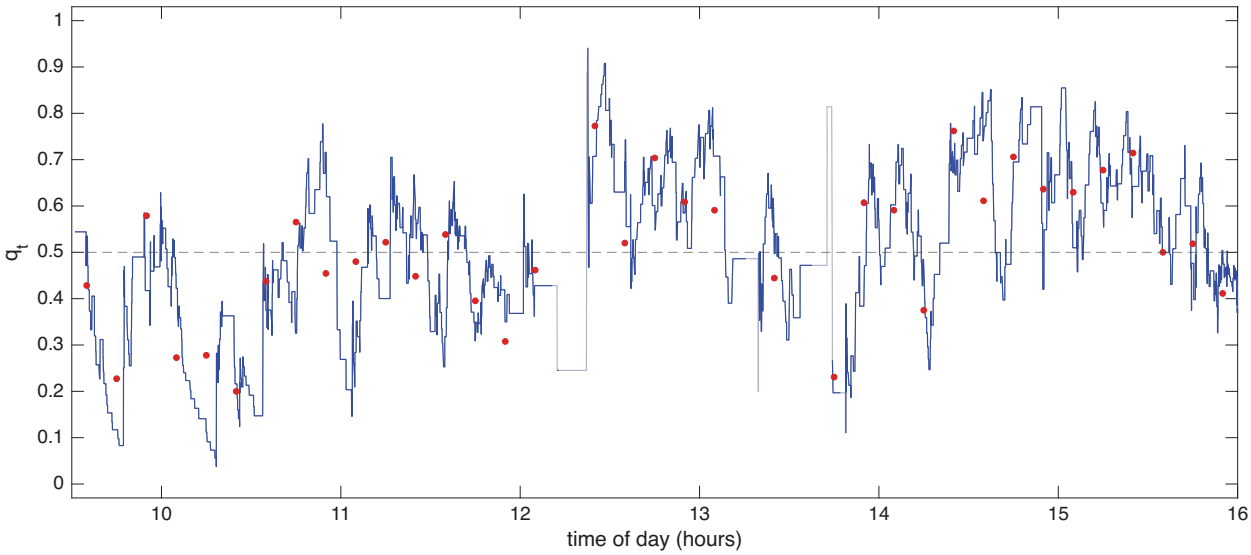

Figure 8: The directional conditional probability for `MELI`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 160 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

At the horizons k = 2, 3, 5, 10, 30 and 100 the half-life minimizing the k-trade-ahead Huber forecast loss is 4, 4, 5, 5, 5 and 5 trades, within one grid point of the end of the 1 to 5 trade grid searched.
The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
The tail index of the residual cannot be estimated for every 10th trade, on 11 of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

CALM 1 to 16 October 2024

trades on 16 October 2024	12,512
average trading rate	0.7 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	587 (4.69%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.37%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	177 s [151, 213]
Mandelbrot a minimizing QLIKE	18 trades
a within 0.0005 of the minimum	15 to 22
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	31% $\rightarrow$ 27%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	31% $\rightarrow$ 28%
median spot volatility, Huber / linear input[5]	15.0 / 14.8 %
hyperbolic against HExp, correlation of logs	0.9932
HExp QLIKE advantage, leave-one-day-out[6]	+0.17917
15-minute signature level	15 to 68 %
Huber signature, 1 min over 1 s, 12 day average	1.43

`CALM` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:49; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.641	0.504	0.493	2.204	1.615	0.890	0.851	4.644
2	0.615	0.560	0.562	1.841	1.423	0.969	0.927	3.587
5	0.697	0.625	0.629	1.377	1.307	1.022	0.982	2.418
10	0.726	0.660	0.660	1.136	1.227	1.048	1.024	1.892
25	0.716	0.672	0.667	0.918	1.132	1.043	1.028	1.430
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.22	0.18	0.19	−0.40	2.83	2.97	2.90	3.15
2	0.05	0.04	0.03	−0.08	3.42	2.92	3.23	3.29
5	−0.02	0.01	0.02	0.00	3.60	4.08	4.48	3.66
10	−0.01	0.01	−0.01	0.00	3.74	4.22	5.02	4.55
25	0.01	0.01	0.01	−0.01	5.41	5.29	5.02	5.71

Table 1: Trade-lag differences, `CALM`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.20	2.01	1.98	3.67	23.9	14.1	13.9	79.2
1 s	1.52	1.42	1.39	2.55	22.1	15.9	15.4	71.5
5 s	1.13	1.06	1.04	1.83	20.5	16.3	16.2	60.5
10 s	0.97	0.92	0.91	1.50	21.0	17.4	17.0	51.3
1 min	0.73	0.72	0.72	0.94	21.9	20.3	20.1	31.8
15 min	0.72	0.74	0.75	0.80	22.2	21.6	21.6	25.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.01	0.03	0.03	−0.08	14.81	7.68	6.36	7.53
1 s	−0.15	−0.12	−0.11	−0.18	26.41	16.69	12.00	10.40
5 s	−0.06	−0.03	−0.02	−0.20	37.26	17.86	15.68	13.68
10 s	−0.10	−0.05	−0.06	−0.36	37.28	17.91	15.43	13.92
1 min	−0.06	−0.03	−0.03	−0.28				

Table 2: Clock-lag differences, `CALM`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	1.49	1.81	1.89
s.d. of $\hat{\epsilon}_t$ in cents	3.37	3.17	3.16
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.49
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.20	0.22	0.18
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.07	0.05
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.02	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	0.01	0.00
80% C.I. for $\theta_t - X_t$	—	−1.10,1.26	−1.17,1.37
95% C.I. for $\theta_t - X_t$	—	−2.24,2.37	−2.42,2.59
99% C.I. for $\theta_t - X_t$	—	−4.26,4.27	−4.68,5.01
s.e. of $\theta_t - X_t$ in cents	—	1.13	1.28

Table 3: Residual summaries, `CALM`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.49, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.12	3.04	2.44 to 4.20	2.45 to 4.57
Residual divided by δ_t	2.55	3.48	1.78 to 4.62	2.17 to 3.58
Repricing windows removed	3.34	3.09	2.60 to 4.28	2.45 to 4.57
Calm windows only	3.87	2.98		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	3.55	2.95	2.79 to 4.74	2.46 to 4.74

Table 4: Hill estimator of the tail index of the Huber filter residual, `CALM`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 8 to 12 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.2 to 40.3 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 15.1 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

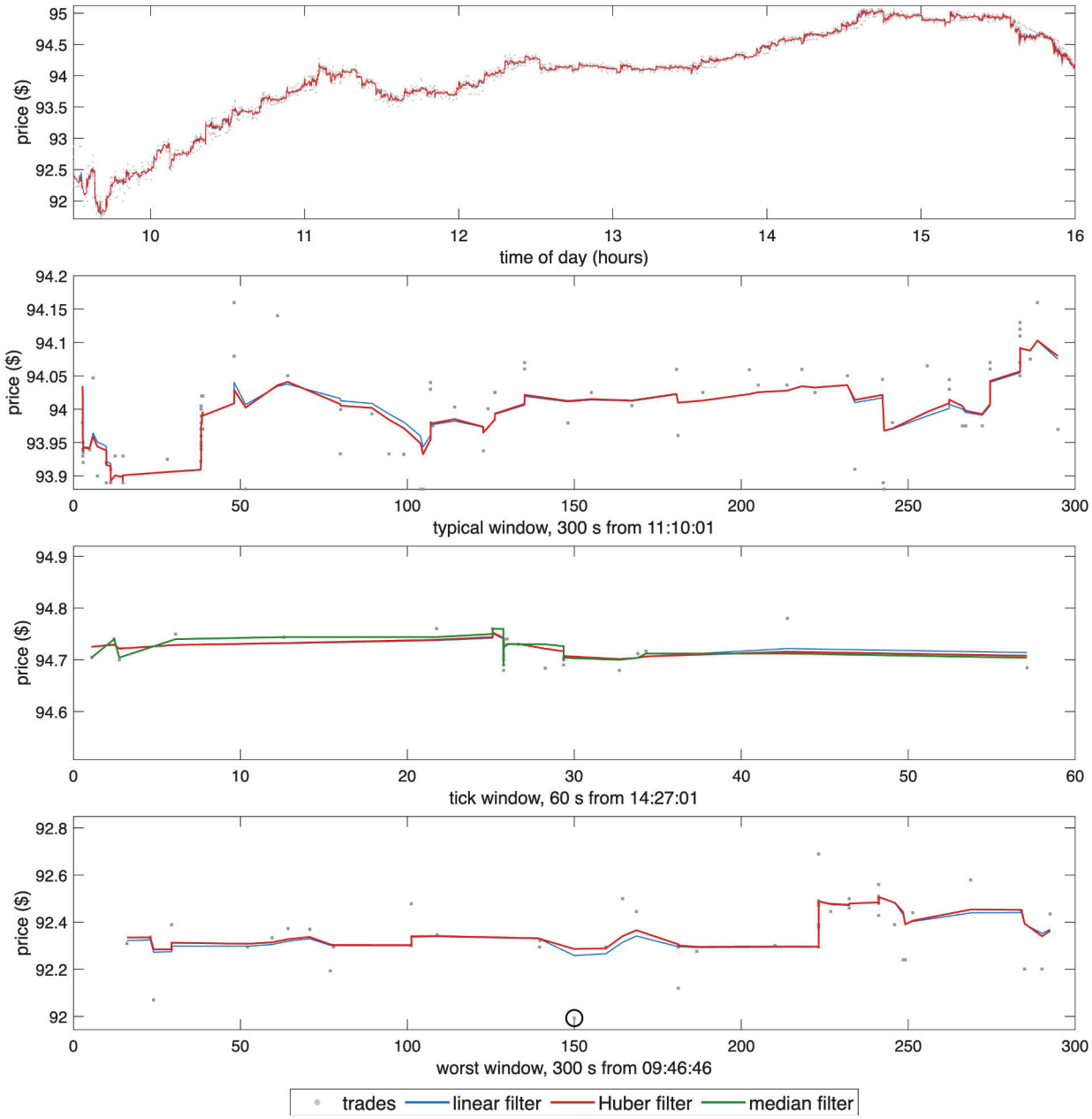

Figure 1: `CALM` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 11:10:01. 3rd: 60-second window at 14:27:01 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

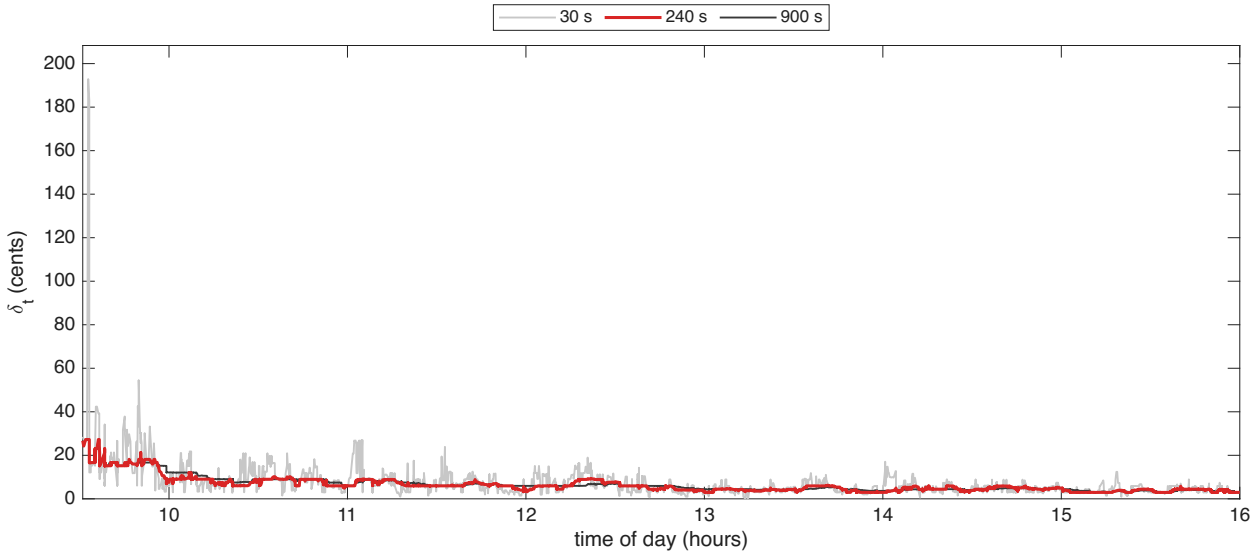

Figure 2: The Huber threshold δ_t in cents for `CALM` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

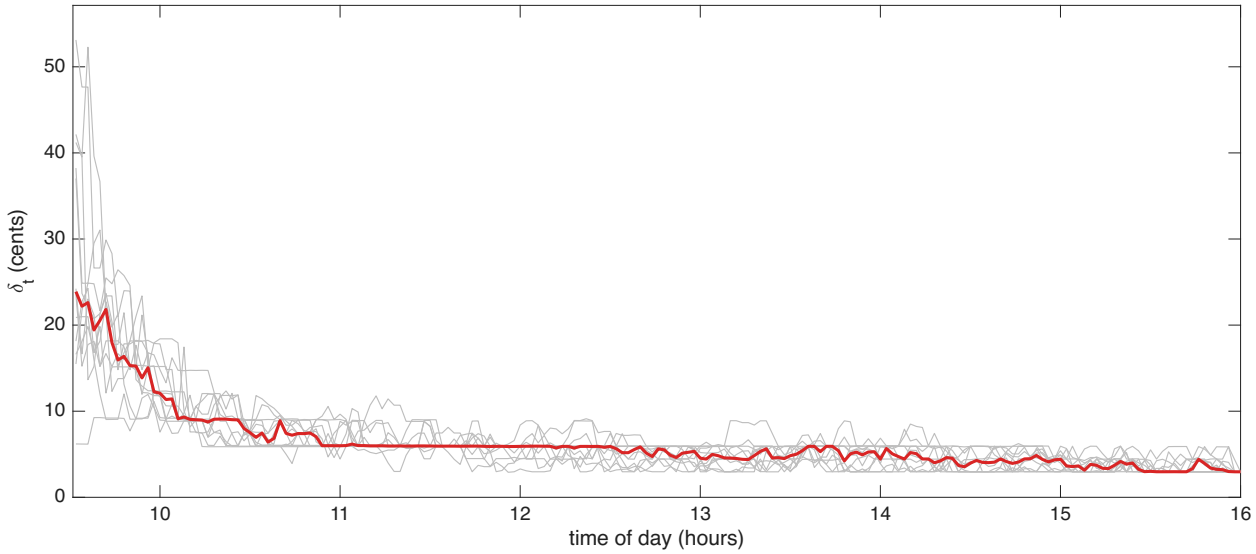

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `CALM` trading days in Oct 2024 (grey lines) and their median (red line).

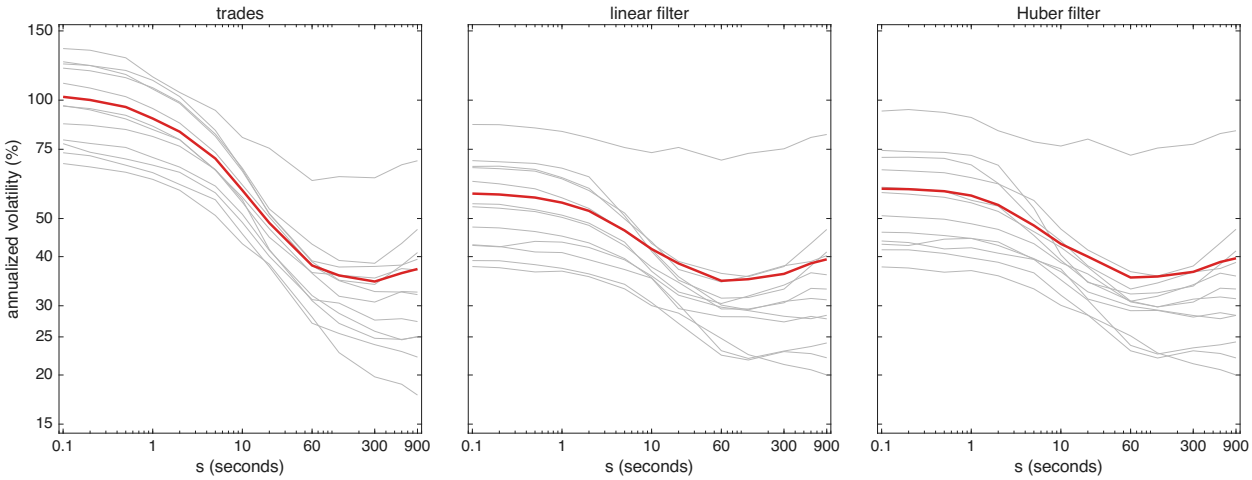

Figure 4: Volatility signature curves for `CALM` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

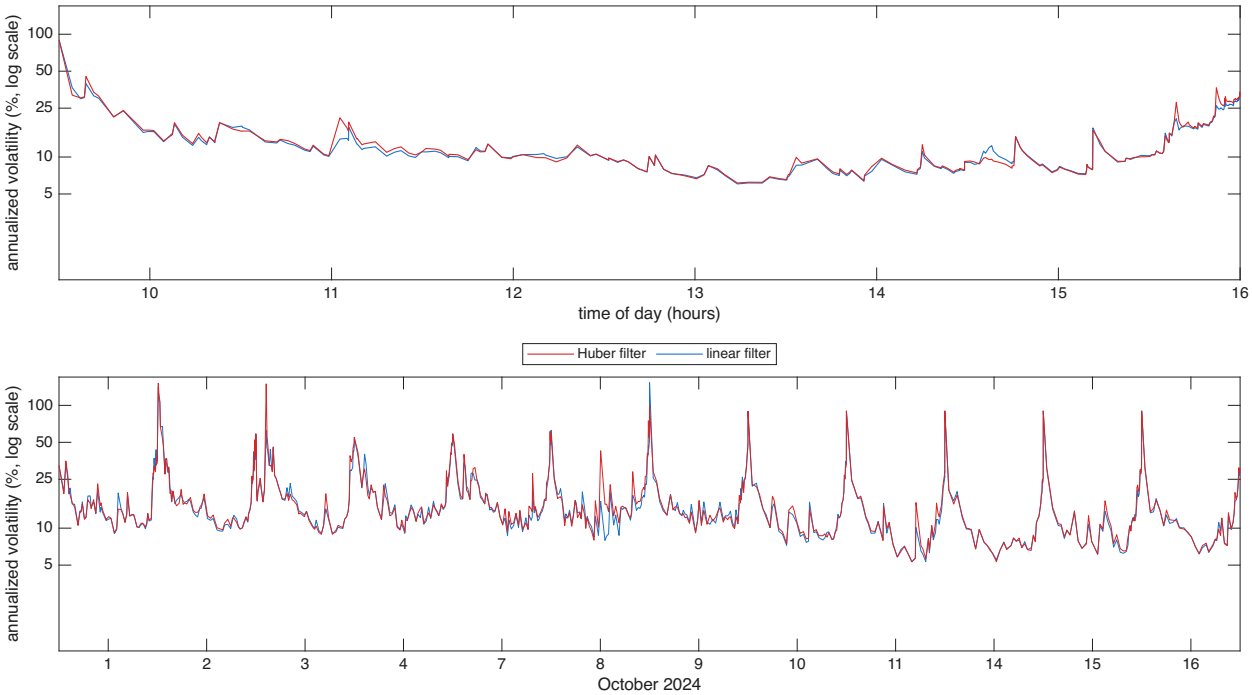

Figure 5: Filtered spot volatility for `CALM` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 18$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

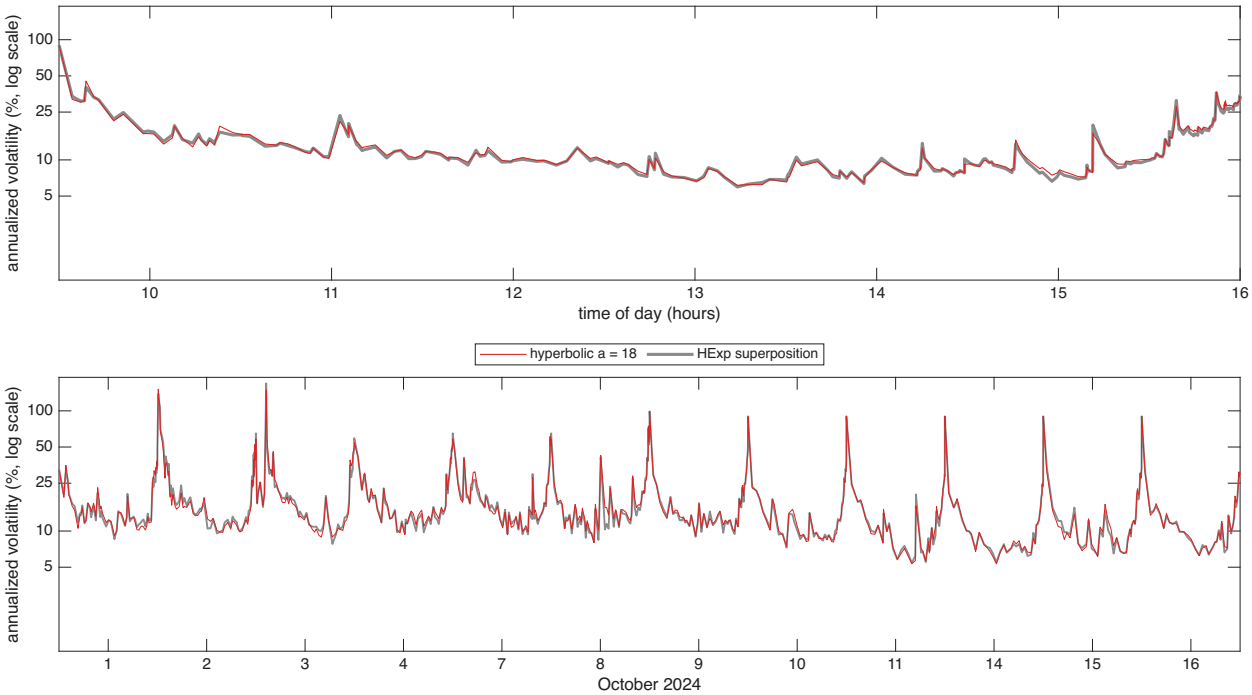

Figure 6: Hyperbolic against superposition weights for `CALM` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 18$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

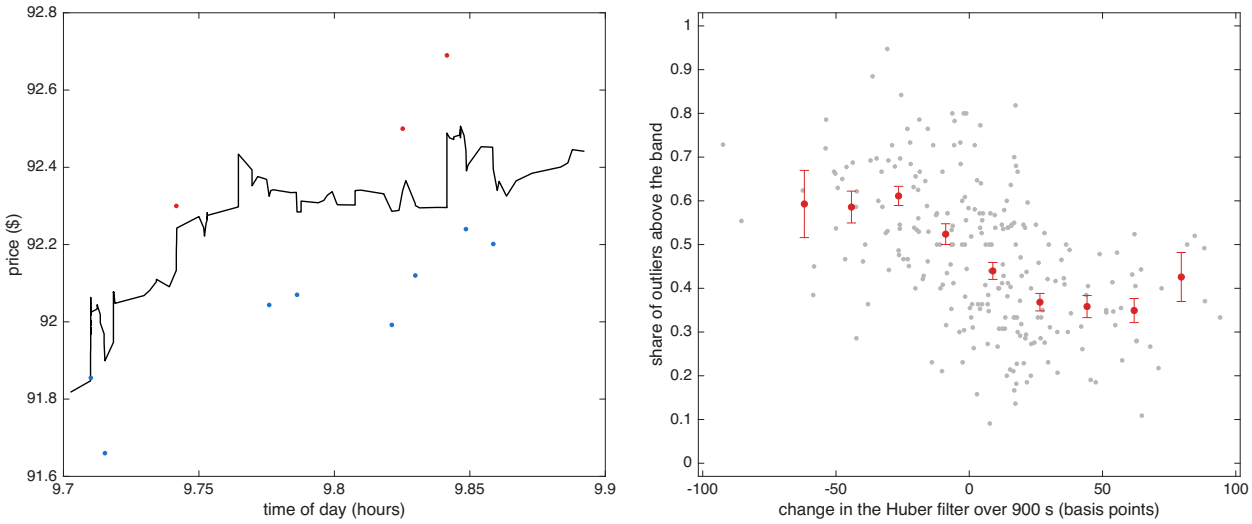

Figure 7: Outliers & robust price for `CALM` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 900 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 900 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

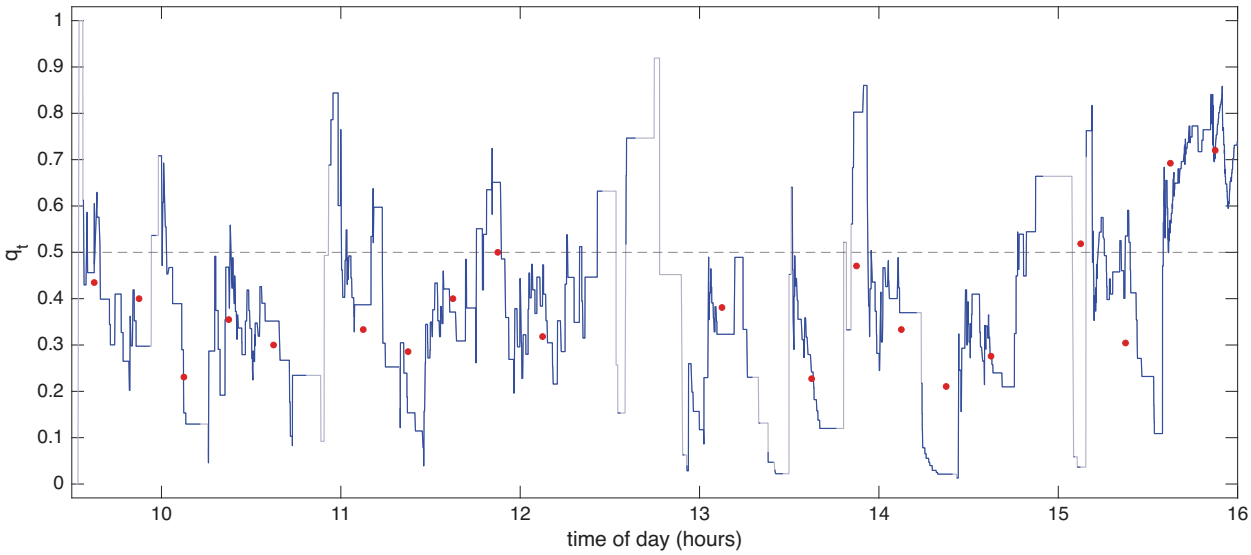

Figure 8: The directional conditional probability for CALM, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 177 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 900 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 3 to 26112 trades, against 18 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 7 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.
On 7 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 47).
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
The largest Huber residual on 16 October 2024 occurs at 09:32, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:49.

Diagnostics.

ORLY 1 to 16 October 2024

trades on 16 October 2024	14,160
average trading rate	0.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,006 (7.10%)
price half-life: run / minimizer[1]	3.0 / 3.0 trades
cost of running 3 trades	+0.18%
scale half-life: run / minimizer[2]	240 / 240 s
cost of running 240 s	+0.00%
direction half-life[3]	335 s [215, 655]
Mandelbrot a minimizing QLIKE	23 trades
a within 0.0005 of the minimum	17 to 32
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	36% $\rightarrow$ 38%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	40% $\rightarrow$ 39%
median spot volatility, Huber / linear input[5]	11.5 / 12.3 %
hyperbolic against HExp, correlation of logs	0.9980
HExp QLIKE advantage, leave-one-day-out[6]	+0.14604
15-minute signature level	15 to 25 %
Huber signature, 1 min over 1 s, 12 day average	1.22

ORLY over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:46; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	8.232	6.296	6.067	29.103	21.373	11.706	10.650	62.829
2	7.541	6.744	6.722	23.526	18.278	12.168	11.133	47.515
5	8.315	7.276	7.274	17.243	16.187	12.353	11.485	30.705
10	8.414	7.558	7.530	14.022	14.777	12.254	11.570	23.520
25	7.825	7.354	7.303	10.753	11.787	10.813	10.574	16.920
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.27	0.08	0.09	−0.43	2.84	2.13	2.49	2.29
2	0.05	−0.02	−0.01	−0.07	3.16	2.39	2.84	2.32
5	0.00	0.05	0.04	0.06	4.29	3.61	3.73	2.65
10	−0.06	−0.01	0.01	0.04	5.03	4.41	5.21	3.00
25	−0.06	−0.04	−0.03	−0.06	4.42	5.77	5.70	2.73

Table 1: Trade-lag differences, `ORLY`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	25.29	22.59	22.41	47.50	27.0	15.7	15.0	93.5
1 s	17.47	15.92	15.73	31.01	24.5	15.7	14.8	84.0
5 s	12.49	11.50	11.50	20.74	21.9	15.3	15.2	63.2
10 s	10.48	9.90	9.98	15.86	21.2	15.8	15.6	49.9
1 min	7.48	7.35	7.32	9.76	18.1	16.6	16.2	30.9
15 min	6.13	6.21	6.20	6.45	16.9	17.1	16.9	20.6
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.01	0.04	0.05	−0.09	4.91	11.64	8.90	2.49
1 s	−0.06	−0.02	−0.01	−0.21	3.79	5.08	5.45	2.95
5 s	−0.18	−0.09	−0.07	−0.40	4.59	8.52	8.29	3.25
10 s	−0.07	−0.03	−0.04	−0.33	3.99	6.94	6.98	3.66
1 min	−0.14	−0.09	−0.11	−0.32				

Table 2: Clock-lag differences, `ORLY`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	18.78	22.40	23.31
s.d. of $\hat{\epsilon}_t$ in cents	43.26	41.26	40.95
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.48
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.10	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.01	−0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.05	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.00	0.01	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	−0.03	−0.03
80% C.I. for $\theta_t - X_t$	—	−12.79,13.35	−13.92,14.52
95% C.I. for $\theta_t - X_t$	—	−23.83,24.93	−26.57,28.01
99% C.I. for $\theta_t - X_t$	—	−44.91,51.94	−50.10,59.62
s.e. of $\theta_t - X_t$ in cents	—	12.75	14.81

Table 3: Residual summaries, `ORLY`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.48, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.51	2.69	2.12 to 3.75	1.92 to 3.93
Residual divided by δ_t	4.72	3.66	2.89 to 6.32	2.71 to 5.10
Repricing windows removed	2.50	2.64	2.14 to 3.75	1.97 to 4.05
Calm windows only	3.02	2.96		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.56	3.18	2.15 to 3.84	1.84 to 4.08

Table 4: Hill estimator of the tail index of the Huber filter residual, `ORLY`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 8 to 11 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 294.6 to 403.1 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 171.3 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

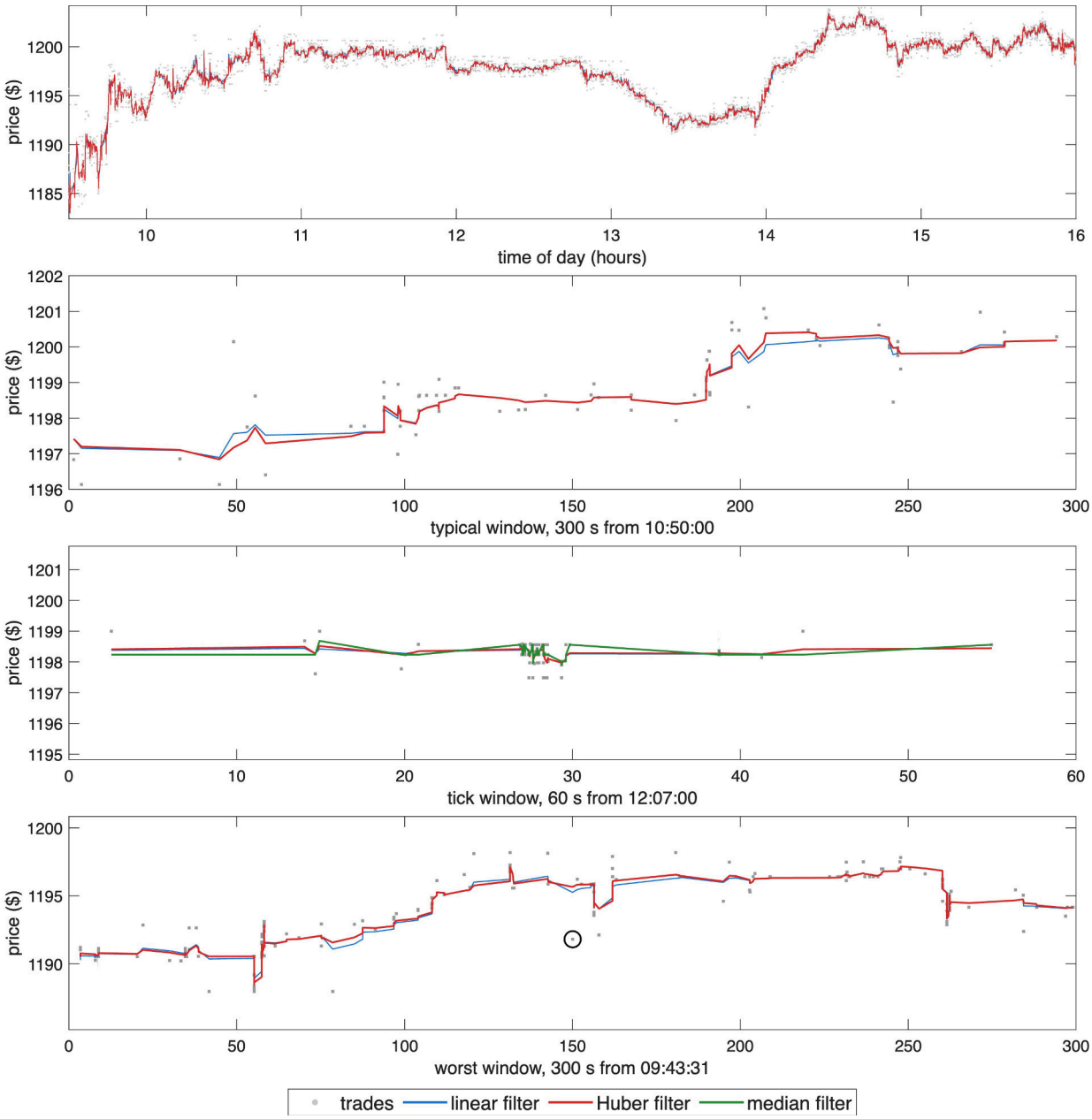

Figure 1: `ORLY` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 10:50:00. 3rd: 60-second window at 12:07:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

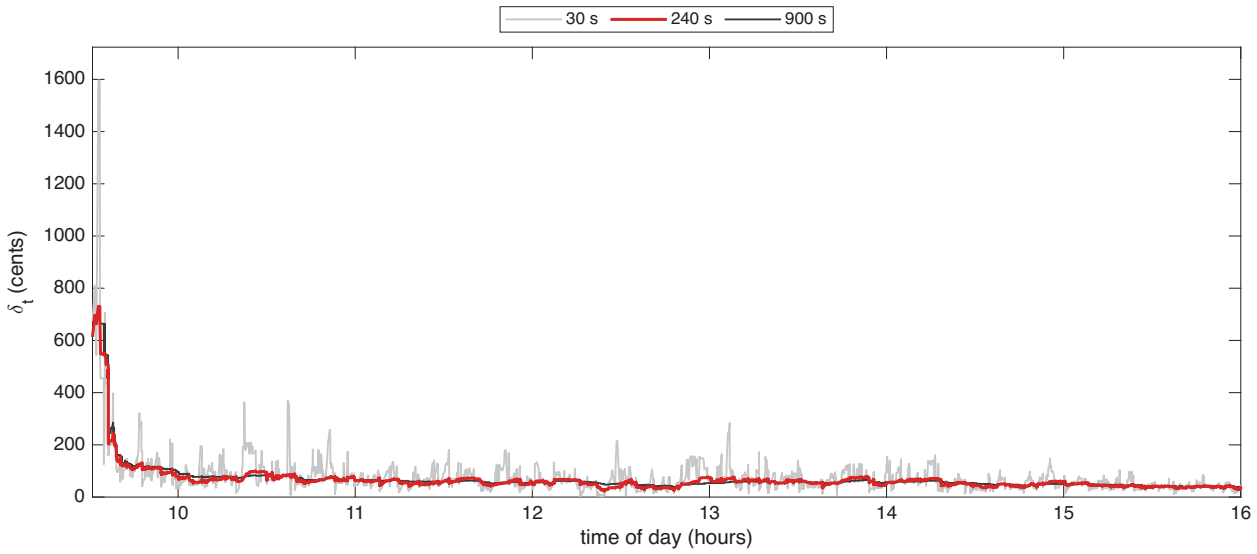

Figure 2: The Huber threshold δ_t in cents for `ORLY` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

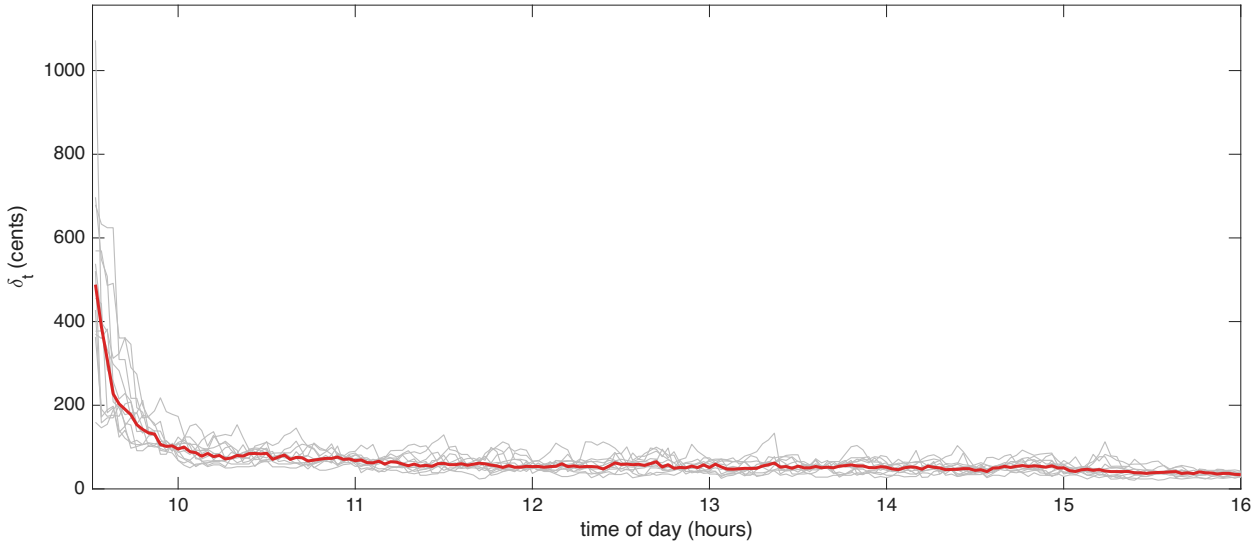

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `ORLY` trading days in Oct 2024 (grey lines) and their median (red line).

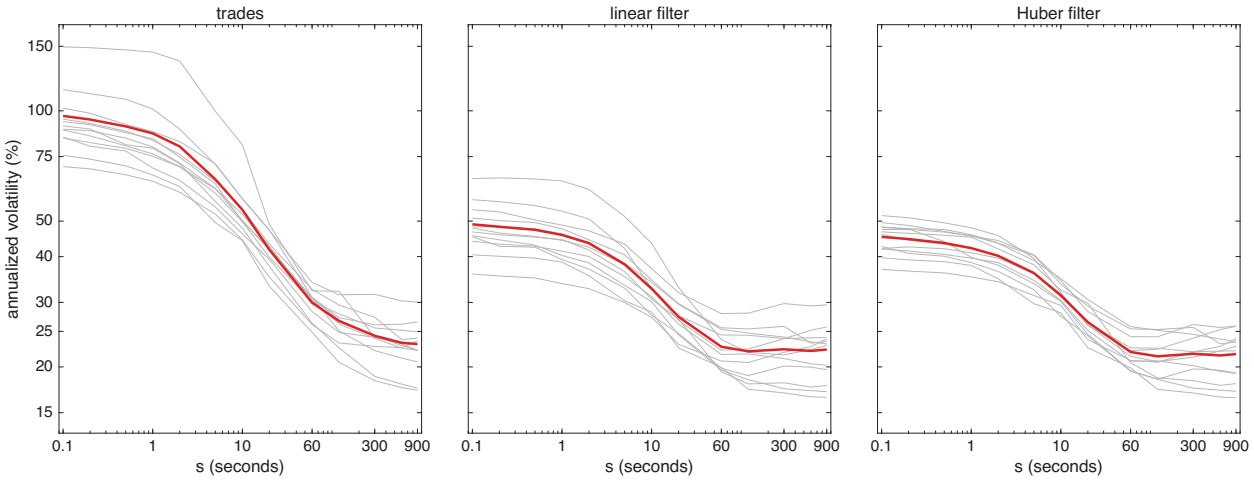

Figure 4: Volatility signature curves for `ORLY` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

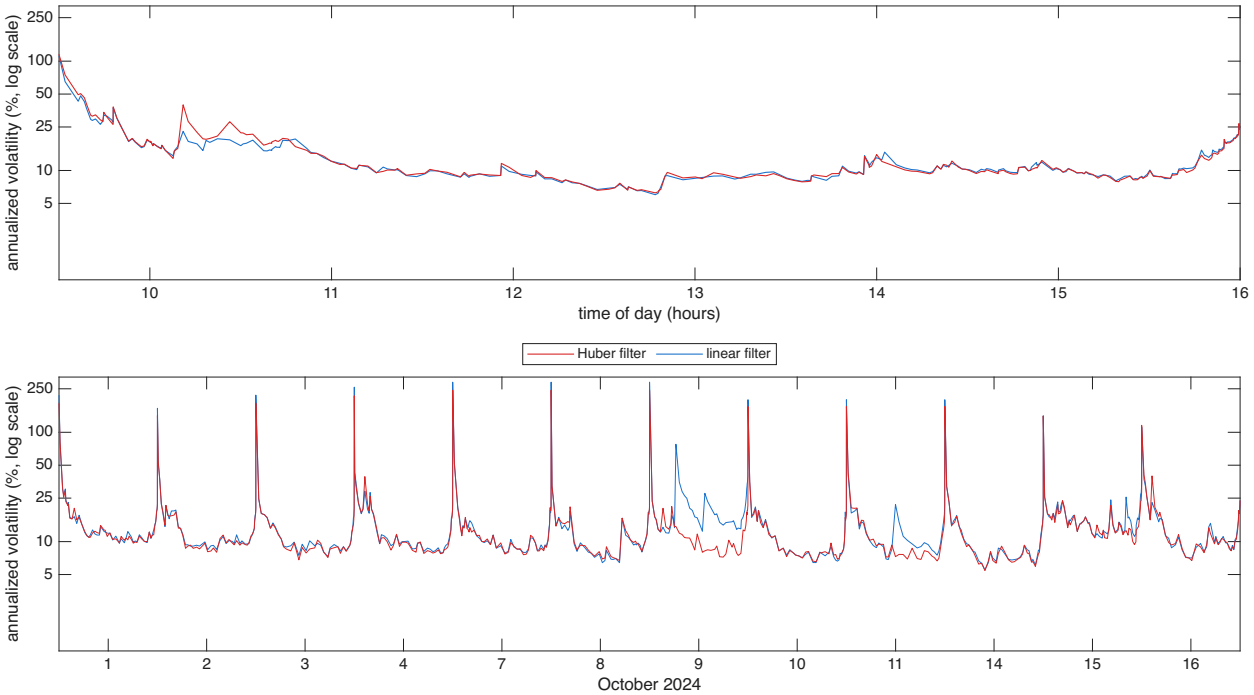

Figure 5: Filtered spot volatility for `ORLY` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 23$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

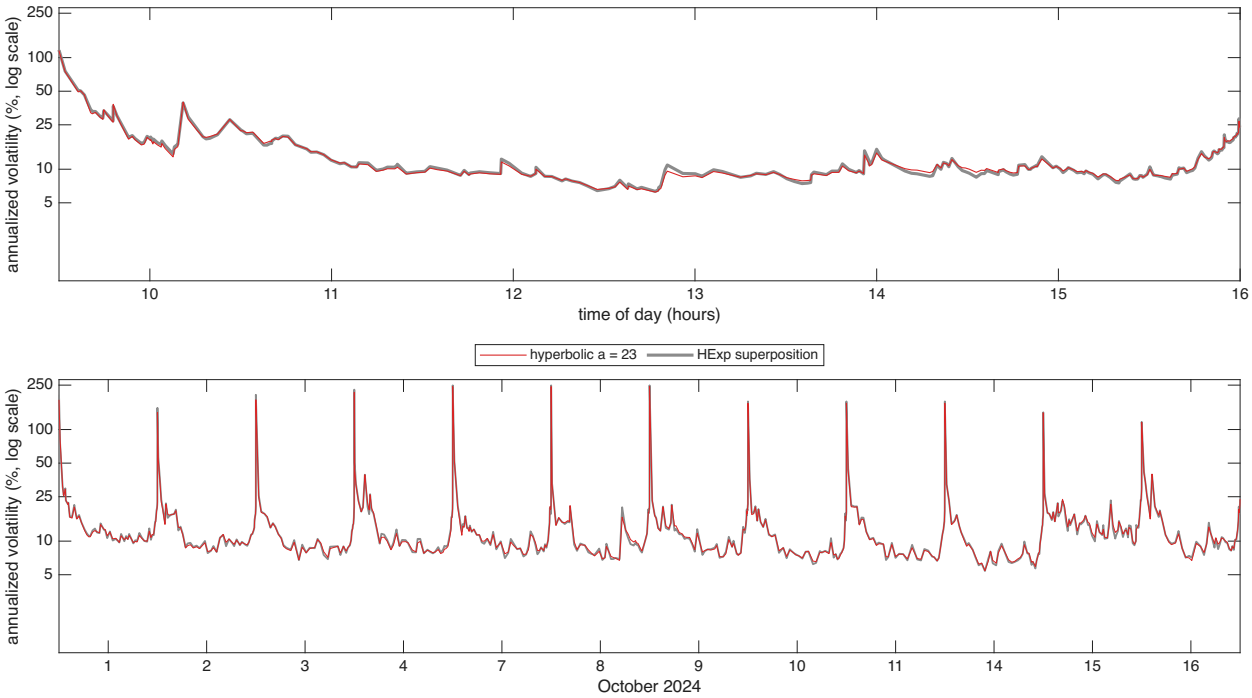

Figure 6: Hyperbolic against superposition weights for `ORLY` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 23$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

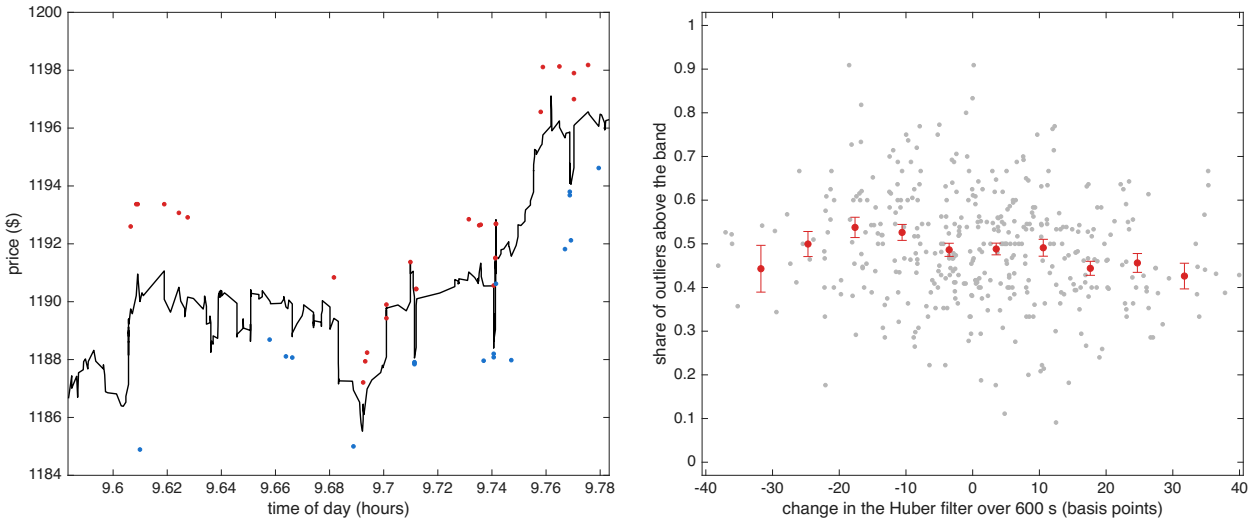

Figure 7: Outliers & robust price for `ORLY` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

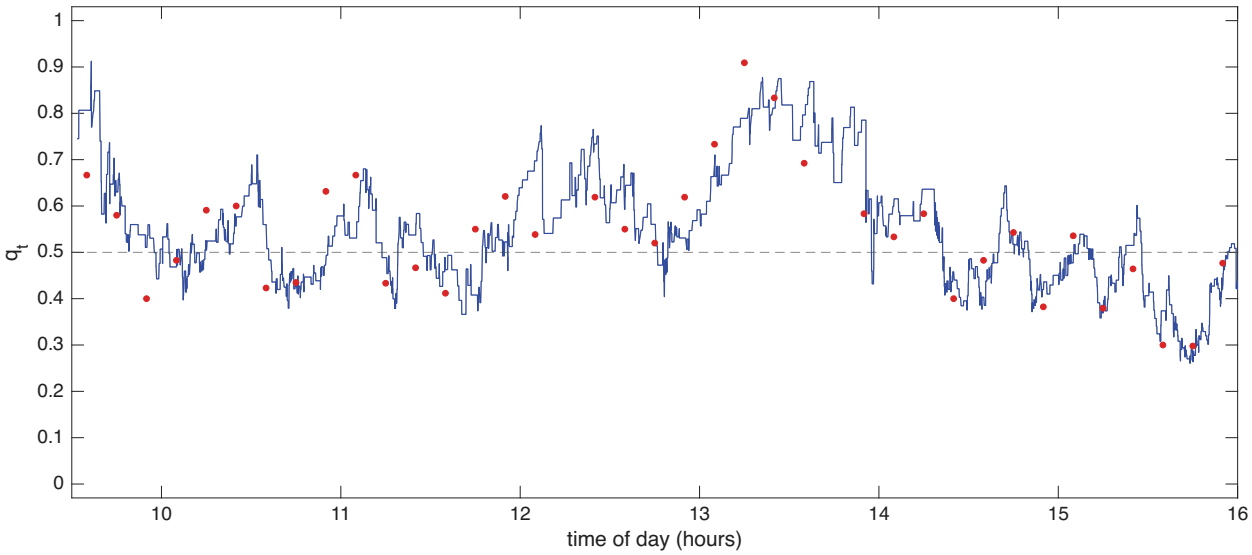

Figure 8: The directional conditional probability for ORLY, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 335 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 9 to 26112 trades, against 23 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 6 trades for this stock.

On 11 October 2024, 35 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 32 per cent on any other day. Every number that day contributes to carries that burst.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:46.

Diagnostics.

WING 1 to 16 October 2024

trades on 16 October 2024	19,408
average trading rate	0.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	1,227 (6.32%)
price half-life: run / minimizer[1]	3.0 / 2.8 trades
cost of running 3 trades	+0.37%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.10%
direction half-life[3]	90 s [79, 106]
Mandelbrot a minimizing QLIKE	37 trades
a within 0.0005 of the minimum	25 to 51
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	53% $\rightarrow$ 61%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	48% $\rightarrow$ 54%
median spot volatility, Huber / linear input[5]	15.8 / 16.0 %
hyperbolic against HExp, correlation of logs	0.9948
HExp QLIKE advantage, leave-one-day-out[6]	+0.18341
15-minute signature level	23 to 44 %
Huber signature, 1 min over 1 s, 12 day average	1.31

WING over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:54; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	3.792	2.982	2.903	13.661	8.896	4.963	5.008	28.904
2	3.626	3.218	3.241	11.203	7.917	5.194	5.174	21.345
5	4.016	3.495	3.533	8.310	7.294	5.467	5.385	14.480
10	4.018	3.582	3.598	6.638	6.604	5.457	5.395	10.923
25	3.830	3.540	3.529	5.138	5.818	5.276	5.254	7.852
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.21	0.09	0.07	−0.45	3.51	3.73	3.10	3.23
2	0.07	0.05	0.04	−0.01	4.13	4.26	4.45	3.40
5	−0.04	−0.03	−0.02	−0.01	4.53	6.44	5.77	3.85
10	0.00	−0.03	−0.02	0.00	4.20	5.36	5.08	4.84
25	0.00	0.03	0.03	0.01	4.55	5.50	5.60	4.61

Table 1: Trade-lag differences, `WING`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	14.85	13.48	13.47	26.33	37.0	23.8	24.9	149.7
1 s	10.73	10.15	10.15	17.99	34.8	23.8	24.5	130.5
5 s	7.93	7.44	7.42	11.74	31.9	25.1	25.2	95.7
10 s	6.69	6.30	6.29	9.55	30.2	25.7	25.6	80.5
1 min	4.58	4.43	4.39	5.53	31.0	29.9	29.6	48.1
15 min	3.90	3.91	3.91	4.12	39.1	39.2	39.1	44.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.05	−0.04	−0.04	−0.13	6.13	6.37	6.31	6.49
1 s	−0.03	−0.03	−0.03	−0.17	7.49	7.72	7.91	9.76
5 s	−0.10	−0.08	−0.07	−0.31	7.69	7.95	8.15	8.58
10 s	−0.10	−0.06	−0.06	−0.37	7.78	7.94	8.13	8.28
1 min	0.11	0.16	0.16	−0.09				

Table 2: Clock-lag differences, `WING`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	9.00	10.76	11.13
s.d. of $\hat{\epsilon}_t$ in cents	20.03	18.99	18.79
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.38
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.17	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.07	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	−0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.03	0.03
80% C.I. for $\theta_t - X_t$	—	−6.73,6.45	−7.30,6.84
95% C.I. for $\theta_t - X_t$	—	−13.24,12.17	−14.71,12.88
99% C.I. for $\theta_t - X_t$	—	−27.01,21.35	−29.20,22.92
s.e. of $\theta_t - X_t$ in cents	—	6.20	6.91

Table 3: Residual summaries, `WING`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.38, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.67	3.55	2.42 to 3.83	2.77 to 4.59
Residual divided by δ_t	4.05	3.59	3.51 to 5.58	2.87 to 4.79
Repricing windows removed	2.63	3.45	2.42 to 3.83	2.77 to 4.59
Calm windows only	2.85	3.41		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.73	4.00	2.42 to 4.16	2.67 to 4.95

Table 4: Hill estimator of the tail index of the Huber filter residual, `WING`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 7 to 12 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 141.1 to 191.4 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 85.8 cents on 16 October 2024, 92% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

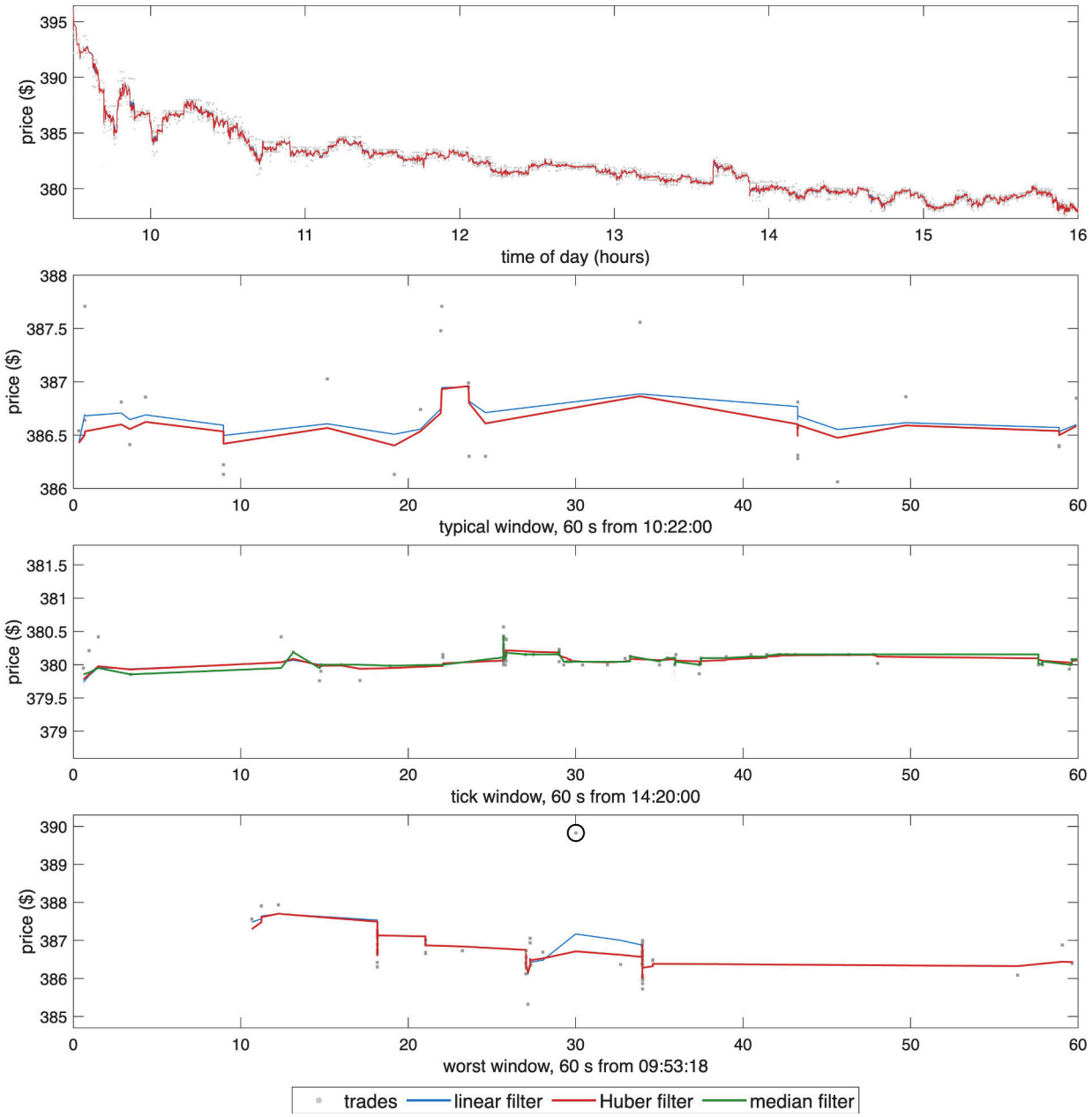

Figure 1: `WING` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 60-second window at the day's median trade count, from 10:22:00. 3rd: 60-second window at 14:20:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 60-second window around the largest Huber residual after 09:45, circled.

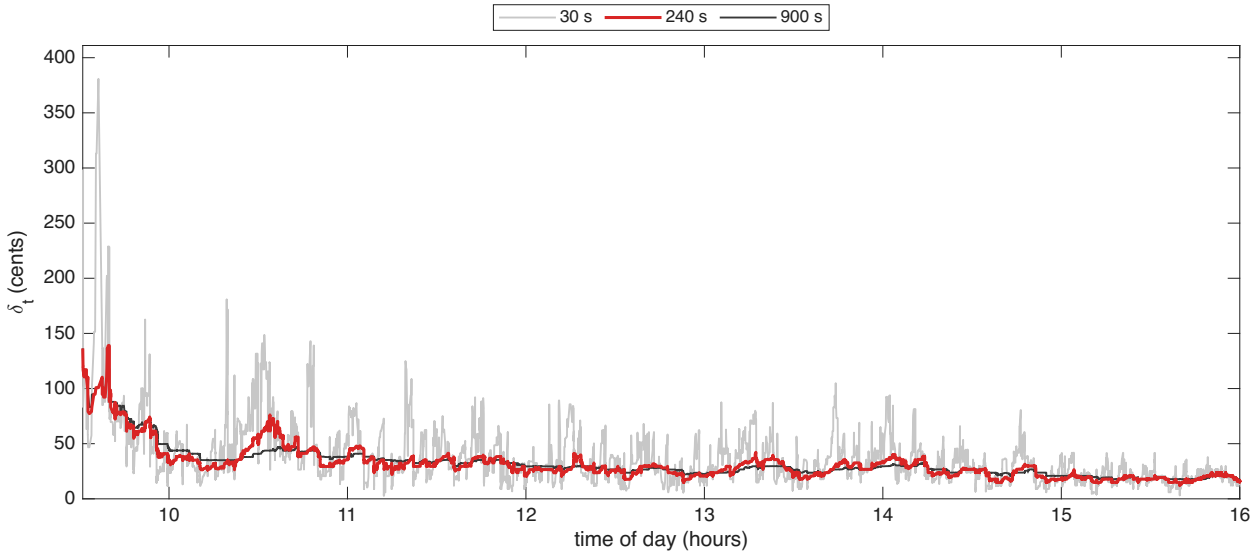

Figure 2: The Huber threshold δ_t in cents for `WING` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

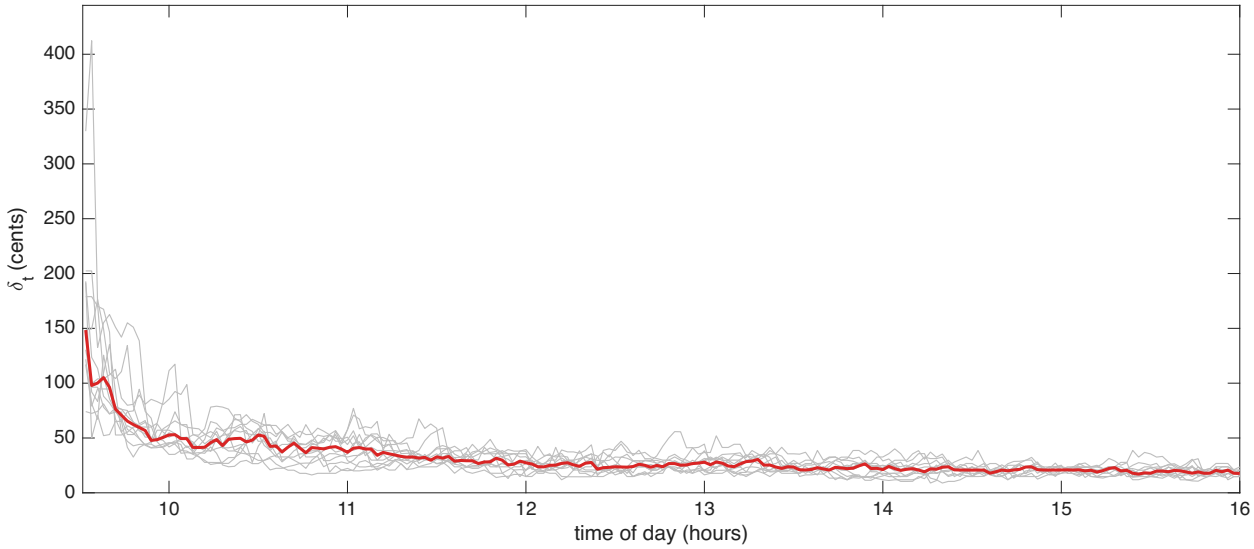

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `WING` trading days in Oct 2024 (grey lines) and their median (red line).

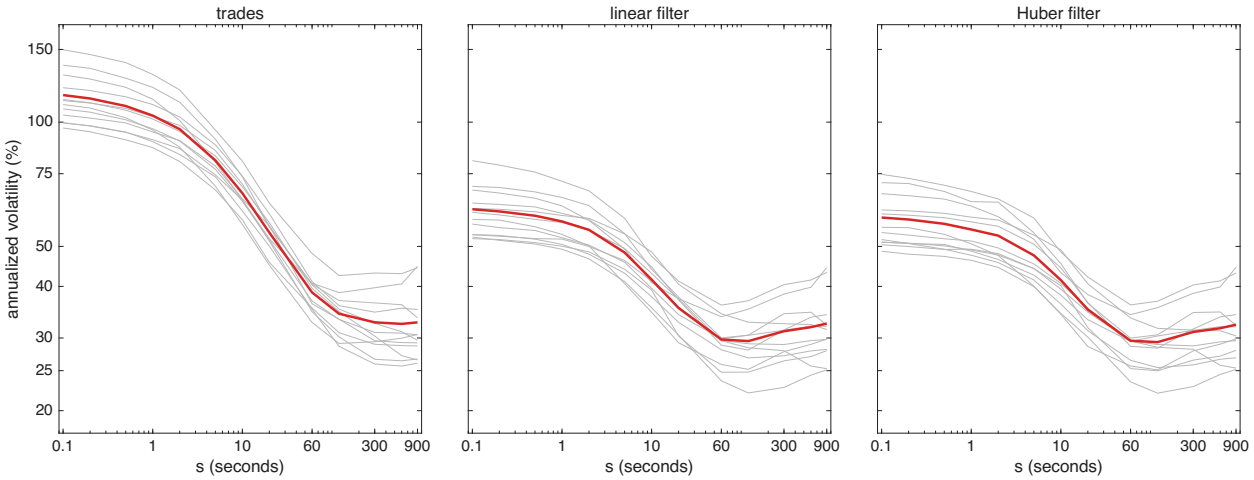

Figure 4: Volatility signature curves for `WING` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

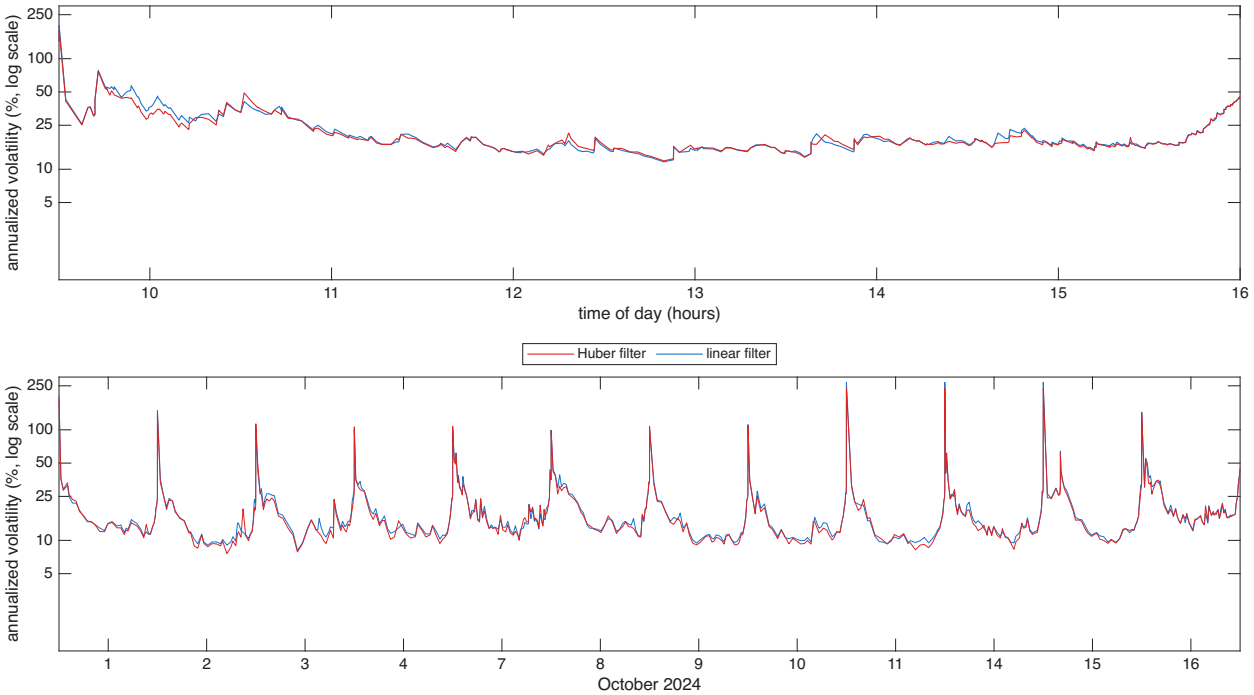

Figure 5: Filtered spot volatility for `WING` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 37$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

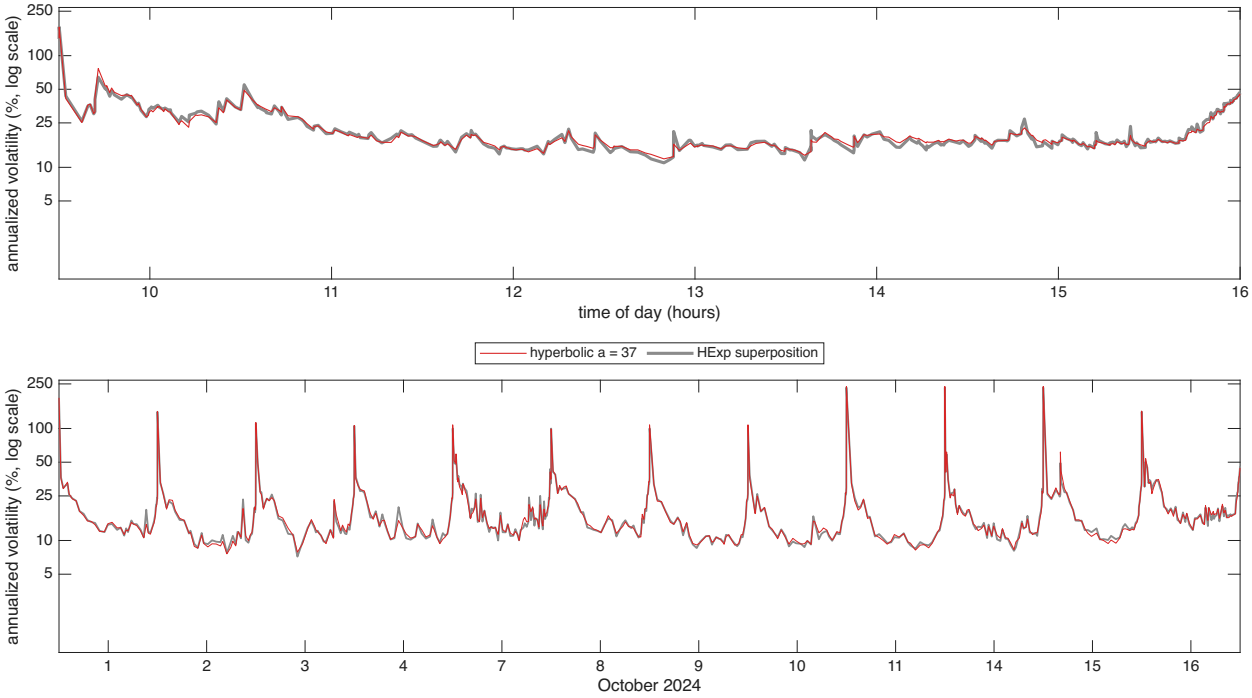

Figure 6: Hyperbolic against superposition weights for `WING` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 37$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

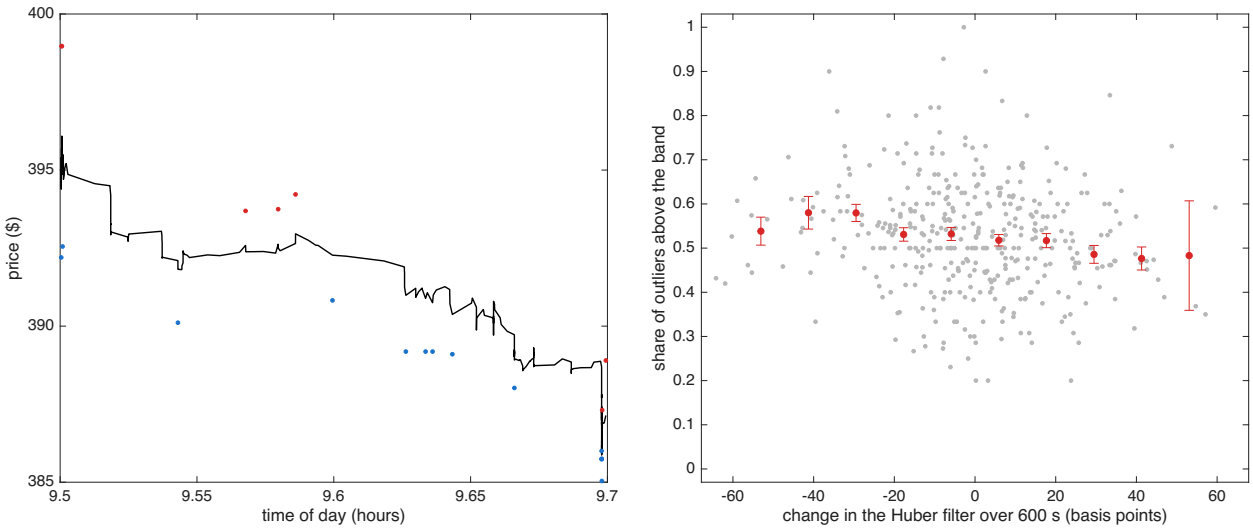

Figure 7: Outliers & robust price for `WING` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

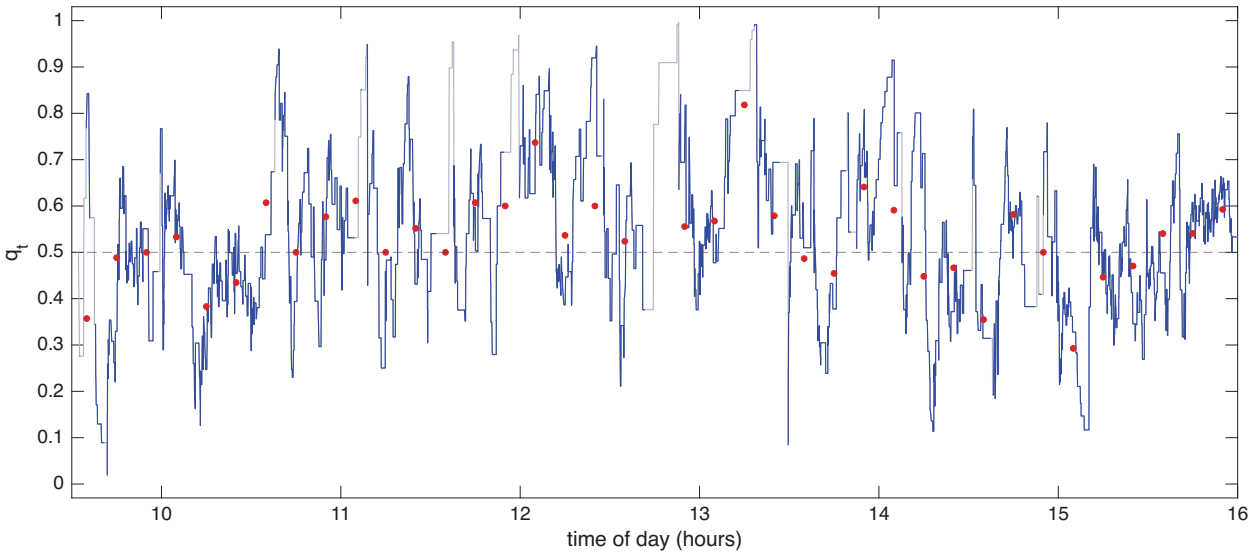

Figure 8: The directional conditional probability for WING, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 90 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 9 to 26112 trades, against 37 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 6 trades for this stock.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:54.

Diagnostics.

BKNG 1 to 16 October 2024

trades on 16 October 2024	11,358
average trading rate	0.6 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	985 (8.67%)
price half-life: run / minimizer[1]	3.0 / 4.0 trades
cost of running 3 trades	+0.81%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.11%
direction half-life[3]	285 s [212, 300]
Mandelbrot a minimizing QLIKE	14 trades
a within 0.0005 of the minimum	10 to 20
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	42% $\to$ 55%
Huber $\hat{\sigma}$, before $\to$ peak	33% $\to$ 45%
median spot volatility, Huber / linear input[5]	16.7 / 18.1 %
hyperbolic against HExp, correlation of logs	0.9986
HExp QLIKE advantage, leave-one-day-out[6]	+0.16295
15-minute signature level	15 to 36 %
Huber signature, 1 min over 1 s, 12 day average	0.98

BKNG over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 09:56; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	48.700	39.499	37.719	190.043	138.955	93.362	71.463	424.512
2	43.212	40.555	41.016	154.929	116.819	87.771	73.421	317.282
5	45.861	41.729	42.543	108.620	95.655	77.235	69.716	201.997
10	44.042	40.131	40.650	83.836	80.254	69.962	65.041	151.682
25	37.994	35.765	35.918	60.166	58.781	53.812	53.036	98.313
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.29	−0.12	0.06	−0.44	1.83	1.50	4.10	2.42
2	−0.02	−0.01	−0.02	−0.01	2.02	1.69	2.57	2.83
5	−0.02	0.04	0.04	0.12	2.53	2.16	3.29	2.94
10	−0.04	−0.15	−0.02	0.00	2.69	2.63	3.71	3.22
25	0.00	0.00	0.01	0.01	2.83	3.25	5.82	11.84

Table 1: Trade-lag differences, `BKNG`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	119.38	105.69	103.90	297.93	37.5	24.9	23.6	139.6
1 s	78.90	74.69	75.81	199.12	32.0	21.5	21.9	120.1
5 s	55.29	53.13	55.19	129.29	25.3	19.2	20.9	91.7
10 s	49.37	48.80	51.09	109.35	24.8	20.1	21.8	73.8
1 min	35.18	37.13	38.84	71.23	19.9	18.9	20.2	41.4
15 min	30.82	34.53	36.03	54.47	20.3	20.5	21.2	28.1
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.12	−0.06	−0.05	−0.15	2.97	2.56	4.36	6.64
1 s	−0.14	−0.07	0.01	−0.15	2.70	2.57	4.00	54.44
5 s	−0.09	−0.01	−0.03	−0.38	3.45	3.20	3.57	4.10
10 s	−0.27	−0.15	−0.17	−0.44	3.05	2.82	4.12	4.66
1 min	−0.13	−0.08	−0.10	−0.37				

Table 2: Clock-lag differences, `BKNG`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	119.93	138.75	144.57
s.d. of $\hat{\epsilon}_t$ in cents	292.27	274.05	271.79
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	3.99
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.19	0.13	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.06	0.01	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.00	0.04
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	−0.01	−0.04	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	−0.02	0.01
80% C.I. for $\theta_t - X_t$	—	−69.81,73.49	−76.90,77.78
95% C.I. for $\theta_t - X_t$	—	−134.97,153.70	−152.42,157.13
99% C.I. for $\theta_t - X_t$	—	−269.20,308.30	−323.47,358.56
s.e. of $\theta_t - X_t$ in cents	—	70.21	86.75

Table 3: Residual summaries, `BKNG`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 3.99, is above two, while the Hill index of the Huber residual itself on this day is 1.90, below two: the two estimates of the noise's index straddle two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	1.90	2.19	1.90 to 4.00	1.81 to 4.63
Residual divided by δ_t	5.47	4.59	3.91 to 5.70	2.90 to 6.15
Repricing windows removed	2.63	2.19	2.63 to 4.64	1.89 to 4.63
Calm windows only	2.54	2.25		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	1.94	2.27	1.94 to 4.45	1.92 to 4.83

Table 4: Hill estimator of the tail index of the Huber filter residual, `BKNG`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 8 to 12 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 1651.1 to 2478.6 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 972.8 cents on 16 October 2024, 98% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

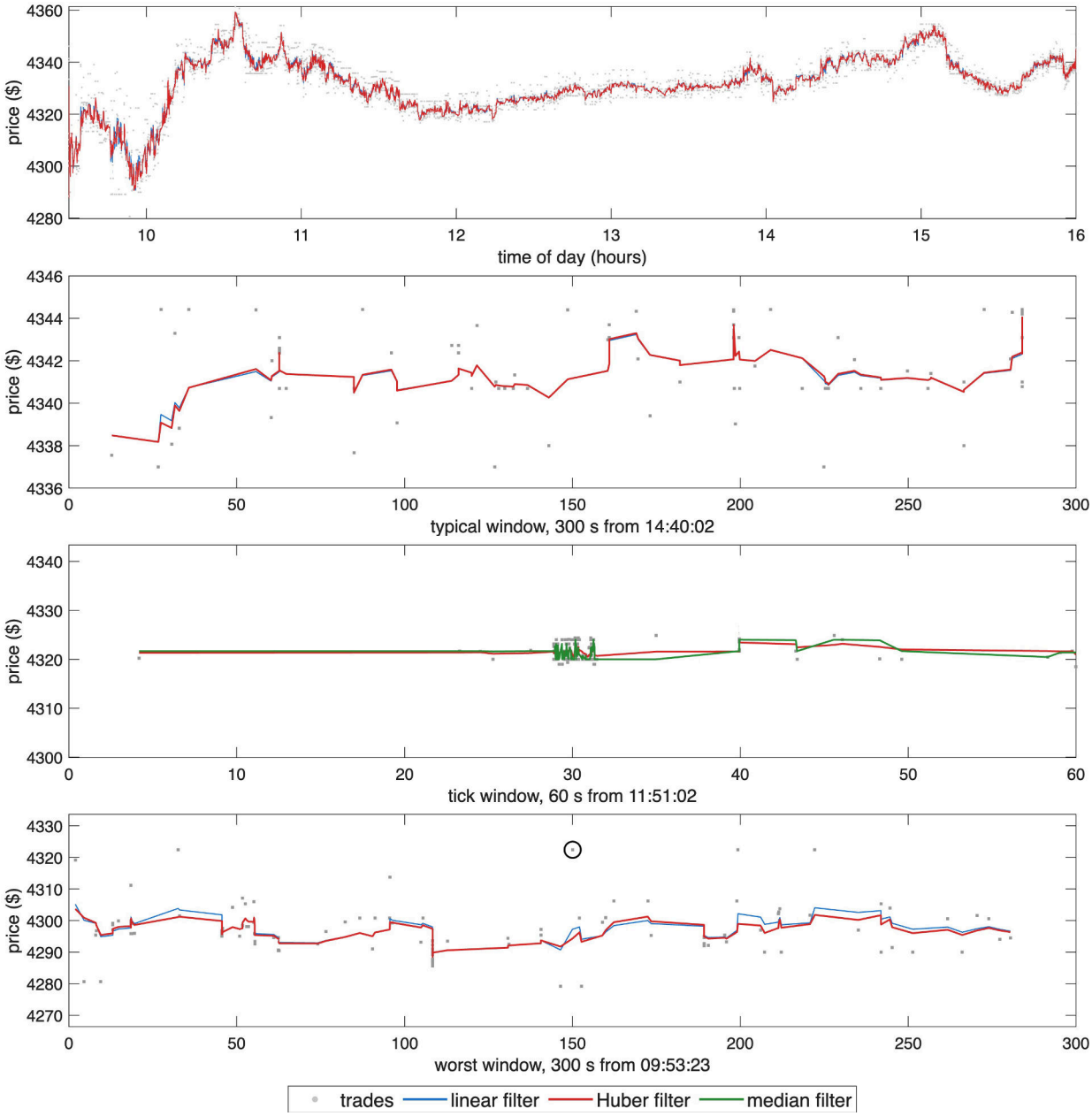

Figure 1: `BKNG` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 14:40:02. 3rd: 60-second window at 11:51:02 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

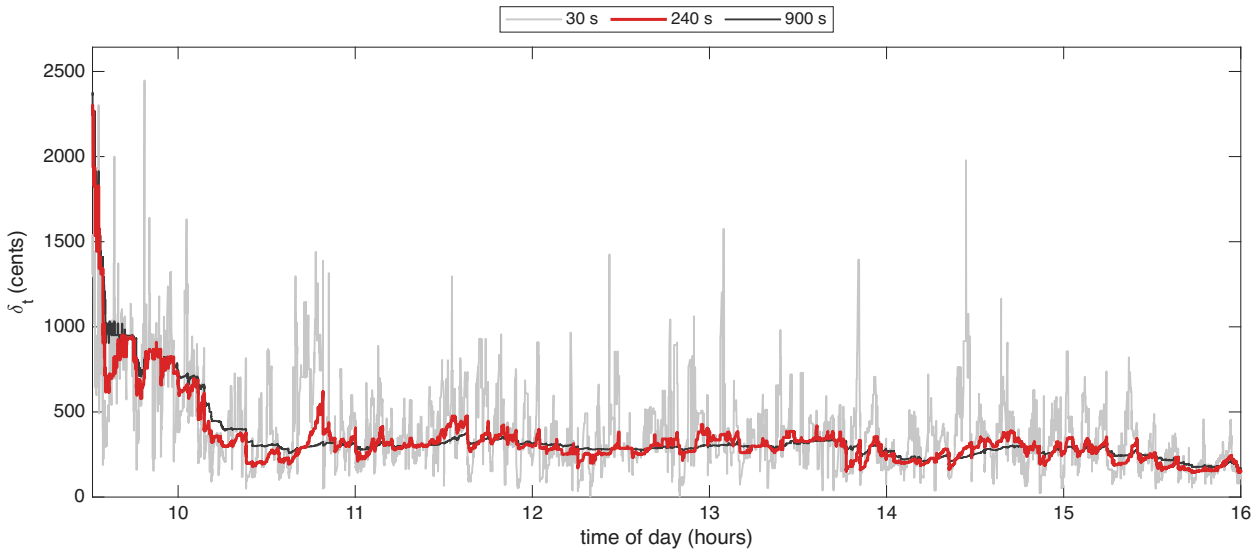

Figure 2: The Huber threshold δ_t in cents for `BKNG` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

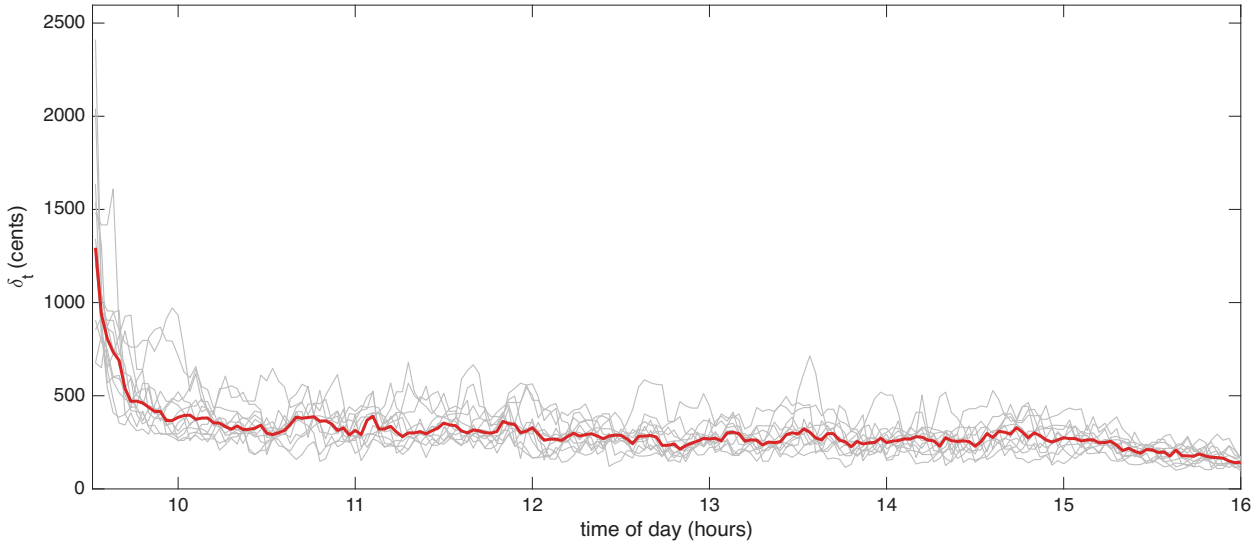

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `BKNG` trading days in Oct 2024 (grey lines) and their median (red line).

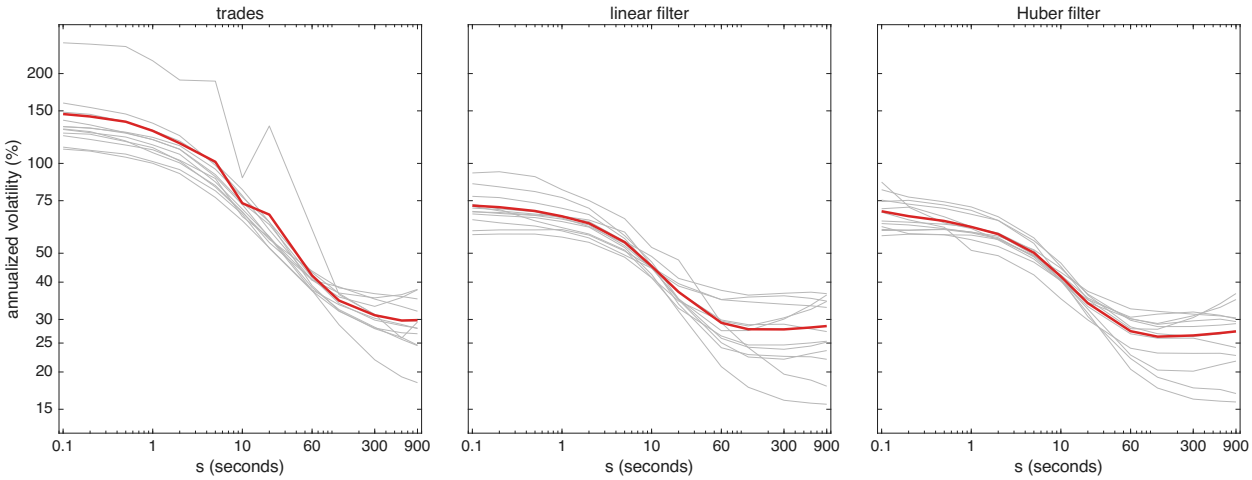

Figure 4: Volatility signature curves for `BKNG` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

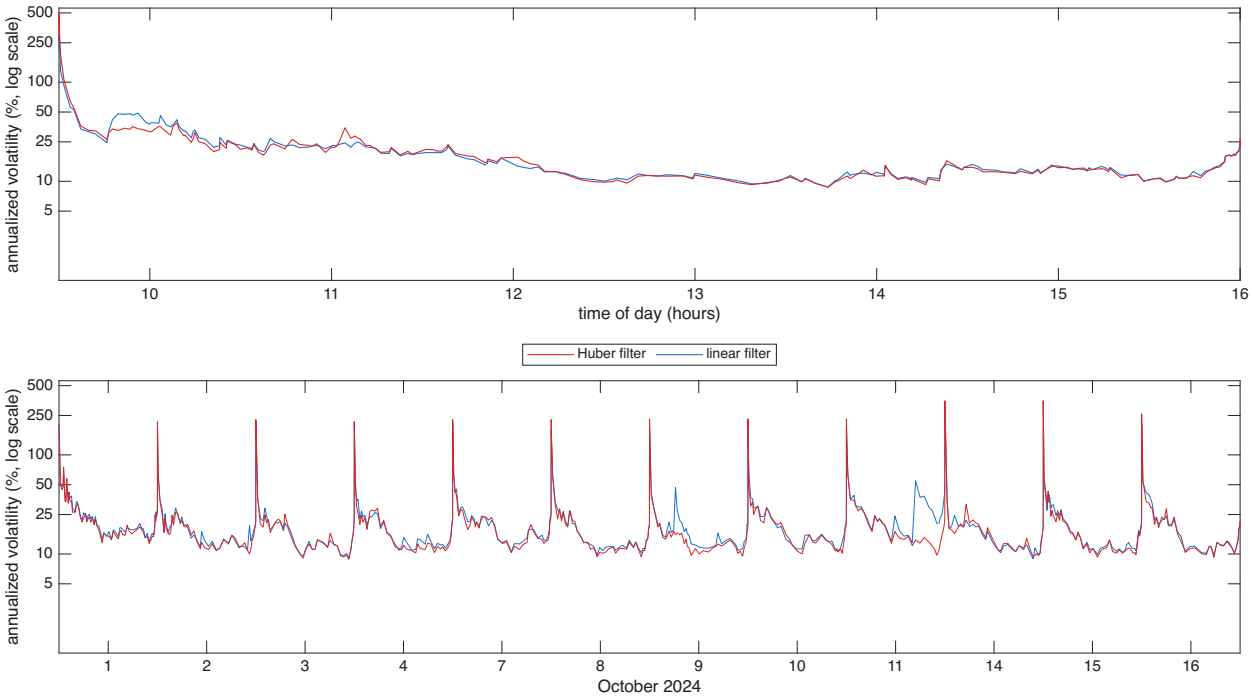

Figure 5: Filtered spot volatility for `BKNG` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 14$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

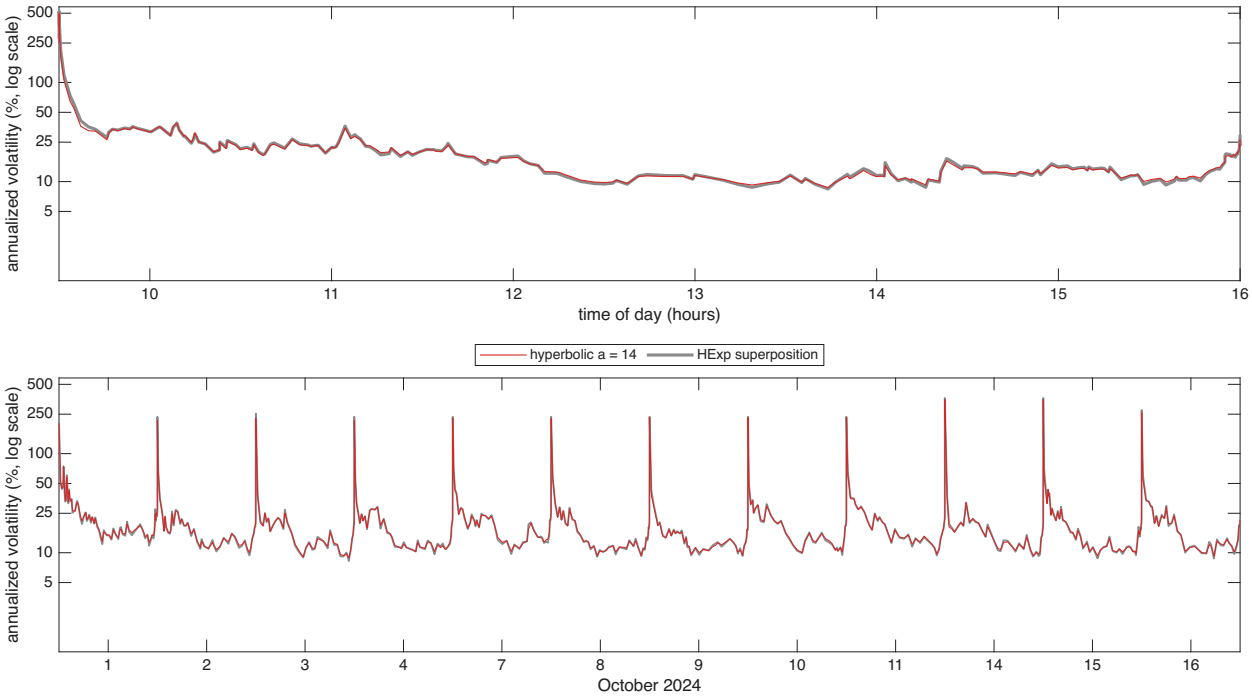

Figure 6: Hyperbolic against superposition weights for `BKNG` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 14$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

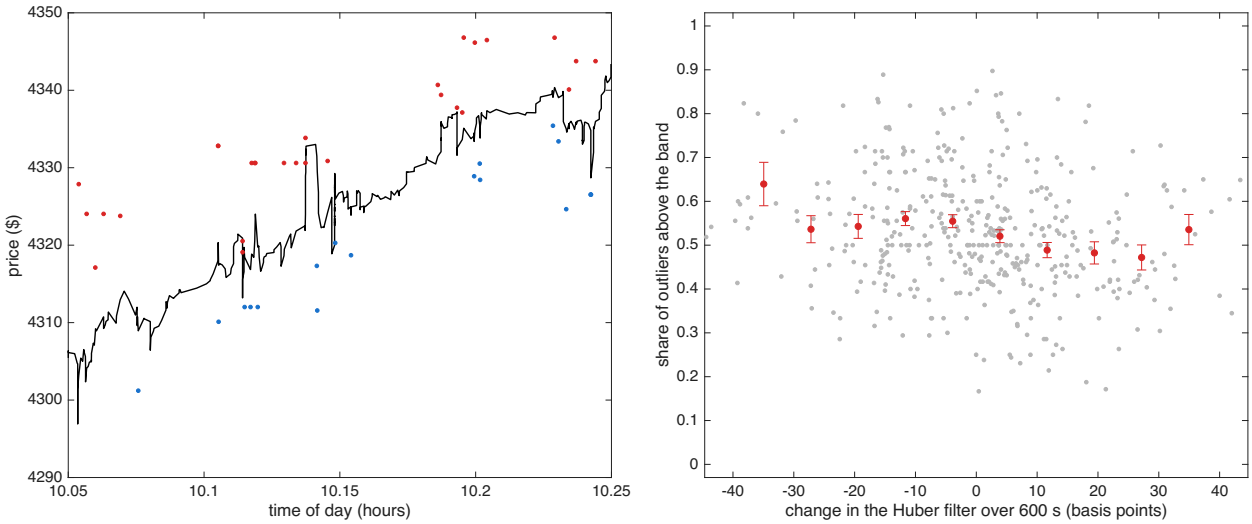

Figure 7: Outliers & robust price for `BKNG` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

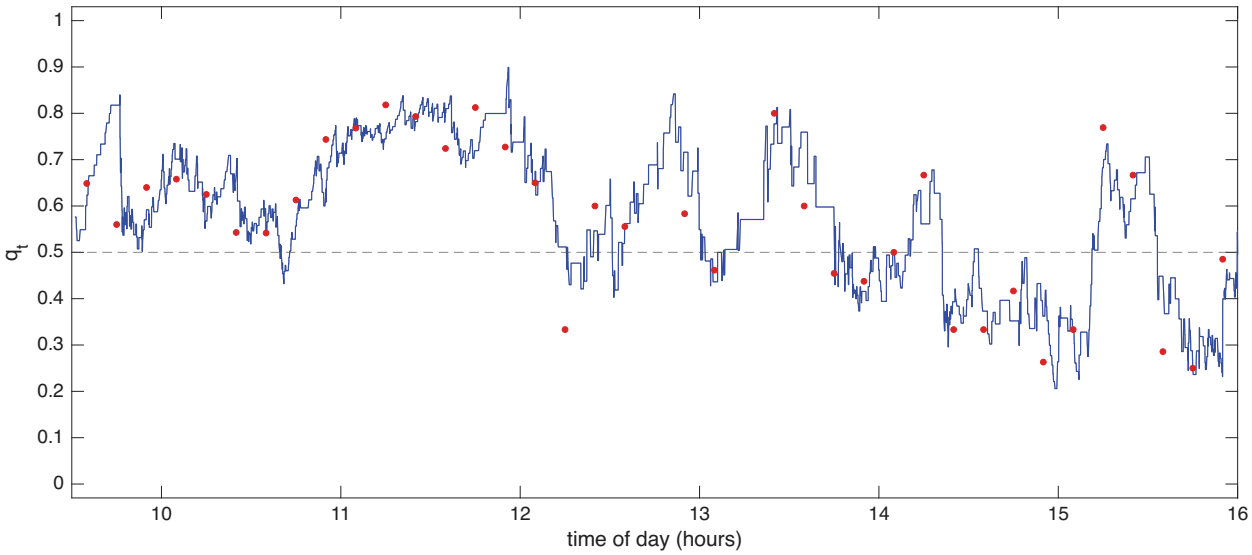

Figure 8: The directional conditional probability for BKNG, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 285 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

At the horizons k = 3, 5, 10 and 30 the half-life minimizing the k-trade-ahead Huber forecast loss is 4, 4, 4 and 4 trades, within one grid point of the end of the 1 to 5 trade grid searched.
The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.
On 16 October 2024, 32 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 30 per cent on any other day. Every number that day contributes to carries that burst.
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 09:56.

Diagnostics.

DUOL 1 to 16 October 2024

trades on 16 October 2024	12,118
average trading rate	0.5 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	791 (6.53%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+0.99%
scale half-life: run / minimizer[2]	240 / 600 s
cost of running 240 s	+0.14%
direction half-life[3]	228 s [170, 300]
Mandelbrot a minimizing QLIKE	25 trades
a within 0.0005 of the minimum	20 to 32
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	21% $\rightarrow$ 24%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	20% $\rightarrow$ 22%
median spot volatility, Huber / linear input[5]	16.4 / 16.7 %
hyperbolic against HExp, correlation of logs	0.9938
HExp QLIKE advantage, leave-one-day-out[6]	+0.23718
15-minute signature level	23 to 45 %
Huber signature, 1 min over 1 s, 12 day average	1.47

DUOL over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:53; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.888	2.264	2.216	9.991	6.949	3.673	3.586	19.984
2	2.773	2.486	2.505	8.265	6.369	4.050	3.885	15.109
5	3.188	2.816	2.838	6.187	6.106	4.553	4.336	10.441
10	3.403	3.043	3.040	5.145	5.828	4.743	4.577	8.131
25	3.536	3.321	3.310	4.323	5.208	4.772	4.731	6.331
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.16	0.22	0.18	−0.43	3.02	3.54	4.23	4.63
2	0.04	0.11	0.10	−0.04	3.10	2.93	3.37	4.66
5	0.00	0.02	0.04	0.00	2.44	2.43	2.69	4.04
10	−0.02	−0.01	0.01	0.02	2.50	2.73	2.90	3.70
25	0.00	0.01	0.02	0.00	3.88	4.96	6.28	5.03

Table 1: Trade-lag differences, `DUOL`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	9.85	8.84	8.62	16.74	30.6	19.1	19.2	108.8
1 s	7.52	6.84	6.77	12.05	29.4	21.5	21.4	97.2
5 s	5.69	5.44	5.41	8.57	28.5	23.5	23.4	77.8
10 s	5.14	4.95	4.90	7.25	30.1	26.2	26.2	67.5
1 min	4.17	4.15	4.16	4.93	36.0	34.6	34.4	46.0
15 min	3.90	4.01	4.01	4.18	38.8	38.5	38.2	40.6
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.02	0.03	0.03	−0.06	4.47	5.12	7.60	11.78
1 s	−0.13	−0.08	−0.07	−0.18	Inf	3.98	4.63	13.23
5 s	−0.04	0.01	−0.01	−0.30	3.84	4.51	5.55	11.32
10 s	−0.03	0.02	0.04	−0.31	4.19	5.22	6.24	7.60
1 min	−0.04	−0.03	−0.03	−0.19				

Table 2: Clock-lag differences, `DUOL`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	6.75	8.12	8.48
s.d. of $\hat{\epsilon}_t$ in cents	14.44	13.62	13.59
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	4.92
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.22	0.21	0.17
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.10	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.04	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.04	0.02	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.02	0.02
80% C.I. for $\theta_t - X_t$	—	−5.76,5.64	−6.39,6.11
95% C.I. for $\theta_t - X_t$	—	−10.73,10.72	−11.85,11.57
99% C.I. for $\theta_t - X_t$	—	−16.78,18.14	−18.40,18.95
s.e. of $\theta_t - X_t$ in cents	—	5.08	5.69

Table 3: Residual summaries, `DUOL`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 4.92, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	4.00	3.11	2.87 to 4.26	2.39 to 3.85
Residual divided by δ_t	3.87	3.10	2.07 to 4.78	2.35 to 4.56
Repricing windows removed	4.05	3.57	2.90 to 4.21	2.68 to 4.00
Calm windows only	4.17	3.38		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	4.12	3.31	3.02 to 4.81	2.67 to 4.35

Table 4: Hill estimator of the tail index of the Huber filter residual, `DUOL`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 7 to 11 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 100.7 to 140.8 cents over the days and removes 0% to 5% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 60.5 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

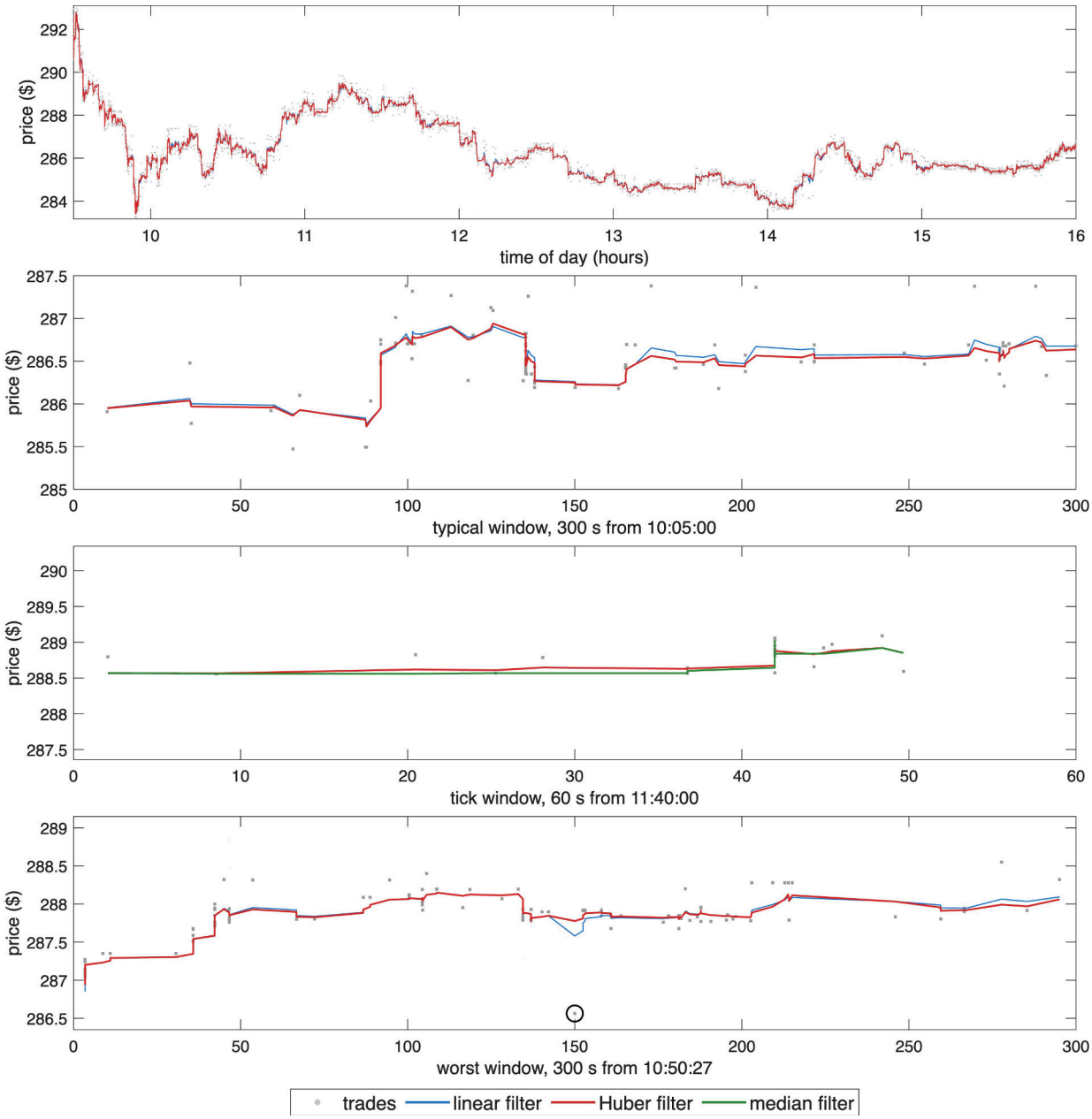

Figure 1: DUOL trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 10:05:00. 3rd: 60-second window at 11:40:00 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

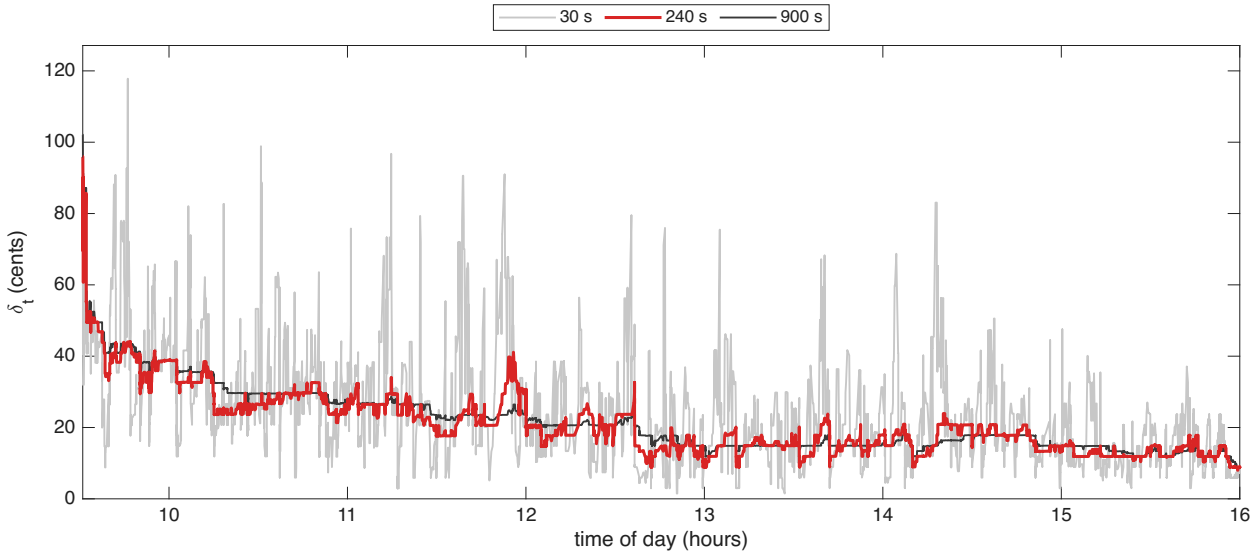

Figure 2: The Huber threshold δ_t in cents for DUOL on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

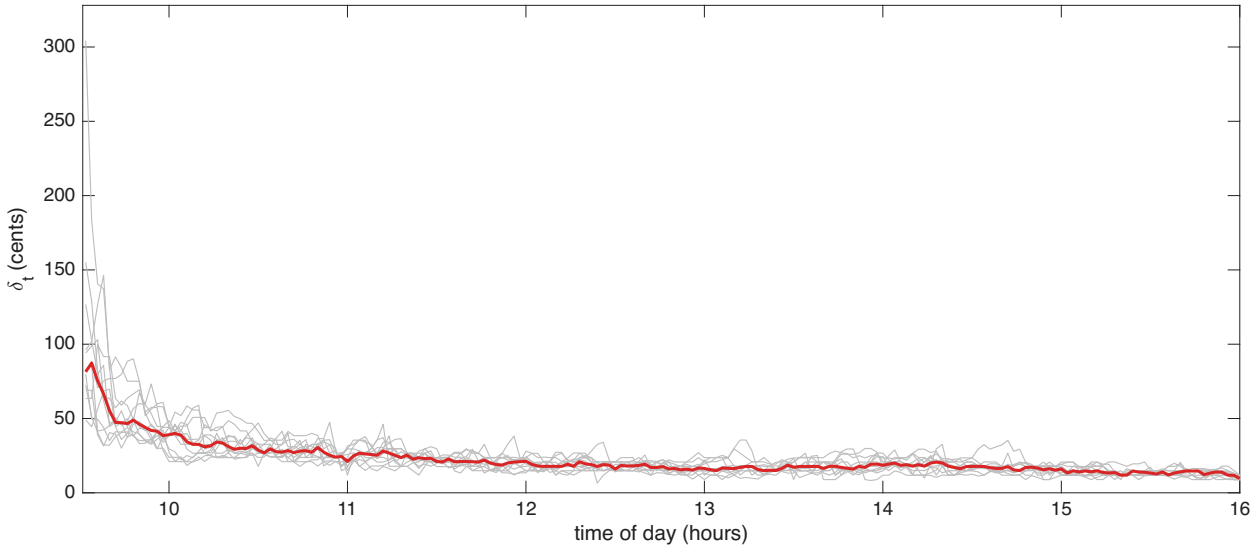

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 DUOL trading days in Oct 2024 (grey lines) and their median (red line).

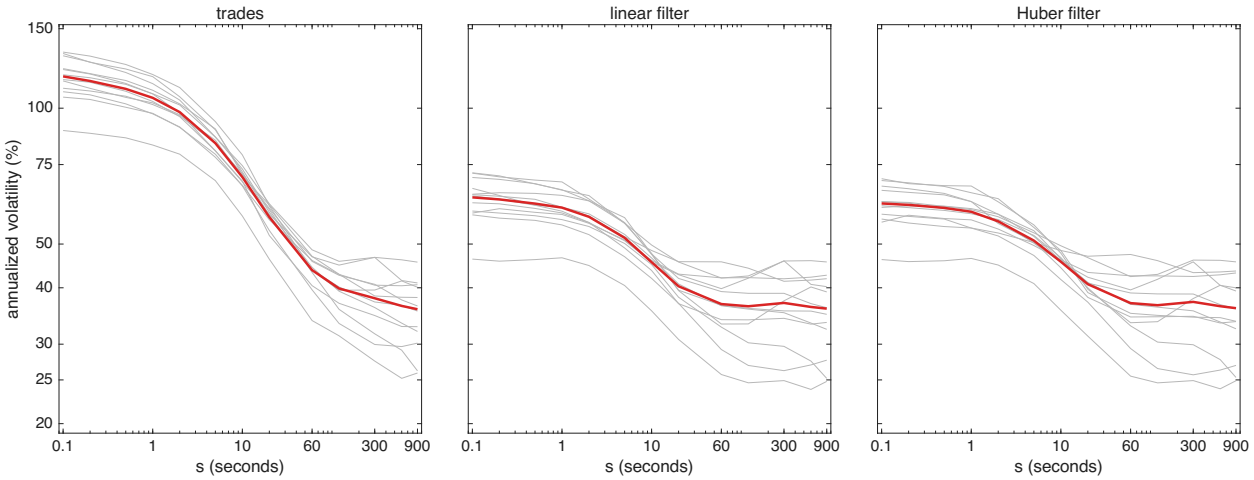

Figure 4: Volatility signature curves for DUOL over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

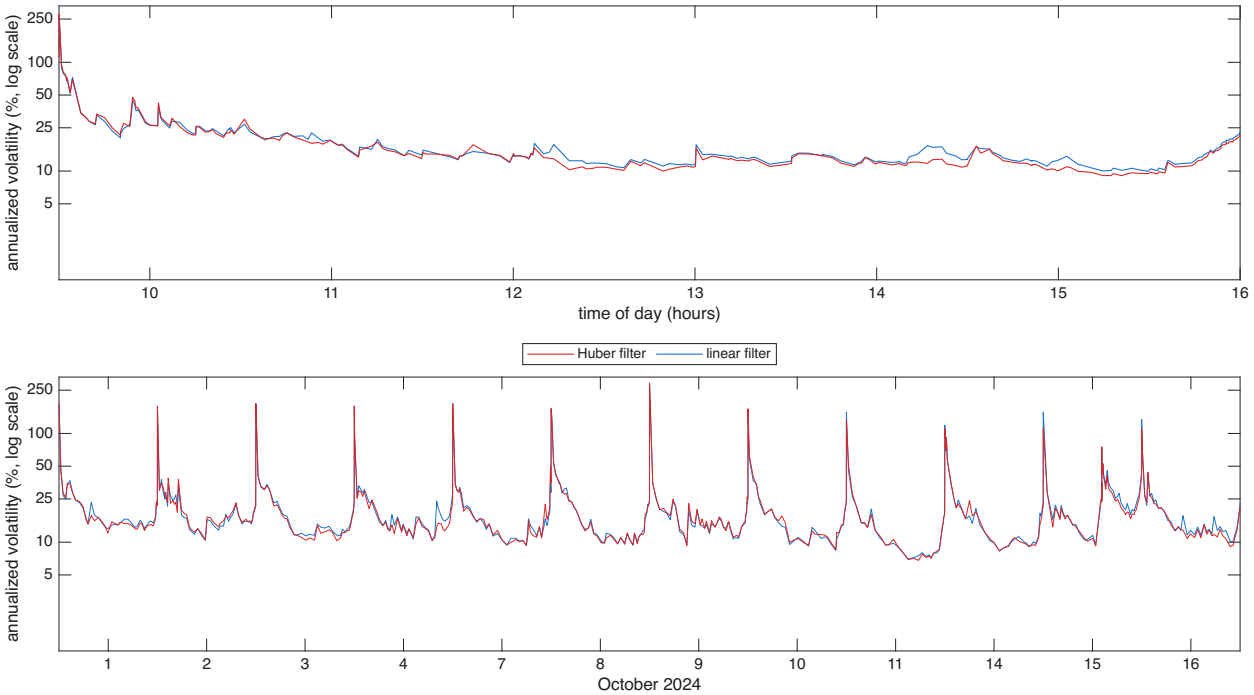

Figure 5: Filtered spot volatility for DUOL over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 25$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

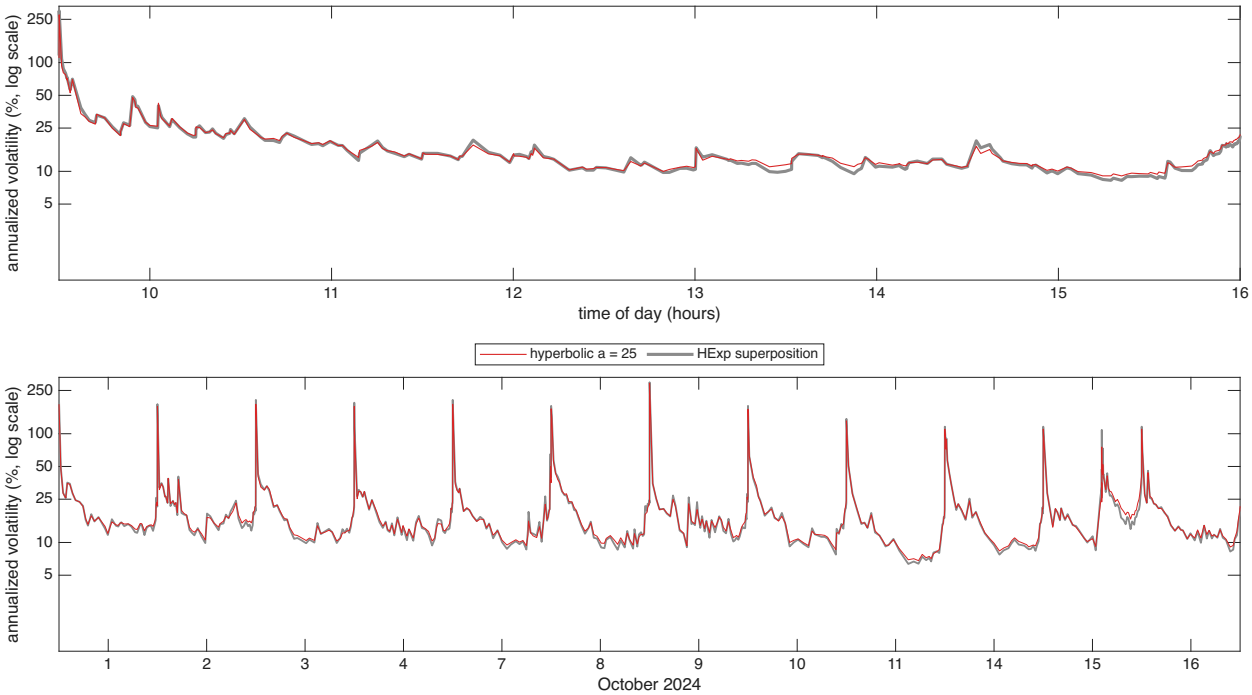

Figure 6: Hyperbolic against superposition weights for `DUOL` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 25$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

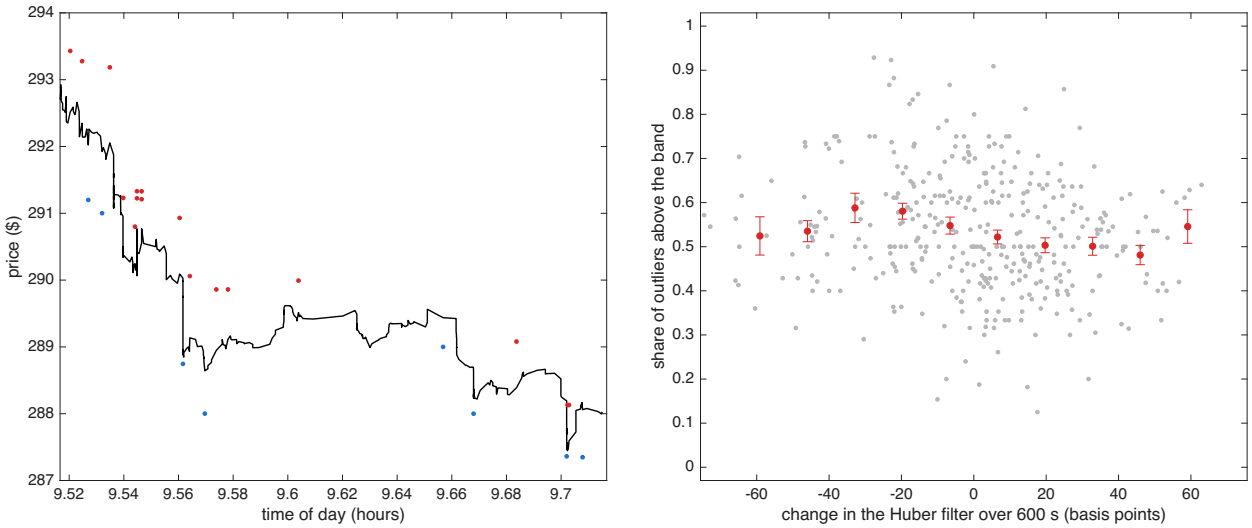

Figure 7: Outliers & robust price for `DUOL` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 600 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 600 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

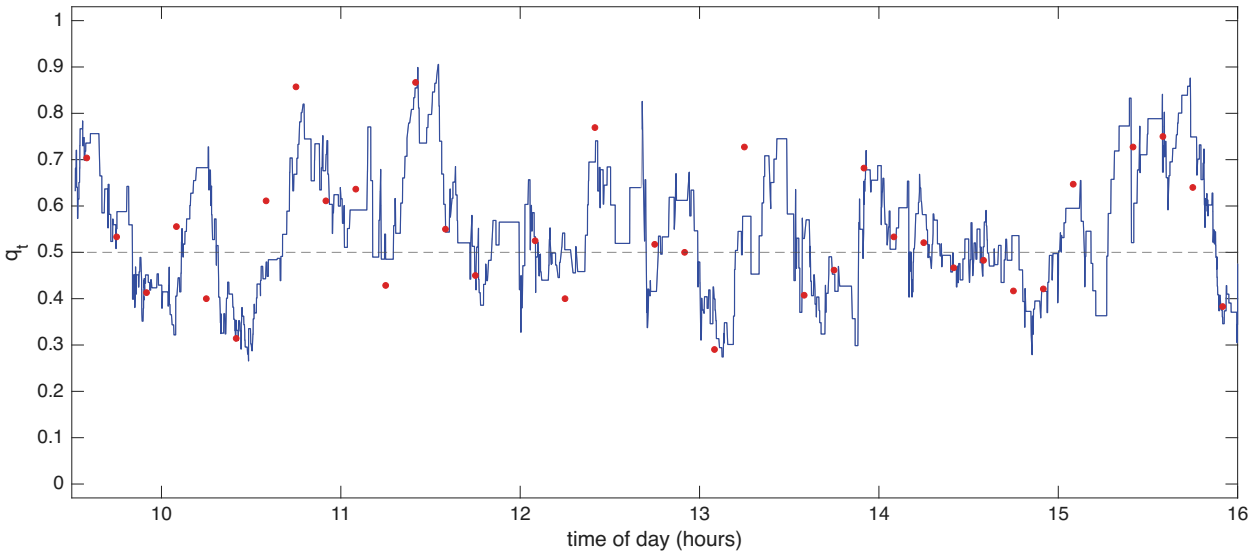

Figure 8: The directional conditional probability for DUOL, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 228 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 600 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 300 seconds, because the log-likelihood was still rising there.

Selecting a on single days gives values from 7 to 26112 trades, against 25 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 5 trades for this stock.

The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.

The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

The largest Huber residual on 16 October 2024 occurs at 09:30, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:53.

Diagnostics.

MIDD 1 to 16 October 2024

trades on 16 October 2024	9,324
average trading rate	0.5 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	467 (5.01%)
price half-life: run / minimizer[1]	3.0 / 2.1 trades
cost of running 3 trades	+2.56%
scale half-life: run / minimizer[2]	240 / 600 s
cost of running 240 s	+0.11%
direction half-life[3]	315 s [251, 410]
Mandelbrot a minimizing QLIKE	21 trades
a within 0.0005 of the minimum	17 to 26
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	7% $\rightarrow$ 12%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	7% $\rightarrow$ 11%
median spot volatility, Huber / linear input[5]	11.0 / 10.7 %
hyperbolic against HExp, correlation of logs	0.9934
HExp QLIKE advantage, leave-one-day-out[6]	+0.35790
15-minute signature level	17 to 36 %
Huber signature, 1 min over 1 s, 12 day average	1.54

MIDD over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:54; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.941	0.687	0.673	2.984	2.378	1.263	1.172	6.354
2	0.877	0.746	0.754	2.488	2.071	1.269	1.204	4.709
5	0.987	0.836	0.841	1.835	1.958	1.390	1.309	3.185
10	1.026	0.905	0.900	1.550	1.763	1.443	1.386	2.534
25	1.079	1.002	0.985	1.301	1.609	1.464	1.428	1.929
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.24	0.01	0.05	−0.45	2.25	2.02	3.08	3.30
2	0.11	0.18	0.13	0.02	2.24	2.41	3.48	5.21
5	−0.07	0.02	0.04	0.03	3.35	2.76	3.33	4.77
10	0.01	0.00	−0.02	−0.03	3.79	3.65	3.79	5.02
25	0.04	0.02	0.02	0.00	4.10	6.17	5.33	3.97

Table 1: Trade-lag differences, `MIDD`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.41	2.92	2.81	4.79	19.9	11.7	11.2	63.8
1 s	2.10	1.85	1.84	3.25	19.1	12.4	11.6	59.5
5 s	1.53	1.35	1.34	2.19	18.3	11.9	11.3	51.3
10 s	1.29	1.17	1.16	1.84	18.2	12.8	12.1	46.4
1 min	1.02	0.95	0.94	1.21	18.8	16.2	15.6	29.8
15 min	0.80	0.79	0.79	0.83	19.2	19.0	18.8	22.8
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.03	0.02	0.02	−0.03	14.42	8.91	18.67	3.81
1 s	−0.05	−0.02	−0.01	−0.19	12.29	8.14	9.26	5.93
5 s	−0.03	0.09	0.07	−0.18	12.35	8.21	9.32	7.00
10 s	−0.05	0.03	0.03	−0.20	14.92	13.72	22.44	10.65
1 min	−0.04	0.02	0.04	−0.20				

Table 2: Clock-lag differences, `MIDD`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	2.01	2.47	2.57
s.d. of $\hat{\epsilon}_t$ in cents	4.50	4.27	4.24
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	4.49
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.17	0.15
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.08	0.12	0.09
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	0.06	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.02	0.00	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.02	0.03	0.03
80% C.I. for $\theta_t - X_t$	—	−1.68,1.56	−1.81,1.63
95% C.I. for $\theta_t - X_t$	—	−2.92,3.01	−3.27,3.21
99% C.I. for $\theta_t - X_t$	—	−4.80,6.23	−5.75,6.69
s.e. of $\theta_t - X_t$ in cents	—	1.52	1.73

Table 3: Residual summaries, `MIDD`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 4.49, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.37	2.64	1.93 to 3.42	2.06 to 3.84
Residual divided by δ_t	2.37	4.28	2.37 to 4.86	2.52 to 4.28
Repricing windows removed	2.31	2.78	2.31 to 3.82	2.14 to 3.97
Calm windows only	2.09	2.81		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.81	2.95	1.99 to 4.13	2.24 to 3.59

Table 4: Hill estimator of the tail index of the Huber filter residual, `MIDD`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 6 to 9 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 40.2 to 50.6 cents over the days and removes 0% to 3% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 20.2 cents on 16 October 2024, 91% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

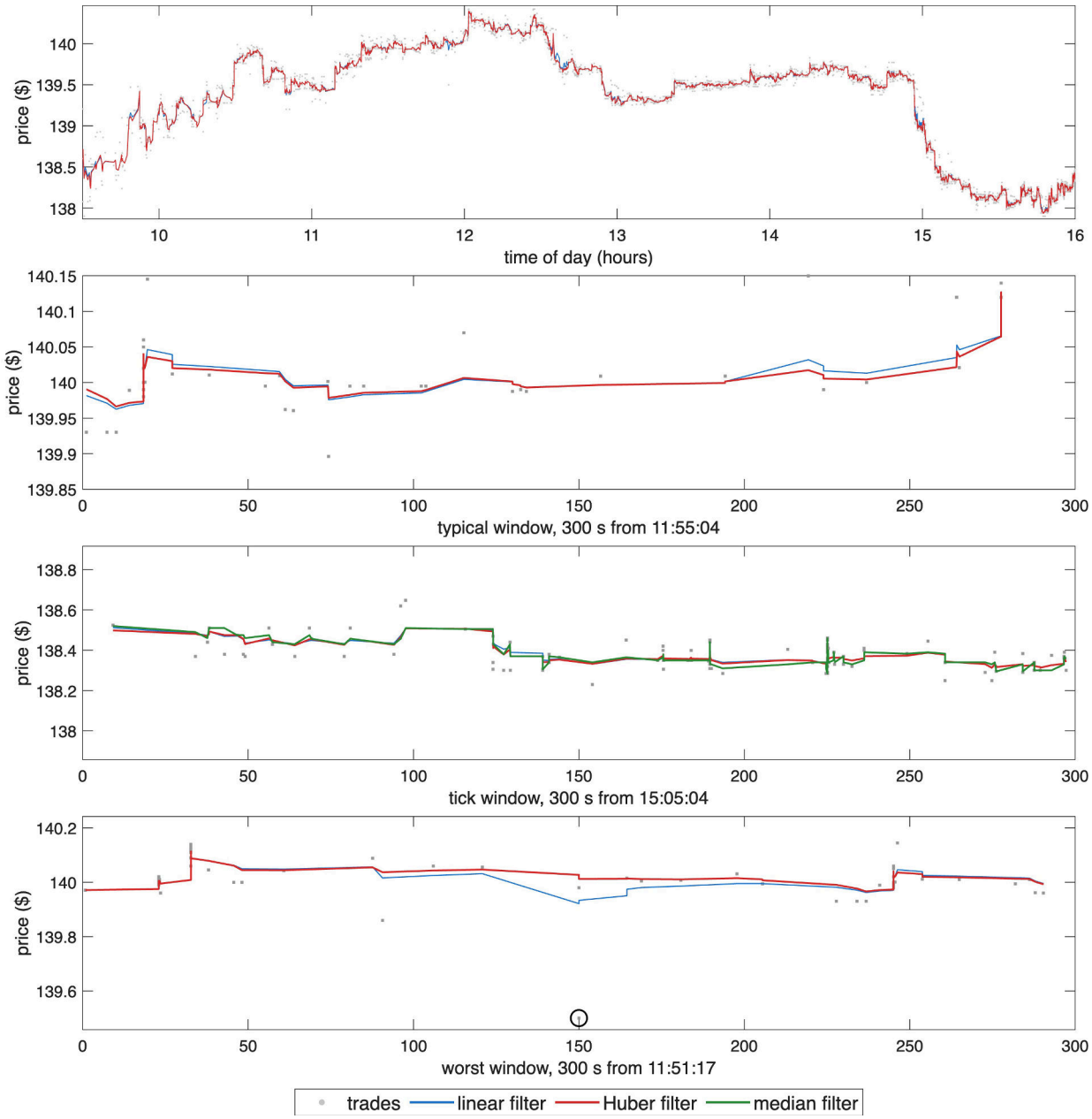

Figure 1: `MIDD` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 11:55:04. 3rd: 300-second window at 15:05:04 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

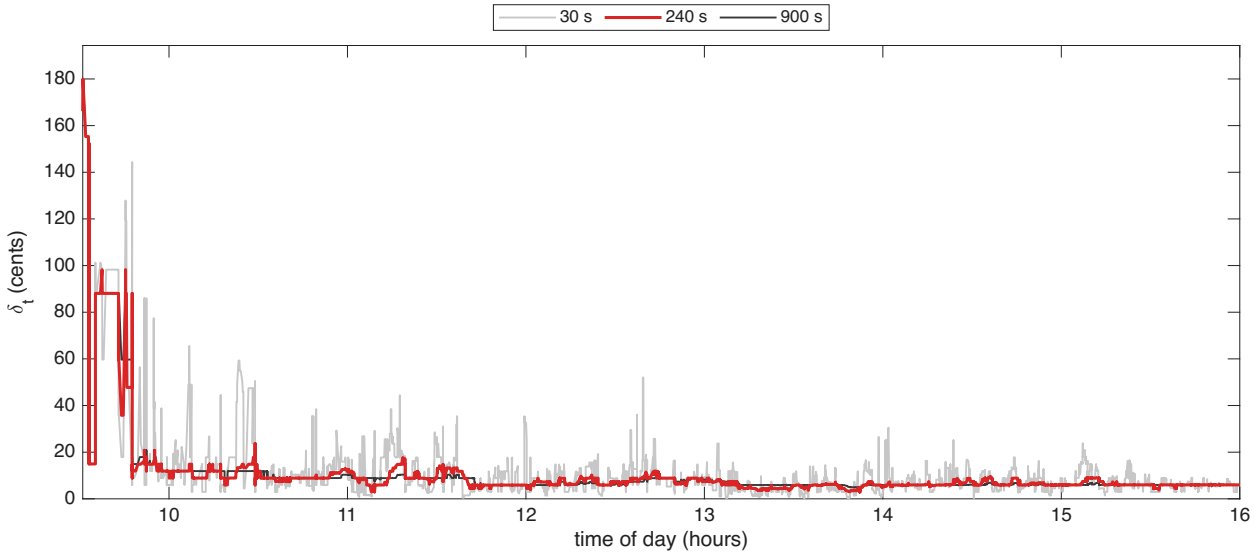

Figure 2: The Huber threshold δ_t in cents for `MIDD` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

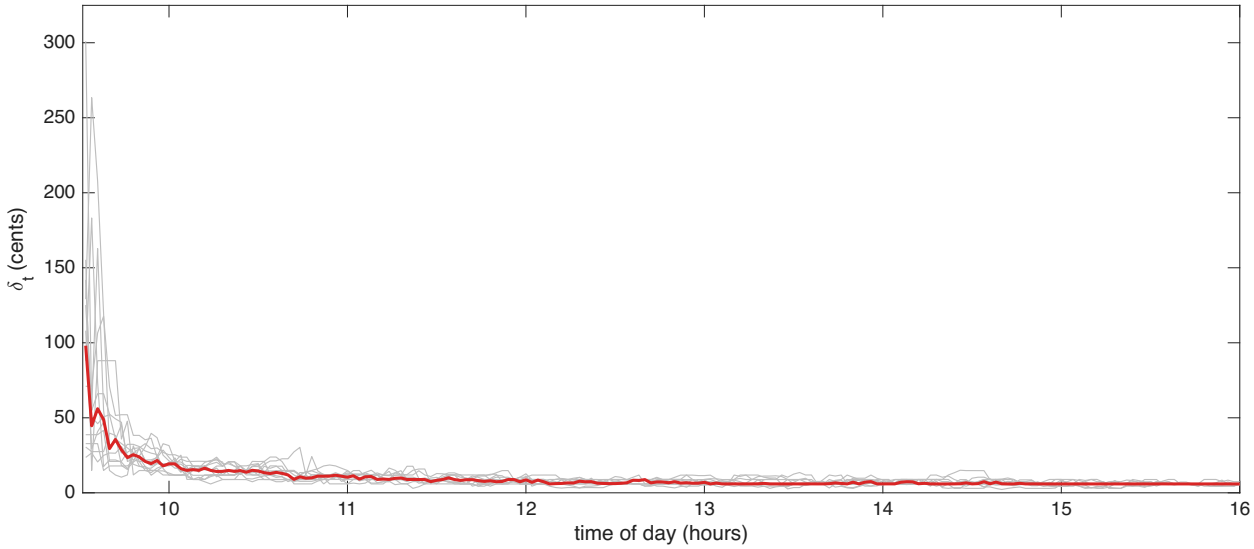

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `MIDD` trading days in Oct 2024 (grey lines) and their median (red line).

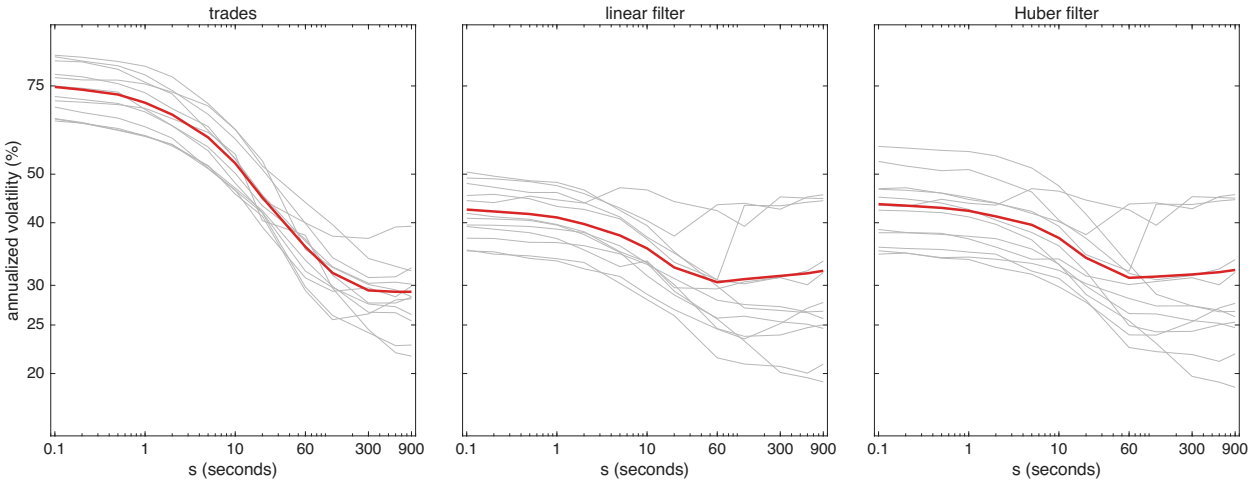

Figure 4: Volatility signature curves for `MIDD` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

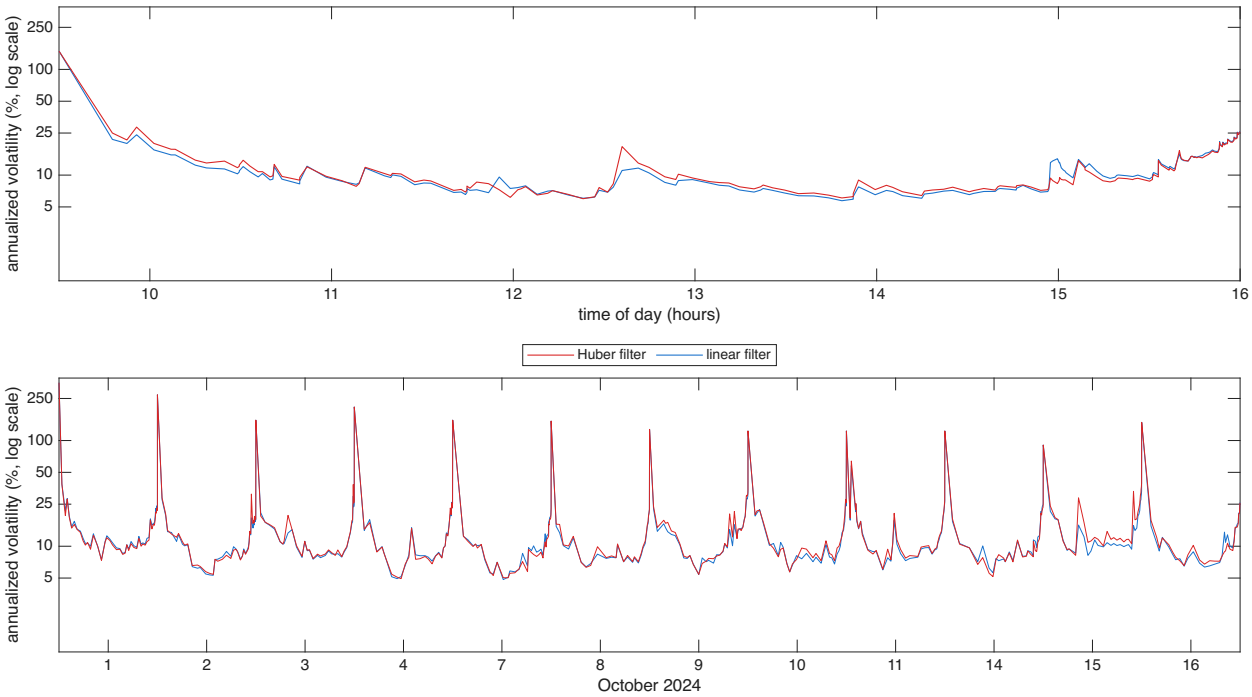

Figure 5: Filtered spot volatility for `MIDD` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 21$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

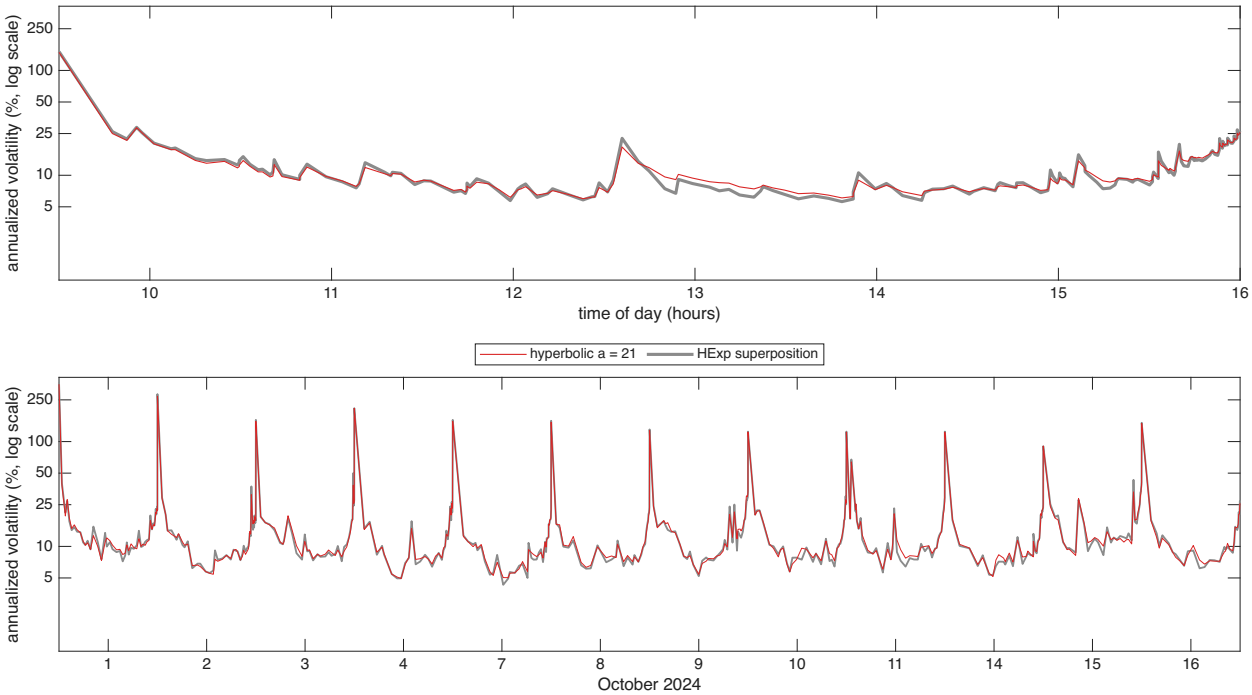

Figure 6: Hyperbolic against superposition weights for `MIDD` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 21$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

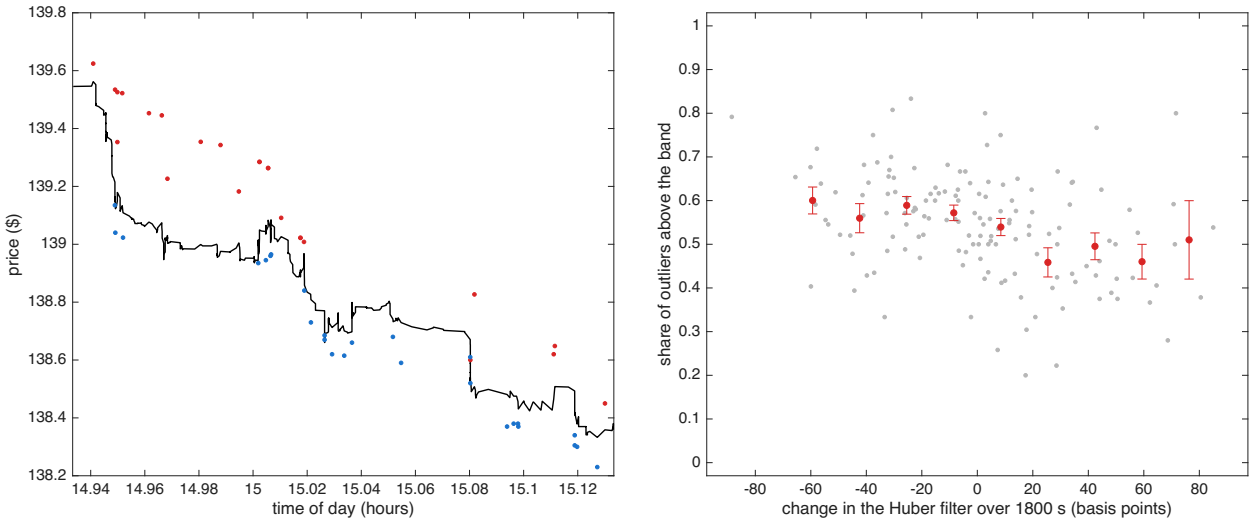

Figure 7: Outliers & robust price for `MIDD` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

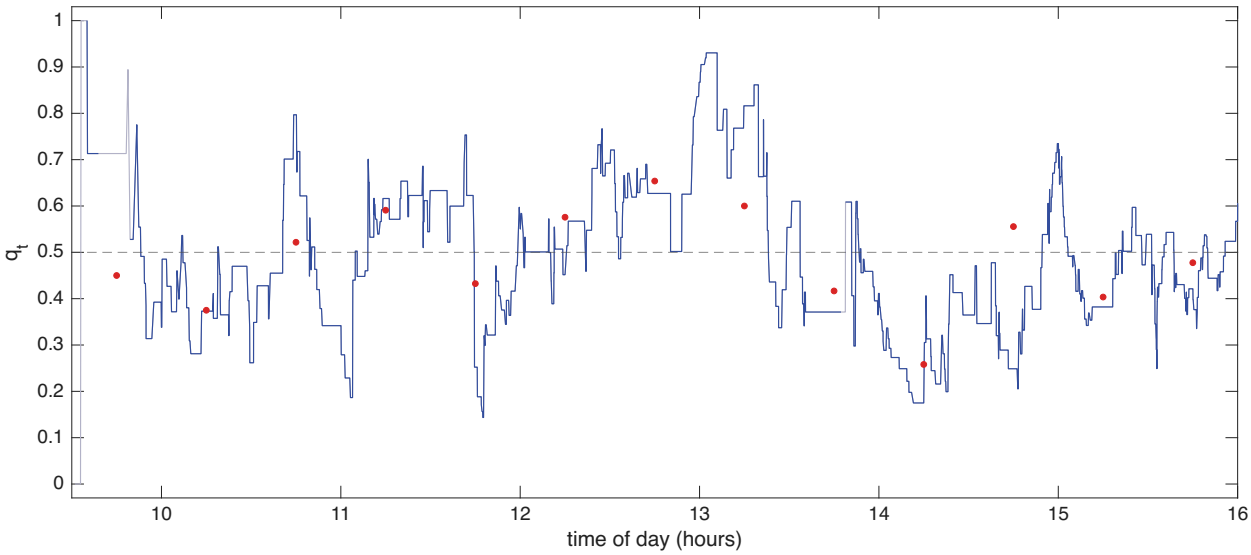

Figure 8: The directional conditional probability for `MIDD`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 315 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 4 to 26112 trades, against 21 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 5 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.
On 12 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 19).
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.

Diagnostics.

SAIA 1 to 16 October 2024

trades on 16 October 2024	8,123
average trading rate	0.4 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	484 (5.96%)
price half-life: run / minimizer[1]	3.0 / 2.5 trades
cost of running 3 trades	+0.70%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.10%
direction half-life[3]	405 s [305, 585]
Mandelbrot a minimizing QLIKE	25 trades
a within 0.0005 of the minimum	19 to 26
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	31% $\rightarrow$ 30%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	29% $\rightarrow$ 30%
median spot volatility, Huber / linear input[5]	16.4 / 16.7 %
hyperbolic against HExp, correlation of logs	0.9962
HExp QLIKE advantage, leave-one-day-out[6]	+0.40386
15-minute signature level	27 to 49 %
Huber signature, 1 min over 1 s, 12 day average	1.41

SAIA over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:01; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	7.387	5.470	5.242	23.276	15.241	8.104	7.791	42.757
2	7.115	6.029	5.914	19.510	13.934	8.959	8.560	32.473
5	7.890	6.672	6.599	14.886	13.068	9.932	9.568	23.097
10	8.064	6.946	6.857	12.152	12.125	10.222	9.972	18.077
25	7.797	7.212	7.157	9.597	11.131	10.443	10.359	13.823
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.16	0.22	0.21	−0.42	3.67	4.84	5.05	5.70
2	0.05	0.10	0.12	0.01	4.55	5.12	5.07	5.61
5	−0.03	0.00	0.01	0.01	5.48	6.11	5.82	4.97
10	−0.03	−0.02	−0.02	−0.02	4.89	4.49	4.72	4.48
25	0.02	0.05	0.05	0.01	9.10	10.71	10.31	7.87

Table 1: Trade-lag differences, `SAIA`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	17.14	14.39	14.15	31.30	32.3	19.9	19.6	118.2
1 s	13.13	11.43	11.31	22.79	30.5	20.4	20.3	108.8
5 s	9.69	8.59	8.49	15.62	28.6	21.0	20.7	94.7
10 s	8.02	7.36	7.30	12.52	29.5	24.3	24.2	82.0
1 min	5.98	5.69	5.66	7.91	31.3	30.2	30.3	50.3
15 min	5.36	5.35	5.32	5.55	39.6	39.6	39.5	40.9
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.05	−0.04	−0.03	−0.13	11.62	4.89	5.08	8.27
1 s	−0.08	−0.02	−0.02	−0.19	11.80	6.07	6.19	7.56
5 s	−0.17	−0.05	−0.05	−0.26	12.25	7.31	7.70	7.23
10 s	−0.07	0.00	−0.01	−0.30	7.56	6.43	7.02	11.08
1 min	−0.01	0.00	−0.01	−0.33				

Table 2: Clock-lag differences, `SAIA`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	15.85	19.40	20.10
s.d. of $\hat{\epsilon}_t$ in cents	30.73	29.51	29.79
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	5.46
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.23	0.24	0.21
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.11	0.13	0.12
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.04	0.01	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.02	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.04	0.06	0.05
80% C.I. for $\theta_t - X_t$	—	−12.63,12.74	−13.47,13.48
95% C.I. for $\theta_t - X_t$	—	−23.27,22.52	−24.74,23.57
99% C.I. for $\theta_t - X_t$	—	−35.73,46.35	−38.56,50.33
s.e. of $\theta_t - X_t$ in cents	—	11.51	12.56

Table 3: Residual summaries, `SAIA`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 5.46, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.55	4.67	2.24 to 3.91	2.09 to 4.67
Residual divided by δ_t	5.23	3.32	2.74 to 5.23	3.32 to 4.70
Repricing windows removed	3.55	4.67	2.24 to 3.88	2.09 to 4.67
Calm windows only	3.48	3.74		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	3.32	5.29	2.44 to 4.33	2.07 to 5.29

Table 4: Hill estimator of the tail index of the Huber filter residual, `SAIA`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 6 to 8 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ±1 second, which is 191.8 to 329.5 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ±5 seconds, 151.4 cents on 16 October 2024, 94% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

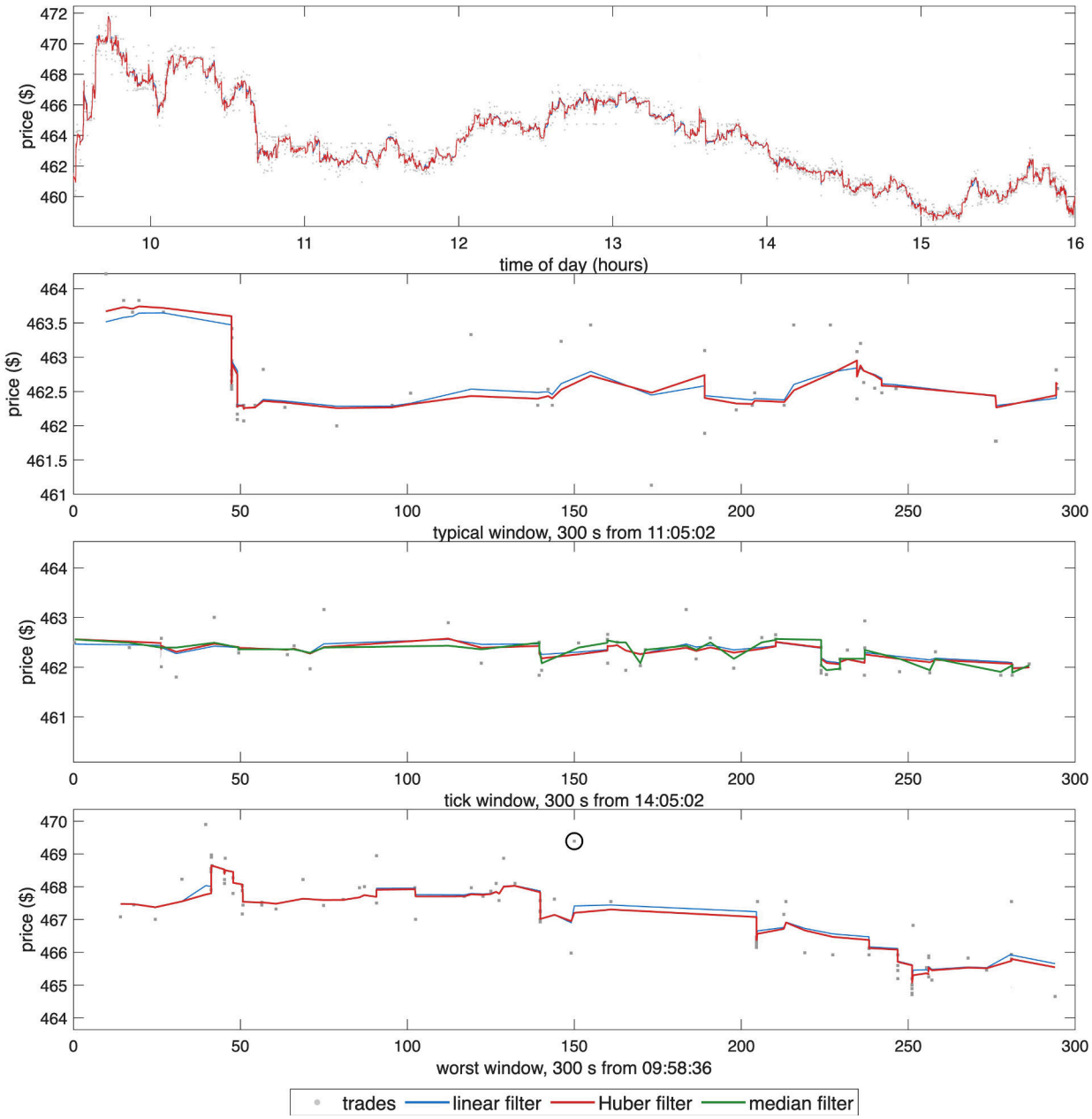

Figure 1: `SAIA` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 11:05:02. 3rd: 300-second window at 14:05:02 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

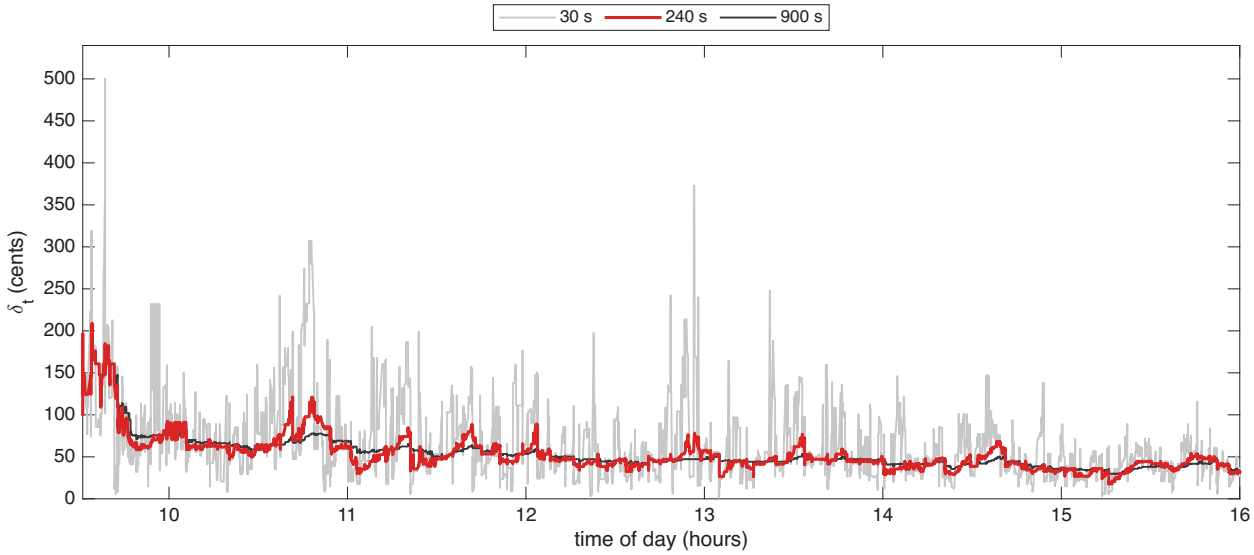

Figure 2: The Huber threshold δ_t in cents for `SAIA` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

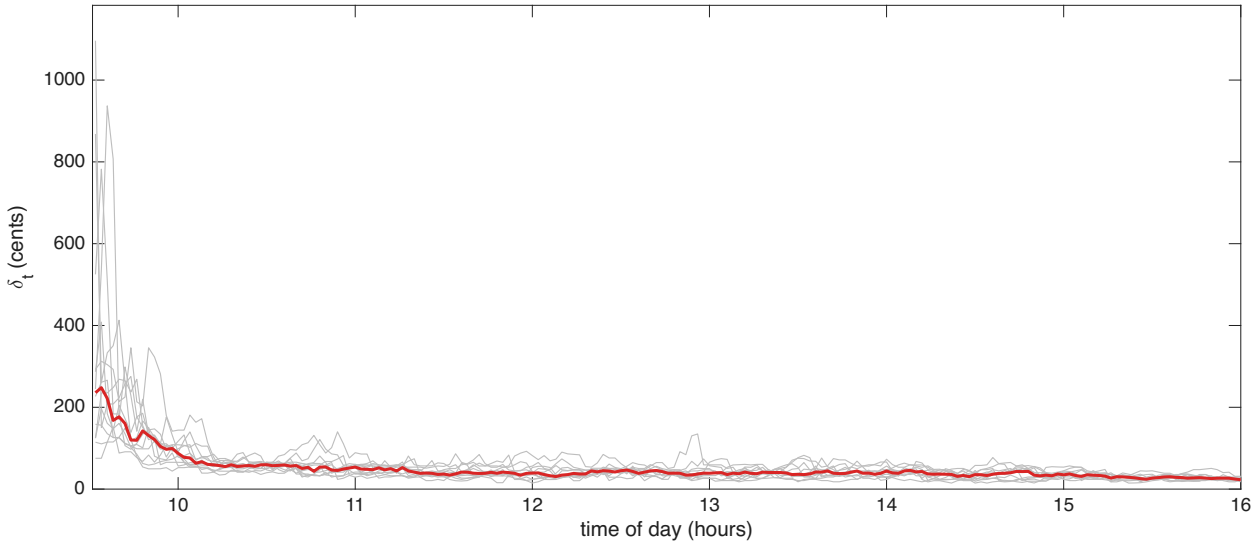

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `SAIA` trading days in Oct 2024 (grey lines) and their median (red line).

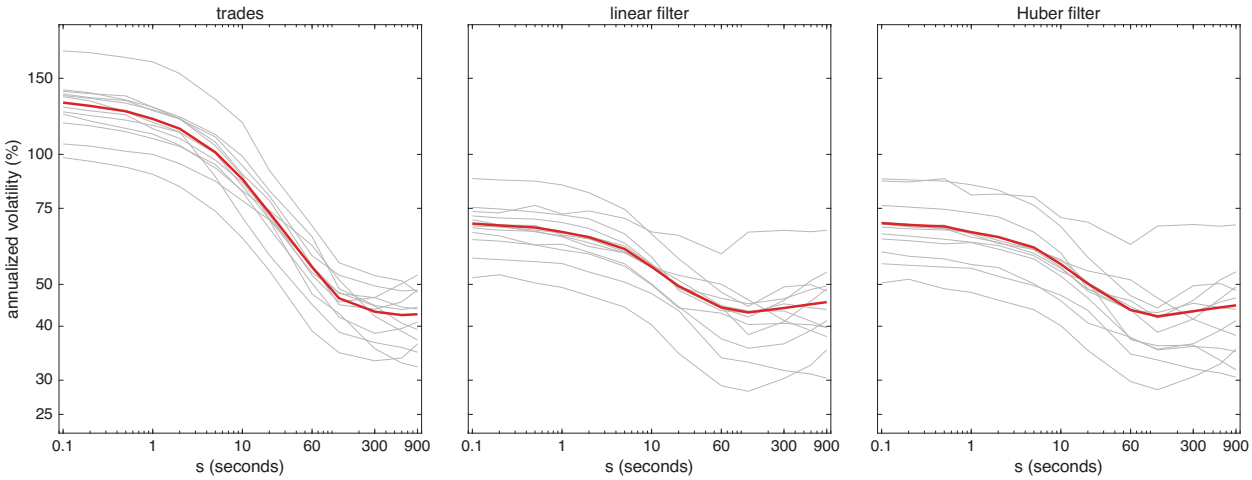

Figure 4: Volatility signature curves for `SAIA` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

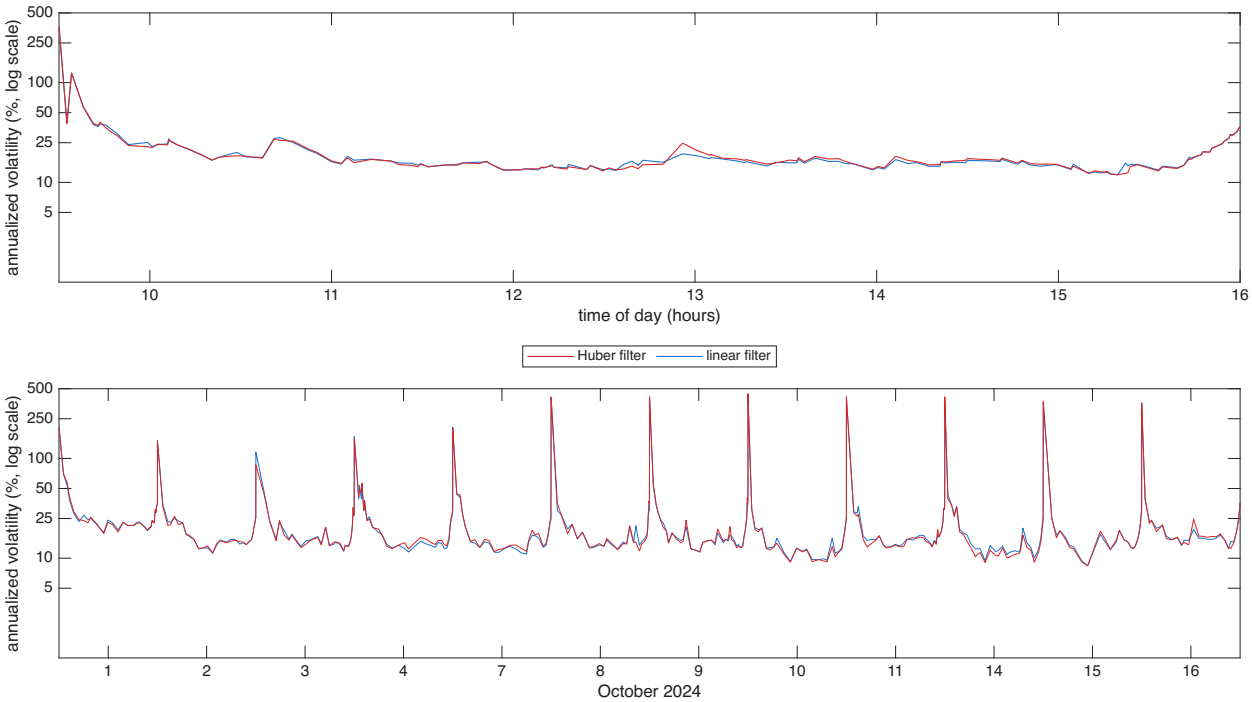

Figure 5: Filtered spot volatility for `SAIA` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 25$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

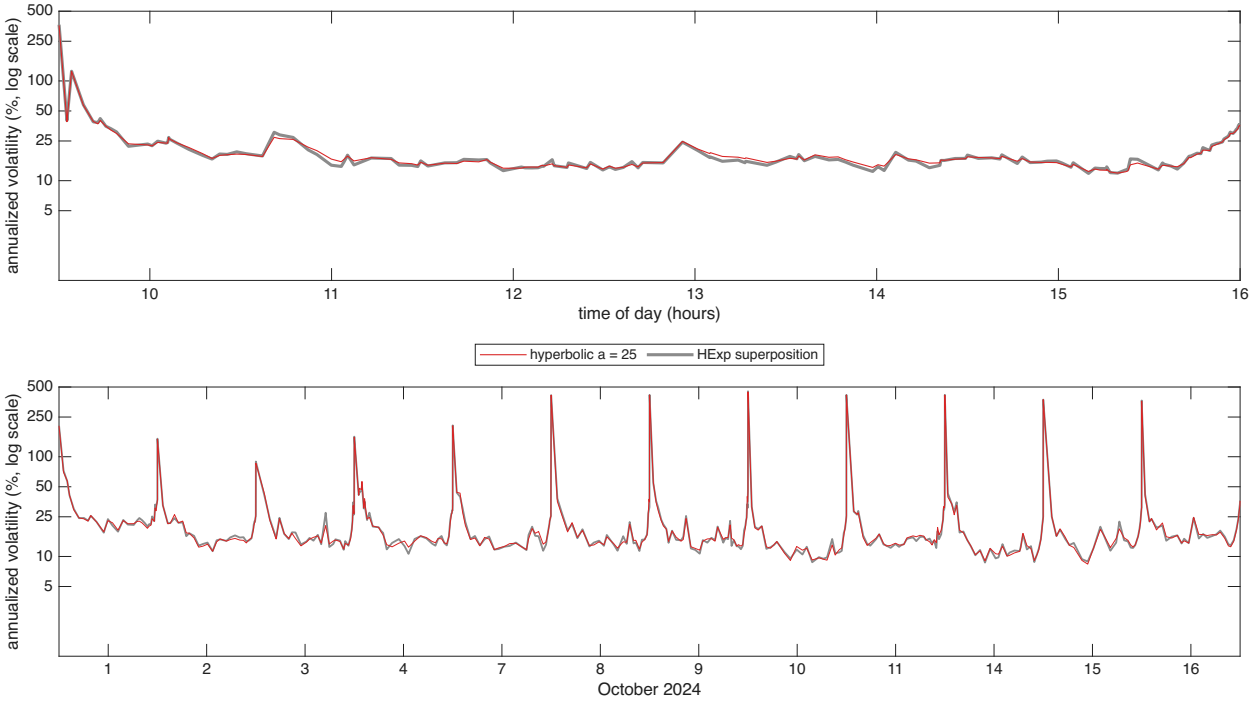

Figure 6: Hyperbolic against superposition weights for `SAIA` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 25$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

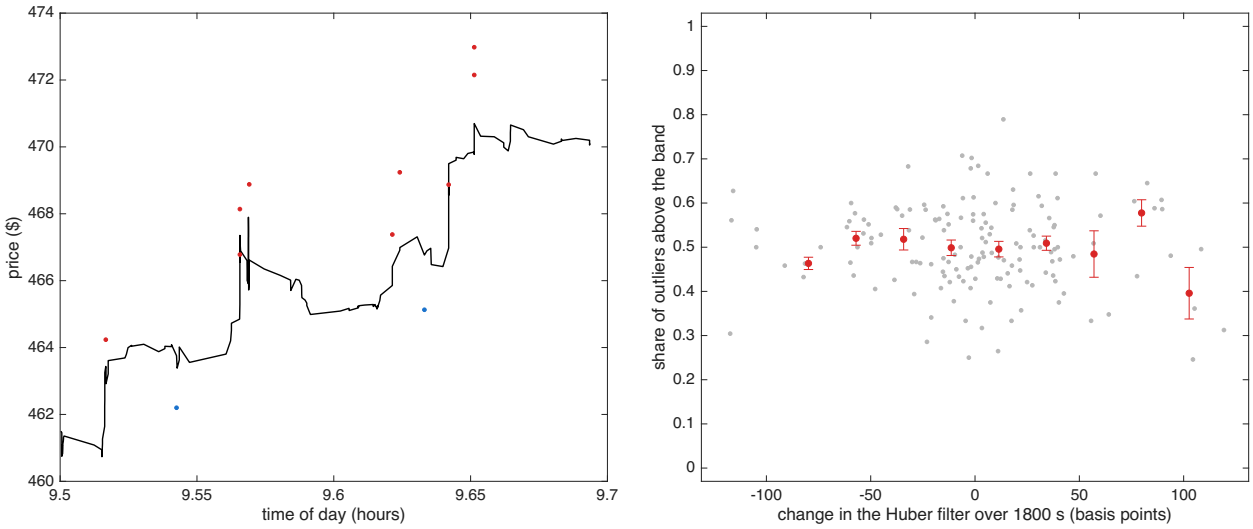

Figure 7: Outliers & robust price for `SAIA` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

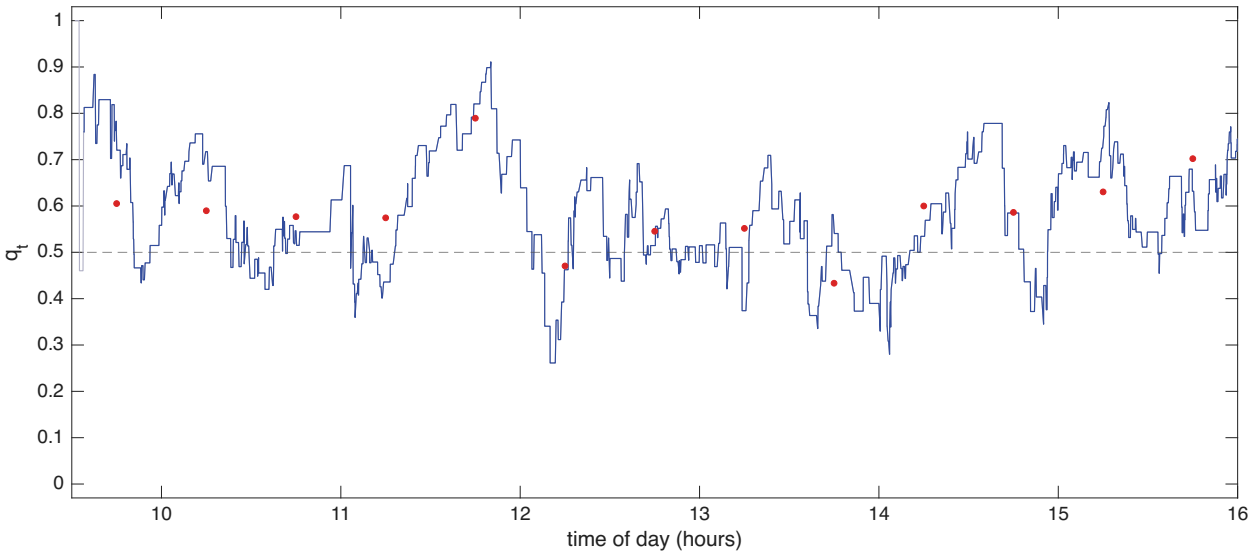

Figure 8: The directional conditional probability for `SAIA`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 405 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 5 to 26112 trades, against 25 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 4 trades for this stock.
On 12 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 35).
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
The largest Huber residual on 16 October 2024 occurs at 09:31, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:01.

Diagnostics.

FIZZ 1 to 16 October 2024

trades on 16 October 2024	5,068
average trading rate	0.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	299 (5.90%)
price half-life: run / minimizer[1]	3.0 / 2.0 trades
cost of running 3 trades	+2.13%
scale half-life: run / minimizer[2]	240 / 360 s
cost of running 240 s	+0.32%
direction half-life[3]	755 s [530, 1225]
Mandelbrot a minimizing QLIKE	12 trades
a within 0.0005 of the minimum	10 to 13
linear $\hat{\sigma}$, before $\to$ at the largest residual[4]	16% $\to$ 17%
Huber $\hat{\sigma}$, before $\to$ peak	16% $\to$ 18%
median spot volatility, Huber / linear input[5]	8.9 / 8.3 %
hyperbolic against HExp, correlation of logs	0.9581
HExp QLIKE advantage, leave-one-day-out[6]	−1738161178573279232 0.00000
15-minute signature level	15 to 28 %
Huber signature, 1 min over 1 s, 12 day average	1.51

FIZZ over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 10:00; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	0.421	0.310	0.295	1.332	1.046	0.514	0.479	2.716
2	0.398	0.343	0.333	1.124	0.896	0.557	0.516	2.102
5	0.448	0.383	0.373	0.850	0.805	0.590	0.555	1.432
10	0.468	0.413	0.403	0.702	0.716	0.594	0.571	1.061
25	0.501	0.466	0.458	0.603	0.673	0.619	0.607	0.837
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.27	0.18	0.16	−0.40	2.89	2.85	2.38	2.95
2	0.04	0.00	0.03	−0.10	4.30	3.20	2.63	2.57
5	−0.07	−0.05	−0.02	−0.04	5.55	3.88	3.19	2.83
10	0.00	0.04	0.06	0.04	8.18	5.50	4.91	3.52
25	0.01	−0.02	−0.01	−0.01	7.41	4.85	4.44	3.04

Table 1: Trade-lag differences, FIZZ, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	0.94	0.79	0.78	1.59	21.0	10.6	10.3	62.9
1 s	0.66	0.58	0.57	1.17	19.9	11.1	10.6	60.1
5 s	0.49	0.43	0.42	0.80	18.6	11.3	10.8	52.6
10 s	0.45	0.41	0.40	0.70	18.5	12.5	12.0	48.8
1 min	0.32	0.31	0.31	0.41	16.9	15.2	15.0	29.4
15 min	0.28	0.28	0.28	0.29	16.0	15.6	15.6	18.4
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	0.05	0.09	0.10	−0.02	7.87	10.39	10.01	3.43
1 s	0.00	0.04	0.05	−0.05	7.95	9.12	8.91	3.74
5 s	−0.14	−0.10	−0.07	−0.22	6.91	7.08	8.13	3.99
10 s	−0.07	−0.01	0.00	−0.31	8.24	7.17	8.16	3.59
1 min	0.07	0.11	0.13	−0.08				

Table 2: Clock-lag differences, FIZZ, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	0.92	1.09	1.13
s.d. of $\hat{\epsilon}_t$ in cents	1.99	1.84	1.84
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.36
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.17	0.16
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.10	0.02	0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.02	−0.05	−0.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.03	0.07	0.06
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.01	−0.01	−0.01
80% C.I. for $\theta_t - X_t$	—	−0.75,0.75	−0.80,0.80
95% C.I. for $\theta_t - X_t$	—	−1.40,1.36	−1.52,1.48
99% C.I. for $\theta_t - X_t$	—	−1.90,2.43	−2.09,2.64
s.e. of $\theta_t - X_t$ in cents	—	0.68	0.73

Table 3: Residual summaries, `FIZZ`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.36, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.22	3.89	1.77 to 3.87	2.39 to 3.89
Residual divided by δ_t	2.72	2.93	2.15 to 3.74	2.12 to 5.31
Repricing windows removed	2.22	3.89	1.72 to 3.87	2.39 to 3.89
Calm windows only	2.17	3.96		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.29	4.21	2.12 to 4.23	2.28 to 4.21

Table 4: Hill estimator of the tail index of the Huber filter residual, `FIZZ`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 3 to 4 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 20.2 to 20.2 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 10.1 cents on 16 October 2024, 95% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

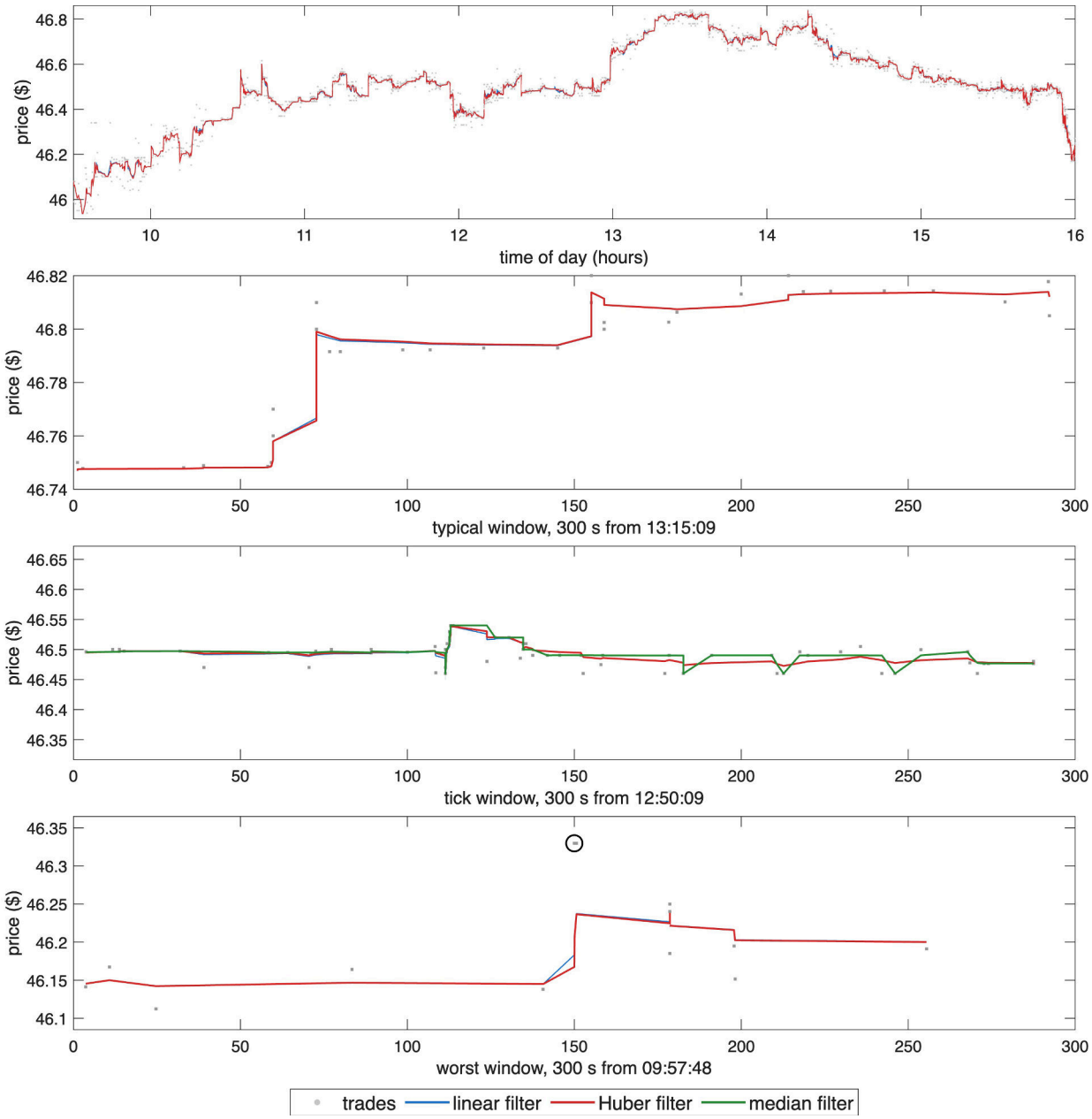

Figure 1: `FIZZ` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 13:15:09. 3rd: 300-second window at 12:50:09 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

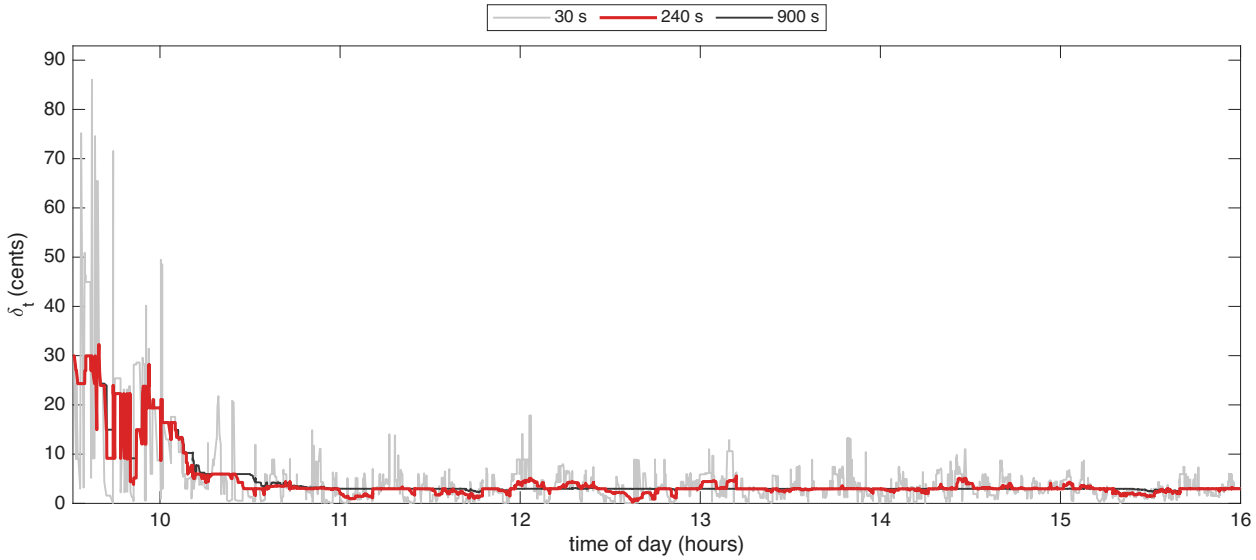

Figure 2: The Huber threshold δ_t in cents for `FIZZ` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

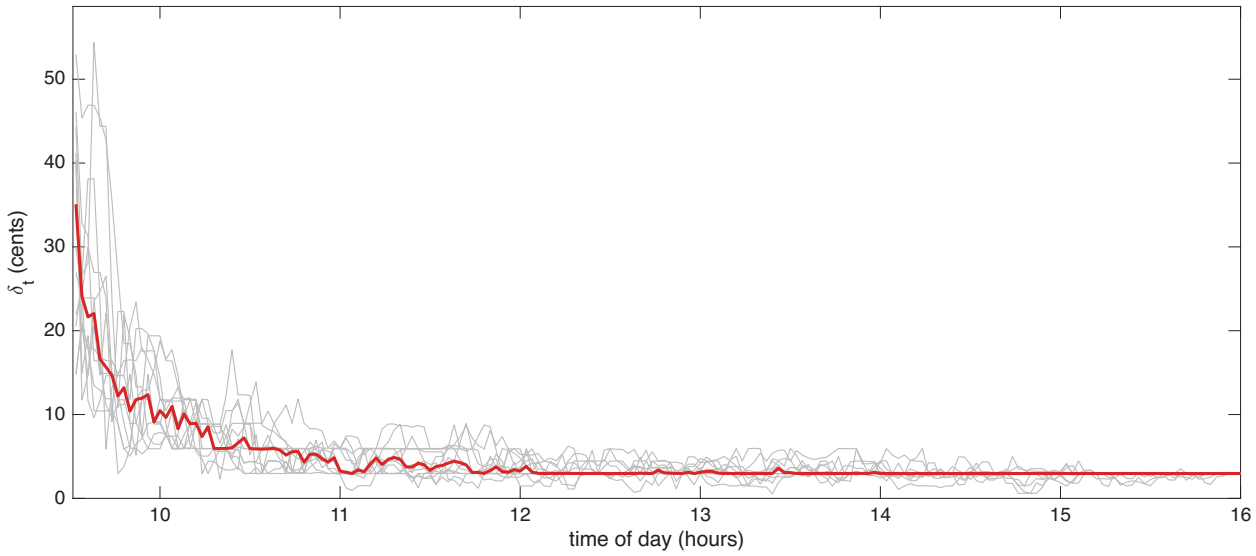

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `FIZZ` trading days in Oct 2024 (grey lines) and their median (red line).

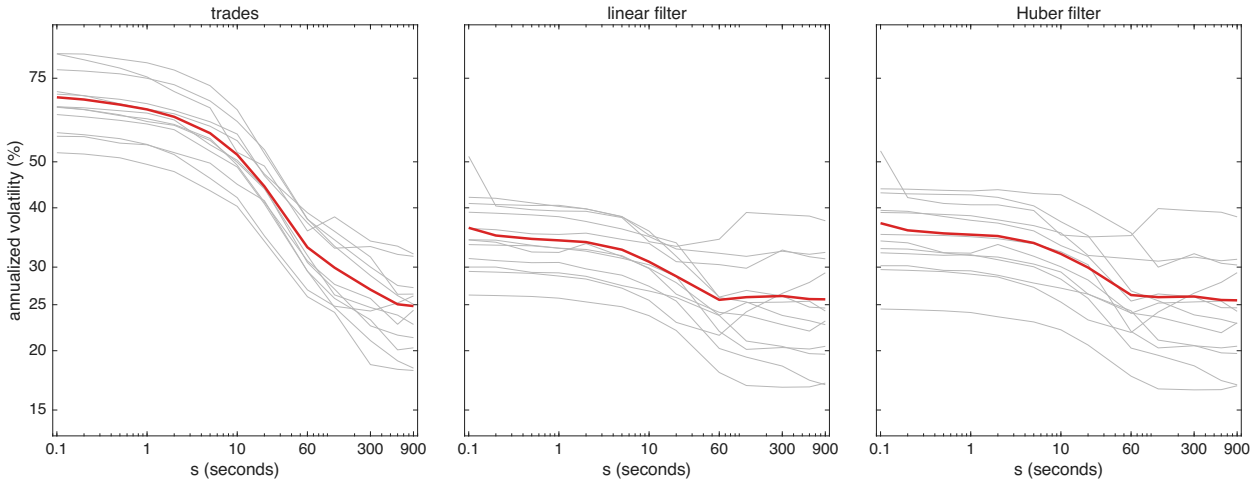

Figure 4: Volatility signature curves for `FIZZ` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

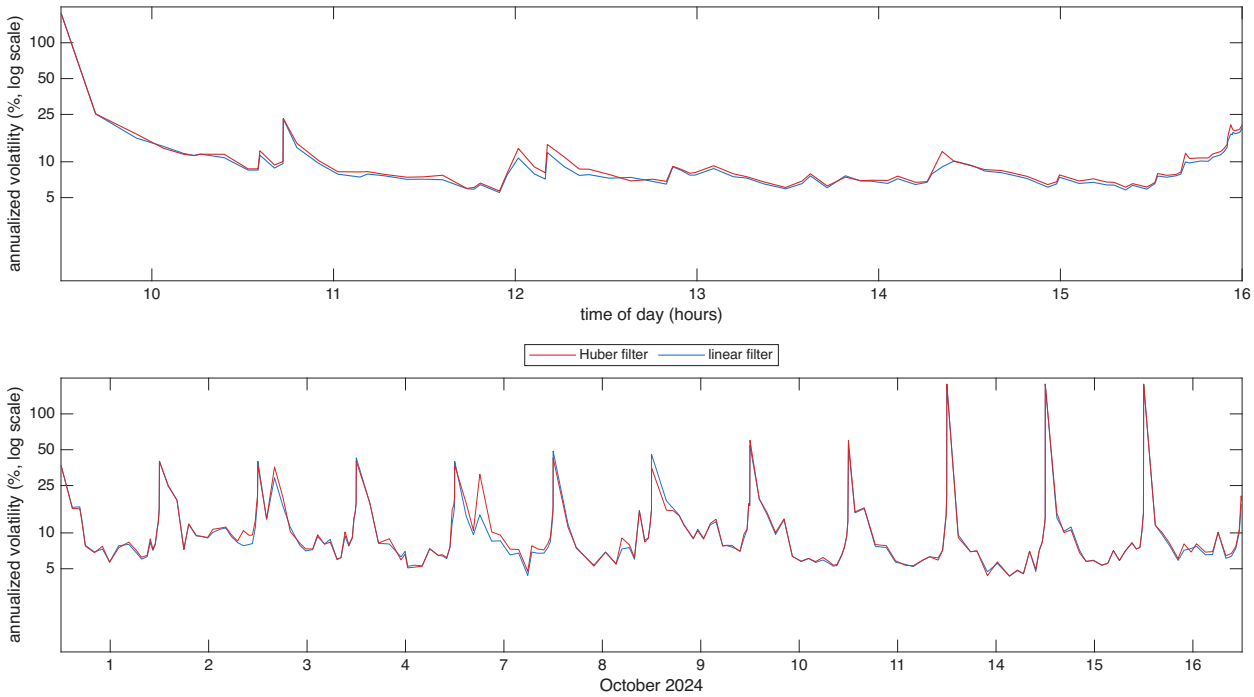

Figure 5: Filtered spot volatility for `FIZZ` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 12$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

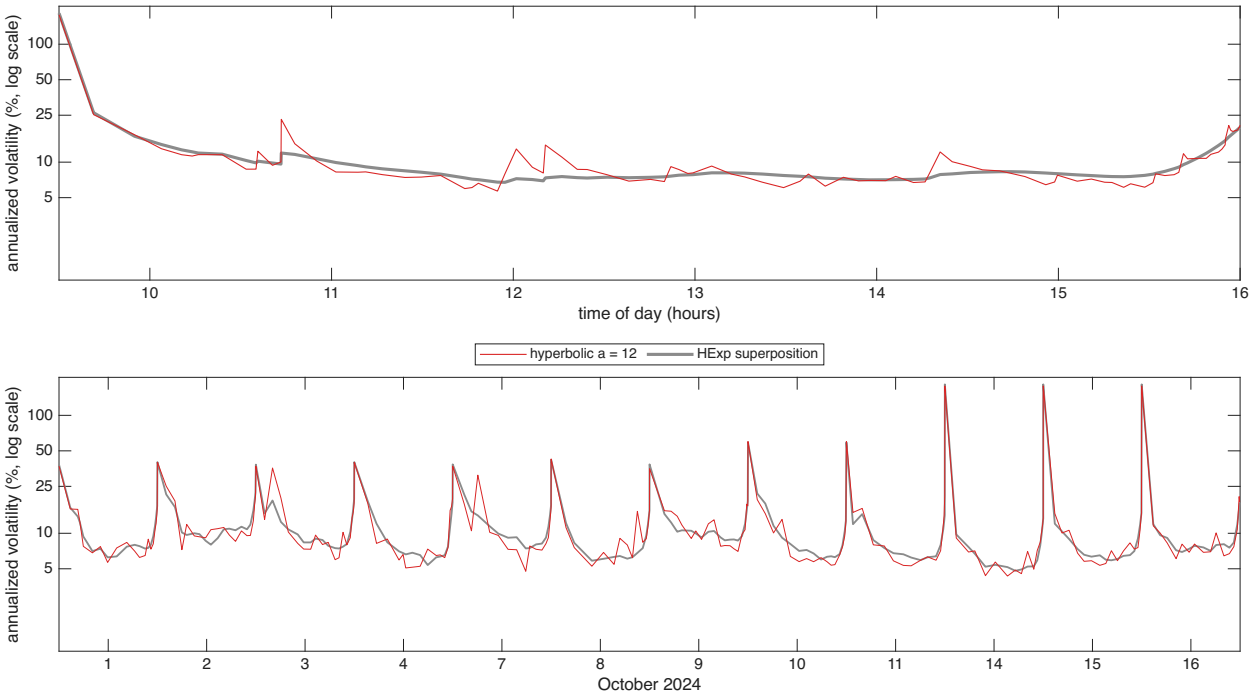

Figure 6: Hyperbolic against superposition weights for `FIZZ` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 12$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

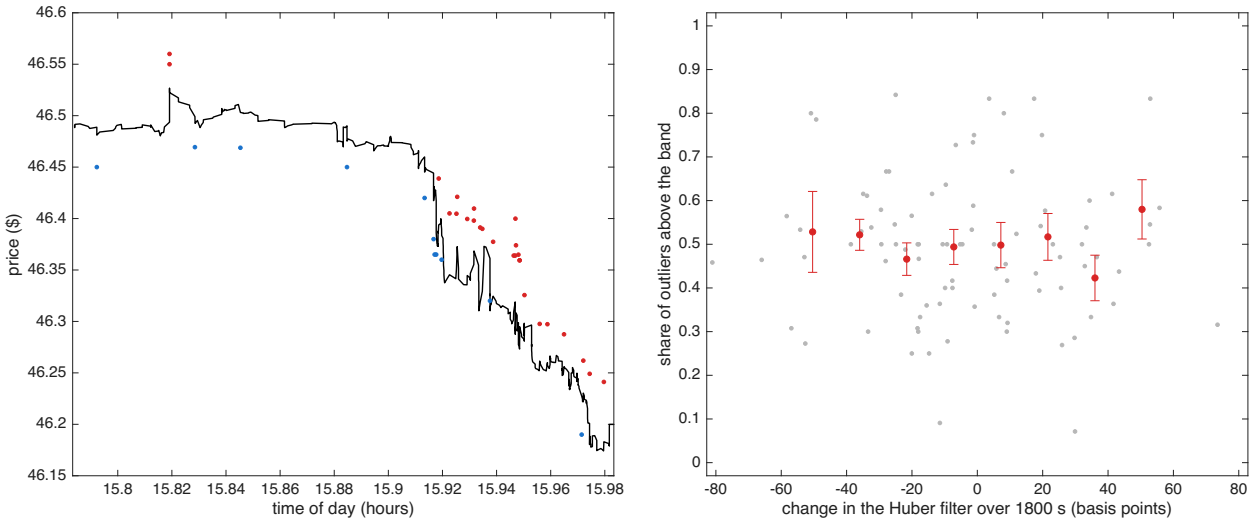

Figure 7: Outliers & robust price for `FIZZ` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

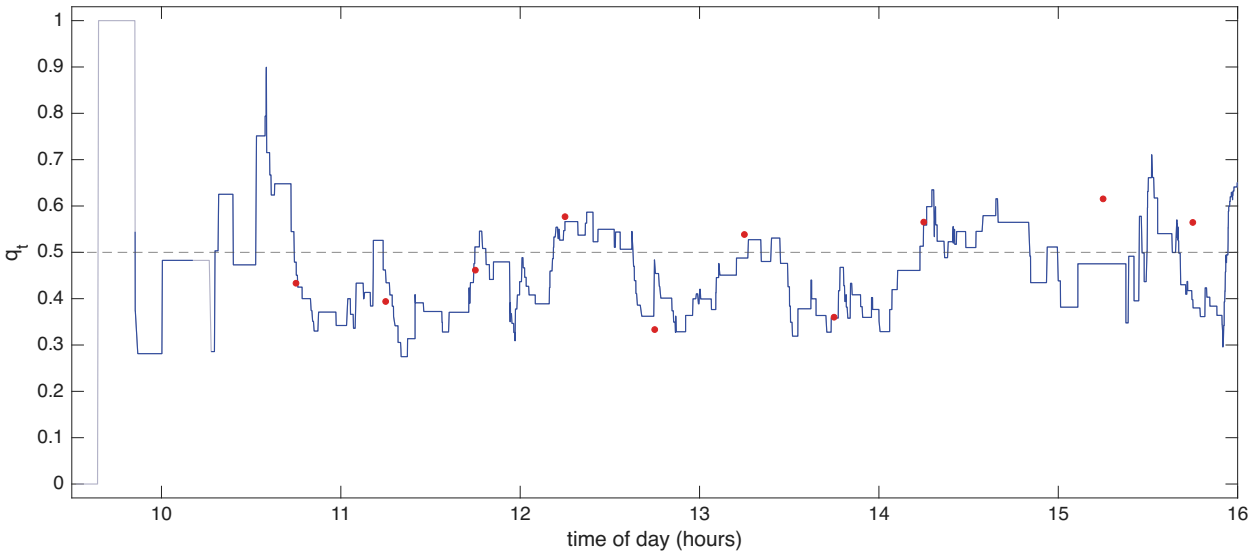

Figure 8: The directional conditional probability for FIZZ, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 755 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

8 of the 22 days built (10 presample and 12 analysis) have fewer than 5,000 trades.
The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 2 to 26112 trades, against 12 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 2 trades for this stock.
The leave-one-day-out selections of a fall outside the values within 0.0005 of the pooled minimum.
5 analysis day(s) set the opening noise scale from fewer than 5 previous openings.
On 12 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 15).
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
Where the tail index can be estimated it averages over as few as 3 depths on one day, because the largest 0.1 to 1 per cent of 4,267 residuals is a short stretch of the tail. It is a much less precise estimate than for a heavily traded name.
The largest Huber residual on 16 October 2024 occurs at 09:37, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 10:00.

Diagnostics.

LANC 1 to 16 October 2024

trades on 16 October 2024	3,945
average trading rate	0.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	211 (5.35%)
price half-life: run / minimizer[1]	3.0 / 3.2 trades
cost of running 3 trades	+0.10%
scale half-life: run / minimizer[2]	240 / 600 s
cost of running 240 s	+0.65%
direction half-life[3]	1130 s [835, 1500]
Mandelbrot a minimizing QLIKE	26112 trades
a within 0.0005 of the minimum	26112 to 26112
linear $\hat{\sigma}$, before → at the largest residual[4]	13% → 12%
Huber $\hat{\sigma}$, before → peak	14% → 14%
median spot volatility, Huber / linear input[5]	12.0 / 11.7 %
hyperbolic against HExp, correlation of logs	0.9752
HExp QLIKE advantage, leave-one-day-out[6]	+0.67302
15-minute signature level	12 to 34 %
Huber signature, 1 min over 1 s, 12 day average	1.20

`LANC` over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 11:13; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	2.432	1.604	1.518	7.427	6.084	2.904	2.664	16.119
2	2.082	1.703	1.657	5.945	4.718	2.902	2.705	11.709
5	2.255	1.848	1.812	4.394	4.132	2.978	2.816	7.986
10	2.213	1.877	1.833	3.538	3.609	2.897	2.752	6.096
25	1.996	1.801	1.763	2.616	2.865	2.534	2.455	3.961
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.40	0.00	0.03	−0.47	2.30	2.24	2.30	1.82
2	0.15	0.01	0.01	−0.03	2.24	2.57	2.95	2.26
5	0.09	−0.06	−0.07	−0.07	2.81	3.14	3.15	2.47
10	−0.01	−0.10	−0.13	−0.19	4.12	4.32	4.09	2.32
25	−0.04	0.01	0.02	0.04	6.39	6.05	5.06	2.85

Table 1: Trade-lag differences, `LANC`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	3.34	2.60	2.52	6.62	29.1	13.8	12.8	80.2
1 s	2.64	2.04	1.98	4.59	28.3	13.6	12.6	78.1
5 s	2.02	1.61	1.56	3.33	25.7	13.1	12.2	70.0
10 s	1.69	1.40	1.36	2.84	25.2	13.5	12.7	67.1
1 min	1.19	1.09	1.07	1.73	20.9	14.7	14.4	45.0
15 min	0.88	0.87	0.86	0.96	12.7	11.8	11.7	20.2
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.01	−0.03	−0.03	−0.09	3.36	23.75	27.47	1.87
1 s	−0.08	−0.01	−0.02	−0.10	3.01	19.88	28.07	1.91
5 s	−0.12	0.05	0.06	−0.21	3.11	12.60	17.34	2.04
10 s	−0.19	−0.11	−0.10	−0.30	3.06	36.07	20.44	2.11
1 min	−0.18	−0.13	−0.07	−0.30				

Table 2: Clock-lag differences, `LANC`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	4.68	5.66	5.81
s.d. of $\hat{\epsilon}_t$ in cents	10.47	10.20	10.19
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	2.27
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.16	0.01	0.03
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.07	0.06	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.03	−0.06	−0.07
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	0.01	−0.13	−0.13
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.04	0.02	0.01
80% C.I. for $\theta_t - X_t$	—	−3.58,3.14	−3.81,3.26
95% C.I. for $\theta_t - X_t$	—	−6.37,6.25	−6.88,6.53
99% C.I. for $\theta_t - X_t$	—	−9.87,12.32	−10.98,15.30
s.e. of $\theta_t - X_t$ in cents	—	3.29	3.55

Table 3: Residual summaries, `LANC`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 2.27, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	2.19	2.12	1.73 to 3.14	2.12 to 3.63
Residual divided by δ_t	3.19	3.62	2.82 to 4.06	2.38 to 4.03
Repricing windows removed	2.19	2.12	1.74 to 3.35	2.12 to 3.63
Calm windows only	2.19	2.12		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	2.15	2.33	1.70 to 3.41	2.33 to 3.80

Table 4: Hill estimator of the tail index of the Huber filter residual, `LANC`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 1 to 5 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 66.3 to 121.0 cents over the days and removes 0% to 1% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 42.1 cents on 16 October 2024, 100% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

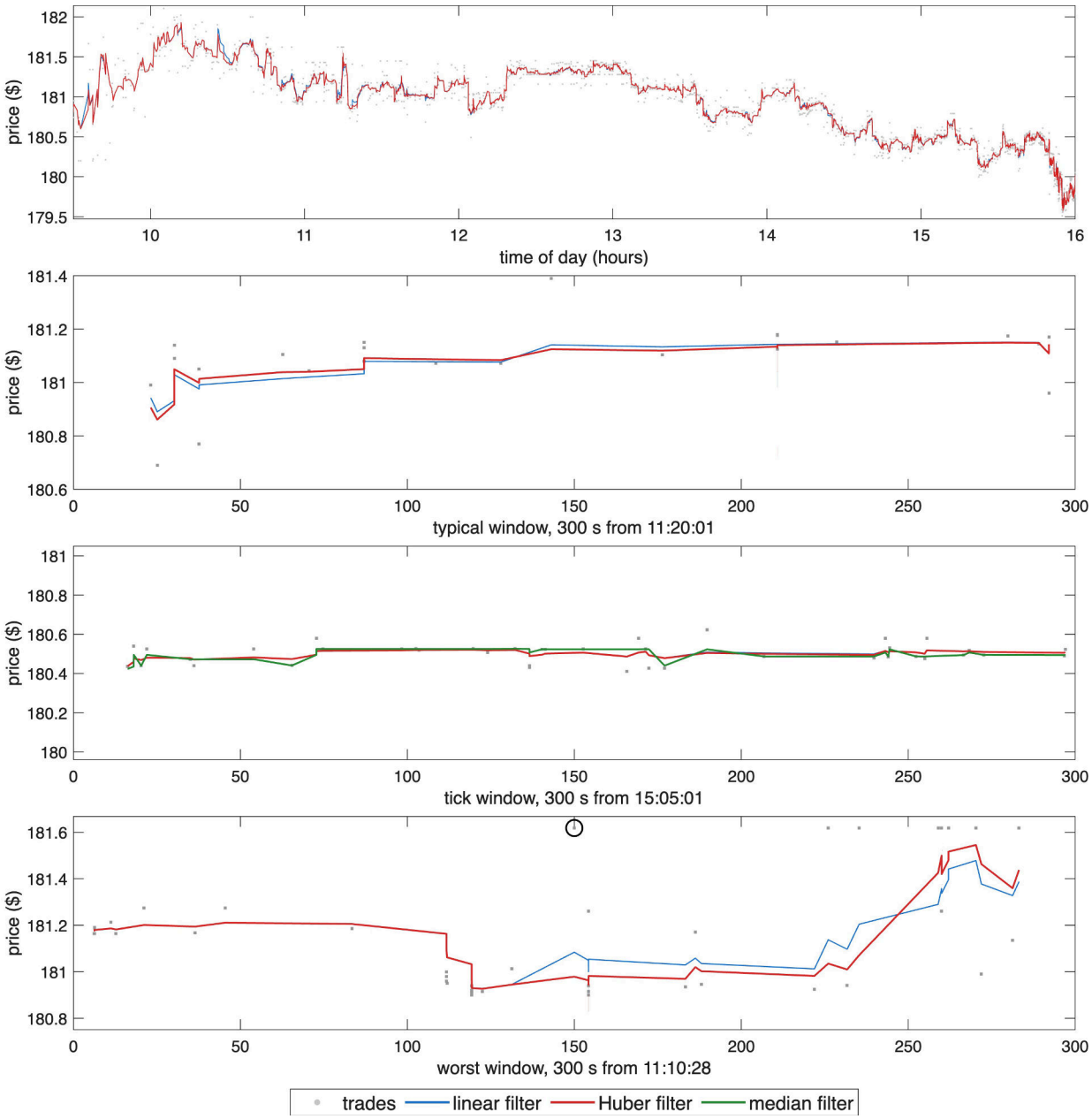

Figure 1: `LANC` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 11:20:01. 3rd: 300-second window at 15:05:01 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

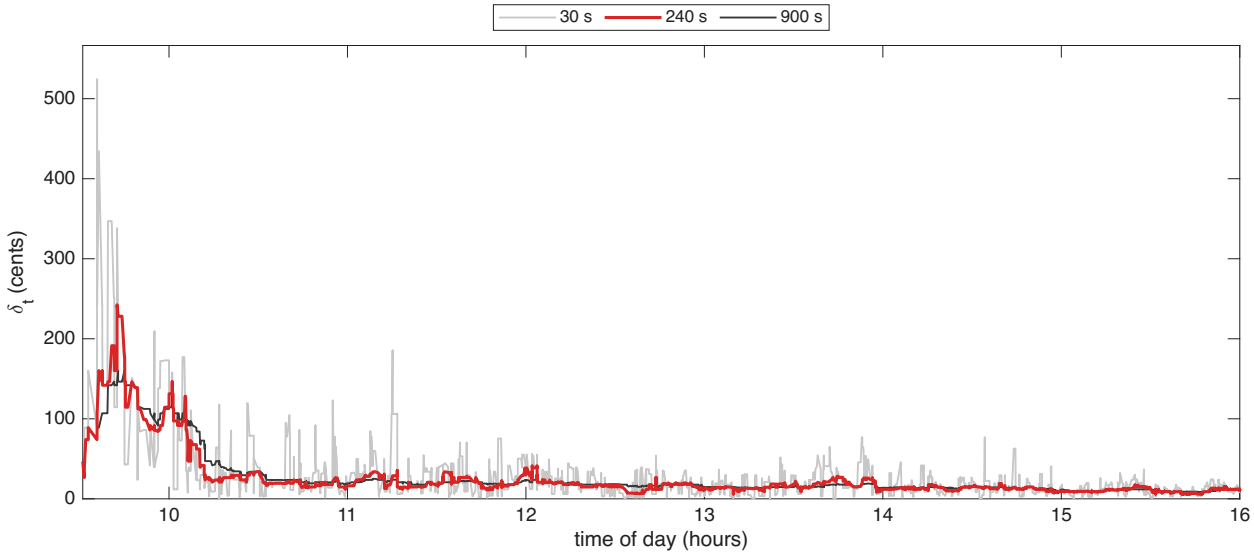

Figure 2: The Huber threshold δ_t in cents for `LANC` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

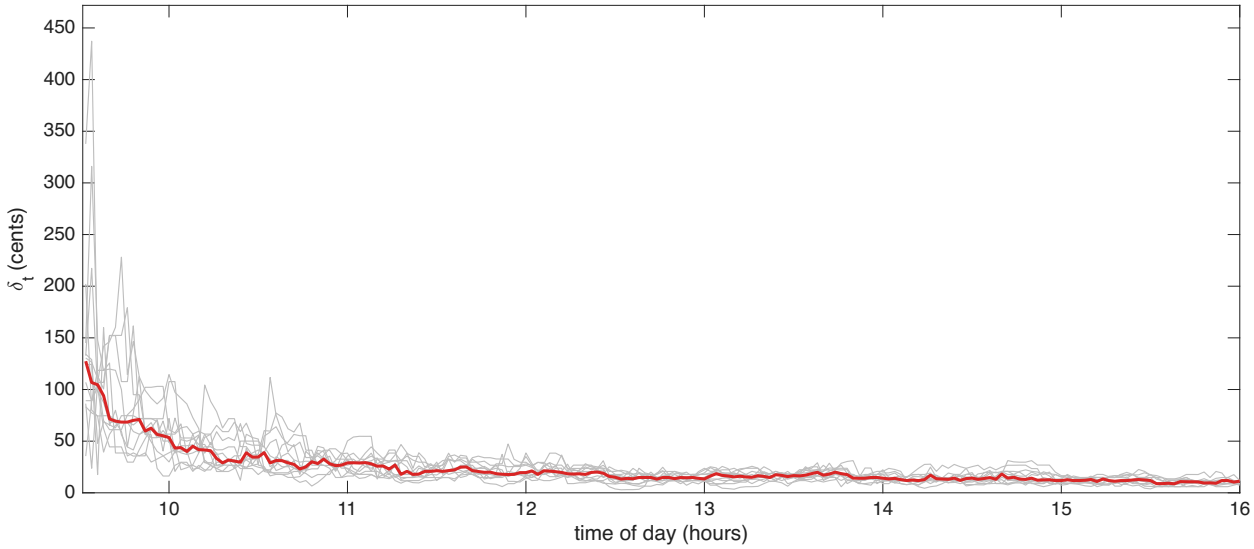

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `LANC` trading days in Oct 2024 (grey lines) and their median (red line).

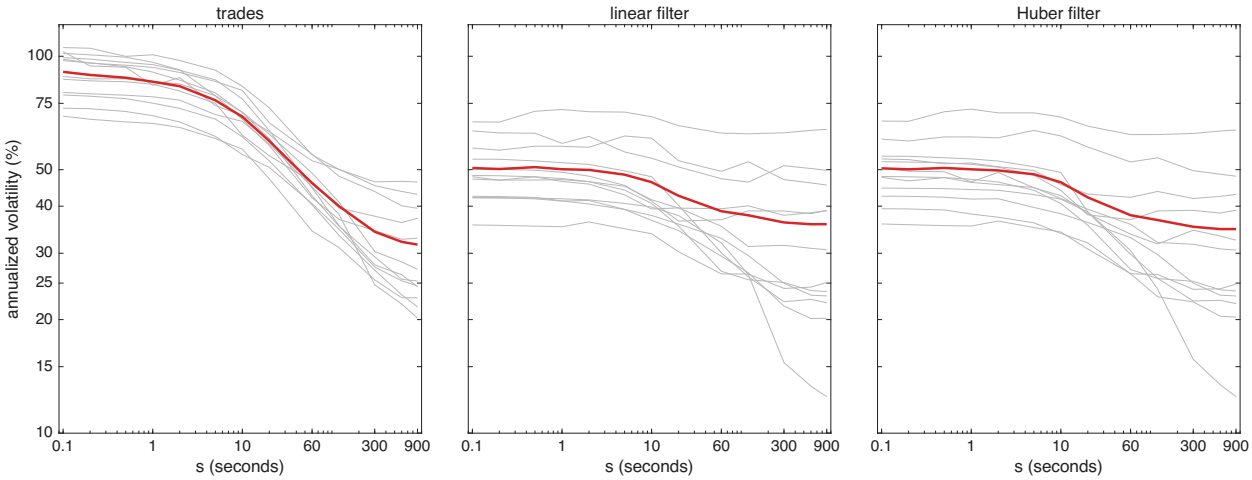

Figure 4: Volatility signature curves for `LANC` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

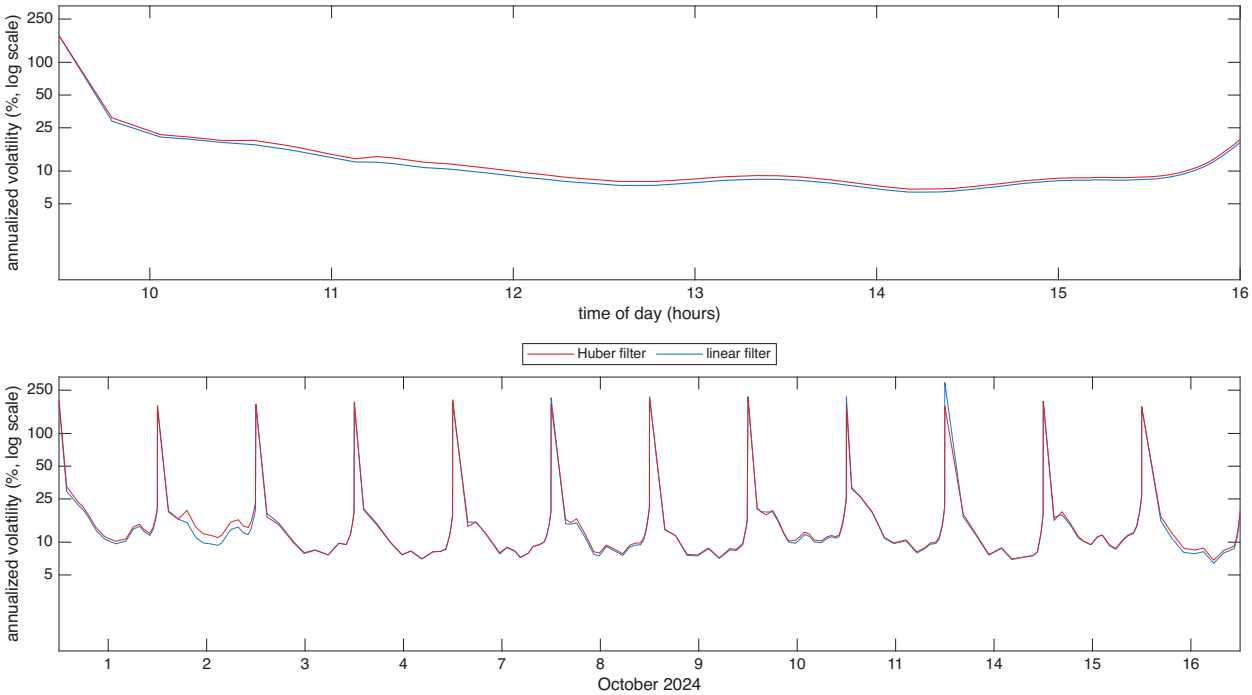

Figure 5: Filtered spot volatility for `LANC` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 26112$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

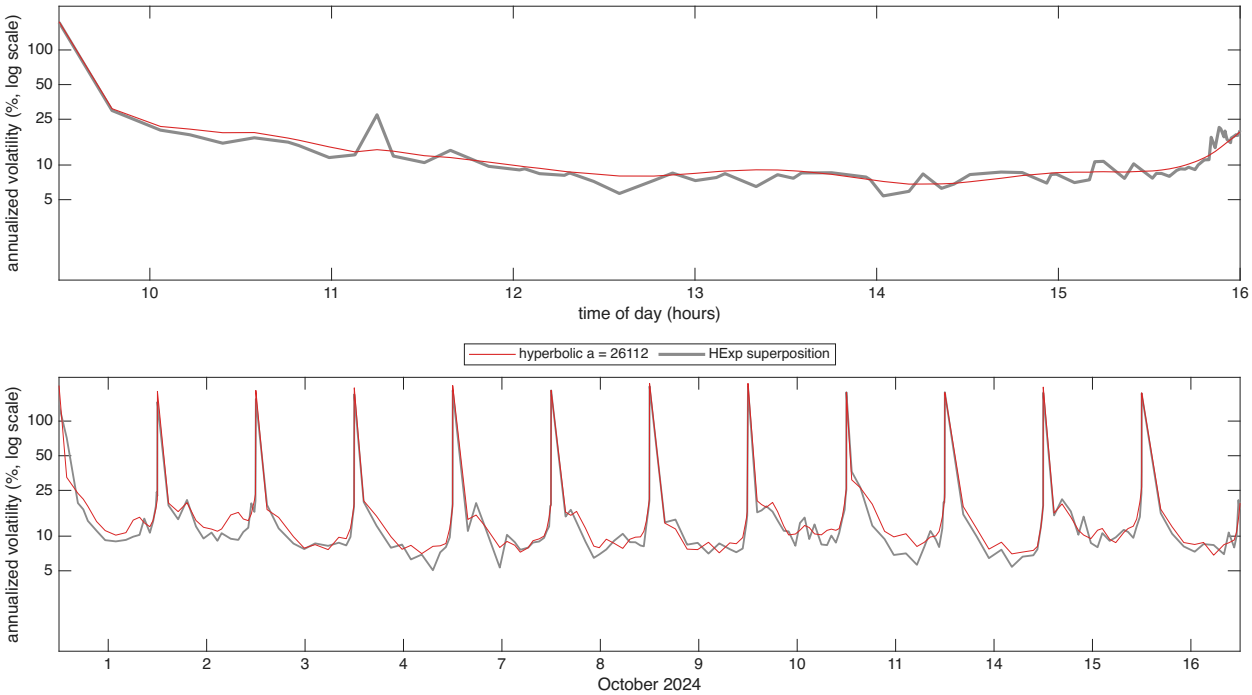

Figure 6: Hyperbolic against superposition weights for `LANC` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 26112$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 60, 600, 3600, 23400 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

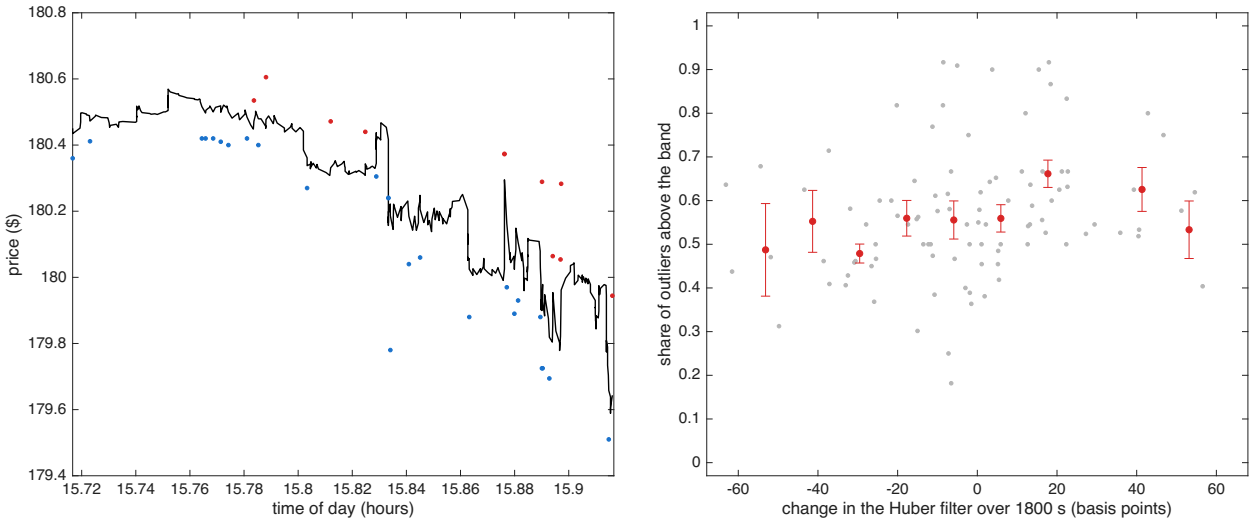

Figure 7: Outliers & robust price for `LANC` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

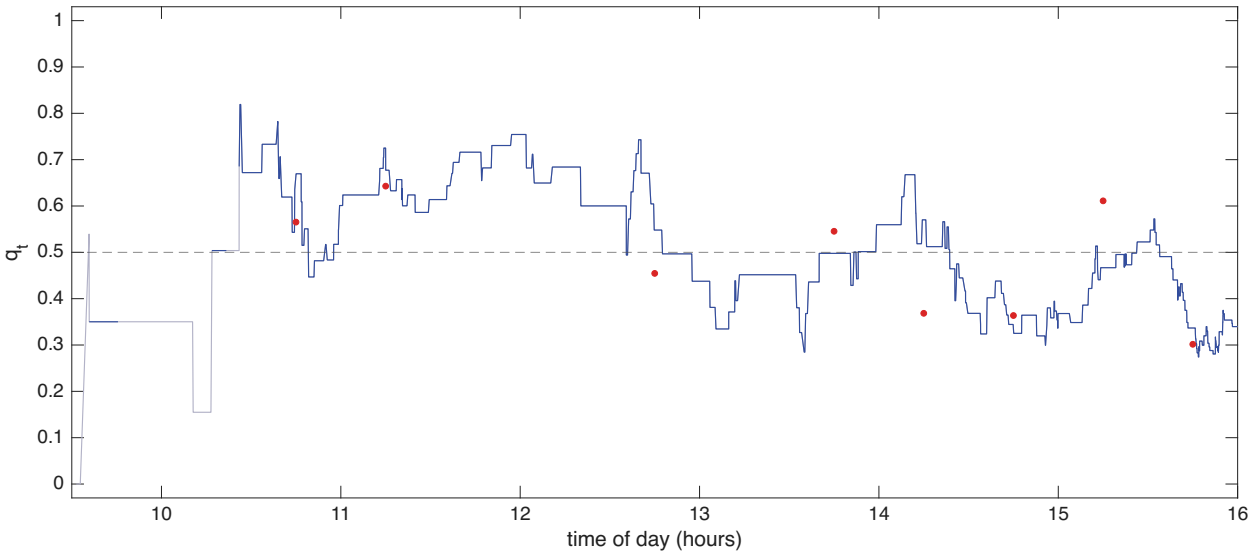

Figure 8: The directional conditional probability for LANC, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 1130 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

9 of the 22 days built (10 presample and 12 analysis) have fewer than 5,000 trades.
The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
The QLIKE-minimizing a lies at the boundary of the range searched.
2 analysis day(s) set the opening noise scale from fewer than 5 previous openings.
On 12 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 30).
On 11 October 2024, 45 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 26 per cent on any other day. Every number that day contributes to carries that burst.
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
Where the tail index can be estimated it averages over as few as 1 depth on one day, because the largest 0.1 to 1 per cent of 3,435 residuals is a short stretch of the tail. It is a much less precise estimate than for a heavily traded name.
The largest Huber residual on 16 October 2024 occurs at 09:36, within the opening minutes; the volatility comparison window is placed at the largest residual after 09:45, at 11:13.

Diagnostics.

IPAR 1 to 16 October 2024

trades on 16 October 2024	4,498
average trading rate	0.2 per second
presample for the diurnal profile	17 to 30 September 2024
outliers on 16 October 2024, o	244 (5.42%)
price half-life: run / minimizer[1]	3.0 / 2.2 trades
cost of running 3 trades	+1.78%
scale half-life: run / minimizer[2]	240 / 600 s
cost of running 240 s	+0.09%
direction half-life[3]	545 s [450, 690]
Mandelbrot a minimizing QLIKE	26 trades
a within 0.0005 of the minimum	26 to 26
linear $\hat{\sigma}$, before $\rightarrow$ at the largest residual[4]	8% $\rightarrow$ 12%
Huber $\hat{\sigma}$, before $\rightarrow$ peak	8% $\rightarrow$ 10%
median spot volatility, Huber / linear input[5]	12.8 / 12.5 %
hyperbolic against HExp, correlation of logs	0.9880
HExp QLIKE advantage, leave-one-day-out[6]	+0.96749
15-minute signature level	20 to 41 %
Huber signature, 1 min over 1 s, 12 day average	1.40

IPAR over the 12 trading days 1 to 16 October 2024, and on 16 October 2024. [1]the median over k from 2 to 100 of the half-life minimizing the k-trade-ahead Huber forecast loss, and the largest cost over those k of running the common 3 trades. [2]the 1-minute-ahead check loss of the noise scale. [3]the half-life maximizing the mean predictive log-likelihood per scored trade on 16 October 2024, with the range within 10^{-4} of the maximum. [4]the largest Huber residual after 09:45, at 14:58; before is the median over $t^* - 15$ to $t^* - 3$ minutes and at is the maximum over $t^* - 6$ to $t^* + 12$ minutes. [5]over the 12 days, sampled every 300 trades. [6]hyperbolic minus HExp, so positive favours HExp.

	Mean of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert/\sqrt{k}$ in cents				Sqrt of Mean of $(\hat{X}_t - \hat{X}_{t-k})^2/k$ in cents			
k	median	Huber	linear	trades	median	Huber	linear	trades
1	1.316	1.027	0.995	4.363	2.855	1.579	1.519	8.291
2	1.284	1.144	1.131	3.633	2.580	1.725	1.663	6.365
5	1.528	1.316	1.302	2.848	2.528	1.928	1.870	4.480
10	1.635	1.418	1.405	2.360	2.428	2.025	1.988	3.496
25	1.611	1.459	1.439	1.907	2.255	2.048	2.015	2.715
	ACF of $\hat{X}_t - \hat{X}_{t-1}$ at lag k				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t-k}\rvert$			
1	−0.18	0.19	0.20	−0.41	4.38	5.19	5.05	3.74
2	0.10	0.10	0.11	−0.04	5.04	3.59	5.23	4.26
5	−0.03	0.00	0.02	−0.02	5.07	5.07	5.88	5.37
10	0.02	0.01	0.00	0.00	5.04	6.29	5.86	5.45
25	0.04	−0.01	0.00	−0.01	4.95	4.74	4.53	5.37

Table 1: Trade-lag differences, `IPAR`, 16 October 2024.

	Mean of $\frac{\lvert\hat{X}_t - \hat{X}_{t,s}\rvert}{\sqrt{\tau_t - \tau_{t,s}}}$ in cents				Annualized volatility: $\hat{\sigma}_{\text{Annual};s}$			
s	median	Huber	linear	trades	median	Huber	linear	trades
0.1 s	2.47	2.16	2.11	4.45	18.6	11.4	11.4	66.2
1 s	2.00	1.75	1.73	3.23	17.6	11.1	11.2	62.7
5 s	1.57	1.40	1.38	2.51	17.5	12.3	12.1	56.2
10 s	1.41	1.28	1.27	2.16	17.8	13.8	13.6	50.6
1 min	1.03	0.93	0.92	1.38	18.9	16.9	16.6	36.3
15 min	0.89	0.87	0.87	0.91	20.2	19.7	19.5	22.0
	$\text{Cor}(\hat{X}_t - \hat{X}_{t,s}, \hat{X}_{t,s} - \hat{X}_{t,2s})$				Hill tail index of $\lvert\hat{X}_t - \hat{X}_{t,s}\rvert$			
0.1 s	−0.04	−0.02	−0.01	−0.10	7.58	5.33	4.64	8.09
1 s	−0.02	0.06	0.06	−0.09	7.90	5.55	4.89	8.20
5 s	−0.07	0.03	0.02	−0.23	7.14	10.49	9.53	10.19
10 s	−0.03	0.04	0.04	−0.18	7.81	13.97	10.80	5.39
1 min	−0.19	−0.15	−0.15	−0.39				

Table 2: Clock-lag differences, `IPAR`, 16 October 2024.

	median	Huber	linear
Mean of $\lvert\hat{\epsilon}_t\rvert$ in cents	3.04	3.68	3.80
s.d. of $\hat{\epsilon}_t$ in cents	6.14	5.80	5.77
Tail index of $\psi(\hat{\epsilon}_t)$	∞	∞	5.02
ACF of $\psi(\hat{\epsilon}_t)$, lag 1	0.21	0.24	0.21
ACF of $\psi(\hat{\epsilon}_t)$, lag 2	0.09	0.13	0.10
ACF of $\psi(\hat{\epsilon}_t)$, lag 5	0.01	0.02	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 10	−0.02	0.00	0.01
ACF of $\psi(\hat{\epsilon}_t)$, lag 25	0.00	0.00	0.00
80% C.I. for $\theta_t - X_t$	—	−2.64,2.42	−2.73,2.57
95% C.I. for $\theta_t - X_t$	—	−5.00,4.57	−5.31,4.80
99% C.I. for $\theta_t - X_t$	—	−8.87,7.93	−8.84,8.50
s.e. of $\theta_t - X_t$ in cents	—	2.29	2.47

Table 3: Residual summaries, `IPAR`, 16 October 2024. Median C.I. and s.e. cells are omitted (price discreteness); the linear s.e. is printed because the Hill index of its score, 5.02, is above two.

	16 October 2024		12 days, min to max	
	upper	lower	upper	lower
All trades (as in the paper)	3.40	3.14	1.71 to 3.65	1.93 to 3.76
Residual divided by δ_t	2.34	3.11	2.07 to 4.12	2.78 to 4.32
Repricing windows removed	3.40	3.14	1.72 to 3.65	1.93 to 3.99
Calm windows only	3.39	2.95		
Every 10th trade	—	—	— to —	— to —
Six-trade half-life	3.58	3.42	1.59 to 4.00	2.15 to 4.67

Table 4: Hill estimator of the tail index of the Huber filter residual, `IPAR`, 12 analysis days. The index is the mean of Hill estimates over the largest 0.1% to 1% of residuals, each tail separately, which for this stock is an average over 1 to 5 depths a day in row 1; the masked and thinned rows use fewer. Second row: divides the residual by its own threshold δ_t. Third row: drops trades where the filtered price moves more than 6.8 times the day's median δ_t within ± 1 second, which is 60.4 to 100.6 cents over the days and removes 0% to 2% of a day; the fourth keeps only trades where it moves less than 3.4 times that median within ± 5 seconds, 30.3 cents on 16 October 2024, 96% of that day. Fifth row: keeps every 10th trade. Sixth row: doubles the filter's half-life with the thresholds held fixed. The ranges run over the days on which the row can be estimated.

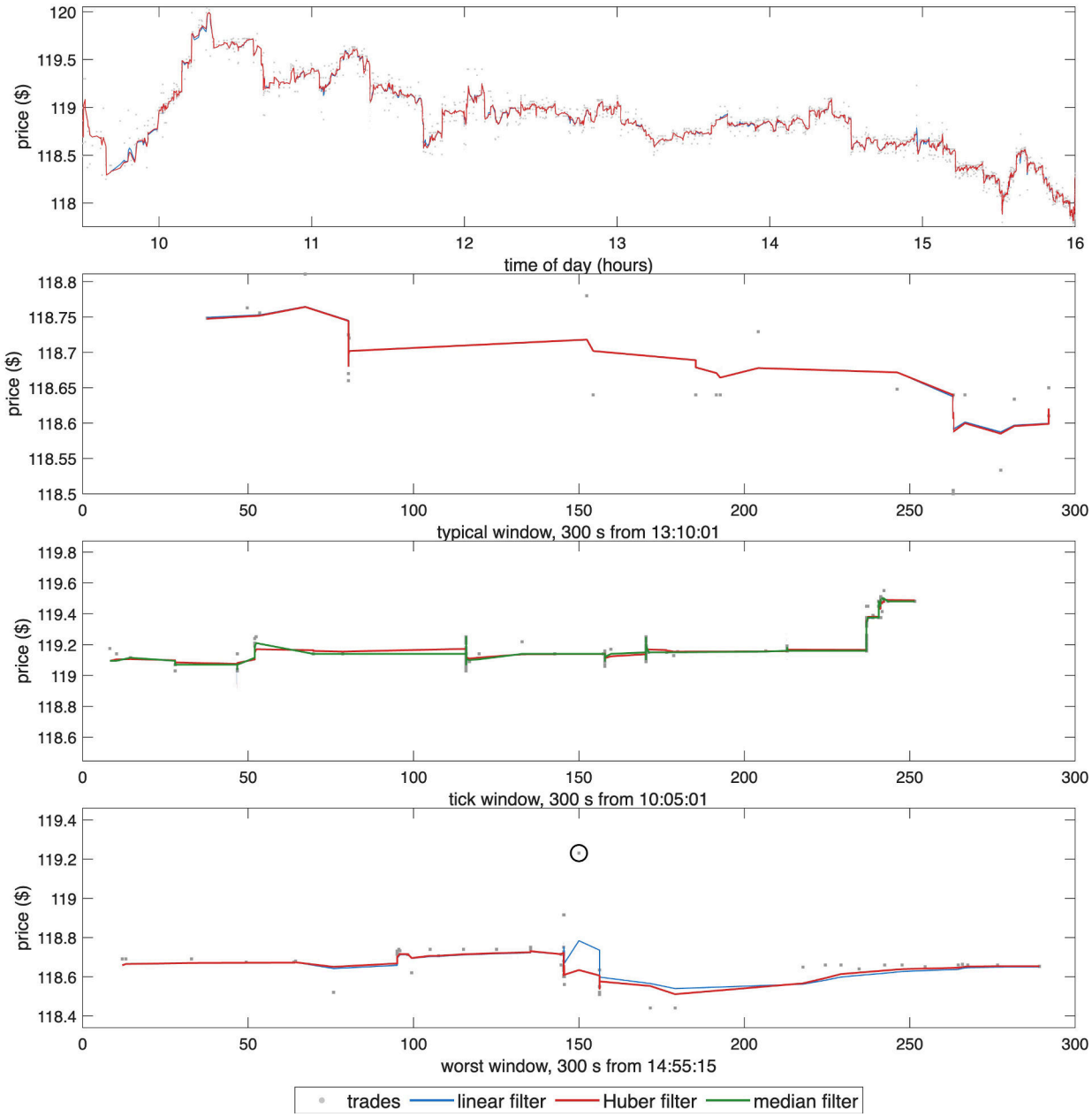

Figure 1: `IPAR` trades (grey points), the linear filter and the Huber filter, 16 October 2024. Top: full trading day. 2nd: 300-second window at the day's median trade count, from 13:10:01. 3rd: 300-second window at 10:05:01 in which the price reverses direction most often, with the median filter m_t added. Bottom: 300-second window around the largest Huber residual after 09:45, circled.

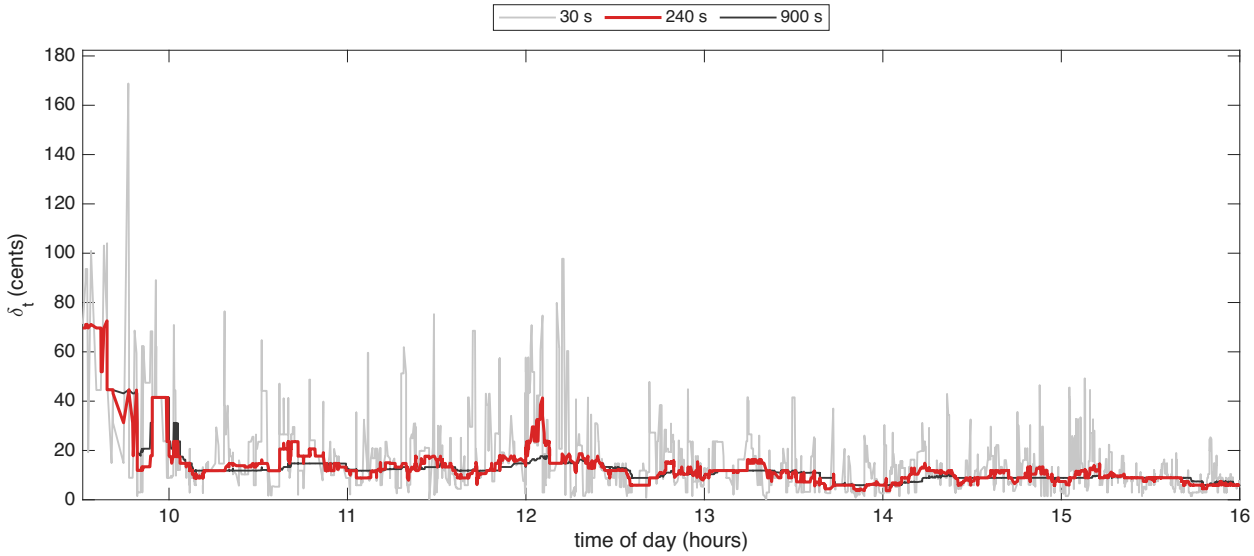

Figure 2: The Huber threshold δ_t in cents for `IPAR` on 16 October 2024, with half-lives of 30, 240, 900 seconds. The 240-second half-life, in red, is used.

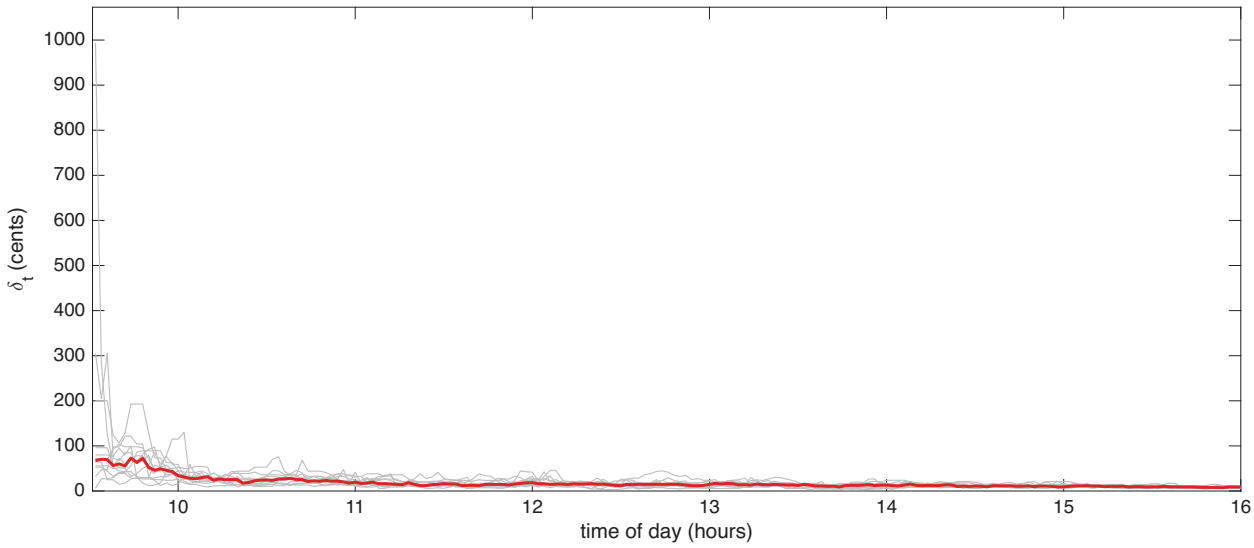

Figure 3: Huber threshold δ_t in cents through the trading day, for 12 `IPAR` trading days in Oct 2024 (grey lines) and their median (red line).

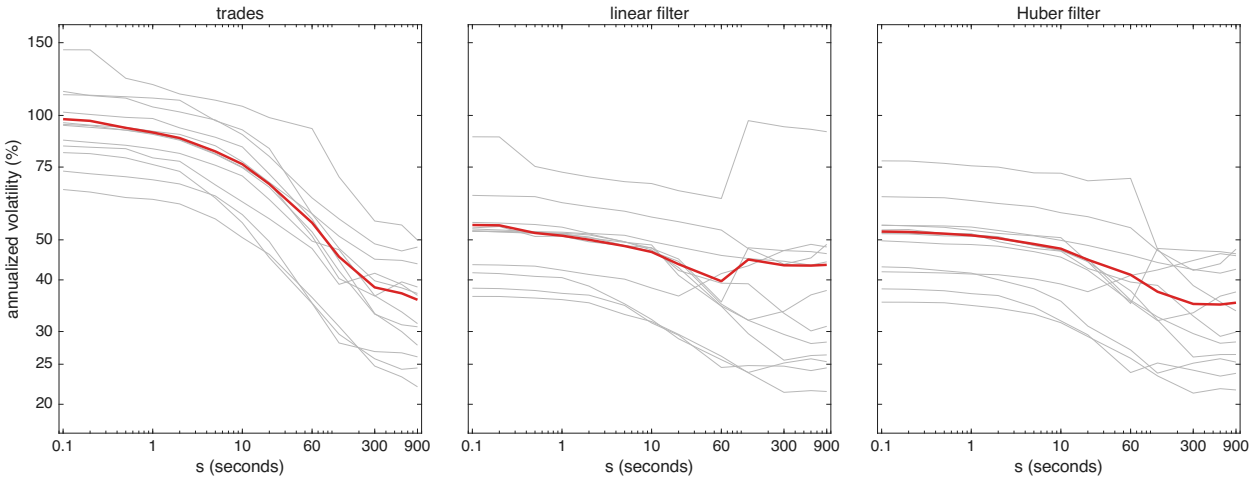

Figure 4: Volatility signature curves for `IPAR` over the 12 trading days in Oct 2024. The annualized vol $\hat{\sigma}100\sqrt{252}$ from the returns $\hat{X}_t - \hat{X}_{t,s}$ of Table 2, against the lag s in seconds, both axes on log scales, for: raw trades, linear filter & Huber filter. Grey lines are the individual trading days & the red line is the square root of the 12 day average of their squares. The linear & Huber returns carry the finite-sample correction $c_{t,s}$ of the paper; the raw trades are unaffected by it.

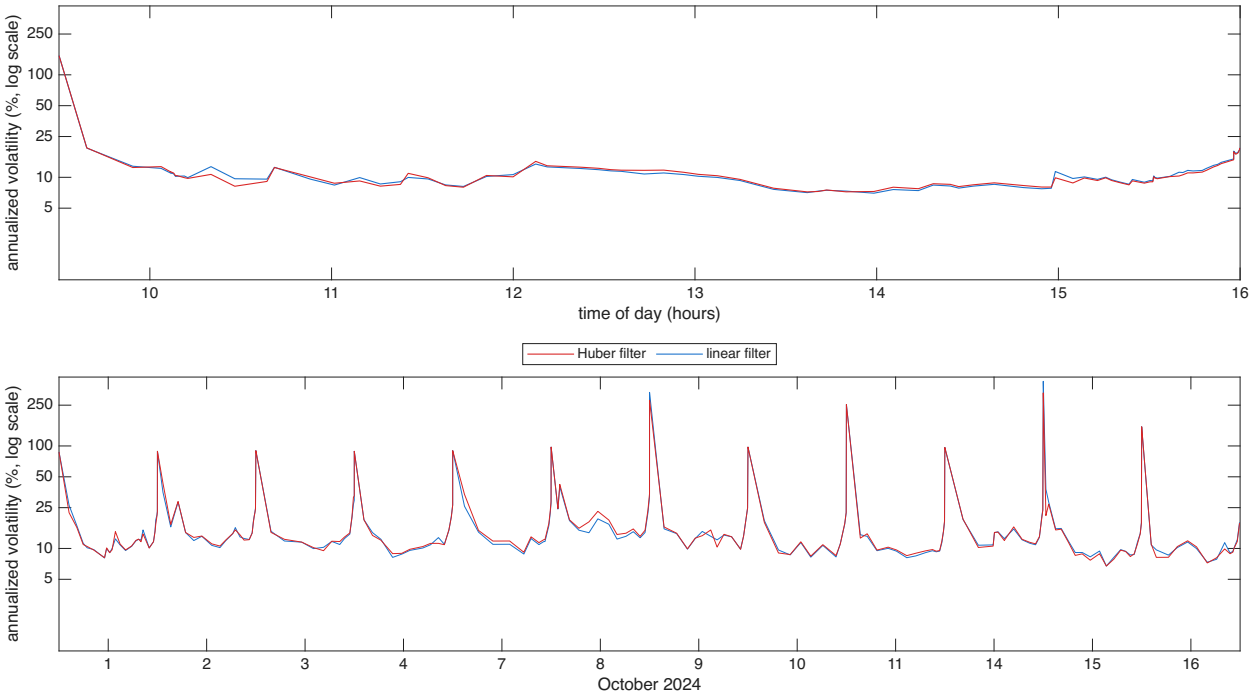

Figure 5: Filtered spot volatility for `IPAR` over 12 trading days in Oct 2024, the diurnally adjusted estimator $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second with hyperbolic weights, Mandelbrot parameter $a = 26$, with the input prices computed using the Huber & linear filters, shown as annualized volatility in percent on a log10 scale. Top: 16 October 2024, bottom: 12 days.

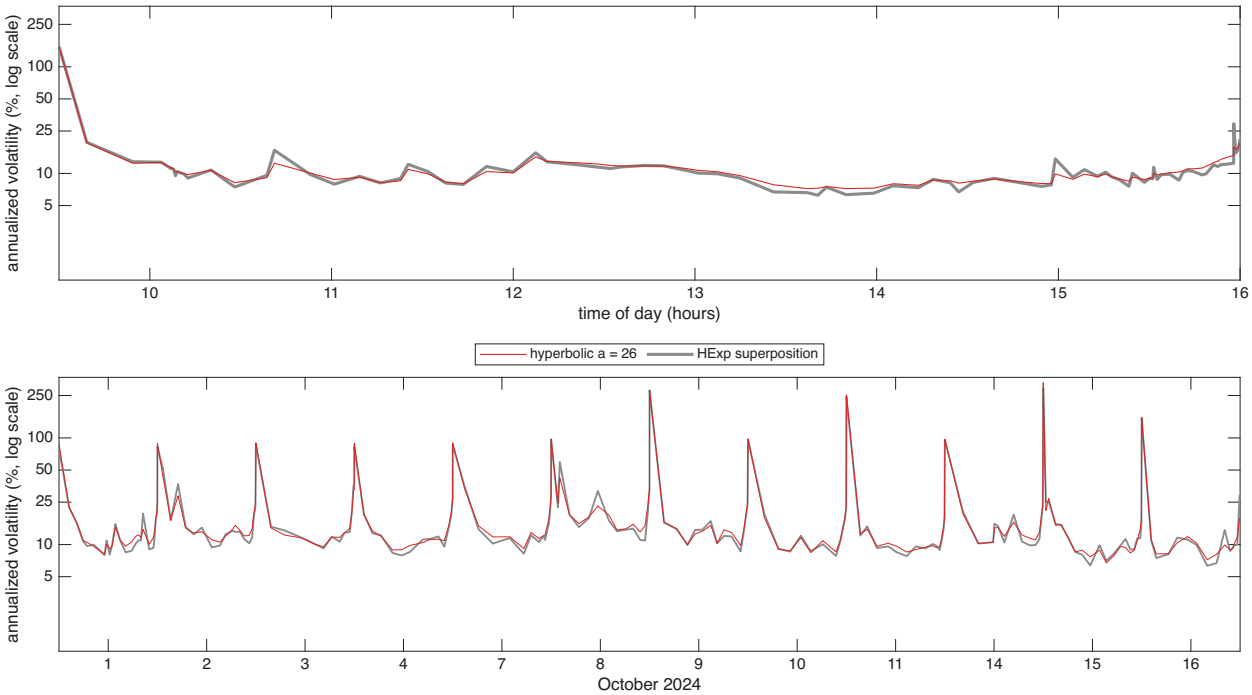

Figure 6: Hyperbolic against superposition weights for `IPAR` over 12 trading days in Oct 2024, the diurnally adjusted $\hat{\sigma}(u)100\sqrt{252}$ at $s = 1$ second from the Huber filter, shown as annualized vol in % on a log10 scale. The only difference between the 2 paths is the weight function: the hyperbolic weights with $a = 26$ in red, drawn on top, and the `HExp` superposition in grey, 4 exponential components with half lives of 10, 100, 1000, 10000 seconds. Hyperparameters minimize the QLIKE objective. Top: 16 October 2024, bottom: 12 days.

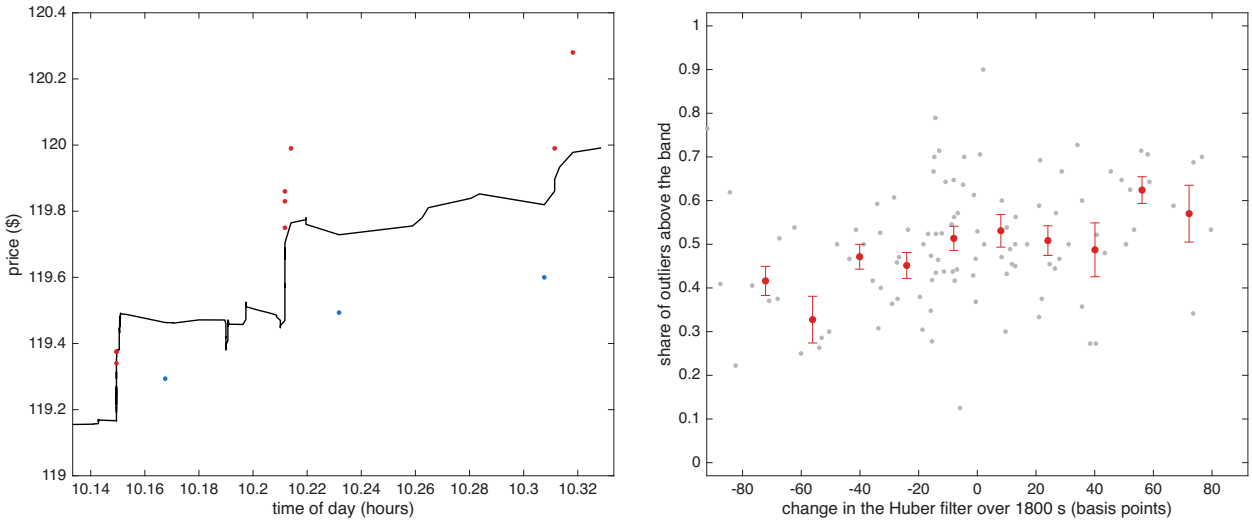

Figure 7: Outliers & robust price for `IPAR` on 16 October 2024. Left: the 12 minutes with the largest move in the Huber filter that day; outliers above $Y_t - \theta_t > \delta_{t-1}$ in red, outliers below $Y_t - \theta_t < -\delta_{t-1}$ in blue, Huber filter in black. Right: over the 12 trading days, the 1800 second empirical conditional probability that outliers are above (rather than below) the Huber band against the 1800 second change in the Huber filter in basis points: light points are single bins, the line & bars are means & s.e.

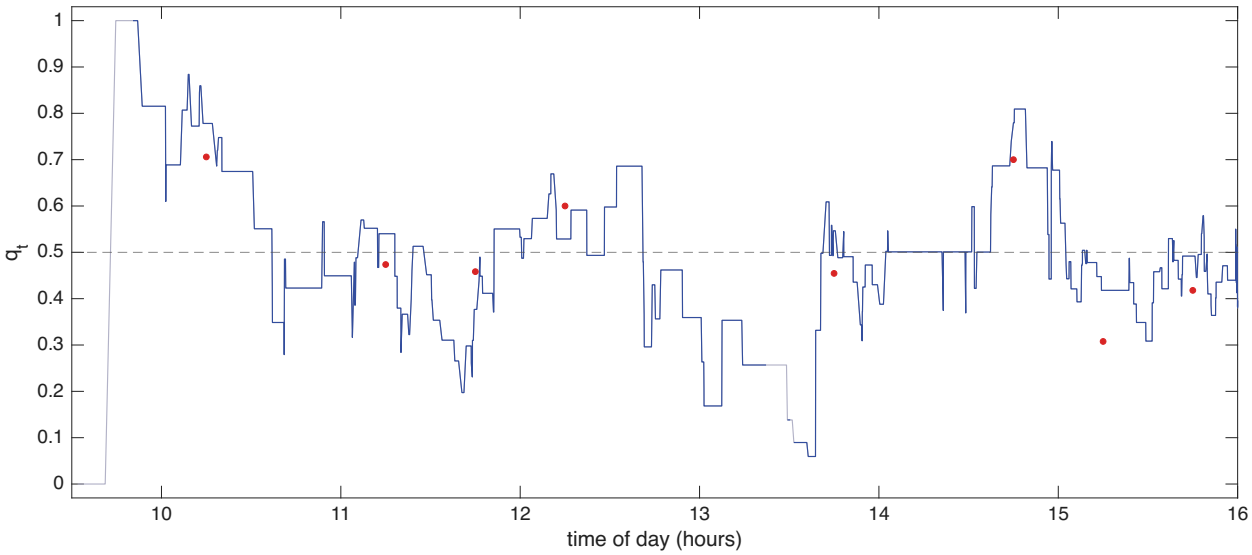

Figure 8: The directional conditional probability for `IPAR`, 16 October 2024. The filtered conditional probability q_t that an outlier lies above the robust Huber band, under exponential weights decaying in clock time with half life of 545 seconds; where fewer than two outliers remain in the weights it is drawn in grey. The dashed line marks 0.5. The red points are the empirical conditional probability using the 1800 second bins of the previous figure.

14 of the 22 days built (10 presample and 12 analysis) have fewer than 5,000 trades.
The direction half-life grid was extended beyond its usual end to 1500 seconds, because the log-likelihood was still rising there.
Selecting a on single days gives values from 4 to 26112 trades, against 26 trades pooled over the 12 days; the 10-second estimate of the spot volatility behind the QLIKE objective averages roughly 2 trades for this stock.
The Flexible Fourier fit of the diurnal profile has R-squared below 0.2 on 1 day(s) (lowest 0.18).
On 12 analysis day(s) the opening noise scale rests on fewer than 100 nonzero residuals from the previous openings (fewest 25).
The spot volatility filter has no return to work with for the first 25 seconds of the slowest opening, because the first continuous trade comes late and its reference must be 1 second old.
On 2 October 2024, 37 per cent of the day's largest one per cent of residuals fall inside its first minute, against at most 29 per cent on any other day. Every number that day contributes to carries that burst.
The tail index of the residual cannot be estimated for every 10th trade, on none of the 12 days: what the row's mask or thinning leaves is too few residuals for the largest 0.1 to 1 per cent to hold enough order statistics.
Where the tail index can be estimated it averages over as few as 1 depth on one day, because the largest 0.1 to 1 per cent of 3,152 residuals is a short stretch of the tail. It is a much less precise estimate than for a heavily traded name.

Diagnostics.