Noise-Robust Realized Variance: Simulation with Jumps and Noise Dependence, and Realized GARCH Application

Ton Vossen
Maastricht University

June 12, 2025

Abstract

This paper evaluates the performance of several realized variance (RV) estimators in the presence of market microstructure noise and jumps. We first replicate the original simulation design of Zhang et al. (2005) and confirm that their ZMA estimator delivers nearly unbiased and low-variance estimates under ideal conditions. We then extend the framework by introducing compound-Poisson jumps and varying the structure of microstructure noise to include heteroskedasticity and serial correlation. Our simulations show that while ZMA remains robust to jumps and heteroskedastic noise, its performance deteriorates when noise exhibits serial dependence. Finally, we apply the ZMA and 5-minute estimators to one year of one-second S&P 500 data and evaluate them within a Realized GARCH(1,1) framework. The ZMA-based model significantly outperforms the naive alternative in out-of-sample density forecasting, as confirmed by Vuong test statistics. These findings support the use of ZMA in practical settings, provided that noise autocorrelation is not severe.

1 Introduction

Accurate measurement of volatility is essential for risk management, portfolio allocation, and derivative pricing. Since Andersen and Bollerslev (1998) demonstrated the benefits of high-frequency data, realized variance (RV)—the sum of squared intraday returns—has become the benchmark estimator of integrated variance. However, naive RV is biased in the presence of market microstructure noise and jumps (Zhang et al., 2005).

This paper makes three contributions:

1. We replicate the original Monte Carlo design of Zhang et al. (2005) to validate the performance of the ZMA estimator under their assumptions.

2. We extend the simulation framework in two directions: first by introducing compound-Poisson jumps, and second by dissecting the effects of microstructure noise assumptions (homoskedasticity vs. serial correlation). We find that ZMA remains robust to jumps but deteriorates sharply when the noise exhibits serial dependence.

3. We embed each RV proxy in a Realized GARCH(1,1) model (Hansen et al., 2011) and evaluate density forecasts on one year of one-second S&P 500 data. The ZMA-based model significantly outperforms a 5-minute sampled RV in out-of-sample likelihood.

The code is available online[1].

[1]https://github.com/KoperSloper/EE2

2 Issues with Realized Variance

Let $X_t = \log S_t$, where S_t is the asset price at time t. For a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, the realized variance of X over $[0, T]$ is

$$[X]_T^{\Pi} = \sum_{i=1}^{n} \big(X_{t_i} - X_{t_{i-1}}\big)^2,$$

which simply sums the squared increments of the log-price.

In the continuous-time limit (as the mesh $\|\Pi\| \to 0$), $[X]_T^{\Pi}$ converges (when it exists) to the quadratic variation: $\langle X \rangle_T = \lim_{\|\Pi\| \to 0} [X]_T^{\Pi}$. Under the Itô process $dX_t = \mu_t\,dt + \sigma_t\,dW_t$, where W_t is Brownian motion, one shows

$$\langle X \rangle_T = \int_0^T \sigma_t^2\,dt.$$

In real data we don't observe the true log-price but a noisy version: $Y_{t_i} = X_{t_i} + \varepsilon_{t_i}$, where ε_{t_i} is zero-mean, i.i.d. microstructure noise. If we define the filtration $\mathcal{F}_T^X = \sigma\big(X_s : 0 \le s \le T\big)$, then conditioning on the true path of X gives

$$\begin{aligned}
\mathbb{E}\big[[Y]_T^{\Pi} \mid \mathcal{F}_T^X\big] &= \sum_{i=1}^{n} \mathbb{E}\big[(\Delta Y_{t_i})^2 \mid \mathcal{F}_T^X\big] = \sum_{i=1}^{n} \mathbb{E}\big[(\Delta X_{t_i} + \Delta\varepsilon_{t_i})^2 \mid \mathcal{F}_T^X\big] \\
&= \sum_{i=1}^{n} (\Delta X_{t_i})^2 \;+\; 2\sum_{i=1}^{n} \Delta X_{t_i}\,\mathbb{E}[\Delta\varepsilon_{t_i} \mid \mathcal{F}_T^X] \;+\; \sum_{i=1}^{n} \mathbb{E}[(\Delta\varepsilon_{t_i})^2 \mid \mathcal{F}_T^X] \\
&= [X]_T^{\Pi} \;+\; 2n\ \mathrm{Var}(\varepsilon) \;\approx\; \langle X \rangle_T \;+\; 2n\ \mathrm{Var}(\varepsilon).
\end{aligned}$$

Here the cross-term drops out since $\mathbb{E}[\Delta\varepsilon_{t_i} \mid \mathcal{F}_T^X] = 0$, and each noise increment $\Delta\varepsilon_{t_i}$ has variance $2\ \mathrm{Var}(\varepsilon)$. Even without noise, the discrete sum $[X]_T^{\Pi}$ differs from $\langle X \rangle_T$ by a discretization error that vanishes as $\|\Pi\| \to 0$.

A more detailed large-n analysis, from the main paper on which this work is based on (Zhang et al., 2005), shows that the noise-contaminated estimator $[Y]_T^{\Pi}$ satisfies in distribution

$$[Y]_T^{\Pi} \overset{\mathcal{L}}{\approx} \langle X \rangle_T \;+\; 2n\ \mathrm{Var}(\varepsilon) \;+\; \sqrt{4n\,\mathbb{E}[\varepsilon^4] \;+\; \frac{2T}{n}\int_0^T \sigma_t^4\,dt}\; Z,$$

where $Z \sim N(0,1)$. The naive realized-variance estimator $[Y]_T^{\Pi}$ is therefore biased upward by $2n\ \mathrm{Var}(\varepsilon)$, and this bias grows with the sampling frequency n. In fact, as $n \to \infty$, the estimator becomes dominated by the microstructure noise rather than the true variation $\langle X \rangle_T$.

3 The Four Estimators

In this paper we compare four methods to estimate the integrated variance $\langle X \rangle_T$ from noisy high-frequency observations. In addition to the usual full–sample Naïve RV, practitioners often compute a sparse-sampled (e.g. 5-minute) RV to mitigate microstructure noise. We therefore include that "5-min Naïve" as its own estimator, alongside Boosted (Subsampled) RV and ZMA.

3.1 Naïve Realized Variance (Full-Sample)

The simplest estimator is

$$[Y]_T^{\Pi} \;=\; \sum_{i=1}^{n} (\Delta Y_{t_i})^2, \qquad \Delta Y_{t_i} = Y_{t_i} - Y_{t_{i-1}},$$

where $\{t_0, \dots, t_n\}$ are all observed tick times. This "tick-by-tick" RV, as we already saw, is heavily biased in the presence of microstructure noise.

3.2 Naïve Realized Variance (5-Minute Sampling)

A common practical workaround is to sample at a fixed interval (e.g. every 5 minutes) rather than every tick. Let $\{s_0, s_1, \dots, s_m\}$ be equally spaced calendar-time points (e.g. $s_j = j \cdot 5\,\text{minutes}$). Then a sparse-sampled Naïve RV is

$$[Y]_T^{5\text{min}} = \sum_{j=1}^{m} \big(Y_{s_j} - Y_{s_{j-1}}\big)^2.$$

By discarding most microstructure noise at ultra-high frequency, practitioners often find that 5-minute RV performs better.

3.3 Boosted (Subsampled) RV

Partition the full tick grid Π into K disjoint subsamples $\Pi^{(k)}$. For example, take every Kth tick with offsets $k = 1, \dots, K$. Each subsample yields

$$[Y]_T^{(k)} = \sum_{t_i \in \Pi^{(k)}} (\Delta Y_{t_i})^2.$$

Zhang et al. (2005) show asymptotically

$$[Y]_T^{(k)} \approx \langle X \rangle_T + 2\,n_k\ \mathrm{Var}(\epsilon) + \sqrt{4\,n_k\,\mathbb{E}[\epsilon^4] + \frac{2T}{n_k}\int_0^T \sigma_t^4\,dt}\ Z_k, \qquad Z_k \sim N(0,1),$$

where $n_k \approx n/K$ is the number of observations in subsample k. Averaging across the K subsamples gives

$$[Y]_T^{\text{avg}} = \frac{1}{K}\sum_{k=1}^{K}[Y]_T^{(k)} \approx \langle X \rangle_T + 2\,\bar{n}\ \mathrm{Var}(\epsilon) + \sqrt{\frac{4\,\bar{n}}{K}\,\mathbb{E}[\epsilon^4] + \frac{4T}{3\,\bar{n}}\int_0^T \sigma_t^4\,dt}\ Z,$$

where $\bar{n} \approx n/K$. Averaging cuts both bias and variance, but leaves a residual bias of $2\,\bar{n}\ \mathrm{Var}(\epsilon)$. Balancing this bias against discretization error yields the optimal

$$\bar{n}^* = \Big(\frac{T}{6\ \mathrm{Var}(\epsilon)^2}\int_0^T \sigma_t^4\,dt\Big)^{1/3}, \qquad K^* \approx \frac{n}{\bar{n}^*}.$$

3.4 ZMA Bias-Corrected Estimator

We now have these two estimators:

$$[Y]_T^{\text{full}} = \sum_{i=1}^{n}(\Delta Y_{t_i})^2 \approx \langle X \rangle_T + 2\,n\ \mathrm{Var}(\epsilon), \quad [Y]_T^{\text{avg}} \approx \langle X \rangle_T + 2\,\bar{n}\ \mathrm{Var}(\epsilon).$$

Solving

$$\frac{1}{\bar{n}}\,[Y]_T^{\text{avg}} \;-\; \frac{1}{n}\,[Y]_T^{\text{full}} \;\approx\; \Big(\tfrac{1}{\bar{n}} - \tfrac{1}{n}\Big)\langle X \rangle_T$$

for $\langle X \rangle_T$ leads to the ZMA estimator

$$[Y]_T^{\text{ZMA}} := \frac{1}{n - \bar{n}}\Big(n\,[Y]_T^{\text{avg}} \;-\; \bar{n}\,[Y]_T^{\text{full}}\Big).$$

Under optimal sampling $K = c\,n^{2/3}$, Zhang et al. (2005) show

$$n^{1/6}\big([Y]_T^{\text{ZMA}} - \langle X\rangle_T\big) \xrightarrow{d} N\Big(0,\ 8\,c^{-2}\ \text{Var}(\epsilon)^2\ +\ c\,\eta^2\,T\Big),\quad \eta^2 = \tfrac{4}{3}\int_0^T \sigma_t^4\,dt.$$

Minimizing the asymptotic variance over c yields

$$c^* = \left(\frac{16\ \text{Var}(\epsilon)^2}{T\,\eta^2}\right)^{1/3}, \qquad K^* = \lfloor c^*\,n^{2/3}\rceil.$$

These four estimators—(1) Naïve full-sample RV, (2) Naïve 5-minute RV, (3) Boosted/Subsampled RV, and (4) ZMA—will be used in subsequent parts, each implemented with its theoretically optimal tuning.

4 Monte Carlo Simulations

We perform three rounds of simulations. First, we replicate the setup of Zhang et al. (2005) to confirm the performance of the ZMA estimator under idealized conditions with i.i.d. microstructure noise and no jumps. Then, we extend the design by adding compound-Poisson jumps to assess how the estimator behaves in the presence of price discontinuities. Finally, we explore the role of microstructure noise assumptions by isolating two key features: time-varying noise variance (heteroskedasticity) and serial correlation. This allows us to identify which assumption is critical for ZMA's robustness and where its performance begins to deteriorate.

4.1 First Simulation Design

Following Zhang et al. (2005), the log-price follows a Heston (Heston, 1993) stochastic-volatility model:

$$\begin{aligned} d\log X_t &= \Big(\mu - \tfrac{1}{2}\,v_t\Big)\,dt + \sqrt{v_t}\,dW_t^X, \\ dv_t &= \kappa(\alpha - v_t)\,dt + \gamma\sqrt{v_t}\,dW_t^v, \\ d\langle W^X, W^v\rangle_t &= \rho\,dt. \end{aligned}$$

The authors' original parameters $\mu = 0.05$, $\kappa = 5$, $\alpha = 0.04$, $\gamma = 0.5$, and $\rho = -0.5$ are used. Simulations run over one trading day ($T = 1/252$), with one-second time steps. Market microstructure noise is added as i.i.d. Gaussian errors with standard deviation 0.0005 (0.05% of price). We draw $M = 1{,}000$ paths and estimate the integrated variance $\langle X\rangle_T$ via the four estimators.

4.2 Results for Design 1

Table 1 reports bias, variance, and RMSE for each estimator across the 1,000 trials. The Naive estimator remains both highly biased and noisy; the 5-minute Naïve sampling sharply reduces bias and variance compared to full-sample Naive but is still outperformed by Subsample; Subsample further cuts both bias and variance; and ZMA achieves essentially zero bias with the lowest variance and RMSE overall. These results closely match those reported in Zhang et al. (2005), confirming the validity of their original simulation findings.

4.3 Second Simulation Design

Next, we enrich the model with compound-Poisson jumps. The jump component follows a Merton (1976) jump-diffusion model, while the stochastic volatility component is (again) based on the Heston (1993) model. Applying Itô's formula extended to jump processes, the log-price evolves as

$$d\log X_t\ =\ \Big(\mu\ -\ \tfrac{1}{2}\,v_t\Big)\,dt\ +\ \sqrt{v_t}, dW_t^X\ +\ \log(J)\,dN_t,$$

Table 1: Estimator performance (Design 1)

Estimator	Bias	Variance	RMSE
Naive	$1.170395e{-2}$	$1.729824e{-8}$	0.011705
FiveMin	$4.017480e{-5}$	$1.150369e{-9}$	0.000053
Subsample	$1.652576e{-5}$	$8.053548e{-10}$	0.000033
ZMA	$-2.351594e{-7}$	$7.073966e{-11}$	0.000008

$$dv_t \;=\; \kappa(\alpha - v_t)\,dt + \gamma\sqrt{v_t}\,dW_t^v, \quad d\langle W^X, W^v\rangle_t = \rho\,dt.$$

Here, $N_t \sim \text{Poisson}(\lambda t)$ is a homogeneous Poisson process with intensity λ, and each jump size $J \sim \text{LogNormal}(\mu_J, \sigma_J^2)$.

The true quadratic variation is then

$$[X]_T \;=\; \int_0^T v_t\,dt \;+\; \sum_{0<t\leq T} \bigl(\log J_t\bigr)^2.$$

Including jumps is important because asset prices often exhibit sudden, discrete changes (e.g., due to earnings announcements or macroeconomic news), which cannot be captured by pure diffusion models. Although the original authors assert that "as far as the effect of jumps on the discretization effect is concerned, it should be noted that consistency holds", they do not offer a simulation-based validation of this claim. This motivates this investigation in the presence of jumps.

4.4 Results for Design 2

Table 2 shows the bias, variance, and RMSE of the four estimators under the model with jumps. As before, the Naive estimator remains heavily biased and yields the largest RMSE. Sampling every 5 minutes (FiveMin) sharply reduces bias but still has higher variance and RMSE than Subsample. Finally, the ZMA estimator delivers essentially zero bias and an order-of-magnitude lower variance than Subsample, resulting in the smallest RMSE overall.

In other words, even after adding realistic features such as jumps, the relative ranking of the estimators is unchanged: ZMA is the most robust, Subsample offers a useful compromise, FiveMin improves over full-sample Naive, and Naive performs poorly once market frictions and discontinuities are present.

Table 2: Estimator Performance (Design 2)

Estimator	Bias	Variance	RMSE
Naive	$1.170390e{-2}$	$1.949298e{-8}$	0.011705
FiveMin	$3.799871e{-5}$	$9.748148e{-8}$	0.000314
Subsample	$7.648784e{-5}$	$2.611008e{-8}$	0.000179
ZMA	$4.046063e{-7}$	$2.799186e{-9}$	0.000053

4.5 Third Simulation Design

The original ZMA assumes

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i}, \quad \mathbb{E}[\varepsilon_{t_i}] = 0, \quad \text{Var}(\varepsilon_{t_i}) = \sigma^2, \quad \text{Cov}(\varepsilon_{t_i}, \varepsilon_{t_j}) = 0 \;(i \neq j),$$

so that the noise is homoskedastic and serially independent. In practice, microstructure noise can exhibit time-varying variance and serial correlation (Hansen and Lunde, 2006). To diagnose the sensitivity of the ZMA estimator, we run three simulations using the same Heston price process (as in design 1) for the efficient price:

1. Heteroskedastic only: the noise variance oscillates intraday according to a sinusoidal pattern, but the noise increments are uncorrelated (no serial correlation).
2. Serially correlated only: the noise follows an AR(1) process with constant variance, with the autocorrelation coefficient set (rather arbitrarily) to -0.5. The negative sign reflects the well-documented effects of bid–ask bounce, which induces negatively correlated high-frequency returns (Bandi and Russell, 2008).
3. Both: the noise has time-varying variance and AR(1) serial dependence.

This allows us to pinpoint exactly where (and if) the ZMA estimator breaks down. Which assumption is critical for its performance: homoskedasticity, independence, or both?

4.6 Results for Design 3

Table 3 summarizes bias, variance, and RMSE of each estimator in the three scenarios.

Table 3: Estimator Performance under Different Noise Designs

Scenario	Estimator	Bias	Variance	RMSE
Heteroskedastic	Naive	1.1695×10^{-2}	2.0220×10^{-8}	1.1696×10^{-2}
	FiveMin	4.1193×10^{-5}	1.0210×10^{-9}	5.2058×10^{-5}
	Subsample	1.6814×10^{-5}	8.4393×10^{-10}	3.4330×10^{-5}
	ZMA	1.3806×10^{-7}	8.0487×10^{-11}	9.0132×10^{-6}
Serially Correlated	Naive	1.7532×10^{-2}	6.7012×10^{-8}	1.7534×10^{-2}
	FiveMin	3.9387×10^{-5}	1.1118×10^{-9}	5.2822×10^{-5}
	Subsample	1.4866×10^{-5}	7.8142×10^{-10}	3.2402×10^{-5}
	ZMA	-1.5859×10^{-4}	1.9826×10^{-10}	1.5867×10^{-4}
Both	Naive	1.7553×10^{-2}	7.2976×10^{-8}	1.7556×10^{-2}
	FiveMin	3.8717×10^{-5}	1.0828×10^{-9}	5.2255×10^{-5}
	Subsample	1.6667×10^{-5}	7.7193×10^{-10}	3.2400×10^{-5}
	ZMA	-1.5958×10^{-4}	2.1038×10^{-10}	1.5963×10^{-4}

The results reveal a clear pattern. When the noise is heteroskedastic but uncorrelated, ZMA remains highly accurate, with near-zero bias and minimal RMSE. In contrast, introducing serial correlation (while keeping the variance constant) leads to a pronounced deterioration in performance: ZMA develops a large negative bias and its RMSE increases by an order of magnitude. Combining both effects yields results nearly identical to the serial-only case.

We conclude that it is primarily the *serial correlation* in the microstructure noise, and not time-varying variance, that is crucial to the accuracy of the ZMA estimator.

5 Empirical Application to S&P 500 Tick Data

We apply the FiveMin and ZMA estimators to one-year of one-second S&P 500 data (May 2024–May 2025). After constructing the log-close series Y_t, we estimate for each trading day:

$$\widehat{\mathbb{E}[\varepsilon^2]} = \frac{1}{2n} \sum_{i=1}^{n} (\Delta Y_i)^2 = \frac{1}{2n} [Y, Y]_T^{(\mathrm{all})},$$

the microstructure-noise variance (Zhang et al., 2005), and the realized quad-power quarticity (Barndorff-Nielsen and Shephard, 2004)

$$\widehat{\mathrm{RQQ}} \;=\; \frac{M\pi^2}{4}\sum_{t=4}^{M}\bigl|\Delta Y_{t-3}\,\Delta Y_{t-2}\,\Delta Y_{t-1}\,\Delta Y_t\bigr|,$$

which consistently estimates $\int_0^T \sigma_t^4\,dt$. These feed into the data-driven choices

$$\bar{n}^* = \Bigl(\frac{T}{6\,\widehat{\mathbb{E}[\varepsilon^2]}^{\,2}}\,\widehat{\mathrm{RQQ}}\Bigr)^{1/3}, \quad K_{\text{boost}} = \lceil n/\bar{n}^* \rfloor,$$

$$\eta^2 = \tfrac{4}{3}\,\widehat{\mathrm{RQQ}}, \quad c^* = \Bigl(\frac{16\,\widehat{\mathbb{E}[\varepsilon^2]}^{\,2}}{T\,\eta^2}\Bigr)^{1/3}, \quad K_{\text{ZMA}} = \lceil c^*\,n^{2/3} \rfloor.$$

This procedure yields two daily time series of realized variance estimates—one from the FiveMinute estimator and one from ZMA—spanning the entire sample period. These series allow us to compare the estimators in a real high-frequency setting

5.1 Realized GARCH with Naïve 5min vs. ZMA RV

Benchmarking realized-variance estimators on real data is challenging because the true integrated variance is unobservable. Instead, we compare the two proxies—5-minute Naïve RV and ZMA RV—by plugging each into a Realized GARCH(1,1) model and evaluating one-day-ahead variance forecasts. The idea is simple: if one proxy tracks actual volatility more closely, then the model using that proxy should produce higher out-of-sample log-likelihoods on daily returns.

Specifically, we fit for each proxy the specification of Hansen et al. (2011):

$$\begin{aligned} r_t &= \sqrt{h_t}\,z_t, \\ h_t &= \omega + \beta\,h_{t-1} + \gamma\,x_{t-1}, \\ x_t &= \xi + \varphi\,h_t + \tau(z_t) + u_t, \end{aligned}$$

where r_t is the daily return, x_t the observed realized variance, $z_t \sim \text{i.i.d.}(0,1)$, $u_t \sim \text{i.i.d.}(0,\sigma_u^2)$.

According to the specification $\mathbb{E}[r_t \mid \mathcal{F}_{t-1}] = 0$ and $\mathrm{Var}(r_t \mid \mathcal{F}_{t-1}) = h_t$. We use the leverage function $\tau(z) = \tau_1 z + \tau_2(z^2-1)$ and assume Gaussian errors.

Using one year of one-second S&P 500 data, we construct daily returns r_t and two realized-variance series, FiveMin and ZMA. We generate one-day-ahead variance forecasts over a 150-day rolling window refitted every 10 days. Each day t in the evaluation window yields forecasted $\widehat{\sigma}_t^2$, and we compute the log likelihood score

$$\ell_t = \log\bigl[\varphi(r_t \mid 0, \widehat{\sigma}_t^2)\bigr].$$

Defining $d_t = \ell_t^{(\text{Naïve})} - \ell_t^{(\text{ZMA})}$, the average $\bar{d}$ measures which proxy produces better density forecasts.

Under the null of equal predictive ability, the Vuong test (Vuong, 1989) yields

$$Z_{\text{iid}} = -5.554, \quad p =< 0.001, \qquad Z_{\text{NW}} = -3.463, \quad p = 5\times 10^{-4},$$

where the latter uses a Newey–West adjustment for serial correlation and heteroskedasticity (Newey and West, 1987). The significantly negative statistics imply that the Realized GARCH model with ZMA realized variance produces substantially better out-of-sample log-scores—and hence more accurate density forecasts—than the model with the Naïve proxy.

6 Discussion

This paper compared several realized-variance estimators in both simulation and empirical settings, with a focus on their robustness to microstructure noise and jumps. First, we replicated the original model specification of Zhang et al. (2005) and confirmed their findings: ZMA outperforms both Naïve and Subsampled estimators under i.i.d. noise with no jumps. We then showed that ZMA also remains highly accurate when compound-Poisson jumps are added to the price process, demonstrating its robustness to discontinuities.

A key contribution of this study is to clarify the assumptions underpinning ZMA's effectiveness. Through controlled simulations, we find that ZMA remains accurate under heteroskedastic noise but breaks down when the noise exhibits serial correlation. Specifically, introducing AR(1) dependence (even with constant variance) induces large negative bias and substantially higher RMSE. These results underscore that ZMA's robustness relies critically on the assumption of serial independence in the noise process.

In the empirical application, we used one year of S&P 500 data to compare the 5-minute and ZMA estimators within a Realized GARCH model. Since true volatility is unobservable, we evaluated model performance via out-of-sample log likelihoods. The ZMA-based model delivered significantly higher scores, and the Vuong test confirmed that the improvement was statistically meaningful. This suggests that, despite its theoretical sensitivity to serial correlation (or perhaps because such dependence is limited in real data) ZMA continues to extract more informative variance signals from high-frequency observations in practice.

Overall, our findings indicate that while 5-minute sampling remains a simple and popular choice, noise-robust estimators like ZMA can offer substantial gains in accuracy and forecasting performance, provided the underlying noise is not strongly autocorrelated.

References

Andersen, T. G. and Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, 39(4):885.

Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, 75(2):339–369.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Hansen, P. R., Huang, Z., and Shek, H. H. (2011). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24(2):127–161.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 6(2):327–343.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.

Newey, W. K. and West, K. D. (1987). A simple, positive Semi-Definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and Non-Nested hypotheses. *Econometrica*, 57(2):307.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales. *Journal of the American Statistical Association*, 100(472):1394–1411.