The University of Chicago
Department of Statistics

TECHNICAL REPORT SERIES

Are Volatility Estimators Robust with Respect to Modeling Assumptions?

Yingying Li and Per A. Mykland

TECHNICAL REPORT NO. 564

February 15, 2006

The University of Chicago
Department of Statistics
5734 S. University Avenue
Chicago, IL 60637

Yingying Li is Doctoral Student, Department of Statistics, The University of Chicago, Chicago, IL 60637. E-mail: yyli@galton.uchicago.edu. Per A. Mykland is Professor, Department of Statistics, The University of Chicago, Chicago, IL 60637. E-mail: mykland@galton.uchicago.edu. We gratefully acknowledge the support of the National Science Foundation under grant DMS-0204639.

Are Volatility Estimators Robust with Respect to Modeling Assumptions? *

Yingying Li and Per A. Mykland

February 2006.
This version: February 15, 2006

Abstract

We consider microstructure as an arbitrary contamination of the underlying latent securities price, through a Markov kernel Q. Special cases include additive error, rounding, and combinations thereof. Our main result is that, subject to smoothness conditions, the two scales realized volatility (TSRV) is robust to the form of contamination Q. To push the limits of our result, we show what happens for some models involving rounding (which is not, of course, smooth) and see in this situation how the robustness deteriorates with decreasing smoothness. Our conclusion is that under reasonable smoothness, one does not need to consider too closely how the microstructure is formed, while if severe non-smoothness is suspected, one needs to pay attention to the precise structure and also to what use the estimator of volatility will be put.
KEY WORDS: Bias Correction; Local Time; Market Microstructure; Martingale; Measurement Error; Robustness; Realized Volatility; Subsampling; Two Scales Realized Volatility (TSRV).

1 INTRODUCTION

Recent years has seen an explosion of literature on the problem of estimating integrated volatility and similar objects with the help of high frequency data. For a sample of recent literature, see Hull and White (1987), Jacod and Protter (1998), Gallant, Hsu, and Tauchen (1999), Chernov and Ghysels (2000), Gloter (2000), Andersen, Bollerslev, Diebold, and Labys (2001), Dacorogna, Gençay, Müller, Olsen, and Pictet (2001), Barndorff-Nielsen and Shephard (2002), and Mykland and Zhang (2002) among others. An important realization has been that log prices do not appear to be semimartingales, but rather like semimartingales observed with error. Main hypotheses proposed

*Yingying Li is Doctoral Student, Department of Statistics, The University of Chicago, Chicago, IL 60637. E-mail: yyli@galton.uchicago.edu. Per A. Mykland is Professor, Department of Statistics, The University of Chicago, Chicago, IL 60637. E-mail: mykland@galton.uchicago.edu. We gratefully acknowledge the support of the National Science Foundation under grant DMS-0204639.

in the literature is that this error occurs by rounding (Delattre and Jacod (1997), Jacod (1994), Zeng (2003)) or by additive error (Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2004), Aït-Sahalia, Mykland, and Zhang (2005), Bandi and Russell (2003), Hansen and Lunde (2006)). More complex (and descriptive) models for microstructure are also available, see for example Hasbrouck (1996) and, coming from a very different perspective, by Farmer, Gillemot, Iori, Krishnamurthy, Smith, and Daniels (2005).

The multiplicity of ways in which errors can be modeled raises the question of how sensitive inference is to modeling assumptions. This is the topic of this paper.

We shall be making the assumption that there is a latent log price process X_t which is a continuous semimartingale on the form

$$dX_t = \mu_t dt + \sigma_t dB_t, \tag{1}$$

where μ_t and σ_t are continuous random processes, and B_t is a Brownian motion. Transactions at times $0 = t_0 < t_1 < ... < t_n = T$ give rise to log prices Y_{t_i} which are contaminated versions of X_{t_i} as follows. We suppose that there is a family $Q(x, dy)$ of conditional distributions so that given X_{t_i}, the law of Y_{t_i} is

$$P(Y_{t_i} \leq y \mid X \text{ process}) = P(Y_{t_i} \leq y \mid X_{t_i}) = Q(X_{t_i}, y). \tag{2}$$

In other words, Y_{t_i} is distributed around X_{t_i} in a way that only depends on the latter.

A simple example of such contamination Q is additive error on the log scale. If $Y = X + \epsilon$ where ϵ has density g and is independent of X, then

$$Q(x, dy) = g(y - x)dy. \tag{3}$$

Another example is rounding or truncation. In this case, $Q(x, dy)$ is a nonrandom distortion of x. We shall look at yet another form of contamination in Section 3.

This paper has two pieces of news, one good and one bad. We shall see that for reasonable types of contaminations Q, one can act as if the error was simply of additive type, and we shall see that the two scales realized volatility (TSRV) of Zhang, Mykland, and Aït-Sahalia (2005) is substantially robust to arbitrary contamination. This is our plan for Section 2.

There are, however, cases when one has to exercise care. We shall see one such case in Section 3, where we shall show that it is not always quite clear what is meant by volatility, and one has to consider carefully what quantity one actually wishes to estimate. This occurs in cases involving rounding.

We are mainly using TSRV by way of example of a volatility estimator, and believe that similar conclusions will apply to, for example, the multi-scale realized volatility (MSRV) of Zhang (2004). With caveats about bias and additional variance, similar conclusions will also apply to traditional realized volatility (RV).

2 ROBUSTNESS and SMOOTHNESS of CONTAMINATION

2.1 Setup

Suppose the latent log price process X follows (1). Let Y be the logarithm of the transaction price, which is observed at times $0 = t_0 < t_1 < ... < t_n = T$. We assume that at these sampling times, Y is related to the latent log price process X through (2). Let $f(X_t)$ be the conditional expectation of Y_t given the X process:

$$f(X_t) = E_Q(Y_t|X_t). \tag{4}$$

We shall assume that

$$f(x) \text{ is twice continuous differentiable.} \tag{5}$$

Note that under the assumption (5), $f(X_t)$ is a continuous semimartingale.

Now, we have two volatilities $\langle X, X\rangle_T = \int_0^T \sigma_t^2 dt$ and $\langle f(X), f(X)\rangle_T = \int_0^T f'(X_t)^2\sigma_t^2 dt$. One interesting question arises immediately – which quantity are we estimating? When we make use of the observations Y_{t_i} to estimate the volatility, we might think that we are estimating $\langle X, X\rangle_T$ because Y_{t_i} is just the contaminated version of X_{t_i}; but in fact, given the X process, Y_{t_i} is centered at $f(X_{t_i})$, rather than X_{t_i}. We note that since both X_t and $f(X_t)$ are Itø processes, without further model assumptions, we have nothing in the model that can answer the question of which volatility is the true underlying one.

These two volatilities $\langle X, X\rangle_T$ and $\langle f(X), f(X)\rangle_T$ are often similar quantities if $f(x) \approx x$, which makes it not so crucial to think about the above question; but this may not always be the case. Our first objective is to make it clear which volatility are the estimators estimating, and how well are the approximations.

TSRV is a typical example of estimators of volatilities. For the moment, we focus on determining the properties of TSRV. We shall make use of some of the notations from Zhang, Mykland, and Aït-Sahalia (2005):

Let $\mathcal{G} = \{t_0, \cdots, t_n\}$ be the full grid containing all the observation points. Suppose $\mathcal{G}$ is partitioned into K non-overlapping subgrids $\mathcal{G}^{(k)}, k = 1, \cdots, K$. As introduced by the authors, a typical example of selecting the subgrids is to use the regular allocation:

$$\mathcal{G}^{(k)} = \{t_{k-1}, t_{k-1+K}, t_{k-1+2K}, \cdots, t_{k-1+n_k K}\}.$$

Let the realized volatility based on all observations be

$$[Y, Y]_T^{(all)} = \sum_{t_i \in \mathcal{G}} (Y_{t_i+1} - Y_{t_i})^2;$$

let the realized volatility based on the subsampled observations $Y_t, t \in \mathcal{G}^{(k)}$ be

$$[Y,Y]_T^{(k)} = \sum_{t_j, t_{j,+} \in \mathcal{G}^{(k)}} (Y_{t_{j,+}} - Y_{t_j})^2,$$

where $t_{j,+}$ denotes the element following t_j in $\mathcal{G}^{(k)}$, when $t_j \in \mathcal{G}^{(k)}$. Let

$$n_k = |\mathcal{G}^{(k)}|, \text{ the integer making } t_{k-1+n_k K} \text{ the last element in } \mathcal{G}^{(k)};$$

and

$$\bar{n} = \frac{1}{K}\sum_{k=1}^{K} n_k = \frac{n-K+1}{K}.$$

The TSRV is given by

$$\widehat{\langle X, X\rangle}_T = [Y,Y]_T^{(avg)} - \frac{\bar{n}}{n}[Y,Y]_T^{(all)}, \tag{6}$$

where $[Y,Y]_T^{(avg)} = \frac{1}{K}\sum_{k=1}^{K}[Y,Y]_T^{(k)}$. We shall for simplicity assume that

$$\Delta t_i = T/n \tag{7}$$

(constant step size), and that

$$\text{as } n \to \infty, n/K \to \infty. \tag{8}$$

Note that our results generalize quite predictably if we allow Δt_i to vary (cf. the theory in Zhang, Mykland, and Aït-Sahalia (2005)).

2.2 Estimators of Volatilities - Estimators of $\langle f(X), f(X)\rangle_T$.

Denote

$$\epsilon_{t_i} = Y_{t_i} - f(X_{t_i});$$

note that under (2), the conditional moments of ϵ_{t_i} only depend on the value of X_{t_i}. We assume the conditional 2nd moment of ϵ_{t_i} is continuous, and there exists $\delta_0 > 0$, such that the conditional $4+2\delta_0$'th moment of ϵ_{t_i} is bounded on compact sets; that is:

$$g(x) := E(\epsilon_{t_i}^2 | X_{t_i} = x) \text{ is continuous}, \tag{9}$$

$$\forall l > 0, \exists M_{(4+2\delta_0,l)}, s.t. E(|\epsilon_{t_i}|^{4+2\delta_0} | X_{t_i} = x) \le M_{(4+2\delta_0,l)}, \text{ when } x \in [-l, l]. \tag{10}$$

Theorem 1. When we take $K = cn^{2/3}$ (the best possible order of TSRV), under the assumptions (5), (9) and (10),

$$\begin{aligned} n^{\frac{1}{6}}(\widehat{\langle X, X\rangle}_T - \langle f(X), f(X)\rangle_T) =& n^{\frac{1}{6}}(\widehat{\langle X, X\rangle}_T - \int_0^T f'(X_t)^2\sigma_t^2 dt) \\ &\to_{\mathcal{L}} (\frac{8}{Tc^2}\int_0^T g(X_t)^2 dt + c\xi^2 T)^{1/2} N(0,1), \end{aligned}$$

where

$$\xi^2 = \frac{4}{3}\int_0^T (f'(X_t)\sigma_t)^4 dt. \tag{11}$$

∎

It is clear from this result what changes and what doesn't change for this more general contamination, compared to the case of independent additive error studied in Zhang, Mykland, and Aït-Sahalia (2005).

• The volatility which is being estimated is that of $f(X_t)$. (In Zhang, Mykland, and Aït-Sahalia (2005), $f(x) = x$.)

• The rate of convergence $n^{1/6}$ is the same as for independent additive error.

• The asymptotic variance changes to reflect the more complex form of contamination.

In summary, if we are happy to estimate the volatility of $f(X_t)$, the TSRV is exceedingly robust. The point about asymptotic variance is only an issue if one wishes to set an interval around the observation. As can be seem from Zhang, Mykland, and Aït-Sahalia (2005), this is difficult even with straight additive contamination.

2.3 Proof of Theorem 1

We need to do some preparations before proving Theorem 1.

First, note that under the assumption (10), the following are true:

$$\forall \theta < 4 + 2\delta_0, E(|\epsilon_{t_i}|^\theta | X_{t_i} = x) \text{ is bounded on } [-l, l]; \text{(we write the bound as } M_{(\theta,l)}); \tag{12}$$

$$Var(\epsilon_{t_i}^2 | X_{t_i} = x) = E(\epsilon_{t_i}^4 | X_{t_i} = x) - E^2(\epsilon_{t_i}^2 | X_{t_i} = x) \text{ is bounded on } [-l, l]; \text{(say, by } M_{(Var,l)}). \tag{13}$$

And we shall use these notations in the proof:

$$M_T^{(1)} = \frac{1}{\sqrt{n}} \sum_{t_i \in \mathcal{G}} (\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | X));$$

$$M_T^{(2)} = \frac{1}{\sqrt{n}} \sum_{t_i \in \mathcal{G}} \epsilon_{t_i} \epsilon_{t_{i-1}};$$

$$M_T^{(3)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{K} \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_{i,-}},$$

where $t_{i,-}$ is the previous element in $\mathcal{G}^{(k)}$ when $t_i \in \mathcal{G}^{(k)}$. $\epsilon_{t_{i,-}} = 0$ for $t_i = \min \mathcal{G}^{(k)}$.

Proposition 1. If a sequence of random variables $\{A_n\}$ has conditional second moment $E(A_n^2|X)$ be of order $O_P(1)$, then A_n is of order $O_P(1)$.

Proof.

$E(A_n^2|X) \sim O_P(1)$ implies

$$\forall \epsilon > 0,\ \exists K,\ s.t.\ P(E(A_n^2|X) > K) < \epsilon. \text{ for big n's}$$

Define the sets D_n to be:

$$D_n = \{E(A_n^2|X) > K\}.$$

Note that I_{D_n} and $I_{D_n^c}$ are both functions of X.

$$E(A_n^2 I_{D_n^c}|X) = E(A_n^2|X) I_{D_n^c} < K \text{ pointwise .}$$

Hence

$$E(A_n^2 I_{D_n^c}) < K.$$

For $M > 0$, and for big n's,

$$\begin{aligned} P(|A_n| > M) &= P(A_n^2 I_{D_n} > M^2) + P(A_n^2 I_{D_n^c} > M^2) \\ &\leq \epsilon + \frac{E(A_n^2 I_{D_n^c})}{M^2} \\ &\leq \epsilon + \frac{K}{M^2}, \end{aligned}$$

which proves that

$$A_n \sim O_p(1).$$

Lemma 1.

$$[Y,Y]_T^{(all)} = [\epsilon,\epsilon]_T^{(all)} + O_P(1) \tag{14}$$

$$[Y,Y]_T^{(avg)} = [\epsilon,\epsilon]_T^{(avg)} + [f(X), f(X)]_T^{(avg)} + O_P(\frac{1}{\sqrt{K}}) \tag{15}$$

Proof.

$$\forall\, l, \tau_l := \inf\{t : |X_t| \geq l\}. \tag{16}$$

$$[Y,Y]_T^{(all)} = [f(X), f(X)]_T^{(all)} + [\epsilon,\epsilon]_T^{(all)} + 2[f(X),\epsilon]_T^{(all)}$$

By (12), $g(X_t) = E(\epsilon_t^2|X_t),\ t \leq T$ is bounded by $M_{(2,l)}$ on $\{\tau_l > T\}$.

$$
\begin{aligned}
&E(([f(X),\epsilon]_T^{(all)})^2 I_{\{\tau_l>T\}}|X)\\
=&I_{\{\tau_l>T\}}\sum_{i=1}^{n-1}(\Delta f(X_{t_{i-1}})-\Delta f(X_{t_i}))^2 E(\epsilon_{t_i}^2|X)+\Delta f(X_{t_{n-1}})^2 E(\epsilon_{t_n}^2|X)+\Delta f(X_{t_0})^2 E(\epsilon_{t_0}^2|X)\\
\le&I_{\{\tau_l>T\}}M_{(2,l)}[\sum_{i=1}^{n-1}(\Delta f(X_{t_{i-1}})-\Delta f(X_{t_i}))^2+\Delta f(X_{t_{n-1}})^2+\Delta f(X_{t_0})^2]\\
=&2I_{\{\tau_l>T\}}M_{(2,l)}([f(X),f(X)]_T-\sum_{i=1}^{n-1}\Delta f(X_{t_{i-1}})\Delta f(X_{t_i}))\\
\le&4I_{\{\tau_l>T\}}M_{(2,l)}[f(X),f(X)]_T\\
=&O_P(1)
\end{aligned}
$$

Hence by proposition 1,

$$[f(X),\epsilon]_T^{(all)} I_{\{\tau_l>T\}}=O_P(1)$$

Also note that, for any $\epsilon>0$, there exists L such that for any $l>L$, $P(I_{\{\tau_l\le T\}}>0)<\epsilon$. This implies

$$[f(X),\epsilon]_T^{(all)}=[f(X),\epsilon]_T^{(all)}I_{\{\tau_l>T\}}+[f(X),\epsilon]_T^{(all)}I_{\{\tau_l\le T\}}=O_P(1),$$

which proves (14).

Parallel argument shows that

$$E(([f(X),\epsilon]_T^{(avg)})^2 I_{\{\tau_l>T\}}|X)=O_p(\frac{1}{K})$$

Hence

$$[f(X),\epsilon]_T^{(avg)}=O_P(\frac{1}{\sqrt{K}}).$$

Equation (15) is proved by writing

$$[Y,Y]_T^{(avg)}=[f(X),f(X)]_T^{(avg)}+[\epsilon,\epsilon]_T^{(avg)}+2[f(X),\epsilon]_T^{(avg)}.$$

Lemma 2. $(M_T^{(2)},M_T^{(3)})$ are asymptotically independent normal, both with variance $\frac{1}{T}\int_0^T g(X_t)^2dt$.

Proof.

$(M_T^{(2)}, M_T^{(3)})$ are the end points of martingales with respect to filtration $\mathcal{F}_i = \sigma(\epsilon_{t_j}, j \leq i, X_t, \text{ all } t)$.

$$\begin{aligned}\langle M^{(2)}, M^{(2)}\rangle_T =& \frac{1}{n}\sum_{t_i\in\mathcal{G}} Var(\epsilon_{t_i}\epsilon_{t_{i-1}}|\mathcal{F}_{t_{i-1}})\\ =&\frac{1}{n}\sum_{t_i\in\mathcal{G}} \epsilon_{t_{i-1}}^2 g(X_{t_i})\\ =&\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i}) + \frac{1}{T}\sum_{t_i\in\mathcal{G}} g(X_{t_{i-1}})g(X_{t_i})\Delta t\end{aligned}$$

The first term $\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i})$ converges to 0 in probability:

$$\begin{aligned}&E((\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i})I_{\{\tau_l>T\}})^2|X)\\ =&Var(\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i})I_{\{\tau_l>T\}}|X)\\ =&\frac{1}{n^2}\sum_{t_i\in\mathcal{G}} Var(\epsilon_{t_{i-1}}^2|X)g^2(X_{t_i})I_{\{\tau_l>T\}}\\ \leq&\frac{1}{n^2}\sum_{t_i\in\mathcal{G}} M_{(Var,l)}M_{(2,l)}\\ =&O_P(1/n)\end{aligned}$$

This implies

$$\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i})I_{\{\tau_l>T\}} \to_P 0.$$

Also note that one can make $P\{\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i})I_{\{\tau_l\leq T\}} > 0\}$ arbitrarily small by choosing large l. Hence,

$$\frac{1}{n}\sum_{t_i\in\mathcal{G}}(\epsilon_{t_{i-1}}^2 - g(X_{t_{i-1}}))g(X_{t_i}) \to_P 0.$$

Therefore, by (9),

$$\langle M^{(2)}, M^{(2)}\rangle_T = \frac{1}{T}\sum_{t_i\in\mathcal{G}} g(X_{t_{i-1}})g(X_{t_i})\Delta t + o_p(1) \to_P \frac{1}{T}\int_0^T g(X_t)^2 dt.$$

Parallel argument shows

$$\langle M^{(3)}, M^{(3)}\rangle_T \to_P \frac{1}{T}\int_0^T g(X_t)^2 dt.$$

On the other hand,

$$\langle M^{(2)}, M^{(3)}\rangle_T = \frac{1}{n}\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}} Cov(\epsilon_{t_i}\epsilon_{t_{i-1}}, \epsilon_{t_i}\epsilon_{t_{i,-}}|\mathcal{F}_{t_{i-1}})$$
$$= \frac{1}{n}\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}} \epsilon_{t_{i-1}}\epsilon_{t_{i,-}}E(\epsilon_{t_i}^2|X)$$

$$\begin{aligned}
&E((\langle M^{(2)}, M^{(3)}\rangle_T)^2 I_{\{\tau_l>T\}}|X)\\
=&\frac{1}{n^2}\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}} E^2(\epsilon_{t_i}^2|X)E(\epsilon_{t_{i-1}}^2\epsilon_{t_{i,-}}^2 I_{\{\tau_l>T\}}|X)\\
\le&\frac{1}{n^2}\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}} E^2(\epsilon_{t_i}^2|X)\sqrt{E(\epsilon_{t_{i-1}}^4 I_{\{\tau_l>T\}}|X)E(\epsilon_{t_{i,-}}^4 I_{\{\tau_l>T\}}|X)}\\
\le&\frac{1}{n}M_{(2,l)}^2 M_{(4,l)}\\
=&O_P(\frac{1}{n})
\end{aligned}$$

By proposition 1, $\langle M^{(2)}, M^{(3)}\rangle_T I_{\{\tau_l>T\}} = O_P(\frac{1}{\sqrt{n}})$. Hence,

$$\langle M^{(2)}, M^{(3)}\rangle_T \to_P 0.$$

One can use the Martingale central limit Theorem (cf. page 58, Corollary 3.1, Hall and Heyde (1980)) to get the conclusion as soon as the conditional Lindeberg conditions are verified:

$$\forall\epsilon>0, \sum_{t_i\in\mathcal{G}} E[(\frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i-1}})^2 I(|\frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i-1}}|>\epsilon)|\mathcal{F}_{n,i-1}] \to_P 0$$

and

$$\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}} E[(\frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i,-}})^2 I(|\frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i-1}}|>\epsilon)|\mathcal{F}_{n,i,-}] \to_P 0$$

In fact, for $Z_{n,i} := \frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i-1}}$ and $Z'_{n,i} := \frac{1}{\sqrt{n}}\epsilon_{t_i}\epsilon_{t_{i,-}}$, a stronger condition, the (conditional) Lyapounov's condition (cf. page 362 Billingsley (1995)) is satisfied:

$$plim\sum_{i=1}^{n} E[|Z_{n,i}|^{2+\delta}|\mathcal{F}_{n,i-1}] = plim\sum_{i=1}^{n} E[|Z'_{n,i}|^{2+\delta}|\mathcal{F}_{n,i-1}] = 0, \text{ for some positive } \delta.$$

The conditional Lyapounov's condition is satisfied for $Z_{n,i}$ because, by assumption (10) and proposition 1, for any $l>0$,

$$\frac{1}{n}\sum_{i=1}^{n}|\epsilon_{t_{i-1}}|^{2+\delta_0}E(\epsilon_{t_i}^{2+\delta_0}|\mathcal{F}_{n,i-1})I_{\{\tau_l>T\}} = O_p(1),$$

which implies,

$$\frac{1}{n}\sum_{i=1}^{n}|\epsilon_{t_{i-1}}|^{2+\delta_0}E(\epsilon_{t_i}^{2+\delta_0}|\mathcal{F}_{n,i-1}) = O_p(1).$$

Hence

$$plim\sum_{i=1}^{n}E[|Z_{n,i}|^{2+\delta_0}|\mathcal{F}_{n,i-1}] = plim\frac{1}{n^{\frac{\delta_0}{2}}}(\frac{1}{n}\sum_{i=1}^{n}|\epsilon_{t_{i-1}}|^{2+\delta_0}E(\epsilon_{t_i}^{2+\delta_0}|\mathcal{F}_{n,i-1})) = 0.$$

Parallel argument shows $plim\sum_{i=1}^{n}E[|Z'_{n,i}|^{2+\delta_0}|\mathcal{F}_{n,i-1}] = 0$, which finishes the proof.

Proof of Theorem 1.

$$\begin{aligned}
\widehat{\langle X, X\rangle}_T &= [Y,Y]_T^{(avg)} - \frac{\bar{n}}{n}[Y,Y]_T^{(all)} \\
\text{(by Lemma 1)} &= [f(X),f(X)]_T^{(avg)} + [\epsilon,\epsilon]_T^{(avg)} + O_p(\frac{1}{\sqrt{K}}) - \frac{\bar{n}}{n}[\epsilon,\epsilon]_T^{(all)} - O_p(\frac{\bar{n}}{n}) \\
&= [f(X),f(X)]_T^{(avg)} + [\epsilon,\epsilon]_T^{(avg)} - \frac{\bar{n}}{n}[\epsilon,\epsilon]_T^{(all)} + O_p(\frac{1}{\sqrt{K}}) \qquad (17)
\end{aligned}$$

$$\begin{aligned}
&[\epsilon,\epsilon]_T^{(all)} \\
=&2\sum_{t_i\in\mathcal{G}}(\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2|X)) - 2\sum_{t_i>0}\epsilon_{t_i}\epsilon_{t_{i-1}} - (\epsilon_{t_0}^2 - E(\epsilon_{t_0}^2|X)) - (\epsilon_{t_n}^2 - E(\epsilon_{t_n}^2|X)) + 2\sum_{t_i\in\mathcal{G}}E(\epsilon_{t_i}^2|X) \\
=&2\sqrt{n}(M^{(1)} - M^{(2)}) + 2\sum_{t_i\in\mathcal{G}}E(\epsilon_{t_i}^2|X) + O_P(1)
\end{aligned}$$

Similarly, one can prove

$$K[\epsilon,\epsilon]_T^{(avg)} = 2\sqrt{n}(M^{(1)} - M^{(3)}) + O_P(K^{1/2}) + 2\sum_{k=1}^{K}\sum_{t_i\in\mathcal{G}^{(k)}}E(\epsilon_{t_i}^2|X)$$

Therefore,

$$\begin{aligned}
&\frac{K}{\sqrt{n}}([\epsilon,\epsilon]_T^{(avg)} - \frac{\bar{n}}{n}[\epsilon,\epsilon]_T^{(all)}) \\
\approx&\frac{1}{\sqrt{n}}(K[\epsilon,\epsilon]_T^{(avg)} - [\epsilon,\epsilon]_T^{(all)}) \\
=&(2(M^{(2)} - M^{(3)})) + O_P(\sqrt{\frac{K}{n}}) \\
\to_{\mathcal{L}}&\ N(0, \frac{8}{T}\int_0^T g(X_t)^2 dt). \qquad (18)
\end{aligned}$$

Note that $\frac{K}{\sqrt{n}} \sim \frac{K}{\sqrt{K\bar{n}}} \sim \sqrt{\frac{K}{\bar{n}}}$, one can re-write (17) and make use of (18) and get

$$\sqrt{\frac{K}{\bar{n}}}(\widehat{\langle X,X\rangle}_T - [f(X),f(X)]_T^{(avg)}) \to_{\mathcal{L}} \sqrt{\frac{8}{T}\int_0^T g(X_t)^2 dt} \cdot Z_{noise}, \tag{19}$$

where $Z_{noise} \sim N(0,1)$.

$f(X_t)$ is a semimartingale,

$$df(X_t) = (f'(X_t)\mu_t + \frac{1}{2}f''(X_t)\sigma_t^2)dt + f'(X_t)\sigma_t dB_t.$$

By Zhang, Mykland, and Aït-Sahalia (2005),

$$\sqrt{\frac{n}{K}}([f(X),f(X)]_T^{(avg)} - \langle f(X),f(X)\rangle_T) \to_{\mathcal{L}} \xi\sqrt{T} \cdot Z_{discrete}, \tag{20}$$

where ξ is defined as in (11), and $Z_{discrete} \sim N(0,1)$, independent of Z_{noise}.

By combining (19), (20), one has

$$\begin{aligned}
&\widehat{\langle X,X\rangle}_T - \langle f(X),f(X)\rangle_T \\
=&(\widehat{\langle X,X\rangle}_T - [f(X),f(X)]_T^{(avg)}) + ([f(X),f(X)]_T^{(avg)} - \langle f(X),f(X)\rangle_T) \\
=&O_p(\frac{\bar{n}^{1/2}}{K^{1/2}}) + O_p(\bar{n}^{-1/2}).
\end{aligned}$$

The error is minimized when $K = O(n^{2/3})$.

If we take $K = cn^{2/3}$, we have

$$\begin{aligned}
n^{\frac{1}{6}}(\widehat{\langle X,X\rangle}_T - \langle f(X),f(X)\rangle_T) &\to_{\mathcal{L}} N(0, \frac{8}{Tc^2}\int_0^T g(X_t)^2 dt) + \xi\sqrt{T}N(0,c) \\
&= (\frac{8}{Tc^2}\int_0^T g(X_t)^2 dt + c\xi^2 T)^{1/2} N(0,1).
\end{aligned}$$

3 A CASE STUDY

3.1 Another Form of Contamination

Two types of errors – additive errors and rounding errors has been proposed to be candidates of market micro structure errors. Results of when one of them plays the role are available. Now, we are going to consider the case when both types of errors are present.

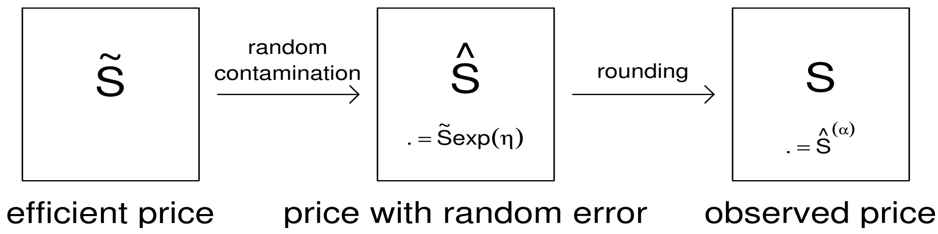

Figure 1: Two stage contamination: random error followed by rounding

Suppose at the transaction times, the latent return process X_{t_i} is contaminated by an independent error process η_{t_i}, and then rounded to reflect that prices are quotes on a grid (typically in multiples of one cent). We thus envisage a two stage procedure where a latent efficient price $\tilde{S} = \exp(X)$ is first subjected to multiplicative random error: $\hat{S} = \tilde{S}\exp(\eta)$. The actual price S is then the rounded value of $\hat{S}$. If we take, as usual, $Y = \log S$, our final product is the observed process Y_{t_i} of rounded contaminated prices:

$$Y_{t_i} = \log((\exp(X_{t_i} + \eta_{t_i}))^{(\alpha)}),$$

where $s^{(\alpha)} = \alpha[s/\alpha]$ is the value of s rounded to the nearest multiple of α. The model is somewhat similar to that used by Large (2005). It can be illustrated as in Figure 1.

For practical purposes, we further assume that the smallest observation of the security price is α, which makes the observations of the log prices have the form

$$Y_{t_i} = \log\alpha \vee \log((\exp(X_{t_i} + \eta_{t_i}))^{(\alpha)}). \tag{21}$$

We consider the case when the random errors are *i.i.d.* normally distributed, with mean 0 and positive variance; that is

$$\eta_{t_i} \sim_{i.i.d.} N(0, \gamma^2), \quad \gamma > 0. \tag{22}$$

In this case,

$$f(x) = E(Y_{t_i}|X_{t_i} = x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\gamma}(\log\alpha \vee \log((e^z)^{(\alpha)}))e^{-\frac{(z-x)^2}{2\gamma^2}}dz \tag{23}$$

is a twice continuous differentiable function. And the assumptions (9) and (10) hold. Therefore, by theorem 1, the TSRV is a robust estimator of $\langle f(X), f(X)\rangle_T$.

3.2 Robustness Works: When γ is Big:

When the latent log price process X_t follows (1), and when $\int_0^T \sigma_t^2 dt$ is small compared with $|\log(\alpha)|$ (so that it guarantees that the latent price $\tilde{S} = \exp(X_t)$ almost doesn't reach α in the time interval we are concerning), one has, under model (21) and assumption (22):

$$f(X_t) \approx X_t, \text{ and } f'(X_t) \approx 1 \text{ for } t \in [0,T], \quad \text{for suitably big } \gamma\text{'s}$$

By "suitably big γ's", we mean the size of the random error is big enough so that the possibility that it pull the observations up or down several grid points is not negligible; and of course it can't be incredibly large compared with the size of the log prices. By the normal assumption, in this case, the expected log transactions Y_{t_i} is pulled back to X_{t_i}.

In this case, $\langle f(X), f(X)\rangle_T \approx \langle X, X\rangle_T$. Therefore, the TSRV, which is a robust estimator of $\langle f(X), f(X)\rangle_T$, is a good estimator of $\langle X, X\rangle_T$ as well.

These relationships are illustrated in section 3.4.

3.3 How Things can Go Wrong: When $\gamma \to 0$:

When γ is small but not 0, by Theorem 1, we know that the TSRV goes to the limit $\langle f(X), f(X)\rangle_T$ robustly. But this volatility $\langle f(X), f(X)\rangle_T$ is no longer close to $\langle X, X\rangle_T$.

To study the limiting behavior of $\langle f(X), f(X)\rangle_T$, we relate it to the local time L_T^a of the semimartingale X_t. For any real number a, the local time of a continuous semimartingale defined in Revuz and Yor (1999) (cf. page 222) is, an increasing continuous process L^a such that,

$$\begin{aligned}
|X_t - a| &= |X_0 - a| + \int_0^t sgn(X_s - a)dX_s + L_t^a, \\
(X_t - a)^+ &= (X_0 - a)^+ + \int_0^1 1_{\{X_s > a\}} dX_s + \frac{1}{2}L_t^a, \\
(X_t - a)^- &= (X_0 - a)^- - \int_0^1 1_{\{X_s \le a\}} dX_s + \frac{1}{2}L_t^a.
\end{aligned}$$

By Revuz and Yor (1999) page 224 Corollary 1.6, for any positive Borel function Φ,

$$\int_0^t \Phi(X_s) d\langle X, X\rangle_s =_{w.p.1} \int_{-\infty}^{\infty} \Phi(a) L_t^a da. \tag{24}$$

And by Corollary 1.8 page 226 in the same book, one sees that the family L^a may be chosen such that

$$w.p.1, \ \ a \to L_t^a \text{ is Hölder continuous of order } \beta, \forall \beta < 1/2, \tag{25}$$

and uniformly in t on every compact interval. We shall consider only the version of the local time L_t^a which satisfies these conditions.

Relating $\langle f(X), f(X)\rangle_T$ to L_T^a by (24), one has the following result:

Theorem 2. As $\gamma \to 0$, almost surely,

$$\gamma \langle f(X), f(X)\rangle_T \to \frac{1}{2\sqrt{\pi}} \sum_{k=1}^{\infty} L_T^{\log((k+\frac{1}{2})\alpha)} (\log \frac{k+1}{k})^2,$$

where L_t^a is the local time of the continuous semimartingale X, defined as above. ∎

In other words, the "target" $\langle f(X), f(X)\rangle_T$ which we are estimating blows up as γ goes to zero, and is of order $1/\gamma$. This raises questions of whether $\langle f(X), f(X)\rangle_T$ is, in this case, the quantity that we are really after.

3.4 Illustration

We take a typical sample path to illustrate the situation: suppose the latent log price X_t follows (1) with $\mu_t = 0$ and $\sigma_t = 0.2$, $\forall t \in [0, \infty)$ (one can think of them as simplified annualized parameters). Suppose at the observation times, the log price process is contaminated by a random error distributed as $N(0, \gamma^2)$, and then rounded to the nearest multiple of $\alpha = 0.01$ (one cent). We estimate $\langle X, X\rangle_T$ over $T = 1/252$ (i.e., one day). We assume that a day consists of 6.5 hours of open trading, and the price process is observed once every second (n=23400).

A sample path of the latent log price process X_t, $t \in [0, T]$ is plotted in Figure 2, together with three expected observed log price processes, conditional on the latent price process X_t, under different modeling assumptions. The solid line is under the assumption that there is no random contamination, or $\gamma = 0$. The "$*$"'s and "o"'s are the conditional expectations of the observed log prices, under the assumptions that the i.i.d. normal contaminations at the observation times have standard error $\gamma = 0.001$ and $\gamma = 0.005$ respectively.

Figure 3 records the TSRV of this particular sample path X_t, $t \in [0, T]$ in Figure 2, under random contamination of different sizes (with standard error γ ranges from 0.0002 to 0.006). The solid line is the volatility $\langle X, X\rangle_T$.

One sees from Figure 2 that for this process, when γ is as big as 0.005, the $f(X)$ is close to X. While when γ is smaller, the expected process diverges from X_t; in fact, it goes closer to the (discontinuous) pure rounded process. Figure 3 shows that when γ is suitably big, the TSRV can be a good estimator of $\langle X, X\rangle_T$, but when γ is too small, the estimator is not estimating $\langle X, X\rangle_T$, but rather a much larger quantity.

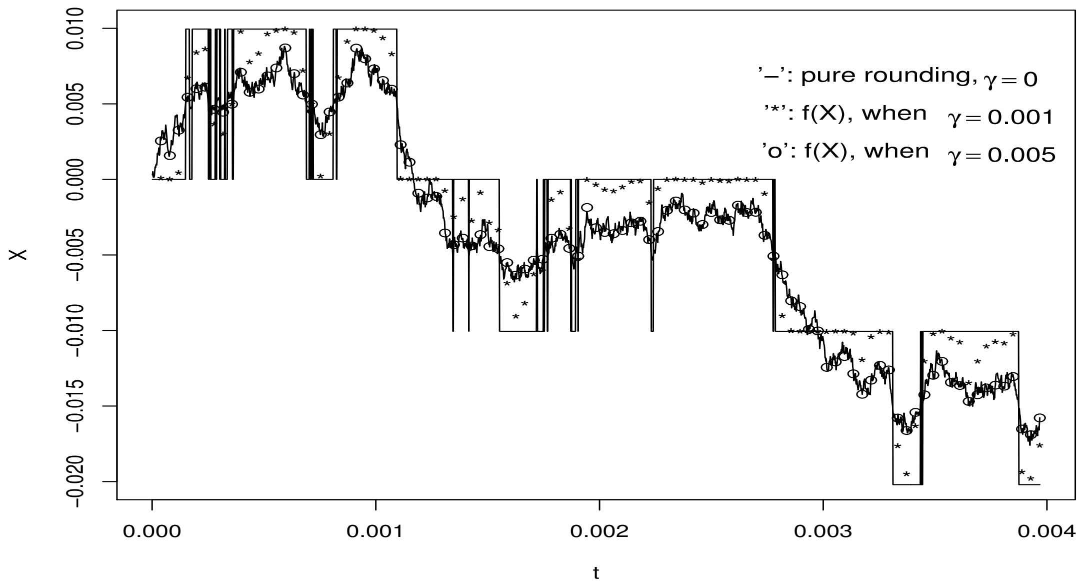

Figure 2: Relationship between $f(X)$ and X on one (random) sample path.

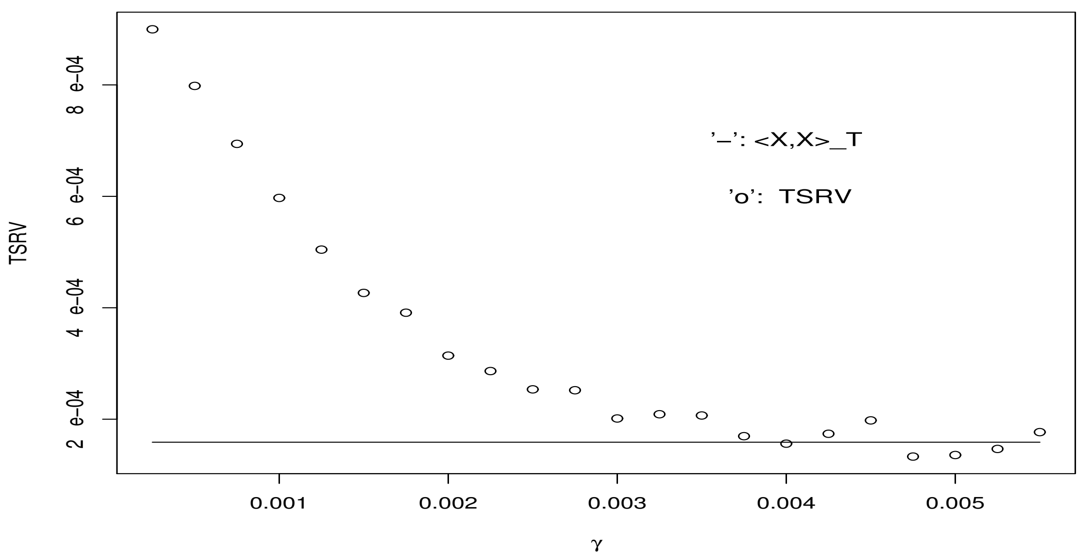

Figure 3: TSRV *vs* size of the random contamination, based on one (random) latent log price process ($\langle X, X\rangle_T \approx 1.59 * 10^{-4}$).

3.5 How Error and Sample Size Relate to Each Other – Comparison to Case When $\gamma = 0$:

When $\gamma = 0$, the additive error is gone, only the rounding error is present. In this case, the observations are themselves the conditional expectations, and $f(x)$ is no longer continuous.

$$Y_{t_i} = f(X_{t_i}) = E(Y_{t_i}|X_{t_i}) = \log\alpha \vee \log((exp(X_{t_i}))^{(\alpha)}). \tag{26}$$

A modification of Jacod (1994)'s proof gives the following result:

Theorem 3. When $X = \sigma W$, where W is a standard Brownian Motion, one has

$$plim_{n\to\infty}\frac{1}{\sqrt{\bar{n}}}\widehat{\langle X, X\rangle}_T = \frac{1}{\sigma\sqrt{T}}\sqrt{\frac{2}{\pi}}\sum_{k=1}^{\infty} L_T^{\log(k+\frac{1}{2}\alpha)}(\log(1+\frac{1}{k}))^2,$$

where $\widehat{\langle X, X\rangle}_T$ is the TSRV and $\bar{n}$ is the optimal average number of elements in the subgrids. ∎

We can see from Theorems 2 and 3 that to first order,

$$\begin{aligned}&\text{TSRV under pure rounding and no contamination}\\ &= \sqrt{\frac{8\bar{n}\gamma^2}{\sigma^2 T}} \times \text{TSRV under rounding after contamination of size } \gamma\end{aligned} \tag{27}$$

Thus, in a sense, contamination plays a rôle slightly similar to sample size when there is no contamination:

$$\gamma^{-2} \text{ under random contamination } \approx \bar{n}\frac{8}{\sigma^2 T} \text{ under no random contamination.} \tag{28}$$

In both cases, the size of of γ^{-2} and $\bar{n}$ have similar functions in quantifying the ill-posedness of the respective estimation problems. The deeper meaning of this remains, for the moment, a little mysterious even to the authors.

3.6 Proofs of Theorem 2 and Theorem 3

Proof of Theorem 2.

One has, by (24),

$$\langle f(X), f(X)\rangle_T = \int_0^T (f'(X_t))^2 d\langle X, X\rangle_t =^{w.p.1} \int_{-\infty}^{\infty} L_T^a \mu(da),$$

where $\mu(da) = (f'(a))^2 da$.

Recall that

$$f(x)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\gamma}(\log\alpha\vee\log((e^z)^{(\alpha)}))e^{-\frac{(z-x)^2}{2\gamma^2}}dz,$$

hence,

$$\begin{aligned}f'(x)&=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\gamma}(\log\alpha\vee\log((e^z)^{(\alpha)}))\frac{z-x}{\gamma^2}e^{-\frac{(z-x)^2}{2\gamma^2}}dz\\&=\int_{-\infty}^{\infty}\frac{1}{\gamma}\log\alpha\vee\log(\exp(x+\gamma v)^{(\alpha)})\frac{v}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}dv.\end{aligned}$$

For $k=1,2,3,\cdots$, one has, $\forall y\in\mathbb{R}$,

$$\begin{aligned}&\gamma f'(\log((k+\frac{1}{2})\alpha)+y\gamma)\\=&\int_{-\infty}^{\infty}\log\alpha\vee\log(\exp(\log((k+\frac{1}{2})\alpha)+y\gamma+v\gamma)^{(\alpha)})\frac{v}{\sqrt{2\pi}}e^{-\frac{v^2}{2}}dv\\=&E_V(\log\alpha\vee\log(((k+\frac{1}{2})\alpha\cdot e^{\gamma(y+V)})^{(\alpha)})\cdot V),\quad V\sim N(0,1).\end{aligned}$$

By the Dominated Convergence Theorem,

$$\begin{aligned}&\lim_{\gamma\to 0}\gamma f'(\log((k+\frac{1}{2})\alpha)+y\gamma)\\=&E_V(\log((k+1)\alpha)\cdot VI\{y+V>0\})+E_V(\log(k\alpha)\cdot VI\{y+V<0\})\\=&\log((k+1)\alpha)\frac{1}{\sqrt{2\pi}}\int_{-y}^{\infty}ze^{-\frac{z^2}{2}}dz+\log(k\alpha)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-y}ze^{-\frac{z^2}{2}}dz\\=&\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}(\log((k+1)\alpha)-\log(k\alpha))\\=&\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\log(\frac{k+1}{k}).\end{aligned}\tag{29}$$

A similar argument shows that

$$\forall y\in\mathbb{R},\quad\gamma f'(\log\alpha+y\gamma)\to 0\text{ as }\gamma\to 0.\tag{30}$$

By (29), for $k = 1, 2, 3 \cdots$,

$$\begin{aligned}
&\lim_{\gamma\to 0}\int_{\log((k+\frac{1}{2})\alpha)-n\gamma}^{\log((k+\frac{1}{2})\alpha)+n\gamma}\gamma(f'(x))^2dx\\
=&\lim_{\gamma\to 0}\int_{-n}^{n}(\gamma f'(\log((k+\frac{1}{2})\alpha)+y\gamma))^2dy\\
=&\int_{-n}^{n}(\frac{1}{\sqrt{2\pi}}(\log((\frac{k+1}{k}))^2e^{-y^2}dy\\
=&\frac{1}{2\sqrt{\pi}}(\log\frac{k+1}{k})^2\frac{1}{\sqrt{2\pi}}\int_{-\sqrt{2}n}^{\sqrt{2}n}e^{-\frac{z^2}{2}}dz.
\end{aligned}$$

For any $a \in [\log\alpha, \infty)$, $b \in (a, \infty]$, suppose the set $\{k : \log((k+\frac{1}{2})\alpha) \in (a, b)\}$ is not empty. Denote $k_a^0 = \lceil \frac{e^a}{\alpha} - \frac{1}{2}\rceil$, the smallest integer k such that $\log((k+\frac{1}{2})\alpha) > a$; $k_b^1 = \lfloor \frac{e^b}{\alpha} - \frac{1}{2}\rfloor$, the biggest integer k such that $\log((k+\frac{1}{2})\alpha) < b$.

Note that for any $N \le \min(\frac{\log((k_a^0+\frac{3}{2})\alpha)-\log((k_a^0+\frac{1}{2})\alpha)}{2\gamma}, \frac{\log((k_a^0+\frac{1}{2})\alpha)-a}{\gamma}, \frac{b-\log((k_b^1+\frac{1}{2})\alpha)}{\gamma})$,

$$\begin{aligned}
\gamma\mu(a,b) \ge& \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\gamma\mu\big(\log((k+\frac{1}{2})\alpha)-N\gamma, \log((k+\frac{1}{2})\alpha)+N\gamma\big)\\
=& \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\int_{\log((k+\frac{1}{2})\alpha)-N\gamma}^{\log((k+\frac{1}{2})\alpha)+N\gamma}\gamma(f'(x))^2dx.
\end{aligned}$$

Hence,

$$\begin{aligned}
\liminf_{\gamma\to 0}\gamma\mu(a,b) \ge& \lim_{\gamma\to 0}\sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\int_{\log((k+\frac{1}{2})\alpha)-N\gamma}^{\log((k+\frac{1}{2})\alpha)+N\gamma}\gamma(f'(x))^2dx\\
=& \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\frac{1}{2\sqrt{\pi}}(\log\frac{k+1}{k})^2(\Phi(\sqrt{2}N)-\Phi(-\sqrt{2}N)),
\end{aligned}$$

where Φ is the distribution function of a standard normal distribution.

As $\gamma \to 0$, N can be chosen to be arbitrarily large, therefore,

$$\begin{aligned}
\liminf_{\gamma\to 0}\gamma\mu(a,b) \ge& \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\frac{1}{2\sqrt{\pi}}(\log\frac{k+1}{k})^2(\Phi(\infty)-\Phi(-\infty))\\
=& \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)}\frac{1}{2\sqrt{\pi}}(\log\frac{k+1}{k})^2. \qquad (31)
\end{aligned}$$

On the other hand, for any $\gamma > 0$, there exists M, such that

$$\begin{aligned}\gamma\mu(a,b) &\leq \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \gamma\mu\big(\log((k+\frac{1}{2})\alpha) - M\gamma, \log((k+\frac{1}{2})\alpha) + M\gamma\big) \\ &= \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \int_{\log((k+\frac{1}{2})\alpha)-M\gamma}^{\log((k+\frac{1}{2})\alpha)+M\gamma} \gamma(f'(x))^2 dx \\ &= \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \int_{-M}^{M} (\gamma f'(\log((k+\frac{1}{2})\alpha) + y\gamma))^2 dy \\ &\leq \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \int_{-\infty}^{\infty} (\gamma f'(\log((k+\frac{1}{2})\alpha) + y\gamma))^2 dy,\end{aligned}$$

which implies,

$$\begin{aligned}\limsup_{\gamma\to 0} \gamma\mu(a,b) &\leq \lim_{\gamma\to 0} \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \int_{-\infty}^{\infty} (\gamma f'(\log((k+\frac{1}{2})\alpha) + y\gamma))^2 dy \\ &= \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \frac{1}{2\sqrt{\pi}} (\log\frac{k+1}{k})^2.\end{aligned} \tag{32}$$

By (31) and (32),

$$\lim_{\gamma\to 0} \gamma\mu(a,b) = \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \frac{1}{2\sqrt{\pi}} (\log\frac{k+1}{k})^2,$$

for any $(a,b) \subset (\log\alpha, \infty)$ $s.t.\{k : \log((k+\frac{1}{2})\alpha) \in (a,b)\} \neq \emptyset$.

By linearity of measures, this result implies that

$$\lim_{\gamma\to 0} \gamma\mu(a,b) = 0 \text{ for } (a,b) \subset (\log\alpha, \infty) \text{ and } \{k : \log((k+\frac{1}{2})\alpha) \in (a,b)\} = \emptyset.$$

By (30) and a similar argument as above, one has,

$$\lim_{\gamma\to 0} \gamma\mu(a,b) = 0 \text{ for any } (a,b) \subset (-\infty, \log(\frac{3}{2}\alpha)).$$

In summary,

$$\lim_{\gamma\to 0} \gamma\mu(a,b) = \sum_{k:\log((k+\frac{1}{2})\alpha)\in(a,b)} \frac{1}{2\sqrt{\pi}} (\log\frac{k+1}{k})^2, \text{ for any } (a,b) \subset \mathbb{R}.$$

Therefore, as $\gamma \to 0$, $\gamma\mu$, as a measure on $\mathbb{R}$, converges to a finite measure ν, which has point mass $\frac{1}{2\sqrt{\pi}}(\log\frac{k+1}{k})^2$ on $\log((k+\frac{1}{2})\alpha)$, $k = 1, 2, 3\cdots$.

On the other hand, by (24),

$$\int_{-\infty}^{\infty} L_T^a da =_{w.p.1} \int_0^T d\langle XX\rangle_t < \infty. \tag{33}$$

The integrability (33), together with the Hölder continuity (25) of L_T^a gives,

$$w.p.1,\ L_T^a(X) \text{ is bounded.} \tag{34}$$

As a consequence, almost surely,

$$\lim_{\gamma\to 0} \gamma\langle f(X), f(X)\rangle_T = \lim_{\gamma\to 0} \gamma \int_{-\infty}^{\infty} L_T^a \mu(da) = \frac{1}{2\sqrt{\pi}} \sum_{k=1}^{\infty} L_T^{\log((k+\frac{1}{2})\alpha)} (\log \frac{k+1}{k})^2.$$

Proof of Theorem 3.:

We borrow the notations from Jacod (1994):

$$\begin{aligned}
\text{For } t_i &= \frac{iT}{n}, i = 0, 1, \cdots n,\\
\xi_i^n &= X_{t_i} - X_{t_{i-1}};\\
\chi_i^n &= (f(X_{t_i}) - f(X_{t_{i-1}}))^2, \text{ where} f(x) \text{ is defined in (26)};\\
R_n^k &= \{(x,y) : [\frac{e^x}{\alpha}] = k, [\frac{e^y}{\alpha}] = k+1; \text{ or } [\frac{e^x}{\alpha}] = k+1, [\frac{e^y}{\alpha}] = k.\},\ k = 1, 2, 3, \cdots;\\
R_n &= \cup_{k=1}^{\infty} R_n^k;\\
S_n &= \mathbb{R}^2 \setminus R_n\\
T(a) &= \{(x,y) : x < a \le y \text{ or } y < a \le x\};\\
T_n &= \cup_{k=1}^{\infty} T(\log((k+\frac{1}{2})\alpha));\\
\hat{W}_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} 1_{R_n}(X_{t_{i-1}}, X_{t_i});\\
W_n &= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \chi_i^n 1_{S_n}(X_{t_{i-1}}, X_{t_i}).
\end{aligned}$$

Recall τ_l defined in (16): $\forall\, l, \tau_l = \inf\{t : |X_t| \ge l\}$. Define $k_l = \lfloor \frac{e^l}{\alpha} \rfloor$, the biggest integer k such that $\log((k+\frac{1}{2})\alpha) < l$.

Note that if $(X_{t_{i-1}}, X_{t_i}) \in R_n^k$, then $\chi_i^n = (\log \frac{k+1}{k})^2$, one has

$$\frac{1}{\sqrt{n}}[Y,Y]^{(all)}1_{\{\tau_l>T\}} = (\hat{W}_n \sum_{k=1}^{\infty} 1_{R_n^k}(X_{t_{i-1}}, X_{t_i})\chi_i^n + W_n)1_{\{\tau_l>T\}}$$
$$= (\sum_{k=1}^{k_l}(\hat{W}_n 1_{R_n^k}(X_{t_{i-1}}, X_{t_i})) \cdot (\log\frac{k+1}{k})^2 + W_n)1_{\{\tau_l>T\}}. \tag{35}$$

We have that

$$W_n \to_P 0. \tag{36}$$

and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} I_{\{|\exp(X_{t_i})-\exp(X_{t_{i-1}})|\geq\alpha\}} \to_P 0. \tag{37}$$

In fact, if $(x,y) \in S_n$, then

$$\begin{aligned}&\text{either} \quad f(x) = f(y),\\ &\text{or} \quad |\alpha \vee (\exp(x))^{(\alpha)} - \alpha \vee (\exp(y))^{(\alpha)}| > \alpha.\end{aligned} \tag{38}$$

In the case of (38), without lost of generality, one can assume $\exp(x) > \exp(y) \geq \alpha$, which implies,

$$\begin{aligned}|f(x) - f(y)| &= \log((\exp(x))^{(\alpha)}) - \log((\exp(y))^{(\alpha)})\\ &= \log\frac{(\exp(x))^{(\alpha)}}{(\exp(y))^{(\alpha)}}\\ &\leq \log\frac{\exp(x)+\alpha/2}{\exp(y)-\alpha/2}\\ &\leq \log\frac{\exp(x)+\exp(x)/2}{\exp(y)-\exp(y)/2}\\ &= \log 3 + |x-y|.\end{aligned} \tag{39}$$

If in addition we know that both $\exp(x)$ and $\exp(y)$ are bounded by $M > 0$, then, there exists ξ, $(x \vee y) \leq \xi \leq (x \wedge y)$, such that $\alpha \leq |\exp(x) - \exp(y)| = \exp(\xi)|x-y| \leq M|x-y|$. This implies that $|x-y| \geq \frac{\alpha}{M}$, hence, $\log 3 \leq \frac{\alpha \log 3}{M}|x-y|$. By (39),

$$|f(x) - f(y)| \leq (\frac{\alpha \log 3}{M} + 1)|x-y|. \tag{40}$$

And it is easy to see that, in fact, (40) holds for any $(x,y) \in S_n$, such that $\exp(x), \exp(y)$ are bounded by M.

Therefore,

$$(W_n I_{\{\tau_{logM}<T\}})^2 \leq n\sum_{i=1}^{n}(\frac{1}{\sqrt{n}}\chi_i^n 1_{S_n}(X_{t_{i-1}},X_{t_i})I_{\{\tau_{logM}<T\}})^2 \leq n\sum_{i=1}^{n}\frac{1}{n}(\frac{\alpha\log 3}{M}+1)^4(\xi_i^n)^4. \quad (41)$$

Note that $E((\xi_i^n)^4)=3\sigma^4T^2/n^2$, one has, $W_n I_{\{\tau_{logM}<T\}} \to 0$ in $\mathbb{L}^2$. One can make $P(I_{\{\tau_{logM}<T\}}>0)$ arbitrarily small by letting M be large. Hence, (36) is true.

The following inequality proves (37):

$$\begin{aligned}
&E(I_{\{|\exp(X_{t_i})-\exp(X_{t_{i-1}})|\geq\alpha\}})\\
=&P(|\exp(X_{\frac{iT}{n}})-\exp(X_{\frac{(i-1)T}{n}})|\geq\alpha)\\
\leq&E(\exp(X_{\frac{iT}{n}})-\exp(X_{\frac{(i-1)T}{n}}))^2/\alpha^2\\
=&(E\exp(2\sigma W_{\frac{iT}{n}})+E\exp(2\sigma W_{\frac{(i-1)T}{n}})-2E\exp(2\sigma W_{\frac{(i-1)T}{n}})E\exp(\sigma W_{\frac{T}{n}}))/\alpha^2\\
=&\frac{1}{\alpha^2}(\exp(\frac{4\sigma^2(iT)}{2n})+\exp(\frac{4\sigma^2(i-1)T}{2n})-2\exp(\frac{4\sigma^2(i-1)T}{2n}+\frac{\sigma^2T}{2n}))\\
=&\frac{1}{\alpha^2}\exp(\frac{2\sigma^2(i-1)T}{n})(\exp(\frac{2\sigma^2T}{n})+1-2\exp(\frac{\sigma^2T}{2n}))\\
=&O(\frac{1}{n}).
\end{aligned}$$

One has $R_n\subset T_n$ and $|\exp(x)-\exp(y)|\geq\alpha$ when $(x,y)\in T_n\setminus R_n$. By (35), (36) and (37), one sees that, for $\hat{W}'_n=\frac{1}{\sqrt{n}}\sum_{i=1}^n 1_{T_n}(X_{t_{i-1}},X_{t_i})$,

$$\begin{aligned}
plim_{n\to\infty}\frac{1}{\sqrt{n}}[Y,Y]^{(all)}1_{\{\tau_l>T\}} &= plim_{n\to\infty}\sum_{k=1}^{k_l}\hat{W}'_n 1_{T(\log((k+\frac{1}{2})\alpha))}(X_{t_{i-1}},X_{t_i})(\log\frac{k+1}{k})^2 1_{\{\tau_l>T\}}\\
&= plim_{n\to\infty}\sum_{k=1}^{k_l}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}1_{T(\log((k+\frac{1}{2})\alpha))}(X_{t_{i-1}},X_{t_i})(\log\frac{k+1}{k})^2 1_{\{\tau_l>T\}}\\
&= \sum_{k=1}^{k_l}[(\log\frac{k+1}{k})^2 plim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}1_{T(\log((k+\frac{1}{2})\alpha))}(X_{t_{i-1}},X_{t_i})]1_{\{\tau_l>T\}}.
\end{aligned}$$

For $k=1,2,3,\cdots$, by Jacod (1994),

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}1_{T(\log((k+\frac{1}{2})\alpha))}(X_{t_{i-1}},X_{t_i}) \to_{\mathbb{L}^2} \frac{1}{\sigma\sqrt{T}}\sqrt{\frac{2}{\pi}}L_T^{\log((k+\frac{1}{2})\alpha)}.$$

Hence, by letting l go to ∞, one has,

$$plim_{n\to\infty}\frac{1}{\sqrt{n}}[Y,Y]^{(all)}=\frac{1}{\sigma\sqrt{T}}\sqrt{\frac{2}{\pi}}\sum_{k=1}^{\infty}L_T^{\log((k+\frac{1}{2})\alpha)}(\log\frac{k+1}{k})^2,$$

and

$$plim_{\bar{n}\to\infty}\frac{1}{\sqrt{\bar{n}}}[Y,Y]^{(avg)} = \frac{1}{\sigma\sqrt{T}}\sqrt{\frac{2}{\pi}}\sum_{k=1}^{\infty} L_T^{\log((k+\frac{1}{2})\alpha)}(\log\frac{k+1}{k})^2.$$

Applying this result to the TSRV (6), we have, when we take (the optimal order) $K \sim O(n^{2/3})$, or, equivalently, $\bar{n} \sim O(n^{1/3})$,

$$\frac{1}{\sqrt{\bar{n}}}\frac{\bar{n}}{n}[Y,Y]^{(all)} \to_P 0;$$

and hence,

$$\frac{1}{\sqrt{\bar{n}}}\widehat{\langle X, X\rangle}_T = \frac{1}{\sqrt{\bar{n}}}([Y,Y]_T^{(avg)} - \frac{\bar{n}}{n}[Y,Y]_T^{(all)})$$
$$\to_P \frac{1}{\sigma\sqrt{T}}\sqrt{\frac{2}{\pi}}\sum_{k=1}^{\infty} L_T^{\log((k+\frac{1}{2})\alpha)}(\log\frac{k+1}{k})^2.$$

4 CONCLUSION

We have shown in this paper that the robustness of the two scales realized volatility (TSRV) depends crucially on the deterministic part of the distortion through the function f defined in (4). On the other hand, in terms of consistency and order of convergence, the TSRV is always robust to the random part of the error $(Y - f(X))$. In Section 3, we have studied a particular model of contamination, involving random error followed by rounding, and we have seen that in this case, depending on parameters, the nonrandom distortion can be benign or problematic.

A lesson from our study is that there are really two candidates for the term "volatility", namely $\langle X, X\rangle_T$, and $\langle f(X), f(X)\rangle_T$, and that in some cases these can diverge substantially. To further investigate what quantity one wishes to estimate, there is need for more research into the use of realized volatility estimates in such applications as portfolio management, options trading and forecasting.

REFERENCES

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005), "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, 18, 351–416.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001), "The Distribution of Exchange Rate Realized Volatility," *Journal of the American Statistical Association*, 96, 42–55.

Bandi, F. M. and Russell, J. R. (2003), "Microstructure Noise, Realized Volatility and Optimal Sampling," Tech. rep., University of Chicago Graduate School of Business.

Barndorff-Nielsen, O. E. and Shephard, N. (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, B*, 64, 253–280.

Billingsley, P. (1995), *Probability and Measure*, New York: Wiley, 3rd ed.

Chernov, M. and Ghysels, E. (2000), "A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics*, 57, 407–458.

Dacorogna, M. M., Gençay, R., Müller, U., Olsen, R. B., and Pictet, O. V. (2001), *An Introduction to High-Frequency Finance*, San Diego: Academic Press.

Delattre, S. and Jacod, J. (1997), "A Central Limit Theorem for Normalized Functions of the Increments of a Diffusion Process, in the Presence of Round-Off Errors," *Bernoulli*, 3, 1–28.

Farmer, J. D., Gillemot, L., Iori, G., Krishnamurthy, S., Smith, D. E., and Daniels, M. G. (2005), "A Random Order Placement Model of Price Formation in the Continuous Double Auction," in *The Economy as an Evolving Complex System, III*, eds. Blume, L. and Durlauf, S., Oxford, England: Oxford University Press, pp. 133–173.

Gallant, A. R., Hsu, C.-T., and Tauchen, G. T. (1999), "Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance," *The Review of Economics and Statistics*, 81, 617–631.

Gloter, A. (2000), "Estimation des Paramètres d'une Diffusion Cachée," Ph.D. thesis, Université de Marne-la-Vallée.

Hall, P. and Heyde, C. C. (1980), *Martingale Limit Theory and Its Application*, Boston: Academic Press.

Hansen, P. R. and Lunde, A. (2006), "Realized Variance and Market Microstructure Noise," *Forthcoming in the Journal of Business and Economic Statistics.*

Hasbrouck, J. (1996), "Modeling Market Microstructure Time Series," in *Handbook of Statistics, volume 14*, eds. Rao, C. R. and Maddala, G. S., Amsterdam: North-Holland, pp. 647–692.

Hull, J. and White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42, 281–300.

Jacod, J. (1994), "Limit of Random Measures Associated with the Increments of a Brownian Semimartingale," Tech. rep., Université de Paris VI.

Jacod, J. and Protter, P. (1998), "Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations," *Annals of Probability*, 26, 267–307.

Large, J. (2005), "Quadratic Variation When Quotes Change One Tick at a Time," Tech. rep., Oxford University.

Mykland, P. A. and Zhang, L. (2002), "ANOVA for Diffusions," *The Annals of Statistics, forthcoming*, –, –.

Revuz, D. and Yor, M. (1999), *Continuous Martingales and Brownian Motion*, Berlin, Germany: Springer-Verlag, 3rd ed.

Zeng, Y. (2003), "A partially-observed model fo0r micro-movement of asset process with Bayes estimation via filtering for a class of micro-movement models of asset price," *Mathematical Finance*, 13, 411–444.

Zhang, L. (2004), "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach," Tech. rep., Carnegie-Mellon University.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 472, 1394–1411.