LARGE DEVIATIONS OF REALIZED VOLATILITY

By

Shin Kanaya and Taisuke Otsu

May 2011

COWLES FOUNDATION DISCUSSION PAPER NO. 1798

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Large deviations of realized volatility*

Shin Kanaya† University of Oxford

Taisuke Otsu‡ Yale University

May 2011

Abstract

This paper studies large and moderate deviation properties of a realized volatility statistic of high frequency financial data. We establish a large deviation principle for the realized volatility when the number of high frequency observations in a fixed time interval increases to infinity. Our large deviation result can be used to evaluate tail probabilities of the realized volatility. We also derive a moderate deviation rate function for a standardized realized volatility statistic. The moderate deviation result is useful for assessing the validity of normal approximations based on the central limit theorem. In particular, it clarifies that there exists a trade-off between the accuracy of the normal approximations and the path regularity of an underlying volatility process. Our large and moderate deviation results complement the existing asymptotic theory on high frequency data. In addition, the paper contributes to the literature of large deviation theory in that the theory is extended to a high frequency data environment.

1 Introduction

Realized volatility and its related statistics have become standard tools to explore the behavior of financial data and to evaluate financial theoretical models including stochastic volatility models.[1] This increase in popularity has been propelled by recent developments of probability and statistical theory

*The authors wish to thank Ilze Kalnina, Gechun Liang, Per Mykland, Neil Shephard, David Steinsaltz, Lan Zhang, Matthias Winkel, and seminar participants at Oxford-Man Institute and the Statistics Department, University of Oxford, for helpful comments. The authors are also grateful to the editor Thomas Mikosch, an associate editor, and two anonymous referees whose comments significantly improved the paper.

†Department of Economics, Nuffield College and Oxford-Man Institute. Address: Oxford-Man Institute, Eagle House, Walton Well Road, Oxford, OX2 6ED, UK. E-mail: shin.kanaya@economics.ox.ac.uk. Phone: +44 (0) 1865-616637. Fax: +44 (0) 1865-616601.

‡Cowles Foundation and Department of Economics. Address: P.O. Box 208281, New Haven, CT 06520-8281, USA. E-mail: taisuke.otsu@yale.edu. Website: http://cowles.econ.yale.edu/faculty/otsu.htm. Phone: +1-203-432-9771. Fax: +1-203-432-6167. Financial support from the National Science Foundation (SES-0720961) is gratefully acknowledged.

[1]See, e.g., Andersen, Bollerslev and Diebold (2009), Bandi and Russell (2007), and Shephard and Andersen (2009) for reviews on realized volatility and stochastic volatility models.

and by the increasing availability of high frequency financial data. Using the asymptotic framework where the number of high frequency observations in a fixed time interval (say, a day) increases to infinity, Barndorff-Nielsen and Shephard (2002) established a law of large numbers and a central limit theorem for realized volatility, which were extended to more general setups and statistics by Barndorff-Nielsen *et al.* (2006a,b). Also, Gonçalves and Meddahi (2009) investigated higher-order properties of the realized volatility statistic and its bootstrap analog based on Edgeworth expansions. These central limit theorem and Edgeworth expansion results are useful to explore asymptotic behaviors of realized volatility, in particular around the center of its distribution. On the other hand, tail behaviors of the realized volatility statistic, such as large and moderate deviation properties, have not been explored yet in the literature.

This paper studies the large and moderate deviation properties of realized volatility of high frequency data. We establish the large deviation principle in the sense of Dembo and Zeitouni (1998, Section 1.2) for the realized volatility statistic when the number of high frequency observations in a fixed time interval increases to infinity. The large deviation result can be used to evaluate and approximate tail probabilities of realized volatility. We also derive a moderate deviation result for a standardized realized volatility statistic, which fills the gap between the central limit theorem and the large deviation one. The moderate deviation result is useful for assessing the validity of normal approximations based on the central limit theorem. In particular, it clarifies that there exists a trade-off between the accuracy of the normal approximations and the path regularity of an underlying volatility process. Our large and moderate deviation results complement the existing asymptotic theory on high frequency data.

This paper also contributes to the literature of large deviation theory.[2] In particular, we extend the strategy of the proof of Gärtner and Ellis' large deviation theorem (Gärtner, 1977 and Ellis, 1984) for general dependent processes to our high frequency data environment. It should be noted that since we cannot determine the limiting behavior of the cumulant generating function at some boundary point, the proof strategy of Gärtner and Ellis' large deviation theorem is not directly applicable to our case. To deal with this technical difficulty, we modify an approach by Bercu, Gamboa and Rouault (1997) and Bryc and Dembo (1997), where they established the large deviation principle for quadratic forms of Gaussian processes.

In Section 2, we present our baseline model and derives the large and moderate deviation results for the realized volatility statistic. In particular, we derive the exponential convergence rate function for the *conditional* tail probabilities of the estimation errors of the realized volatility statistic *given an underlying volatility process*. By this conditioning combined with the no leverage effect assumption (i.e., independence between the volatility and innovation processes), we concentrate on characterizing the estimation errors driven by the innovation process which is assume to be a standard Brownian motion. Since the underlying volatility process is unobservable, our large deviation results should be considered as a contribution to probability theory (where we derive implications from the assumed

[2]See, e.g., Dembo and Zeitouni (1998) for a review on large deviation theory in the probability and statistics literature.

probabilistic model) rather than theory of statistical inference (where we utilize observed data to infer on the unknown data generating process). We note that Gonçalves and Meddahi (2009) adopt a similar conditioning strategy to derive the second-order properties of the realized volatility statistic. Section 3.1 discusses some technical issues to derive the unconditional version of our large deviation result. Section 3.2 extends the baseline model to allow some specific form of leverage effects and derive analogous large and moderate deviation results. Our large deviation analysis can be considered as a starting point to derive more general properties (e.g., the unconditional large deviation theorems) or to compare the realized volatility statistic with other estimators for the integrated volatility, such as the ones by Barndorff-Nielsen *et al.* (2008), Christensen, Oomen and Podokskij (2010), and Zhang, Mykland and Aït-Sahalia (2005a).

2 Main Results

We first introduce our basic setup. Let us consider a univariate continuous time process Y on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P\right)$, which satisfies the following assumption.

Assumption. *Y follows a Brownian semimartingale:*

$$Y_t := Y_0 + \int_0^t a_u du + \int_0^t \sigma_{u-} dW_u, \tag{1}$$

where Y_0 is a real-valued (initial) random variable, W is a standard Brownian motion, a is a predictable and bounded drift process, and σ is an adapted càdlàg volatility process which is locally bounded away from zero and independent of W.[3]

We adopt this setup as a benchmark and later discuss some possible extensions to more general setups. There are at least two limitations in this assumption. First, the condition of independence between σ and W excludes the presence of so-called leverage effects. A negative correlation between asset returns and volatilities is referred to as the leverage effect, which is often observed in stock price data (see Black, 1976). In our context, a leverage effect corresponds to a negative contemporaneous correlation between the volatility σ and the Brownian innovation W. While assuming the absence of leverage effects is restrictive for stock returns, it is empirically reasonable for some exchange rate data (see Andersen, Bollerslev and Meddahi, 2005). In Section 3.2, we consider an extension of the baseline model to allow some leverage effects. Second, the above assumption does not allow jumps in the process Y. Some previous studies argue that the presence of jumps is a prominent feature of some high frequency financial data.[4] On the other hand, it should be noted that several existing theoretical

[3]This assumption guarantees that the integrals on the right-hand side of (1) are well-defined, which implies the existence and uniqueness of the process Y (see, e.g., Section 3.2 of Karatzas and Shreve, 1991). Also note that the stochastic integral $\int_0^t \sigma_{u-} dW_u$ is a (local) martingale, provided that σ is adapted and càdlàg (see the reference above).

[4]For example, Barndorff-Nielsen and Shephard (2006) tested the presence of jumps in exchange rate data for each day in 10 years. They found that the null of no jump is rejected at the 5% significance level on about 20% of days, while

studies on realized volatility have imposed assumptions similar to ours. For example, Bandi and Russell (2008) excluded leverage effects and jumps to derive an optimal sampling frequency under the presence of market microstructure noises. Gonçalves and Meddahi (2009) also excluded them to investigate higher-order properties of realized volatility and its bootstrap counterpart. Additionally, we do not explicitly take into account the presence of measurement errors in Y, so-called market microstructure noises. Some papers, such as Andersen *et al.* (2003), suggest the use of realized volatility with sampling at lower frequencies to alleviate biases caused by the market microstructure noises. Our large deviation analysis may give a reasonable description of the tail behavior under such a sampling method.[5]

Throughout the paper, let $\Pr\{\cdot\}$ and $E[\cdot]$ be the conditional probability and conditional expectation given the path of σ, respectively. Given the independence of σ and W, $\Pr\{\cdot\}$ and $E[\cdot]$ are taken with respect to the innovation process W. We employ a càglàd version of the volatility process and write the Brownian martingale component as $\int_0^t \sigma_{u-} dW_u$ in (1).[6] We may write this component as $\int_0^t \sigma_u dW_u$ since both are almost surely equal (as long as $\int_0^t \sigma_u dW_u$ is martingale).[7] In this paper, we use the expression $\int_0^t \sigma_{u-} dW_u$ to facilitate some technical argument below (see, e.g., (28) in the proof of Lemma A.1). Also note that $\{\sigma_{u-}^2\}$ is locally Riemann integrable by the càdlàg condition of σ. In particular, the integrated volatility $\int_0^1 \sigma_{u-}^2 du$ is well-defined, which is of our interest in estimation by the realized volatility statistic.[8]

Suppose that from the process Y in (1), we obtain n high frequency observations of asset returns on the time interval $[0, 1]$ (say, one day), that is

$$\Delta_i^n Y := Y_{i/n} - Y_{(i-1)/n},$$

for $i = 1, \dots, n$. Based on the high frequency returns, the realized volatility statistic is defined as their

at the 1% significance level on about 10% of days. See also Aït-Sahalia and Jacod (2009) and Andersen, Bollerslev and Dobrev (2007) for analyses of jumps in individual stock returns and S&P500 future returns, respectively.

[5]See also Bandi and Russell (2008), Barndorff-Nielsen *et al.* (2008), Hansen and Lunde (2005), and Zhang, Mykland and Aït-Sahalia, (2005a,b). Some of these papers suggest alternative volatility statistics which are robust to the market microstructure noises. It is interesting to investigate large deviation properties of such statistics. While it is uncertain if we can obtain sensible large deviation results with allowing for the noises, an asymptotic assumption in Zhang, Mykland and Aït-Sahalia (2011) may help us to proceed, where they established Edgeworth expansions for realized volatility and some related statistics under the presence of market microstructure noises by using small-noise asymptotics.

[6]A càglàd process is a process whose paths are left-continuous with right limits almost surely.

[7]The almost sure equivalence holds because W is a Brownian motion. To see this point, observe that

$$E\left[\left|\int_0^t (\sigma_u - \sigma_{u-})\, dW_u\right|^2\right] = E\left[\int_0^t (\sigma_u - \sigma_{u-})^2\, du\right] = 0,$$

where the first equality holds by the Ito isometry, and the second by the càdlàg condition of σ (note that a realized path of the càdlàg process may have infinitely many jumps on any finite interval but its number is at most countable). Therefore, it holds that $\int_0^t (\sigma_u - \sigma_{u-})\, dW_u = 0$ in the L_2 sense, and so it does in the almost sure sense.

[8]By the same token, the integrated volatility may be written as $\int_0^1 \sigma_u^2 du$, which is equal to $\int_0^1 \sigma_{u-}^2 du$ almost surely (as well as everywhere conditional on the realized path of σ).

squared sum:

$$RV_n := \sum_{i=1}^{n} (\Delta_i^n Y)^2 .$$

Given the increasing availability of high frequency financial data, the realized volatility statistic RV_n has been used as a fundamental tool to investigate volatility in financial markets (e.g., Andersen *et al.*, 2003 and Barndorff-Nielsen and Shephard, 2002). For example, realized volatility has been employed as a descriptive measure of volatility of stock returns in financial markets and a basic diagnostic to evaluate stochastic volatility models. Despite its fundamental importance and long history of empirical applications, theoretical studies on the realized volatility statistic have started only recently. Barndorff-Nielsen and Shephard (2002) established the weak law of large numbers and the central limit theorem for the realized volatility statistic RV_n when the data frequency n increases to infinity (see also Barndorff-Nielsen and Shephard, 2004, and Barndorff-Nielsen *et al.*, 2006a,b, for more general results). Letting $\overline{\sigma^q} := \int_0^1 \sigma_{u-}^q du$, these limit theorems are stated as

$$RV_n \xrightarrow{p} \overline{\sigma^2},$$

$$\frac{\sqrt{n}\left(RV_n - \overline{\sigma^2}\right)}{\sqrt{\overline{\sigma^4}}} \xrightarrow{d} N(0,2) . \tag{2}$$

The weak law of large numbers says that the realized volatility statistic RV_n converges in probability to the integrated volatility $\overline{\sigma^2}$, and the central limit theorem tells us that an asymptotic approximation to the estimation error $RV_n - \overline{\sigma^2}$ is given by the normal distribution. In other words, the large deviation probability $\Pr\left\{\left|RV_n - \overline{\sigma^2}\right| > c\right\}$ converges to zero for any positive constant c, and the local deviation probability $\Pr\left\{\left|RV_n - \overline{\sigma^2}\right| > c/\sqrt{n}\right\}$ is approximated by the normal distribution. Due to the localization, the central limit theorem is useful for approximating the finite sample distribution of RV_n particularly around the center of its distribution. On the other hand, it can be less precise when one wishes to capture the tail behavior of RV_n. As a complement to the above limiting theorems, this paper considers an asymptotic approximation to the large deviation probability $\Pr\left\{\left|RV_n - \overline{\sigma^2}\right| > c\right\}$ with a given constant c, which is able to provide a more accurate description of the tail behavior of RV_n. Also, to describe the tail behavior between the local and large deviations, we consider the moderate deviation probability $\Pr\left\{\left|RV_n - \overline{\sigma^2}\right| > c/m_n\right\}$ with $m_n \to \infty$ and $m_n/\sqrt{n} \to 0$.

More specifically, we establish the large deviation principle in the sense of Dembo and Zeitouni (1998, Section 1.2) for the realized volatility statistic. We say that a sequence $\{Z_n\}_{n\in\mathbb{N}}$ of random variables on $\mathcal{Z}$ satisfies the large deviation principle (LDP) with speed $s_n \searrow 0$ and good rate function $I : \mathcal{Z} \to [0,\infty]$, if

(i) I is a good rate function: I is lower semicontinuous and level compact (i.e., $I^{-1}([0,c])$ is compact for all $c \in (0,\infty)$),

(ii) for any closed set $F \subset \mathbb{R}$,

$$\limsup_{n\to\infty} s_n \log \Pr\{Z_n \in F\} \le -\inf_{x\in F} I(x) , \tag{3}$$

(iii) for any open set $G \subset \mathbb{R}$,

$$\liminf_{n\to\infty} s_n \log \Pr\{Z_n \in G\} \geq -\inf_{x\in G} I(x). \tag{4}$$

We establish the LDP by setting Z_n as the realized volatility statistic RV_n and its normalized version.

To present our main results, we introduce some notation. Let

$$\bar{\lambda} := \frac{1}{2\sup_{u\in[0,1]}\sigma_{u-}^2}, \tag{5}$$

$$\Lambda(\lambda) := -\frac{1}{2}\int_{[0,1]} \log\left(1 - 2\lambda\sigma_{u-}^2\right) du, \quad \text{for } \lambda \in \left(-\infty, \bar{\lambda}\right), \tag{6}$$

$$\Lambda^*(x) := \sup_{\lambda\in\left(-\infty,\bar{\lambda}\right)} \{\lambda x - \Lambda(\lambda)\}, \quad \text{for } x \in \mathbb{R}, \tag{7}$$

$$\Lambda^{**}(x) := \sup_{\lambda\in\mathbb{R}} \left\{\lambda x - \lambda^2\overline{\sigma^4}\right\} = \frac{x^2}{4\overline{\sigma^4}}, \quad \text{for } x \in \mathbb{R}. \tag{8}$$

In Lemma A.1, we show that $\Lambda(\lambda)$ is the (pointwise) limit for the normalized cumulant generating function $n^{-1}\Lambda_n(\lambda n)$ for each $\lambda \in \left(-\infty, \bar{\lambda}\right)$, where $\Lambda_n(u) := \log E[\exp(uRV_n)]$. The function Λ^* is the Fenchel-Legendre transform of Λ. The function Λ^{**} corresponds to the Fenchel-Legendre transform of the cumulant generating function of the normal distribution $N\left(0, 2\overline{\sigma^4}\right)$.[9]

Our main theorems are presented as follows.

Theorem 1. *[Large deviation] Suppose that Assumption holds. Then the sequence of the realized volatility $\{RV_n\}_{n\in\mathbb{N}}$ satisfies the LDP with speed n^{-1} and good rate function Λ^*.*

Theorem 2. *[Moderate deviation] Suppose that Assumption holds and that for the given path of σ, the sequence $\{m_n\}_{n\in\mathbb{N}}$ satisfies $m_n \to \infty$, $m_n^2/n \to 0$,*

$$m_n\left(\frac{1}{n}\sum_{i=1}^n \sigma_{(i-1)/n}^2 - \overline{\sigma^2}\right) \to 0, \quad \text{and} \quad \frac{m_n}{n}\sum_{i=1}^n \sup_{u\in((i-1)/n,i/n]} \left|\sigma_{u-} - \sigma_{(i-1)/n}\right| \to 0, \tag{9}$$

as $n \to \infty$. Then the normalized sequence of the realized volatility $\left\{m_n\left(RV_n - \overline{\sigma^2}\right)\right\}_{n\in\mathbb{N}}$ satisfies the LDP with speed m_n^2/n and good rate function Λ^{**}.

Proofs of these theorems are provided in Appendix. Some remarks on the theorems are in order.

Remarks:

1. First of all, we emphasize that the probability $\Pr\{\cdot\}$ and expectation $E[\cdot]$ are conditional ones given the path of the volatility process σ. Thus, the above theorems describe the *conditional* large and moderate deviation probabilities for the realized volatility statistic RV_n *given the path*

[9] Except for Section 3.1, all of our arguments and proofs are made under the conditional expectations and probabilities given a realized path of the volatility process σ. Since we assume that σ is càdlàg and locally bounded away from zero, each realized path of σ is uniformly bounded from above and away from zero over $[0, 1]$. This implies that $\sup_{u\in[0,1]}\sigma_{u-}^2$ is finite for each realized path of σ and $\bar{\lambda}$ cannot be null.

of σ. Since the volatility process σ is unobservable, these theorems should be considered as contributions to probability theory (where we derive implications from the assumed probabilistic model) rather than theory of statistical inference (where we utilizes observed data to infer on the unknown data generating process). Section 3.1 discusses technical issues to extend these results to the unconditional probabilities.

2. Theorem 1 is on the large deviation property. This theorem says that the large deviation probability of realized volatility decays at an exponential rate, whose exponent is characterized by the Fenchel-Legendre transform Λ^* of Λ. This form of the rate function is analogous to the ones obtained in the existing large deviation theorems, such as Cramér's theorem for the sum of iid observations and Gärtner-Ellis' theorem for the sum of possibly dependent heterogeneous observations. Relying on the Brownian semimartingale assumption, we can obtain a specific form of the limiting moment generating function Λ. Note that Λ is the cumulant generating function of the (scaled) χ^2 distribution when the volatility process is constant. Intuitively, given the volatility process σ, we can approximate the large deviation probability $\Pr\{RV_n \in A\}$ by the formula $\exp\{-n \inf_{x\in A} \Lambda^*(x)\}$ for any interval A. For example, the estimation error probability $\Pr\left\{\left|RV_n - \overline{\sigma^2}\right| > c\right\}$ for some $c > 0$ can be approximated by setting $A = \left(\overline{\sigma^2} + c, \infty\right) \cup \left(-\infty, \overline{\sigma^2} - c\right)$.

3. A key to Gärtner-Ellis' theorem is the assumption that the limit of the normalized cumulant, i.e., $\lim_{n\to\infty} n^{-1}\Lambda_n(\lambda n)$, exists and is determinate in the extended real line for each $\lambda \in \mathbb{R}$.[10] However, in the present setup the limit may be finite or infinite at the boundary point $\bar{\lambda}$, depending on the realized path of σ. If we can assume $\lim_{n\to\infty} n^{-1}\Lambda_n\left(\bar{\lambda} n\right) = \infty$, then we can adapt Gärtner-Ellis' theorem to our context. However, such a requirement may be too strong to accommodate some volatility processes.[11] Furthermore, it is not generally easy to check the requirement since $\Lambda_n\left(\bar{\lambda} n\right)$ needs to be evaluated for each realized path of σ. To circumvent this technical difficulty, we adopt a modified approach of Bercu, Gamboa and Rouault (1997) and Bryc and Dembo (1997), where

[10] Additionally, Gärtner-Ellis' theorem requires that the limiting normalized cumulant is essentially smooth, which is also typically violated in our setup.

[11] The condition $\lim_{n\to\infty} n^{-1}\Lambda_n\left(\bar{\lambda} n\right) = \infty$ holds when the Lebesgue measure of the set $\mathcal{S}_\sigma := \left\{u \in [0,1] : \sigma^2_{u-} = \sup_{s\in[0,1]} \sigma^2_{s-}\right\}$ is not zero, i.e., $\xi(\mathcal{S}_\sigma) > 0$, where ξ is the Lebesgue measure on $\mathbb{R}$. This is satisfied if σ is a piecewise constant process for example. On the other hand, if $\xi(\mathcal{S}_\sigma) = 0$, the condition may fail to hold. To see this point let us consider a process $\sigma^2_u = \pi_\theta(u)$, where

$$\pi_\theta(u) := \begin{cases} 1 - \exp\{(1/2-u)^{-\theta}\} & \text{for } u \in [0,1/2), \\ 1 & \text{for } u = 1/2, \\ 1 - \exp\{(u-1/2)^{-\theta}\} & \text{for } u \in (1/2,1], \end{cases}$$

and θ is a non-negative constant parameter. Note that $\log\left(1 - 2\bar{\lambda}\sigma^2_u\right) = \log(1 - \pi_\theta(u))$, which is bounded everywhere except at a single point $u = 1/2$ and $\xi(\mathcal{S}_\sigma) = 0$. In this case it can be seen that $\lim_{n\to\infty} n^{-1}\Lambda_n\left(\bar{\lambda} n\right) = \infty$ for $\theta \geq 1$ (but $< \infty$ for $0 \leq \theta < 1$).

they investigate large deviation behavior of quadratic forms of Gaussian processes.[12]

4. Theorem 2 is on the moderate deviation property. A scaling factor m_n for the normalized statistic $m_n\left(RV_n - \overline{\sigma^2}\right)$ diverges to infinity but with the rate slower than $\sqrt{n}$. Let

$$L_n := \frac{1}{n}\sum_{i=1}^{n} \sup_{u\in((i-1)/n,i/n]} \left|\sigma_{u-} - \sigma_{(i-1)/n}\right|.$$

Both conditions in (9) are satisfied when $m_n L_n \to 0$. This is because

$$\left|\frac{1}{n}\sum_{i=1}^{n} \sigma^2_{(i-1)/n} - \overline{\sigma^2}\right| \leq \frac{1}{n}\sum_{i=1}^{n} \sup_{u\in((i-1)/n,i/n]} \left|\sigma^2_{u-} - \sigma^2_{(i-1)/n}\right|, \quad \text{and}$$
$$\left|\sigma^2_{u-} - \sigma^2_{(i-1)/n}\right| \leq C\left|\sigma_{u-} - \sigma_{(i-1)/n}\right|,$$

where C is some positive constant (uniform over i and u), whose existence is guaranteed by Assumption. We say that the realized path of the volatility process σ is more regular if the decay rate of L_n is faster. The decay rate of L_n may be regarded as a measure of the degree of path continuity/smoothness of the process. Note that the rate function Λ^{**} in (8) is the Fenchel-Legendre transform of the cumulant generating function of the normal distribution $N\left(0, 2\overline{\sigma^4}\right)$, which corresponds to the limiting distribution of $\sqrt{n}\left(RV_n - \overline{\sigma^2}\right)$ in the central limit theorem (2). Therefore, Theorem 2 says that if the path of the volatility process is sufficiently regular to satisfy (9) with a given factor m_n, then the moderate deviation probability of RV_n is still approximated by the normal distribution. We are faced with a trade-off between the degree of regularity of the volatility process and the possible range of rates of m_n as characterized by (9).

5. In connection with the previous remark we provide the following additional discussion of the conditions in (9). First, since the càdlàg property of σ implies $L_n \to 0$ (see, pp. 121-123 of Billingsley, 1999), we can always find a sequence $\{m_n\}$ satisfying (9). Possible rates of m_n depend upon the path property of σ. We below present some examples in order.

   (a) Suppose that σ is a (continuous) Brownian semimartingale written as

$$\sigma_t = \sigma_0 + \int_0^t \theta_s ds + \int_0^t v_s dZ_s, \tag{10}$$

   where Z is a Brownian motion, and θ and v are locally bounded processes whose integral and stochastic integral are respectively well-defined. In this case, we can let $m_n = o\left(\sqrt{n/\log n}\right)$ because there exists a random variable X such that $X(\omega) < \infty$ for almost every $\omega \in \Omega$ and that

$$\Pr_\sigma\left\{\limsup_{n\to\infty}\left\{\frac{\max_{1\leq i\leq n}\sup_{u\in((i-1)/n,i/n]}\left|\sigma_u - \sigma_{(i-1)/n}\right|}{\sqrt{n^{-1}\log n}}\right\} \leq X\right\} = 1,$$

[12] Alternatively it might be possible to apply a perturbed version of Gärtner-Ellis' theorem by Feng, Forde and Fouque (2010).

where $\Pr_\sigma$ is the probability with respect to σ, and the rate $\sqrt{n^{-1}\log n}$ is exact (note that $\sigma_u = \sigma_{u-}$). This result follows from the (local) modulus of continuity of the Brownian motion and the Dambis-Dubins-Schwarz theorem (see Chapters I and V of Revuz and Yor, 1999, respectively). The specification of (10) allows the following type of stochastic differential equation:

$$df(\sigma_t) = \alpha(\sigma_t)\,ds + \beta(\sigma_t)\,dZ_t,$$

where f is a twice continuously differentiable function (e.g., $f(x) = x^2$, $\log x$), and α and β are the drift and diffusion functions. Heston's (1993) model and the continuous-time limit version of the GARCH model in Nelson (1990) are in this category.

(b) For a general continuous process which may not be a semimartingale, it is possible to estimate the degree of continuity of the volatility path by using the Kolmogorov-Čentsov criterion (see Theorem 2.8 in Chapter 2 of Karatzas and Shreve, 1991). Let E_σ be the expectation with respect to σ. If

$$E_\sigma\left[|\sigma_{t+h} - \sigma_t|^p\right] \le C\,|h|^{1+q} \quad \text{for some positive constants } C, p \text{ and } q, \tag{11}$$

then σ_t is almost surely (locally) Hölder continuous with degree of $\gamma \in (0, q/p)$. In this case, we can set $m_n = o(n^\gamma)$ for any $\gamma \in (0, \min\{q/p, 1/2\})$. As an example, consider the case where σ is driven by a fractional Brownian motion (see Chapter 4 of Embrechts and Maejima, 2002):

$$\sigma_t = f(B_H(t)),$$

where f is a twice continuously differentiable function, and $B_H(t) := \int_{-\infty}^t (t-s)^{H-1/2}\,dZ_s$ with the Hurst index $H \in (0,1)$. In this case we can show (11) with $q = pH - 1$ for any $p \in (1/H, \infty)$. Therefore, we can pick any $\gamma \in (0, H)$. The smaller H means the more irregular path of σ, and thus m_n must grow at a slower rate (the range of possible rates is narrower). Note that the Brownian semimartingale in (10) essentially corresponds to the case of $H = 1/2$. Therefore, the normal approximations under (10) may be less accurate than those under volatility processes with $H > 1/2$ (i.e., long-range dependence or long memory process; see, e.g., Comte and Renault, 1996).

(c) We may also work with a volatility process whose paths exhibit jumps. Suppose that σ is described by a process which consists of continuous and jump components, i.e., $\sigma_t = X_t + J_t$ with a continuous process X and a pure jump process J. We can consider two possible cases for J. First, if J is a process whose number of jumps is almost surely finite over any finite interval (e.g., a compound Poisson process), then the possible rate of m_n is completely determined by the path property of the continuous part X. If X is a Brownian semimartingale defined in (10), then we can set $m_n = o\left(\sqrt{n/\log n}\right)$. If $X_t = 0$, we can set $m_n = o\left(n^{1/2}\right)$, which is the fastest divergence rate of m_n among any volatility processes. The second is the case where σ is a process which may have infinitely many (small) jumps over a finite interval, such as a Lévy-type process with the infinite Lévy measure (see, e.g., Barndorff-Nielsen and Shephard, 2001, Barndorff-Nielsen, Shephard and Winkel, 2006, and Todorov, 2008). Although the convergence $L_n \to 0$ for this type of process is guaranteed by the càdlàg condition, its rate is generally unknown.

6. We can naturally think of an extension of our theorems to the realized power variations:

$$R_r := n^{-1+r/2} \sum_{i=1}^{n} \left| \Delta_i^n Y \right|^r ,$$

for $r > 0$. The realized volatility statistic RV_n corresponds to the case of $r = 2$. The law of large numbers and the central limit theorem for this statistic are studied in the literature (Barndorff-Nielsen and Shephard, 2002, 2004, Barndorff-Nielsen *et al.* 2006a,b, and Tauchen and Todorov, 2010). To establish a large deviation theorem for R_r, we can use the same strategy as in our proof in particular for $r \in (0, 2)$. However, the proof of Lemma A.1, which derives the limit of the normalized cumulant generating function, will be completely different. Thus, we need to investigate the limiting behavior of the cumulant generating function individually for each r.

7. When we consider the possibility of jumps in asset returns, a bipower variation is a useful statistic (see, e.g. Barndorff-Nielsen and Shephard, 2004, 2006). Under Assumption (i.e., Y does not exhibit any jump) we conjecture that a large deviation theorem for the bipower variation statistics can be derived by the same strategy. Again, a challenging part is to establish an analogous result of Lemma A.1 to characterize the limiting behavior of the cumulant generating functions. For the case where the process Y contains jumps, it may be possible to proceed by assuming certain stationarity and mixing conditions of the process, as is done in Section 6.4 of Dembo and Zeitouni (1998).

3 Discussions

In this section, we discuss two important directions to extend the results obtained in the previous section. First, we consider the possibility of establishing the large or moderate deviation result for the realized volatility statistic RV_n *without* conditioning on the path of the volatility process σ. Second, we consider a certain class of models which allow leverage effects (i.e., dependence between σ and W), and derive analogous large and moderate deviation theorems to the ones in the last section.

3.1 Unconditional LDP

In the last section, we focus on the *conditional* large or moderate deviation probability of RV_n given the path of the volatility process σ (recall that $\Pr\{\cdot\}$ and $E[\cdot]$ mean the conditional probability and conditional expectation given the path of σ, respectively). Since we do not observe the volatility process, our main results in the last section, which focus on deducing the large or moderate deviation probabilities for each given (hypothetical) path of σ, should be considered as a contribution to probability theory rather than inferential statistics. Gonçalves and Meddahi (2009) adopted a similar approach and characterized higher-order properties of the realized volatility statistic and its bootstrap analog by conditioning on the path of σ. Although a formal analysis is beyond the scope of the paper, this subsection discusses some technical issue to derive the unconditional LDP for RV_n.

As discussed in the last section, a key to establish the LDP is to find the limit of the normalized cumulant generating function of RV_n. For the conditional case, the limit $\Lambda(\lambda)$ is derived in Lemma A.1. For the unconditional case, however, it would not be an easy task to compute such a limit even if we maintain the independence of σ and W. To see this point, let us consider the same setup as the last section and look at the normalized cumulant generating function of RV_n, that is

$$\frac{1}{n}\Lambda_{U,n}(\lambda n) := \frac{1}{n}\log E_\sigma\left[E\left[\exp\left\{\lambda n RV_n\right\}\right]\right],$$

where $E_\sigma[\cdot]$ is the expectation with respect to σ. Suppose that the limit

$$\Lambda_U(\lambda) := \lim_{n\to\infty}\frac{1}{n}\log E_\sigma\left[\prod_{i=1}^{n}\left[1-2\lambda\sigma^2_{(i-1)/n}\right]^{-1/2}\right]$$

exists for all $\lambda \in \mathbb{R}$ in the extended real line and that the volatility process σ is almost surely bounded (in addition to Assumption in the last section). Then a similar argument to the proof of Lemma A.1 yields

$$\lim_{n\to\infty}\frac{1}{n}\Lambda_{U,n}(\lambda n) = \lim_{n\to\infty}\frac{1}{n}\log E_\sigma\left[E\left[\exp\left\{\lambda n\sum_{i=1}^{n}\left(\int_{(i-1)/n}^{i/n}dW_u\right)^2\sigma^2_{(i-1)/n}\right\}\right]\right] = \Lambda_U(\lambda),$$

where the second equality follows from the property of the Brownian motion W. Based on this, Gärtner-Ellis' theorem implies that $\{RV_n\}_{n\in\mathbb{N}}$ satisfies the LDP with speed n^{-1} and good rate function $\Lambda_U^*(x) := \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda_U(\lambda)\}$ for $x \in \mathbb{R}$. However, it seems difficult to establish a general condition to guarantee the existence of the limit $\Lambda_U(\lambda)$. For example, because the sequence $\left\{\left(\int_{(i-1)/n}^{i/n}dW_u\right)^2\sigma^2_{(i-1)/n}\right\}_{i=1}^{n}$ is not independent, we cannot proceed as in (25) for the proof of Lemma A.1 and it is not clear how to proceed with the current strategy of the proof. By the same token, the dependence of $\left\{\left(\int_{(i-1)/n}^{i/n}dW_u\right)^2\sigma^2_{(i-1)/n}\right\}_{i=1}^{n}$ prevents a direct extension of the current strategy of the proof for Theorem 2 to derive a moderate deviation result.

3.2 Leverage Effect

In this subsection, we consider the following specification for the continuous time process Y, which allows some form of leverage effects.

Assumption'. *Y follows a Brownian semimartingale:*

$$Y_t := Y_0 + \int_0^t a_u du + \int_0^t \sigma_{u-}\left(\sqrt{1-\rho^2}dW_u + \rho dZ_u\right), \tag{12}$$

where Y_0 is a real-valued (initial) random variable, W and Z are standard Brownian motions which are independent, a is a predictable and bounded drift process, σ is an adapted càdlàg volatility process which is locally bounded away from zero and independent of W with $\sup_{u\in[0,1]} E_\sigma\left[\sigma_{u-}^4\right]<\infty$, and $\rho\in[0,1)$ is a constant. Also, assume that the filtrations $\{\mathfrak{F}_t^\sigma\}_{t\geq 0}$ and $\left\{\mathfrak{F}_t^Z\right\}_{t\geq 0}$ generated by σ and Z ($\mathfrak{F}_t^\sigma$, $\mathfrak{F}_t^Z\subset\mathcal{F}_t$), respectively, coincide.

This process is one of the most popular specifications in economics and finance literature (see, e.g., Romano and Touzi, 1997). The parameter ρ captures the degree of dependence between σ and W (i.e., the leverage effect). Although the last condition on the filtrations $\{\mathfrak{F}_t^\sigma\}_{t\geq 0}$ and $\left\{\mathfrak{F}_t^Z\right\}_{t\geq 0}$ is restrictive, it allows us to make our conditional argument transparent: given the realized path of σ, we can treat objects which consist of σ and Z (e.g., $\int_{(i-1)/n}^{i/n}\sigma_{u-}dZ_u$) as given (below we discuss how to proceed without this condition). One possible specification for σ satisfying this filtration condition is the process driven by Z with the following form

$$\sigma_t = h\left(t, Z_t\right), \tag{13}$$

where $h:[0,\infty)\times\mathbb{R}\to(0,\infty)$ is a positive-valued non-random function such that for each $t\in[0,\infty)$, $h\left(t,z\right)$ is càdlàg and strictly monotone in z. An example of (13) is the geometric Brownian motion

$$\sigma_t = \mu\sigma_t dt + s\sigma_t dZ_t,$$

with constants $\mu\in\mathbb{R}$ and $s\in(0,\infty)$. This process can be written as $\sigma_t=\sigma_0\exp\left\{\left(\mu-s^2/2\right)t+sZ_t\right\}$ with some initial constant $\sigma_0>0$ and satisfies all conditions on σ in Assumption'.

We define the following objects:

$$\bar{\lambda}_D := \frac{1}{2\left(1-\rho^2\right)\sup_{u\in[0,1]}\sigma_u^2}, \tag{14}$$

$$\begin{aligned}\Lambda_D\left(\lambda\right) &:= -\frac{1}{2}\int_0^1 \log\left[1-2\lambda\left(1-\rho^2\right)\sigma_{u-}^2\right]du \\ &\quad +2\lambda^2\left(1-\rho^2\right)\rho^2\int_0^1\frac{\sigma_{u-}^4}{1-2\lambda\left(1-\rho^2\right)\sigma_{u-}^2}du+\lambda\rho^2\overline{\sigma^2}\quad\text{for }\lambda\in\left(-\infty,\bar{\lambda}_D\right),\end{aligned} \tag{15}$$

$$\Lambda_D^*\left(x\right) := \sup_{\lambda\in\left(-\infty,\bar{\lambda}_D\right)}\left\{\lambda x-\Lambda_D\left(\lambda\right)\right\},\quad\text{for }x\in\mathbb{R}, \tag{16}$$

$$\Lambda_D^{**}\left(x\right) := \sup_{\lambda\in\mathbb{R}}\left\{\lambda x-\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}\right\}=\frac{x^2}{4\left(1-\rho^4\right)\overline{\sigma^4}},\quad\text{for }x\in\mathbb{R}. \tag{17}$$

These objects are counterparts of (5)-(8) to the present context. Note that when there is no leverage effect (i.e., $\rho=0$), these objects coincide with the counterparts in (5)-(8). We obtain the following theorems (the proofs are provided in the Appendix):

Theorem 3. *[Large deviation with leverage effects] Suppose that Assumption' holds. Then the sequence of the realized volatility $\{RV_n\}_{n\in\mathbb{N}}$ satisfies the LDP with speed n^{-1} and good rate function Λ_D^*.*

Theorem 4. *[Moderate deviation with leverage effects] Suppose that Assumption' holds and that for the given path of σ, the sequence $\{m_n\}_{n\in\mathbb{N}}$ satisfies $m_n \to \infty$, $m_n^2/n \to 0$, and the conditions provided in (9). Then the normalized sequence of the realized volatility $\left\{m_n\left(RV_n - \overline{\sigma^2}\right)\right\}_{n\in\mathbb{N}}$ satisfies the LDP with speed m_n^2/n and good rate function Λ_D^{**}.*

Theorem 3 establishes the LDP for the realized volatility under the process in (12). We can interpret this result in a way analogous to Remark 2. When there is no leverage effect (i.e., $\rho = 0$), the rate function Λ_D^* coincides with Λ^* in (7). Note that the limit of the normalized cumulant generating function Λ_D may be regarded as the mixture of the cumulant generating functions of the noncentral χ^2 distributions. This is due to the introduction of the leverage effects. On the other hand, the limiting function Λ in Theorem 1 for the no leverage case can be regarded as the mixture of the cumulant generating functions of the (scaled) χ^2 distributions (see Remark 2).

Theorem 4 derives a moderate deviation result. Similar comments to Remarks 4 and 5 apply here. When there is no leverage effect (i.e., $\rho = 0$), the rate function Λ_D^{**} coincides with Λ^{**} in (8). Note that Λ_D^{**} is the Fenchel-Legendre transform of the cumulant generating function of the normal distribution $N\left(0, 2\left(1-\rho^4\right)\overline{\sigma^4}\right)$. Thus, Theorem 4 says that if the path of the volatility process is sufficiently regular to satisfy (9) with a given factor m_n, then the moderate deviation probability of RV_n is still approximated by the normal distribution $N\left(0, 2\left(1-\rho^4\right)\overline{\sigma^4}\right)$. However, it should be noted that the above results are derive by conditioning on the realized path of σ and the derivation of the unconditional moderate deviation result is beyond the scope of this paper.

We conclude this section by giving some remarks on the possibilities of more general specifications. First, the last filtration condition in Assumption' is imposed for us to proceed with probabilities and expectations conditional on the realized path of σ, or equivalently, on $\{\mathfrak{F}_t^\sigma\}$ (in the same way as we did in Section 2). If we consider an extended filtration $\left\{\mathfrak{F}_t^\sigma \cup \mathfrak{F}_t^Z\right\}_{t\geq 0}$ and work with probabilities and expectations conditional on $\left\{\mathfrak{F}_t^\sigma \cup \mathfrak{F}_t^Z\right\}$, then we are able to develop large and moderate deviation results analogous to Theorems 3 and 4. Second, if we allow the leverage effect in the original process (1), then the estimation error can be written as

$$RV_n - \overline{\sigma^2} = \sum_{i=1}^n \xi_{n,i} \quad \text{with} \quad \xi_{n,i} := \int_{(i-1)/n}^{i/n} \int_{(i-1)/n}^{u} \sigma_{v-} dW_v \sigma_{u-} dW_u,$$

where $\{\xi_{n,i}\}$ is a martingale difference array (regardless of the dependence between σ and W). If some large deviation results on a very general martingale difference array were available, we could import them to our context. Unfortunately, to the best of our knowledge, results currently available in the literature cannot be immediately applied to our case.[13]

[13]For example, Grama (1997), and Lesigne and Volný (2001) provide exponential bounds of tail probabilities for discrete-time martingales; Grama and Haeusler (2000) suppose that martingale differences are almost surely bounded.

4 Conclusion

This paper derives large and moderate deviation theorems for realized volatility of high frequency financial data. Obtained results are natural extensions of conventional large and moderate deviation theorems, such as Cramér's and Gärtner-Ellis' theorems, to high frequency data environments where we increase the number of data frequency to infinity for the asymptotic approximation. Our large deviation result can be used to examine behaviors of the tail probabilities of the realized volatility statistics. Our moderate deviation result is useful for characterizing the validity of the central limit theorem based approximations in the high frequency context. In particular, it clarifies that the accuracy of the normal approximation depends upon the degree of regularity of the volatility process. Our large deviation analysis can be considered as a starting point to derive more general properties (e.g., the unconditional large deviation theorems), to work with more general setups with jumps and/or measurement errors, and to compare various types of estimators for the integrated volatility.

A Appendix

A.1 Auxiliary Lemmas

Here, we present two auxiliary lemmas which are used for the proof of Theorem 1. For each n, define the cumulant generating function for the realized volatility statistic as

$$\Lambda_n(u) := \log E\left[\exp\left\{u\sum_{i=1}^{n}(\Delta_i^n Y)^2\right\}\right].$$

For each n, the domain of Λ_n is the whole real line $\mathbb{R}$. The normalized cumulant generating function is defined as

$$\frac{1}{n}\Lambda_n(\lambda n) = \frac{1}{n}\log E\left[\exp\left\{\lambda n\sum_{i=1}^{n}(\Delta_i^n Y)^2\right\}\right]. \tag{18}$$

We derive the pointwise limit of $\frac{1}{n}\Lambda_n(\lambda n)$ on the extended real line $\mathbb{R}\cup\{-\infty\}\cup\{\infty\}$ for each $\lambda \in \left(-\infty, \bar{\lambda}\right)\cup\left(\bar{\lambda}, \infty\right)$ as n tends to ∞.

Lemma A.1. *Under Assumption, it holds that for each $\lambda \in \left(-\infty, \bar{\lambda}\right)$, the limit of $\frac{1}{n}\Lambda_n(\lambda n)$ as n tends to ∞ is $\Lambda(\lambda) < \infty$, i.e.,*

$$\lim_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda n) = \Lambda(\lambda) < \infty, \tag{19}$$

and for each $\lambda \in \left(\bar{\lambda}, \infty\right)$, the limit of $\frac{1}{n}\Lambda_n(\lambda n)$ as n tends to ∞ is infinity, i.e.,

$$\lim_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda n) = \infty, \tag{20}$$

where the threshold value $\bar{\lambda}$ and the limit function Λ are defined in (5) and (6), respectively.

Proof. First, we derive upper and lower bounds of $(\Delta_i^n Y)^2$. Letting

$$q_n(i) := \sup_{u \in ((i-1)/n, i/n]} \left| \sigma_{u-} - \sigma_{(i-1)/n} \right|, \tag{21}$$

and using the expression (1) for Y_t, we can write as

$$\begin{aligned}
(\Delta_i^n Y)^2 &= \left(\int_{(i-1)/n}^{i/n} a_u du \right)^2 + \left(\int_{(i-1)/n}^{i/n} dW_u \right)^2 \sigma_{(i-1)/n}^2 \\
&\quad + 2 \left(n^{1/4} \int_{(i-1)/n}^{i/n} a_u du \right) \left(n^{-1/4} \int_{(i-1)/n}^{i/n} \sigma_{u-} dW_u \right) \\
&\quad + \left(q_n^{-1/2}(i) \int_{(i-1)/n}^{i/n} \left[\sigma_{u-} - \sigma_{(i-1)/n} \right] dW_u \right) \left(q_n^{1/2}(i) \int_{(i-1)/n}^{i/n} \left[\sigma_{u-} + \sigma_{(i-1)/n} \right] dW_u \right).
\end{aligned}$$

Since the drift process a is assumed to be uniformly bounded, it holds that $\left(\int_{(i-1)/n}^{i/n} a_u du \right)^2 \leq Cn^{-2}$ for some $C > 0$. Thus, by applying the inequality $AB \leq (A^2 + B^2)/2$ (for $A, B \in \mathbb{R}$) to the last two terms on the right-hand side of the above equality,

$$\begin{aligned}
(\Delta_i^n Y)^2 &\leq Cn^{-2} \left(1 + n^{1/2} \right) + \left(\int_{(i-1)/n}^{i/n} dW_u \right)^2 \sigma_{(i-1)/n}^2 + \left(n^{-1/4} \int_{(i-1)/n}^{i/n} \sigma_{u-} dW_u \right)^2 \\
&\quad + \frac{1}{2} \left(q_n^{-1/2}(i) \int_{(i-1)/n}^{i/n} \left[\sigma_{u-} - \sigma_{(i-1)/n} \right] dW_u \right)^2 + \frac{1}{2} \left(q_n^{1/2}(i) \int_{(i-1)/n}^{i/n} \left[\sigma_{u-} + \sigma_{(i-1)/n} \right] dW_u \right)^2 \\
&= \sum_{j=1}^4 M_j^2(i) + Cn^{-2} \left(1 + n^{1/2} \right),
\end{aligned} \tag{22}$$

for $i = 1, \ldots, n$, where

$$M_j(i) := \int_{(i-1)/n}^{i/n} f_{j,u}(i) \, dW_u, \tag{23}$$

and $f_{j,u}(i)$ is a stochastic process on $[(i-1)/n, i/n]$ which is defined by

$$\begin{aligned}
&f_{1,u}(i) := \sigma_{(i-1)/n}, \quad f_{2,u}(i) := n^{-1/4} \sigma_{u-}, \\
&f_{3,u}(i) := (1/\sqrt{2}) q_n^{-1/2}(i) \left[\sigma_{u-} - \sigma_{(i-1)/n} \right], \quad f_{4,u}(i) := (1/\sqrt{2}) q_n^{1/2}(i) \left[\sigma_{u-} + \sigma_{(i-1)/n} \right],
\end{aligned}$$

for each $i = 1, \ldots, n$ and $j = 1, \ldots, 4$. By a similar argument, a lower bound for $(\Delta_i^n Y)^2$ is obtained as

$$(\Delta_i^n Y)^2 \geq M_1^2(i) - \sum_{j=2}^4 M_j^2(i) - Cn^{-2} \left(1 + n^{1/2} \right), \tag{24}$$

for $i = 1, \ldots, n$. We use (22) and (24) to find upper and lower bounds for $\frac{1}{n} \Lambda_n(\lambda n)$. We will show that both the bounds for $\frac{1}{n} \Lambda_n(\lambda n)$ have the same limit given in (19) for $\lambda \in \left(-\infty, \bar{\lambda} \right)$ and the lower bound of $\frac{1}{n} \Lambda_n(\lambda n)$ takes ∞ for $\lambda \in \left(\bar{\lambda}, \infty \right)$, which is sufficient for the conclusion. We split into two cases: (I) $\lambda \geq 0$ and (II) $\lambda < 0$. Recall that for each n, the domain of Λ_n is the whole real line $\mathbb{R}$.

Case (I): $\lambda \geq 0$**.** Note that for any $\lambda \geq 0$,

$$\begin{aligned}
\frac{1}{n}\Lambda_n(\lambda n) &= \frac{1}{n}\log E\left[\exp\left\{\lambda n \sum_{i=1}^{n} (\Delta_i^n Y)^2\right\}\right] \\
&\leq \frac{1}{n}\log E\left[\prod_{i=1}^{n} \exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i) + \lambda C n^{-1}\left(1+n^{1/2}\right)\right\}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \log E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i)\right\}\right] + \lambda C n^{-1}\left(1+n^{1/2}\right), \qquad (25)
\end{aligned}$$

where the inequality follows from (22), and the last equality follows from the independent increments property of W and the independence between $(f_1(i), \ldots, f_4(i))$ and W. Similarly, by using (24),

$$\frac{1}{n}\Lambda_n(\lambda n) \geq \frac{1}{n}\sum_{i=1}^{n} \log E\left[\exp\left\{\lambda n M_1^2(i) - \lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right] - \lambda C n^{-1}\left(1+n^{1/2}\right). \qquad (26)$$

To proceed, we split Case (I) further into two sub-cases: (I-A) $0 \leq \lambda < \bar{\lambda}$ and (I-B) $\bar{\lambda} < \lambda$.

Case (I-A): $0 \leq \lambda < \bar{\lambda}$**.** Pick any $\lambda \in [0, \bar{\lambda})$. Let

$$\begin{aligned}
g_n(i) := n^{-1/2} &\sup_{u\in((i-1)/n, i/n]} \sigma_{u-}^2 + 2q_n^{-1}(i) \sup_{u\in((i-1)/n, i/n]} \left|\sigma_{u-} - \sigma_{(i-1)/n}\right|^2 \\
&+ 2q_n(i) \sup_{u\in((i-1)/n, i/n]} \left|\sigma_{u-} + \sigma_{(i-1)/n}\right|^2. \qquad (27)
\end{aligned}$$

Note that from the definition of $q_n(i)$ in (21),

$$0 \leq \max_{1\leq i\leq n} g_n(i) \leq C_1\left[n^{-1/2} + \max_{1\leq i\leq n} \sup_{u\in((i-1)/n, i/n]} \left|\sigma_{u-} - \sigma_{(i-1)/n}\right|\right] \to 0, \qquad (28)$$

for some $C_1 > 0$, where the convergence follows from the left-continuity of $\{\sigma_{u-}\}$. By applying (i-a) of Lemma A.2 to (25),

$$\frac{1}{n}\Lambda_n(\lambda n) \leq -\frac{1}{2n}\sum_{i=1}^{n} \log\left(1 - 2\lambda\left(\sigma_{(i-1)/n}^2 + g_n(i)\right)\right) + \lambda C n^{-1}\left(1+n^{1/2}\right) \to \Lambda(\lambda), \qquad (29)$$

for all n large enough, where the convergence follows from the càdlàg condition of σ and (28). By applying the Hölder inequality for negative exponents (Theorem 2.2 of Cheung, 2001) to (26),

$$\begin{aligned}
&\frac{1}{n}\Lambda_n(\lambda n) \geq \frac{2}{n}\sum_{i=1}^{n} \log E\left[\exp\left\{\lambda n M_1^2(i)\right\}\right] - \frac{1}{n}\sum_{i=1}^{n} \log E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i)\right\}\right] - \lambda C n^{-1}\left(1+n^{1/2}\right) \\
&\geq -\frac{1}{n}\sum_{i=1}^{n} \log\left[1 - 2\lambda\sigma_{(i-1)/n}^2\right] + \frac{1}{2n}\sum_{i=1}^{n} \log\left[1 - 2\lambda\left(\sigma_{(i-1)/n}^2 + g_n(i)\right)\right] - O\left(\lambda n^{-1/2}\right) \\
&\to \Lambda(\lambda), \qquad (30)
\end{aligned}$$

where the second inequality follows from $M_1(i) \sim N\left(0, n^{-1}\sigma_{(i-1)/n}^2\right)$ (recall the definition of $M_1(i)$ in (23)), and (i-a) of Lemma A.2. Therefore, by (29) and (30), we obtain $\frac{1}{n}\Lambda_n(\lambda n) \to \Lambda(\lambda)$ for all $\lambda \in [0, \bar{\lambda})$.

Case (I-B): $\bar{\lambda} < \lambda < \infty$**.** Pick any $\lambda \in (\bar{\lambda}, \infty)$. By applying the Hölder inequality for negative exponents to (26),

$$\frac{1}{n}\Lambda_n(\lambda n) \geq \frac{1}{n(1-\delta)}\sum_{i=1}^{n}\log E\left[\exp\left\{(1-\delta)\lambda n M_1^2(i)\right\}\right]$$
$$-\frac{\delta}{n(1-\delta)}\sum_{i=1}^{n}\log E\left[\exp\left\{\left(\frac{1-\delta}{\delta}\right)\lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right] - \lambda C n^{-1}\left(1+n^{1/2}\right), \tag{31}$$

for $\delta \in \left(0, 1-\bar{\lambda}/\lambda\right)$ (i.e., $(1-\delta)\lambda > \bar{\lambda}$). Since $M_1(i) \sim N\left(0, n^{-1}\sigma^2_{(i-1)/n}\right)$ and $(1-\delta)\lambda > \bar{\lambda}$, the first term in the right-hand side of (31) takes ∞. On the other hand, the second term satisfies

$$\frac{\delta}{n(1-\delta)}\sum_{i=1}^{n}\log E\left[\exp\left\{\left(\frac{1-\delta}{\delta}\right)\lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right]$$
$$\leq -\frac{\delta}{2n(1-\delta)}\sum_{i=1}^{n}\log\left(1-2\left(\frac{1-\delta}{\delta}\right)\lambda g_n(i)\right) \to 0,$$

where the inequality holds for all n large enough, which follows from (i-b) of Lemma A.2, and the convergence follows from (28). Therefore, we obtain $\frac{1}{n}\Lambda_n(\lambda n) \to \infty$ for all $\lambda \in (\bar{\lambda}, \infty)$, as desired.

Case (II): $-\infty < \lambda < 0$**.** Pick any $\lambda \in (-\infty, 0)$. By an argument similar to that for (25) (but with negative λ) and (ii) of Lemma A.2,

$$\frac{1}{n}\Lambda_n(\lambda n) \geq \frac{1}{n}\sum_{i=1}^{n}\log E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i)\right\}\right] + \lambda C n^{-1}\left(1+n^{1/2}\right)$$
$$\geq \frac{1}{n}\sum_{i=1}^{n}\log\left(\left[1-2\lambda\sigma^2_{(i-1)/n}\right]^{-1/2} - \left[1+2\lambda g_n(i)\right]^{-1/2}+1\right) + O\left(\lambda n^{-1/2}\right)$$
$$\to \Lambda(\lambda),$$

where the convergence follows from the càdlàg condition of σ and (28). Also, by an argument similar to that for (26) (but with negative λ) and (ii) of Lemma A.2,

$$\frac{1}{n}\Lambda_n(\lambda n) \leq \frac{1}{n}\sum_{i=1}^{n}\log E\left[\exp\left\{\lambda n M_1^2(i) - \lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right] + \lambda C n^{-1}\left(1+n^{1/2}\right)$$
$$\leq \frac{1}{n}\sum_{i=1}^{n}\log\left(\left[1-2\lambda\sigma^2_{(i-1)/n}\right]^{-1/2} + \left[1+2\lambda g_n(i)\right]^{-1/2}-1\right) + O\left(\lambda n^{-1/2}\right)$$
$$\to \Lambda(\lambda).$$

Therefore, we obtain $\frac{1}{n}\Lambda_n(\lambda n) \to \Lambda(\lambda)$ for all $\lambda \in (-\infty, 0)$. The proof of Lemma A.1 is completed by showing the following lemma.

Lemma A.2. *Suppose that Assumption holds. Let $M_j(i)$ and $g_n(i)$ be a random variable defined in (23) (for $i = 1, \ldots, n$ and $j = 1, \ldots, 4$), and let $g_n(i)$ be a variable defined in (27) (for $i = 1, \ldots, n$).*
(i-a) For each $\lambda \in [0, \bar{\lambda})$, it holds that for all n large enough,

$$E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i)\right\}\right] \leq \left[1-2\lambda\left(\sigma^2_{(i-1)/n} + g_n(i)\right)\right]^{-1/2}, \tag{32}$$

for all $i = 1, \ldots, n$.
(i-b) For each $\lambda \in (\bar{\lambda}, \infty)$ *and each* $\delta \in (0,1)$*, it holds that for all* n *large enough,*

$$E\left[\exp\left\{[(1-\delta)/\delta]\,\lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right] \le [1 - 2[(1-\delta)/\delta]\,\lambda g_n(i)]^{-1/2}, \tag{33}$$

for all $i = 1, \ldots, n$.
(ii) For each $\lambda \in (-\infty, 0)$*, it holds that for all* n *large enough,*

$$E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^{4} M_j^2(i)\right\}\right] \ge \left[1 - 2\lambda\sigma_{(i-1)/n}^2\right]^{-1/2} - [1 + 2\lambda g_n(i)]^{-1/2} + 1, \tag{34}$$

$$E\left[\exp\left\{\lambda n M_1^2(i) - \lambda n \sum\nolimits_{j=2}^{4} M_j^2(i)\right\}\right] \le \left[1 - 2\lambda\sigma_{(i-1)/n}^2\right]^{-1/2} + [1 + 2\lambda g_n(i)]^{-1/2} - 1, \tag{35}$$

for all $i = 1, \ldots, n$.

Proof. First, we derive some moment bounds of $\sum_{j=1}^{4} M_j^2(i)$. Without loss of generality, we can set $i = 1$. For each $j = 1, \ldots, 4$, define a stochastic process:

$$\mathcal{M}_{j,t} := \int_0^t f_{j,u}(1)\, dW_u,$$

for $t \in [0, 1/n]$. By Ito's formula for continuous semimartingales, it holds that for each $j = 1, \ldots, 4$,

$$\mathcal{M}_{j,t}^p = \int_0^t p\mathcal{M}_{j,s}^{p-1} d\mathcal{M}_{j,s} + \frac{1}{2}\int_0^t p(p-1)\,\mathcal{M}_{j,s}^{p-2} f_{j,u}(1)\, ds \quad \text{for any } p \ge 2.$$

Consequently, for any even integer $p \ge 2$,

$$E\left[\mathcal{M}_{j,t}^p\right] \le \frac{p(p-1)}{2}\mathfrak{g}_j \int_0^t E\left[\mathcal{M}_{j,s}^{p-2}\right] ds, \tag{36}$$

where $\mathfrak{g}_j := \sup_{s\in(0,1/n]} f_{j,s}^2(1)$, and the inequality holds since $f_{j,s}^2 \le \mathfrak{g}_j$ for almost every $s \in [0, 1/n]$ (with respect to the Lebesgue measure). Recall that in this paper the expectation $E[\cdot]$ means the conditional expectation given the realized path of the volatility process σ and that the realized path of σ is uniformly bounded over $[0,1]$. Therefore, the property of the Brownian motion W guarantees that the (conditional) expectation $E\left[\mathcal{M}_{j,t}^p\right]$ exists for any $p \ge 2$. By the same token, we can guarantee that if $p = 2$, then $E\left[\mathcal{M}_{j,u}^2\right] \le \mathfrak{g}_j u$ for each $j = 1, \ldots, 4$. By using (36) repeatedly, we have

$$\begin{aligned}
E\left[\mathcal{M}_{j,v}^4\right] &\le \frac{4\cdot 3}{2}\mathfrak{g}_j \int_0^v (\mathfrak{g}_j u)\, du = \frac{4\cdot 3}{2}\frac{(\mathfrak{g}_j v)^2}{2}, \\
E\left[\mathcal{M}_{j,w}^6\right] &\le \frac{6\cdot 5}{2}\mathfrak{g}_j \int_0^w \left[\frac{4\cdot 3}{2}\frac{(\mathfrak{g}_j v)^2}{2}\right] dv = \frac{6\cdot 5\cdot 4\cdot 3}{2^2}\frac{(\mathfrak{g}_j w)^3}{2\cdot 3}, \\
&\vdots \\
E\left[\mathcal{M}_{j,\tau}^p\right] &\le \frac{p!}{2^{p/2}(p/2)!}(\mathfrak{g}_j\tau)^{p/2} \quad \text{for any even integer } p \ge 2,
\end{aligned}$$

for each $j = 1, \ldots, 4$, which leads to

$$E\left[\mathcal{M}^p_{j,1/n}\right] \leq \frac{p!}{2^{p/2}(p/2)!}\mathfrak{g}_j^{p/2} n^{-p/2} \quad \text{for any even integer } p \geq 2, \tag{37}$$

for each $j = 1, \ldots, 4$. Now, for any non-negative integer k,

$$\begin{aligned}
&E\left[\left(\sum\nolimits_{j=1}^4 \mathcal{M}^2_{j,1/n}\right)^k\right] \\
&= \sum_{\substack{0\leq l_1,l_2,l_3,l_4\leq k\\ l_1+l_2+l_3+l_4=k}} \begin{pmatrix} k \\ l_1, l_2, l_3, l_4 \end{pmatrix} E\left[\mathcal{M}^{2l_1}_{1,1/n}\mathcal{M}^{2l_2}_{2,1/n}\mathcal{M}^{2l_3}_{3,1/n}\mathcal{M}^{2l_4}_{4,1/n}\right] \\
&\leq \sum_{\substack{0\leq l_1,l_2,l_3,l_4\leq k\\ l_1+l_2+l_3+l_4=k}} \begin{pmatrix} k \\ l_1, l_2, l_3, l_4 \end{pmatrix} E\left[\mathcal{M}^{2k}_{1,1/n}\right]^{l_1/k} E\left[\mathcal{M}^{2k}_{2,1/n}\right]^{l_2/k} E\left[\mathcal{M}^{2k}_{3,1/n}\right]^{l_3/k} E\left[\mathcal{M}^{2k}_{J,1/n}\right]^{l_4/k} \\
&= \left(\sum\nolimits_{j=1}^4 E\left[\mathcal{M}^{2k}_{j,1/n}\right]^{1/k}\right)^k \leq \frac{(2k)!}{2^k k!}\left(\sum\nolimits_{j=1}^4 \mathfrak{g}_j\right)^k n^{-k},
\end{aligned} \tag{38}$$

where the equalities follow from the multinomial theorem with $\begin{pmatrix} k \\ l_1, l_2, l_3, l_4 \end{pmatrix} := \frac{k!}{l_1!l_2!l_3!l_4!}$, the first inequality holds by the generalized Hölder inequality (see, e.g., Finner, 1992), and the last inequality holds by (37) with $p = 2k$. Note that $M_j(1) = \mathcal{M}_{j,1/n}$, $\sigma_0^2 = \mathfrak{g}_1$, and $g_n(1) = \mathfrak{g}_2 + \mathfrak{g}_3 + \mathfrak{g}_4$, where $g_n(i)$ is defined in (27). Then, by (38), we have

$$E\left[\left(\sum\nolimits_{j=1}^4 M_j^2(i)\right)^k\right] \leq \frac{(2k)!}{2^k k!}\left(\sigma^2_{(i-1)/n} + g_n(i)\right)^k n^{-k}, \tag{39}$$

for $i = 1$. By the same argument, we can also show that (39) holds for any $i\,(= 1, \ldots, n)$.

Proof of (i-a). Pick any $\lambda \in [0, \bar{\lambda})$. It holds that for each $i = 1, \ldots, n$,

$$\begin{aligned}
E\left[\exp\left\{\lambda n \sum\nolimits_{j=1}^4 M_j^2(i)\right\}\right] &= \sum_{k=0}^{\infty}\frac{(\lambda n)^k}{k!}E\left[\left(\sum\nolimits_{j=1}^4 M_j^2(i)\right)^k\right] \\
&\leq \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\frac{(2k)!}{2^k k!}\left(\sigma^2_{(i-1)/n} + g_n(i)\right)^k \\
&= E\left[\exp\left\{\lambda\left(\sigma^2_{(i-1)/n} + g_n(i)\right)Z^2\right\}\right] \\
&= \left[1 - 2\lambda\left(\sigma^2_{(i-1)/n} + g_n(i)\right)\right]^{-1/2},
\end{aligned} \tag{40}$$

where $Z \sim N(0,1)$, the first equality holds by the Taylor expansion and the monotone convergence theorem, the inequality holds by (39), and the last two equalities holds by the fact that if $Y \sim N(0,\theta)$, then

$$\begin{aligned}
E\left[Y^{2k}\right] &= \frac{(2k)!}{2^k k!}\theta^k \quad \text{for any non-negative integer } k, \\
E\left[\exp\left\{\eta Y^2\right\}\right] &= [1 - 2\eta\theta]^{-1/2} \quad \text{for any } \eta < 1/2\theta,
\end{aligned} \tag{41}$$

and that

$$1-2\lambda\left(\sigma^2_{(i-1)/n}+g_n(i)\right)>0 \text{ for all sufficiently large } n \text{ uniformly over } i. \tag{42}$$

Note that (42) holds since $\lambda\in\left[0,\bar{\lambda}\right)$ and (28). Now, (40) leads to the desired result (32).

Proof of (i-b). Pick any $\lambda\in\left(\bar{\lambda},\infty\right)$ and any $\delta\in(0,1)$. By the same argument to derive (39) (without $M_1^2(i)$), we can see that

$$E\left[\left(\sum\nolimits_{j=2}^4 M_j^2(i)\right)^k\right]\le\frac{(2k)!}{2^k k!}\left(g_n(i)\right)^k n^{-k},$$

for all $i=1,\ldots,n$ and all non-negative integer k. By a similar argument to derive (40),

$$\begin{aligned}
E\left[\exp\left\{\left[(1-\delta)/\delta\right]\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}\right] &= \sum_{k=0}^{\infty}\frac{\left(\left[(1-\delta)/\delta\right]\lambda n\right)^k}{k!}E\left[\left(\sum\nolimits_{j=2}^4 M_j^2(i)\right)^k\right]\\
&\le\sum_{k=0}^{\infty}\frac{\left(\left[(1-\delta)/\delta\right]\lambda\right)^k}{k!}\frac{(2k)!}{2^k k!}\left(g_n(i)\right)^k\\
&= E\left[\exp\left\{\left[(1-\delta)/\delta\right]\lambda g_n(i)Z^2\right\}\right]\\
&= \left[1-2\left[(1-\delta)/\delta\right]\lambda g_n(i)\right]^{-1/2},
\end{aligned}$$

where the last two equalities follow from (41) and $1-2\left[(1-\delta)/\delta\right]\lambda g_n(i)>0$ for all sufficiently large n uniformly over i. Therefore, the conclusion (33) is obtained.

Proof of (ii). Pick any $\lambda\in(-\infty,0)$. By an argument similar to the proof of part (i), we have

$$\begin{aligned}
E\left[\exp\left\{\lambda n\sum\nolimits_{j=1}^4 M_j^2(i)\right\}\right] &= E\left[\exp\left\{\lambda n M_1^2(i)\right\}+\exp\left\{\lambda n M_1^2(i)\right\}\left[\exp\left\{\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}-1\right]\right]\\
&\ge E\left[\exp\left\{\lambda n M_1^2(i)\right\}+\exp\left\{\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}-1\right]\\
&\ge E\left[\exp\left\{\lambda n M_1^2(i)\right\}-\exp\left\{-\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}+1\right]\\
&\ge\left[1-2\lambda\sigma^2_{(i-1)/n}\right]^{-1/2}-\left[1+2\lambda g_n(i)\right]^{-1/2}+1,
\end{aligned}\tag{43}$$

for each $i=1,\ldots,n$, which implies the desired result (34). The first inequality for (43) follows from $\exp\left\{\lambda n M_1^2(i)\right\}\le 1$ and $\exp\left\{\lambda n\sum_{j=2}^4 M_j^2(i)\right\}-1\le 0$, the second inequality follows from $\exp\{x\}+\exp\{-x\}\ge 2$ for $x\in\mathbb{R}$, and the last inequality follows from $M_1(i)\sim N\left(0,n^{-1}\sigma^2_{(i-1)/n}\right)$ and

$$E\left[\exp\left\{|\lambda|\,n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}\right]\le\left[1-2|\lambda|\,g_n(i)\right]^{-1/2}. \tag{44}$$

Note that (44) holds by the same argument as in deriving (40). Similarly, we have

$$\begin{aligned}
&E\left[\exp\left\{\lambda n M_1^2(i)-\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}\right]\\
&= E\left[\exp\left\{\lambda n M_1^2(i)\right\}+\exp\left\{\lambda n M_1^2(i)\right\}\left[\exp\left\{-\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}-1\right]\right]\\
&\le E\left[\exp\left\{\lambda n M_1^2(i)\right\}+\exp\left\{-\lambda n\sum\nolimits_{j=2}^4 M_j^2(i)\right\}-1\right]\\
&\le\left[1-2\lambda\sigma^2_{(i-1)/n}\right]^{-1/2}+\left[1+2\lambda g_n(i)\right]^{-1/2}-1,
\end{aligned}$$

which implies the desired result (35). Now, we have completed the proof.

A.2 Proof of Theorem 1

Proof of the upper bound. First, we show the upper bound (3) for the case where the set F is compact. Pick any compact interval $[x_L, x_U] \subset \mathbb{R}$ with $-\infty < x_L \leq x_U < \infty$ and a constant $\delta > 0$. Denote

$$\Lambda_\delta^*(x) := \min\left\{\Lambda^*(x) - \delta, \delta^{-1}\right\} = \min\left\{\sup_{\lambda\in(-\infty,\bar{\lambda})}\{\lambda x - \Lambda(\lambda)\} - \delta, \delta^{-1}\right\}.$$

Pick any point $x \in [x_L, x_U]$. From the continuity of $\Lambda(\lambda)$ on $\left(-\infty, \bar{\lambda}\right)$, there exists a point $\lambda_x \in \left(-\infty, \bar{\lambda}\right)$ such that

$$\lambda_x x - \Lambda(\lambda_x) \geq \Lambda_\delta^*(x). \tag{45}$$

Also there exists a neighborhood $B_x := \{y : |y - x| < r_x\}$ with center x and radius $r_x > 0$ such that $\delta \geq r_x |\lambda_x|$. For each $\lambda \in \mathbb{R}$,

$$\begin{aligned}\Pr\{RV_n \in B_x\} &\leq \Pr\left\{\lambda(RV_n - x) \geq \inf_{y\in B_x}\lambda(y-x)\right\}\\ &\leq E\left[\exp\{\lambda(RV_n - x)\}\right]\exp\left\{-\inf_{y\in B_x}\lambda(y-x)\right\},\end{aligned}$$

where the first inequality follows from the fact that the inequality $|RV_n - x| < r_x$ implies $\lambda(RV_n - x) \geq -|\lambda| r_x = \inf_{y\in B_x}\{\lambda(y-x)\}$, and the second inequality follows from the Markov inequality. By setting $\lambda = \lambda_x n$,

$$\frac{1}{n}\log\Pr\{RV_n \in B_x\} \leq -\inf_{y\in B_x}\lambda_x(y-x) - \left\{\lambda_x x - \frac{1}{n}\Lambda_n(\lambda_x n)\right\} \leq \delta - \left\{\lambda_x x - \frac{1}{n}\Lambda_n(\lambda_x n)\right\}, \tag{46}$$

where the second inequality holds by the definition of B_x. Since $[x_L, x_U]$ is compact, there exists a finite covering $\left\{B_{x_j}\right\}_{j=1}^J$ such that $[x_L, x_U] \subset \cup_{j=1}^J B_{x_j}$ and we have

$$\begin{aligned}\frac{1}{n}\log\Pr\{RV_n \in [x_L, x_U]\} &\leq \frac{1}{n}\log\left(J\max_{1\leq j\leq J}\Pr\left\{RV_n \in B_{x_j}\right\}\right)\\ &\leq \frac{1}{n}\log J + \delta - \min_{1\leq j\leq J}\left\{\lambda_{x_j}x_j - \frac{1}{n}\Lambda_n\left(\lambda_{x_j}n\right)\right\},\end{aligned}$$

where the first inequality follows from the set inclusion relation, and the second inequality follows from (46). Thus, Lemma A.1 implies

$$\limsup_{n\to\infty}\frac{1}{n}\log\Pr\{RV_n \in [x_L, x_U]\} \leq \delta - \min_{1\leq j\leq J}\left\{\lambda_{x_j}x_j - \Lambda\left(\lambda_{x_j}\right)\right\} \leq \delta - \inf_{x\in[x_L,x_U]}\Lambda_\delta^*(x).$$

Then, by letting $\delta \to 0$, we obtain the upper bound for the compact case:

$$\limsup_{n\to\infty}\frac{1}{n}\log\Pr\{RV_n \in [x_L, x_U]\} \leq -\inf_{x\in[x_L,x_U]}\Lambda^*(x),$$

where we used the fact that $\lim_{\delta\to 0}\inf_{x\in[x_L,x_U]}\Lambda_\delta^*(x) = \inf_{x\in[x_L,x_U]}\Lambda^*(x)$ (see page 6 of Dembo and Zeitouni, 1998).

Next, we show the exponential tightness of the sequence of probability measures of RV_n, i.e., for each $\epsilon \in (0,\infty)$, there exists a compact set $K_\epsilon \subset \mathbb{R}$ such that $\limsup_{n\to\infty} \frac{1}{n} \log \Pr\{RV_n \in K_\epsilon\} \le -\epsilon$. Since $\Lambda(\lambda) < \infty$ for any $\lambda \in \left(-\infty, \bar{\lambda}\right)$ and $\bar{\lambda} > 0$, there exist $\lambda_1 \in (0,\infty)$ and $\lambda_2 \in \left(0, \bar{\lambda}\right)$ such that $\Lambda(-\lambda_1) < \infty$ and $\Lambda(\lambda_2) < \infty$. For each $r > 0$, the Markov inequality and Lemma A.1 yield

$$\lim_{r\to\infty} \limsup_{n\to 0} \frac{1}{n} \log \Pr\{RV_n \le -r\} \le \lim_{r\to\infty} (-\lambda_1 r) + \Lambda(-\lambda_1) = -\infty,$$
$$\lim_{r\to\infty} \limsup_{n\to 0} \frac{1}{n} \log \Pr\{RV_n \ge r\} \le \lim_{r\to\infty} (-\lambda_2 r) + \Lambda(\lambda_2) = -\infty,$$

which implies the exponential tightness of the sequence of measures of RV_n.

Finally, combining the upper bound for the compact case and the exponential tightness of the sequence of measures of RV_n, Lemma 1.2.18 (a) in Dembo and Zeitouni (1998) implies that the upper bound holds for any closed set.

Proof of the lower bound. We now show the lower bound (4). Define

$$\Lambda\left(\bar{\lambda}\right) := \lim_{\lambda \nearrow \bar{\lambda}} \Lambda(\lambda).$$

If $\Lambda\left(\bar{\lambda}\right) = \infty$, Gärtner-Ellis' theorem implies the conclusion. Thus, we hereafter suppose that $\Lambda\left(\bar{\lambda}\right) < \infty$. Let

$$s(\lambda) := \int_{[0,1]} \frac{\sigma_{u-}^2}{1 - 2\lambda\sigma_{u-}^2} du \quad \text{for } \lambda \in \left(-\infty, \bar{\lambda}\right), \quad \text{and} \quad \bar{x} := \lim_{\lambda \nearrow \bar{\lambda}} s(\lambda).$$

Note that $s(\lambda) = (d/d\lambda)\Lambda(\lambda)$ for any $\lambda \in \left(-\infty, \bar{\lambda}\right)$, which can be shown by using the dominated convergence theorem. If $\bar{x} = \infty$, Gärtner-Ellis' theorem implies the conclusion. Therefore, we focus on the case where $\Lambda\left(\bar{\lambda}\right) < \infty$ and $\bar{x} < \infty$. It is sufficient to show that for any open set $G \subset \mathbb{R}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \Pr\{RV_n \in G \cap (-\infty, \bar{x})\} \ge - \inf_{x \in G \cap (-\infty, \bar{x})} \Lambda^*(x), \tag{47}$$
$$\liminf_{n\to\infty} \frac{1}{n} \log \Pr\{RV_n \in G \cap (\bar{x}, \infty)\} \ge - \inf_{x \in G \cap (\bar{x}, \infty)} \Lambda^*(x). \tag{48}$$

Proof of (47). Here, we follow the same steps of the proof of the Gärtner-Ellis theorem in Dembo and Zeitouni (1998, pp. 49-51). Pick any $y \in (0, \bar{x})$, and let $B_{y,\delta}$ be a neighborhood around $y \in (0, \bar{x})$ with radius $\delta > 0$. We first show that

$$\lim_{\delta\to 0} \liminf_{n\to 0} \frac{1}{n} \log \Pr\{RV_n \in B_{y,\delta}\} \ge -\Lambda^*(y). \tag{49}$$

Note that $s(\lambda)$ is strictly increasing in $\lambda \in \left(-\infty, \bar{\lambda}\right)$. Then, since $y \in (0, \bar{x})$, $\lim_{\lambda\to-\infty} s(\lambda) = 0$ and $\lim_{\lambda \nearrow \bar{\lambda}} s(\lambda) = \bar{x}$, there exists a unique solution $\lambda_y \in \left(-\infty, \bar{\lambda}\right)$ to the equation $y = s(\lambda)$. Since $-\infty < \lambda_y < \bar{\lambda}$, it holds that $\Lambda_n(\lambda_y n) < \infty$ for all n large enough by Lemma A.1. Thus, letting μ_n be the probability measure of RV_n given the path of σ, we can define the probability measure $\tilde{\mu}_n$ with the Radon-Nikodym derivative

$$\frac{d\tilde{\mu}_n}{d\mu_n}(z) = \exp\{\lambda_y n z - \Lambda_n(\lambda_y n)\}.$$

Observe that

$$
\begin{aligned}
\lim_{\delta\to0}\liminf_{n\to0}\frac{1}{n}\log\Pr\{RV_n\in B_{y,\delta}\} &= \lim_{\delta\to0}\liminf_{n\to\infty}\frac{1}{n}\log\int_{z\in B_{y,\delta}}\exp\{-\lambda_y nz+\Lambda_n(\lambda_y n)\}\,d\tilde{\mu}_n(z)\\
&\geq -\lambda_y y+\Lambda(\lambda_y)+\lim_{\delta\to0}\liminf_{n\to\infty}\frac{1}{n}\log\int_{z\in B_{y,\delta}}\exp\{-|\lambda_y|\,n\delta\}\,d\tilde{\mu}_n(z)\\
&\geq -\Lambda^*(y)+\lim_{\delta\to0}\liminf_{n\to\infty}\frac{1}{n}\log\tilde{\mu}_n(B_{y,\delta}),
\end{aligned}
$$

where the equality follows from the change of measures, the first inequality follows from Lemma A.1, and the second inequality follows from the definition of Λ^*. Therefore, we can establish (49) if it holds that

$$\lim_{\delta\to0}\liminf_{n\to\infty}\frac{1}{n}\log\tilde{\mu}_n(B_{y,\delta})=0, \tag{50}$$

which we below show. Let $\tilde{\Lambda}(\lambda):=\Lambda(\lambda+\lambda_y)-\Lambda(\lambda_y)$ for $\lambda\in\left(-\infty,\bar{\lambda}-\lambda_y\right)$, and

$$\tilde{\Lambda}^*(x):=\sup_{\lambda\in\left(-\infty,\bar{\lambda}-\lambda_y\right)}\left\{\lambda x-\tilde{\Lambda}(\lambda)\right\}=\Lambda^*(x)-\lambda_y x+\Lambda(\lambda_y), \tag{51}$$

for $x\in\mathbb{R}$. By the change of measures and Lemma A.1,

$$\frac{1}{n}\tilde{\Lambda}_n(\lambda n):=\frac{1}{n}\log\int_{\mathbb{R}}\exp\{\lambda nz\}\,d\tilde{\mu}_n(z)=\frac{1}{n}\Lambda_n((\lambda+\lambda_y)n)-\frac{1}{n}\Lambda_n(\lambda_y n)\to\tilde{\Lambda}(\lambda), \tag{52}$$

for each $\lambda\in\left(-\infty,\bar{\lambda}-\lambda_y\right)$. We here apply the same argument as for the proof of the upper bound to $\tilde{\mu}_n$. Then, by (52), we have the large deviation probability of a closed set $B^c_{y,\delta}$, the complement of $B_{y,\delta}$, bounded as

$$\limsup_{n\to0}\frac{1}{n}\log\tilde{\mu}_n(B^c_{y,\delta})\leq-\inf_{x\in B^c_{y,\delta}}\tilde{\Lambda}^*(x)\quad\text{for each }\delta>0.$$

Note that $\tilde{\Lambda}^*$ is lower semicontinuous and level compact, which follows from the goodness of Λ^* (i.e., Λ^* is lower semicontiniuous and level compact). Then, by the (generalized) Weierstrass Theorem, there exists a point $x_0\in B^c_{y,\delta}$ such that for each $\delta>0$,

$$\inf_{x\in B^c_{y,\delta}}\tilde{\Lambda}^*(x)=\tilde{\Lambda}^*(x_0)=\sup_{\lambda\in\left(-\infty,\bar{\lambda}\right)}\{\lambda x_0-\Lambda(\lambda)\}-\{\lambda_y x_0-\Lambda(\lambda_y)\}=:l_0,$$

where the second equality follows from (51) and the definition of $\Lambda^*(x)$. Note that $\Lambda(\lambda)$ is strictly increasing and convex. Therefore, (a) if $x_0\leq0$, then $\lambda x_0-\Lambda(\lambda)$ is strictly decreasing in λ (and $\lim_{\lambda\to-\infty}\Lambda(\lambda)=-\infty$), thus $l_0=\infty$; (b) if $0<x_0$, then there exists a unique solution $\lambda_{x_0}\,(\neq\lambda_y)$ to $x_0=s(\lambda)$ satisfying $l_0=\{\lambda_{x_0}x_0-\Lambda(\lambda_{x_0})\}-\{\lambda_y x_0-\Lambda(\lambda_y)\}>0$. Now, we have shown $l_0>0$ and thus $\tilde{\mu}_n(B_{y,\delta})\to1$ for all $\delta>0$. This implies the desired result (50).

We are now prepared to prove (47). Pick any open set $G\subset\mathbb{R}$. For each $y\in G\cap(-\infty,\bar{x})$, we can take a neighborhood $B_{y,\delta}\subset G$ for all δ small enough, and (49) implies

$$\liminf_{n\to\infty}\frac{1}{n}\log\Pr\{RV_n\in G\cap(-\infty,\bar{x})\}\geq-\Lambda^*(y),$$

for all $y \in G \cap (-\infty, \bar{x})$. Therefore, the desired result (47) follows.

Proof of (48). Let $i_n^* := \arg\max_{1 \le i \le n} \sigma^2_{(i-1)/n}$ and $RV_n^- := RV_n - \left(\Delta^n_{i_n^*} Y\right)^2$. Pick any $y \in (\bar{x}, \infty)$ and $\epsilon > 0$. Since RV_n^- and $\left(\Delta^n_{i_n^*} Y\right)^2$ are independent,

$$\Pr\{RV_n > y\} \ge \Pr\left\{RV_n^- > \bar{x} - \epsilon\right\} \Pr\left\{\left(\Delta^n_{i_n^*} Y\right)^2 > y - \bar{x} + \epsilon\right\}, \tag{53}$$

for all $n \in \mathbb{N}$. Observe that

$$\liminf_{n\to\infty} \frac{1}{n} \log \Pr\left\{RV_n^- > \bar{x} - \epsilon\right\} \ge \liminf_{n\to\infty} \frac{1}{n} \log \Pr\left\{RV_n^- \in (\bar{x} - \epsilon, \bar{x})\right\} \ge -\inf_{x \in (\bar{x}-\epsilon, \bar{x})} \Lambda^*(x), \tag{54}$$

where the second inequality follows from an argument analogous to the proof of (47) (replace RV_n and G with RV_n^- and $(\bar{x} - \epsilon, \infty)$, respectively). Note also that

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{1}{n} \log \Pr\left\{\left(\Delta^n_{i_n^*} Y\right)^2 > y - \bar{x} + \epsilon\right\} \\
&\ge \liminf_{n\to\infty} \frac{1}{n} \log \Pr\left\{n^{-1}\left(\max_{1\le i\le n} \sigma^2_{(i-1)/n} - g_n\left(i_n^*\right)\right) Z^2 > y - \bar{x} - \epsilon\right\} \\
&\ge -\bar{\lambda}\left(y - \bar{x} + \epsilon\right),
\end{aligned} \tag{55}$$

where $Z \sim N(0,1)$, the first inequality follows from (22) and $\int_{(i_n^*-1)/n}^{i_n^*/n} dW_u \sim N\left(0, n^{-1}\right)$, and the second inequality follows from Bercu, Gamboa and Rouault (1997, Lemma 6) with $\max_{1\le i\le n} \sigma^2_{(i-1)/n} \to 1/\left(2\bar{\lambda}\right) = \sup_{u\in[0,1]} \sigma^2_{u-}$ (by (i) the right continuity of σ, and (ii) the fact that the number of jumps larger than δ is finite for any $\delta > 0$, both of which follow from the càdlàg condition of σ) and $g_n\left(i_n^*\right) \to 0$ (by (28)). Combining (53)-(55) and letting $\epsilon \to 0$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \Pr\{RV_n > y\} \ge -\Lambda^*(\bar{x}) - \bar{\lambda}(y - \bar{x}) = -\inf_{x\in(y,\infty)} \Lambda^*(x),$$

for all $y \in (\bar{x}, \infty)$. By applying an argument analogous to that for $\Pr\{RV_n \in G \cap (\bar{x}, \infty)\}$, we obtain the desired result (48). The proof is now completed.

A.3 Proof of Theorem 2

In this case, since the set of exposed points is the whole real line $\mathbb{R}$, it is sufficient to show that the limiting cumulant generating function is lower semicontinuous and essentially smooth based on (a part of) Gärtner-Ellis' theorem (see, Dembo and Zeitouni, 1998, p. 44). Let $T_n := m_n\left\{RV_n - \overline{\sigma^2}\right\}$ and $s_n := m_n^2/n$. In particular, it is sufficient to show that for each $\lambda \in \mathbb{R}$,

$$\lim_{n\to\infty} s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right] = \lambda^2 \overline{\sigma^4}. \tag{56}$$

First, we consider the case where $\lambda > 0$. Pick any $\lambda \in (0, \infty)$, and let $g_n(i)$ be as defined in (27) in the proof of Lemma A.1. Then,

$$
\begin{aligned}
s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right] &= \frac{m_n^2}{n}\sum_{i=1}^{n}\log E\left[\exp\left\{\frac{1}{m_n}\lambda n\left(\Delta_i^n Y\right)^2\right\}\right] - m_n\lambda\overline{\sigma^2} \\
&\leq -\frac{m_n^2}{2n}\sum_{i=1}^{n}\log\left[1-\frac{2}{m_n}\lambda\left(\sigma_{(i-1)/n}^2 + g_n(i)\right)\right] - m_n\lambda\overline{\sigma^2} + \frac{m_n}{n}\left(1+n^{1/2}\right)\lambda C \\
&= \lambda m_n\left(\frac{1}{n}\sum_{i=1}^{n}\sigma_{(i-1)/n}^2 - \overline{\sigma^2}\right) + \lambda\frac{m_n}{n}\sum_{i=1}^{n}g_n(i) \\
&+ \lambda^2\frac{1}{n}\sum_{i=1}^{n}\frac{\left(\sigma_{(i-1)/n}^2 + g_n(i)\right)^2}{\left[1-\frac{2}{m_n}\lambda t_{i,n}\left(\sigma_{(i-1)/n}^2 + g_n(i)\right)\right]^2} + \frac{m_n}{n}\left(1+n^{1/2}\right)\lambda C,
\end{aligned}
\tag{57}
$$

for some $t_{i,n} \in [0,1]$ $(i = 1, \ldots, n)$ and all n large enough, where the inequality follows from Lemma A.1 (i) and (25), and the equality in the third line follows from the Taylor expansion of $\log(1-w) = -w - \frac{1}{2[1-t_w w]^2}w^2$ for $|w| < 1$ with $t_w \in [0,1]$. From the first condition in (9), the first term on the right-hand side of (57) converges to zero. From the second condition in (9) combined with (28), the second term on the right-hand side of (57) converges to zero. From (28) and the boundedness of $\sigma_{(i-1)/n}^2$, it holds that $\max_{1\leq i\leq n}\frac{2}{m_n}\lambda t_{i,n}\left(\sigma_{(i-1)/n}^2 + g_n(i)\right) \to 0$ and thus the third term on the right-hand side of (57) converges to $\lambda^2\overline{\sigma^4}$. From the condition $m_n^2/n \to 0$, the fourth term on the right-hand side of (57) converges to zero. Combining these results implies that the upper bound in (57) converges to $\lambda^2\overline{\sigma^4}$. We can also show that the lower bound of $s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right]$ converges to $\lambda^2\overline{\sigma^4}$ in the same manner. For the case where $\lambda < 0$, (56) can be shown analogously.

A.4 Proof of Theorem 3

Since the other part of the proof is similar to that of Theorem 1, it is sufficient to show the following counterpart of Lemma A.1.

Lemma A.3. *Under Assumption', it holds that for each $\lambda \in \left(-\infty, \bar{\lambda}_D\right)$, the limit of $\frac{1}{n}\Lambda_n(\lambda n)$ (defined in (18)) as n tends to ∞ is $\Lambda_D(\lambda) < \infty$, i.e.,*

$$
\lim_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda n) = \Lambda_D(\lambda) < \infty, \tag{58}
$$

and for each $\lambda \in \left(\bar{\lambda}_D, \infty\right)$, the limit of $\frac{1}{n}\Lambda_n(\lambda n)$ as n tends to ∞ is infinity, i.e.,

$$
\lim_{n\to\infty}\frac{1}{n}\Lambda_n(\lambda n) = \infty, \tag{59}
$$

where the threshold value $\bar{\lambda}_D$ and the limit function Λ_D are defined in (14) and (15), respectively.

To simplify the presentation, we provide a proof for the case of $a_u = 0$. Similarly to the proof of Lemma A.1, under the boundedness assumption on a, we can show that the the presence of non-zero drift has no impact on the form of the rate function. Let

$$r_n(i) := \sup_{u\in((i-1)/n,i/n]} \left|\sigma_{u-} - \sigma_{(i-1)/n}\right|, \tag{60}$$

$$L(i) := \rho \int_{(i-1)/n}^{i/n} \sigma_{u-} dZ_u \text{ for each } i \in \{1,\dots,n\}. \tag{61}$$

Note that $r_n(i) \to 0$ uniformly over i as $n \to \infty$. By using (12) (with $a_u = 0$), we can write

$$\begin{aligned}
(\Delta_i^n Y)^2 &= \left(1-\rho^2\right)\left(\int_{(i-1)/n}^{i/n} dW_u\right)^2 \sigma_{(i-1)/n}^2 \\
&\quad + \left(1-\rho^2\right)\left(r_n^{-1/2}(i)\int_{(i-1)/n}^{i/n}\left[\sigma_{u-}-\sigma_{(i-1)/n}\right]dW_u\right)\left(r_n^{1/2}(i)\int_{(i-1)/n}^{i/n}\left[\sigma_{u-}+\sigma_{(i-1)/n}\right]dW_u\right) \\
&\quad + 2\sqrt{1-\rho^2}\left(\int_{(i-1)/n}^{i/n} dW_u\right)\sigma_{(i-1)/n}L(i) \\
&\quad + 2\left(\sqrt{1-\rho^2}\,r_n^{-1/2}(i)\int_{(i-1)/n}^{i/n}\left[\sigma_{u-}-\sigma_{(i-1)/n}\right]dW_u\right)r_n^{1/2}(i)\,L(i) + L^2(i).
\end{aligned} \tag{62}$$

Now let

$$N_j(i) := \int_{(i-1)/n}^{i/n} \tilde{f}_{1,u}(i)\, dW_u, \tag{63}$$

for $j = 1,2,3$, where $\tilde{f}_{j,u}(i)$ is a stochastic process on $[(i-1)/n, i/n]$ defined by

$$\begin{aligned}
&\tilde{f}_{1,u}(i) := \sqrt{1-\rho^2}\sigma_{(i-1)/n}, \quad \tilde{f}_{2,u}(i) := 2^{-1/2} r_n^{-1/2}(i)\sqrt{1-\rho^2}\left[\sigma_{u-}-\sigma_{(i-1)/n}\right], \\
&\tilde{f}_{3,u}(i) := 2^{-1/2} r_n^{1/2}(i)\sqrt{1-\rho^2}\left[\sigma_{u-}+\sigma_{(i-1)/n}\right].
\end{aligned}$$

By similar arguments to derive (22) and (24) (we apply the inequality $AB \le A^2/2 + B^2/2$ for $A, B \in \mathbb{R}$ to the second and fourth terms in the right-hand side of (62)), we obtain

$$(\Delta_i^n Y)^2 \le N_1^2(i) + 2N_1(i)L(i) + 3N_2^2(i) + N_3^2(i) + [1 + r_n(i)]L^2(i), \tag{64}$$

$$(\Delta_i^n Y)^2 \ge N_1^2(i) + 2N_1(i)L(i) - 3N_2^2(i) - N_3^2(i) + [1 - r_n(i)]L^2(i). \tag{65}$$

Based on these upper and lower bounds, we can proceed in the same way as in the proof of Lemma A.1, i.e., find upper and lower bounds of the normalized cumulant generating function of $(\Delta_i^n Y)^2$ and then show that these bounds converge to the same limit for each λ. Note that given the path of σ, the terms $r_n(i)$ and $L(i)$ can be treated as given from the filtration condition in Assumption'. The main difference from the proof of Lemma A.1 is that the upper and lower bounds of $(\Delta_i^n Y)^2$ include the term $N_1(i)L(i)$, which is normally distributed for each realized path of σ.

The following notation and results are used later. Let

$$\tilde{g}_n(i) := 3\sup\nolimits_{u\in((i-1)/n,i/n]}\left|\tilde{f}_{2,u}(i)\right|^2 + \sup\nolimits_{u\in((i-1)/n,i/n]}\left|\tilde{f}_{3,u}(i)\right|^2. \tag{66}$$

>From the continuity of σ,

$$0 \leq \max_{1 \leq i \leq n} \tilde{g}_n(i) \leq C \max_{1 \leq i \leq n} \sup_{u \in ((i-1)/n, i/n]} \left| \sigma_{u-} - \sigma_{(i-1)/n} \right| \to 0 \quad (\text{as } n \to \infty), \tag{67}$$

for some $C > 0$. Since $N_1(i) \sim N\left(0, \left(1-\rho^2\right) n^{-1} \sigma^2_{(i-1)/n}\right)$, it holds that for each $\lambda \in \left(-\infty, \bar{\lambda}_D\right)$,

$$\begin{aligned} G_n(i, \lambda) &:= E\left[\exp\left\{\lambda n \left[N_1^2(i) + 2N_1(i) L(i)\right]\right\}\right] \\ &= \frac{1}{\sqrt{1 - 2\lambda\left(1-\rho^2\right)\sigma^2_{(i-1)/n}}} \exp\left\{\frac{2\lambda^2 n \left(1-\rho^2\right)\sigma^2_{(i-1)/n} L^2(i)}{1 - 2\lambda\left(1-\rho^2\right)\sigma^2_{(i-1)/n}}\right\}, \end{aligned} \tag{68}$$

for $i = 1, \ldots, n$.

We now consider three cases to find the limit $\frac{1}{n}\Lambda_n(\lambda n)$: (I-A) $0 \leq \lambda < \bar{\lambda}_D$; (I-B) $\bar{\lambda}_D < \lambda < \infty$; and (II) $\lambda < 0$.

Case (I-A): Pick any $\lambda \in \left[0, \bar{\lambda}_D\right)$. Recall the form of $n^{-1}\Lambda_n(\lambda n)$ in (18). By (64) and Lemma A.4 (i),

$$\begin{aligned} \frac{1}{n}\Lambda_n(\lambda n) &\leq \frac{1}{n} \log\left(E\left[\prod_{i=1}^n \exp\left\{\lambda n \left[N_1^2(i) + 2N_1(i) L(i) + 3N_2^2(i) + N_3^2(i) + \left[1 + r_n(i)\right] L^2(i)\right]\right\}\right]\right) \\ &= \frac{1}{n} \sum_{i=1}^n \log E\left[\exp\left\{\lambda n \left[N_1^2(i) + 2N_1(i) L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right] + \lambda \sum_{i=1}^n \left[1 + r_n(i)\right] L^2(i) \\ &\leq \frac{1}{n} \sum_{i=1}^n \log\left[G_n(i, \lambda) + C\lambda \tilde{g}_n(i)\left[1 - 4l\lambda \tilde{g}_n(i)\right]^{-1/4l}\right] + \lambda\left[1 + \max_{1 \leq i \leq n} r_n(i)\right] \sum_{i=1}^n L^2(i) \\ &\to \Lambda_D(\lambda), \end{aligned} \tag{69}$$

where the convergence follows from the uniform convergence of $\tilde{g}_n(i)$ and $r_n(i)$ (defined in (66) and in (60), respectively), the uniform continuity of σ, and the following fact:

$$\sum_{i=1}^n L^2(i) \to \rho^2 \int_0^1 \sigma^2_{u-} du = \rho^2 \overline{\sigma^2}, \tag{70}$$

whose proof is provided below. For all n large enough, we also have

$$\begin{aligned} \frac{1}{n}\Lambda_n(\lambda n) &\geq \frac{1}{n} \sum_{i=1}^n \log E\left[\exp\left\{\lambda n \left[N_1^2(i) + 2N_1(i) L(i) - 3N_2^2(i) - N_3^2(i)\right]\right\}\right] + \lambda \sum_{i=1}^n \left[1 - r_n(i)\right] L^2(i) \\ &\geq \frac{2}{n} \sum_{i=1}^n \log E\left[\exp\left\{\lambda n \left[N_1^2(i) + N_1(i) L(i)\right]\right\}\right] \\ &\quad - \frac{1}{n} \sum_{i=1}^n \log E\left[\exp\left\{\lambda n \left[N_1^2(i) + 2N_1(i) L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right] + \lambda\left[1 - \max_{1 \leq i \leq n} r_n(i)\right] L^2(i) \\ &\to \Lambda_D(\lambda), \end{aligned} \tag{71}$$

where the first inequality holds by (65), the second inequality follows from the Hölder inequality for negative exponents, and the convergence follows from the same argument to derive (69) combined with (68) and Lemma A.4 (i). Therefore, (69) and (71) imply that $n^{-1}\Lambda_n(\lambda n) \to \Lambda_D(\lambda)$ for all $\lambda \in \left[0, \bar{\lambda}_D\right)$.

Case (I-B): Pick any $\lambda \in \left(\bar{\lambda}_D, \infty\right)$. Analogously to (31), we use (65) and the Hölder inequality for negative exponents, and obtain

$$\begin{aligned}
\frac{1}{n}\Lambda_n(\lambda n) &\geq \frac{1}{n}\sum_{i=1}^{n}\log E\left[\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i) - 3N_2^2(i) - N_3^2(i)\right]\right\}\right] \\
&\geq \frac{1}{n(1-\delta)}\sum_{i=1}^{n}\log E\left[\exp\left\{\lambda n(1-\delta)\left[N_1^2(i) - 2N_1(i)L(i)\right]\right\}\right] \\
&\quad - \frac{\delta}{n(1-\delta)}\sum_{i=1}^{n}\log E\left[\exp\left\{\left[(1-\delta)/\delta\right]\lambda n\left[N_2^2(i) + 3N_3^2(i)\right]\right\}\right], \qquad (72)
\end{aligned}$$

for $\delta \in \left(0, 1 - \bar{\lambda}_D/\lambda\right)$ (i.e., $(1-\delta)\lambda > \bar{\lambda}_D$). Since $N_1(i) \sim N\left(0, n^{-1}\left(1-\rho^2\right)\sigma^2_{(i-1)/n}\right)$ and $(1-\delta)\lambda > \bar{\lambda}_D$, the first term in the right-hand side of (72) takes ∞ for all n large enough. For the second term,

$$\frac{1}{n}\sum_{i=1}^{n}\log E\left[\exp\left\{\left[(1-\delta)/\delta\right]\lambda n\left[3N_2^2(i) + N_3^2(i)\right]\right\}\right] \leq -\frac{1}{2n}\sum_{i=1}^{n}\log\left(1 - 2\left[(1-\delta)/\delta\right]\lambda\tilde{g}_n(i)\right) \to 0,$$

where the inequality follows from the same argument to derive Lemma A.2 (i-b). Therefore, we have $\frac{1}{n}\Lambda_n(\lambda n) \to \infty$ for all $\lambda \in \left(\bar{\lambda}_D, \infty\right)$.

Case (II): $-\infty < \lambda < 0$. The same argument to Case (II) of the proof of Lemma A.1 combined with Lemma A.4 (ii) and (70) yields the conclusion. Thus we omit the proof for this case.

The proof is completed by showing (70). By the definition of $L(i)$ in (61) and Ito's lemma, we have

$$\sum_{i=1}^{n}L^2(i) = \rho^2\sum_{i=1}^{n}\left(\textstyle\int_{(i-1)/n}^{i}\sigma_{u-}dZ_u\right)^2 = \rho^2\sum_{i=1}^{n}\left(\textstyle\int_{(i-1)/n}^{i}\sigma_{u-}^2du + \eta_n(i)\right) = \rho^2\overline{\sigma^2} + \rho^2\sum_{i=1}^{n}\eta_n(i), \quad (73)$$

where

$$\eta_n(i) =: \textstyle\int_{(i-1)/n}^{i}\left(\int_{(i-1)/n}^{s}\sigma_{u-}dZ_u\right)\sigma_{s-}dZ_s. \qquad (74)$$

Let $E_\sigma[\cdot]$ be the expectation with respect to σ (and Z, where we note the filtration condition in Assumption'). The uniform moment bound of $\eta_n(i)$ is obtained as

$$\begin{aligned}
E_\sigma\left[\eta_n(i)^2\right]^{1/2} &= E_\sigma\left[\textstyle\int_{(i-1)/n}^{i}\left(\int_{(i-1)/n}^{s}\sigma_{u-}dZ_u\right)^2\sigma_{s-}^2ds\right]^{1/2} \\
&\leq E_\sigma\left[\left(\max_{s\in[(i-1)/n.i/n]}\left|\textstyle\int_{(i-1)/n}^{s}\sigma_{u-}dZ_u\right|\right)^2\textstyle\int_{(i-1)/n}^{i}\sigma_{s-}^2ds\right]^{1/2} \\
&\leq E_\sigma\left[\left|\max_{s\in[(i-1)/n.i/n]}\textstyle\int_{(i-1)/n}^{s}\sigma_{u-}dZ_u\right|^4\right]^{1/2} + E_\sigma\left[\left|\textstyle\int_{(i-1)/n}^{i}\sigma_{s-}^2ds\right|\right]^{1/2} \\
&\leq CE_\sigma\left[\left|\textstyle\int_{(i-1)/n}^{i}\sigma_{u-}^2du\right|^2\right]^{1/2} \leq CE_\sigma\left[n^{-1}\textstyle\int_{(i-1)/n}^{i}\sigma_{u-}^4du\right]^{1/2} \leq \tilde{C}n^{-1}, \qquad (75)
\end{aligned}$$

for some $C > 0$, where $\tilde{C} := C\sqrt{\sup_{u\in[0,1]}E_\sigma\left[\sigma_{u-}^4\right]}$ (note that $\tilde{C} < \infty$ by Assumption'), the equality follows from the Ito isometry, the second inequality holds by the Minkowski inequality, the third inequality follows from the Burkholder-Davis-Gundy inequality, and the fourth inequality follows from the Jensen inequality. Thus, for each $c > 0$, it holds that

$$\Pr_\sigma \left\{\left|\sum_{i=1}^n \eta_n(i)\right| > c\right\} \leq c^{-2} E_\sigma \left[\left|\sum_{i=1}^n \eta_n(i)\right|^2\right] = c^{-2} E_\sigma \left[\sum_{i=1}^n \eta_n^2(i)\right] \leq c^{-2} \tilde{C} n^{-1}, \tag{76}$$

where $\Pr_\sigma$ is the probability with respect to σ (and Z), the first inequality follows from the Markov inequality, the equality follows from the martingale property of $\eta_n(i)$, and the second inequality follows from (75). Since (76) holds for any constant $c > 0$, the Borel-Cantelli lemma implies that $\sum_{i=1}^n \eta_n(i) \to 0$ almost surely (with respect to the probability measure of σ and Z). In other words, we have $\sum_{i=1}^n \eta_n(i) \to 0$ for each realized path of σ, which, together with (73), implies the desired result (70).

Lemma A.4. *Suppose that Assumption' holds. For $i = 1, \ldots, n$, let $N_j(i)$ be a random variable defined in (61) ($j = 1, 2, 3$), $L(i)$ and $\tilde{g}_n(i)$ be variables defined in (63) and (66), respectively, and $G_n(i, \lambda)$ be a function of $\lambda \in \left(-\infty, \bar{\lambda}_D\right)$ defined in (68).*
(i) For each $\lambda \in \left[0, \bar{\lambda}_D\right)$, there exist some constant $C > 0$ and some positive integer l (both C and l are independent of i, λ, and n) such that for all n large enough,

$$E\left[\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right] \leq G_n(i, \lambda) + C\lambda \tilde{g}_n(i)\left[1 - 4l\lambda \tilde{g}_n(i)\right]^{-1/4l}, \tag{77}$$

for all $i = 1, \ldots, n$.
(ii) For each $\lambda \in (-\infty, 0)$, it holds that for all n large enough,

$$E\left[\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right] \geq G_n(i, \lambda) - \left[1 + 2\lambda \tilde{g}_n(i)\right]^{-1/2} + 1, \tag{78}$$

$$E\left[\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i) - 3N_2^2(i) - N_3^2(i)\right]\right\}\right] \leq G_n(i, \lambda) + \left[1 + 2\lambda \tilde{g}_n(i)\right]^{-1/2} - 1, \tag{79}$$

for all $i = 1, \ldots, n$.

Proof of (i). Pick any $\lambda \in \left(0, \bar{\lambda}_D\right)$. Observe that

$$\begin{aligned}
&E\left[\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right] \\
&= G_n(i, \lambda) + E\left[\left[\exp\left\{\lambda n\left[N_2^2(i) + 3N_3^2(i)\right]\right\} - 1\right]\exp\left\{\lambda n\left[N_1^2(i) + 2N_1(i)L(i)\right]\right\}\right] \\
&\leq G_n(i, \lambda) + E\left[\left|\exp\left\{\left[\lambda n N_2^2(i) + 3N_3^2(i)\right]\right\} - 1\right|^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} E\left[\exp\left\{\lambda n p\left[N_1^2(i) + 2N_1(i)L(i)\right]\right\}\right]^{1/p},
\end{aligned} \tag{80}$$

where the inequality follows from the Hölder inequality with $p \in \left(1, \bar{\lambda}_D/\lambda\right)$. Now, we derive an upper bound for the second term on the right hand side of (80). First, pick any integer $l \geq \frac{p}{p-1}$. For all n

large enough, it holds that

$$
\begin{aligned}
&E\left[\left|\exp\left\{\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\right\}-1\right|^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \le E\left[\left|\exp\left\{\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\right\}-1\right|^{l}\right]^{1/l} \\
&\le E\left[\left|\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\exp\left\{\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\right\}\right|^{l}\right]^{1/l} \\
&\le \lambda n\left\{E\left[\left|3N_2^2(i)+N_3^2(i)\right|^{2l}\right]E\left[\exp\left\{2l\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\right\}\right]\right\}^{1/2l} \\
&\le \left\{\frac{(4l)!}{2^{4l}(2l)!}\right\}^{1/2l}\lambda\tilde{g}_n(i)\left[1-4l\lambda\tilde{g}_n(i)\right]^{-1/4l}, \qquad (81)
\end{aligned}
$$

for any $i=1,\ldots,n$, where the second inequality follow from the inequality: $\exp(x)-1\le x\exp(x)$ for any $x\ge 0$, the third inequality follows from the Cauchy-Schwarz inequality, and the last inequality uses the following facts: for all n large enough,

$$
\begin{aligned}
4l\lambda\max_{1\le i\le n}\tilde{g}_n(i) &< 1, \\
E\left[\left|3N_2^2(i)+N_3^2(i)\right|^{2l}\right] &\le \left\{\frac{(4l)!}{2^{4l}(2l)!}\right\}\left(\tilde{g}_n(i)n^{-1}\right)^{2l}, \\
E\left[\exp\left\{2l\lambda n\left[3N_2^2(i)+N_3^2(i)\right]\right\}\right] &\le \left[1-4l\lambda\tilde{g}_n(i)\right]^{-1/2},
\end{aligned}
$$

for any $i=1,\ldots,n$. The first result follows from (67). The second result follows from the same argument to derive (39). The third result follows from the same argument as for (i-a) of Lemma A.2. Second, from the same argument for (68) and the uniform boundedness of σ and L_i, there exists some constant $\tilde{C}>0$ such that

$$
\limsup_{n\to\infty}\max_{1\le i\le n}E\left[\exp\left\{\lambda np\left[N_1^2(i)+2N_1(i)L(i)\right]\right\}\right]<\tilde{C}. \qquad (82)
$$

Combining (80), (81), and (82), we obtain (77) with $C=\left\{\frac{(4l)!}{2^{4l}(2l)!}\right\}^{1/2l}\tilde{C}$.

Proof of (ii). Since we can use analogous arguments to those used in the proofs for Part (i) of this lemma and Lemma A.2 (ii), we only sketch the proof. Pick any $\lambda\in(-\infty,0)$. Analogously to (80), it holds that for any $p>1$,

$$
\begin{aligned}
&E\left[\exp\left\{\lambda n\left[N_1^2(i)+2N_1(i)L(i)+N_2^2(i)+3N_3^2(i)\right]\right\}\right] \\
&\ge G_n(i,\lambda)-E\left[\left|\exp\left\{\lambda n\left[N_2^2(i)+3N_3^2(i)\right]\right\}-1\right|^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}E\left[\exp\left\{\lambda np\left[N_1^2(i)+2N_1(i)L(i)\right]\right\}\right]^{1/p}.
\end{aligned}
$$

By using the same arguments as for (81) and (82), we can show that the second term on the right-hand side converges to zero uniformly over i. Now, (68) implies the desired result (78). The proof of (79) follows from the same argument and is omitted.

A.5 Proof of Theorem 4

Since the basic idea is similar to the previous proofs, we omit some details and outline only main points. By the same argument to the proof of Theorem 2, it is sufficient to show that for each $\lambda\in\mathbb{R}$,

$$
\lim_{n\to\infty}s_n\log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right]=\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}, \qquad (83)
$$

where $T_n := m_n \left\{ RV_n - \overline{\sigma^2} \right\}$ and $s_n := m_n^2/n$. Below we only consider the case of $\lambda \geq 0$ and show that the limit of an upper bound of $s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right]$ is $\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}$. In the same manner, we can show that the limit of a lower bound is also $\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}$. The proof for the case of $\lambda < 0$ is similar.

Pick any $\lambda \in [0, \infty)$. By using (64),

$$s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right] \leq A_n + B_n, \tag{84}$$

where

$$A_n := \frac{m_n^2}{n}\sum_{i=1}^{n} \log E\left[\exp\left\{\frac{\lambda n}{m_n}\left[N_1^2(i) + 2N_1(i)L(i) + 3N_2^2(i) + N_3^2(i)\right]\right\}\right],$$
$$B_n := \lambda m_n \sum_{i=1}^{n}\left[1 + r_n(i)\right]L^2(i) - \lambda m_n \overline{\sigma^2}.$$

First, we consider A_n. Recall the definitions of $G_n(i,\lambda)$ in (68) and $\tilde{g}_n(i)$ in (66). Observe that

$$\begin{aligned}
A_n &\leq \frac{m_n^2}{n}\sum_{i=1}^{n}\log\left[G_n\left(i,\frac{\lambda}{m_n}\right) + 2C\lambda\frac{\tilde{g}_n(i)}{m_n}\right] \\
&= \frac{m_n^2}{n}\sum_{i=1}^{n}\left\{\log G_n\left(i,\frac{\lambda}{m_n}\right) + 2C\lambda\frac{\tilde{g}_n(i)}{m_n}\left[G_n\left(i,\frac{\lambda}{m_n}\right) + 2C\lambda\frac{\tilde{g}_n(i)}{m_n}\xi_{n,i}\right]^{-1}\right\} \\
&= \frac{m_n^2}{n}\sum_{i=1}^{n}\log G_n\left(i,\frac{\lambda}{m_n}\right) + O(1)\left\{\lambda\frac{m_n}{n}\sum_{i=1}^{n}\sup_{u\in((i-1)/n,i/n]}\left|\sigma_{u-} - \sigma_{(i-1)/n}\right|\right\},
\end{aligned} \tag{85}$$

for some $\xi_{n,i} \in [0,1]$, where the inequality holds by Lemma A.4 (i) and $\max_{1\leq i\leq n}\left[1 - 4l\left(\frac{\lambda}{m_n}\right)\tilde{g}_n(i)\right]^{-1/4l} \leq 2$ (the positive integer l appears in (77)) for all n large enough (since $\max_{1\leq i\leq n}\tilde{g}_n(i)/m_n \to 0$), the first equality follows from the mean value theorem, and the second equality follows from the definition of $\tilde{g}_n(i)$ given in (66) and the result

$$\max_{1\leq i\leq n}\left[G_n\left(i,\frac{\lambda}{m_n}\right) + 2C\lambda\frac{\tilde{g}_n(i)}{m_n}\xi_{n,i}\right]^{-1} = O(1). \tag{86}$$

The result (86) can be shown by noting that $\min_{1\leq i\leq n} G_n\left(i,\frac{\lambda}{m_n}\right) \geq 1/2$ for all n large enough and the assumptions $m_n \to \infty$ and $n/m_n^2 \to 0$. We now consider the first term on the right-hand side of (85). Recall the definition of $L(i)$ in (61) and note that an expansion yields $L^2(i) = \int_{(i-1)/n}^{i}\sigma_{u-}^2 du + \eta_n(i)$, where $\eta_n(i)$ is defined in (74). Thus, we have

$$\begin{aligned}
\frac{m_n^2}{n}\sum_{i=1}^{n}\log G_n\left(i,\frac{\lambda}{m_n}\right) &= -\frac{m_n^2}{2n}\sum_{i=1}^{n}\log\left[1 - 2\frac{\lambda}{m_n}\left(1-\rho^2\right)\sigma_{(i-1)/n}^2\right] \\
&\quad + 2\lambda^2\left(1-\rho^2\right)\rho^2\sum_{i=1}^{n}\frac{\sigma_{(i-1)/n}^2\int_{(i-1)/n}^{i}\sigma_{u-}^2 du + \eta_n(i)\,\sigma_{(i-1)/n}^2}{1 - 2\frac{\lambda}{m_n}\left(1-\rho^2\right)\sigma_{(i-1)/n}^2}.
\end{aligned} \tag{87}$$

By applying the Taylor expansion to the first term on the right-hand side of (87) and using the conditions on m_n and σ (in the same manner as in (57)), we have

$$-\frac{m_n^2}{2n}\sum_{i=1}^{n}\log\left[1 - \frac{2\lambda}{m_n}\left(1-\rho^2\right)\sigma_{(i-1)/n}^2\right] = \lambda\left(1-\rho^2\right)\frac{m_n}{n}\sum_{i=1}^{n}\sigma_{(i-1)/n}^2 - \lambda^2\left(1-\rho^2\right)^2\overline{\sigma^4}\left\{1 + o(1)\right\}.$$

On the other hand, by using the boundedness and continuity properties of σ and a similar argument for (70), we can see that the second term on the right-hand side of (87) is written as $2\lambda^2\left(1-\rho^2\right)\rho^2\overline{\sigma^4}\left\{1+o\left(1\right)\right\}$. Combining these results,

$$\frac{m_n^2}{n}\sum_{i=1}^{n}\log G_n\left(i,\frac{\lambda}{m_n}\right)=\lambda\left(1-\rho^2\right)\frac{m_n}{n}\sum_{i=1}^{n}\sigma_{(i-1)/n}^2+\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}+o\left(1\right). \tag{88}$$

Next, we consider B_n. Since $L^2\left(i\right)=\int_{(i-1)/n}^{i}\sigma_{u-}^2du+\eta_n\left(i\right)$,

$$\begin{aligned}B_n&=\lambda\rho^2m_n\underbrace{\left[\sum_{i=1}^{n}\int_{(i-1)/n}^{i}\sigma_{u-}^2du-\overline{\sigma^2}\right]}_{=0}-\lambda\left(1-\rho^2\right)m_n\overline{\sigma^2}+\lambda\rho^2m_n\sum_{i=1}^{n}\eta_n\left(i\right)\\&\quad+\lambda m_n\sum_{i=1}^{n}r_n\left(i\right)\rho^2\left[\int_{(i-1)/n}^{i}\sigma_{u-}^2du+\eta_n\left(i\right)\right],\\&=-\lambda\left(1-\rho^2\right)m_n\overline{\sigma^2}+O\left(1\right)\left\{\lambda\frac{m_n}{n}\sum_{i=1}^{n}\sup_{u\in((i-1)/n,i/n]}\left|\sigma_{u-}-\sigma_{(i-1)/n}\right|\right\},\\&\quad+\lambda\rho^2m_n\sum_{i=1}^{n}\eta_n\left(i\right)\left[1+r_n\left(i\right)\right],\end{aligned}\tag{89}$$

where the second equality follows from $\int_{(i-1)/n}^{i}\sigma_{u-}^2du=O\left(n^{-1}\right)$ and the definition of $r_n\left(i\right)$ in (60). Now, putting (84), (85), (88), and (89) together, we can write

$$\begin{aligned}s_n\log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right]&\leq\lambda\left(1-\rho^2\right)\frac{m_n}{n}\sum_{i=1}^{n}\left(\sigma_{(i-1)/n}^2-\overline{\sigma^2}\right)+\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}\\&\quad+O\left(1\right)\left\{\lambda\frac{m_n}{n}\sum_{i=1}^{n}\sup_{u\in((i-1)/n,i/n]}\left|\sigma_{u-}-\sigma_{(i-1)/n}\right|\right\}\\&\quad+\lambda\rho^2m_n\sum_{i=1}^{n}\eta_n\left(i\right)\left[1+r_n\left(i\right)\right]+o\left(1\right).\end{aligned}\tag{90}$$

Finally, we show that if we $m_n^2/n\to0$,

$$m_n\sum_{i=1}^{n}\eta_n\left(i\right)\to0\quad\text{and}\quad m_n\sum_{i=1}^{n}r_n\left(i\right)\eta_n\left(i\right)\to0, \tag{91}$$

which implies that the fourth term on the right hand side of (90) converges to zero. Then we can see that the right hand side of (90) converges to $\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}$ under the conditions in (9) as desired. We only consider the former convergence result in (91), since the latter can be shown analogously. Observe that for any positive integers m and n with $m<n$,

$$\begin{aligned}E_\sigma\left[\max_{m<k\leq n}\left|\sum_{i=m+1}^{k}\eta_n\left(i\right)\right|^2\right]&\leq C_1\sum_{i=m+1}^{n}E_\sigma\left[\int_{(i-1)/n}^{i}\left(\int_{(i-1)/n}^{s}\sigma_{u-}dZ_u\right)^2\sigma_{s-}^2ds\right]\\&\leq C_1\tilde{C}^2\left(n-m\right)n^{-2},\end{aligned}\tag{92}$$

for some $C_1 > 0$ and $\tilde{C} > 0$, where E_σ stands for the expectation with respect to σ (and Z), the first inequality holds by the Burkholder-Davis-Gundy inequality (recall the definition of $\eta_n(i)$ in (74)), and the second inequality uses the result in (75). Similarly, we have

$$E_\sigma\left[\left|\sum\nolimits_{i=1}^{m} \eta_n(i)\right|^2\right] = E_\sigma\left[\sum\nolimits_{i=1}^{m} \eta_n^2(i)\right] \leq \tilde{C}^2 m n^{-2} \leq \tilde{C}^2 n^{-1}. \tag{93}$$

Now, let $S_n := n^{1/2}\sum_{i=1}^{n}\eta_n(i)$. >From (92) and (93),

$$\begin{aligned}
E_\sigma\left[\max_{m<k\leq n}|S_k - S_m|^2\right] &= E_\sigma\left[\max_{m<k\leq n}\left|k^{1/2}\sum\nolimits_{i=m+1}^{k}\eta_n(i) + \left(k^{1/2}-m^{1/2}\right)\sum\nolimits_{i=1}^{m}\eta_n(i)\right|^2\right] \\
&\leq 2nE_\sigma\left[\max_{m<k\leq n}\left|\sum\nolimits_{i=m+1}^{k}\eta_n(i)\right|^2\right] + 2\left(n^{1/2}-m^{1/2}\right)^2 E_\sigma\left[\left|\sum\nolimits_{i=1}^{m}\eta_n(i)\right|^2\right] \\
&\leq Cn^{-1}\left[(n-m) + \left(n^{1/2}-m^{1/2}\right)^2\right] \leq 3Cn^{-1}(n-m),
\end{aligned}$$

where the first inequality uses the inequality: $(a+b)^2 \leq 2\left(a^2+b^2\right)$ for $a, b \in \mathbb{R}$, the second inequality follows from (92) and (93) with $C = 2(C_1+1)\tilde{C}^2$, and the last inequality holds since $\left(n^{1/2}-m^{1/2}\right)^2 \leq 2(n-m)$. Defining $d_n(i) := \sqrt{3Cn^{-1}}$ for $i = 1, \ldots, n$ (an array which does not vary with i), we can write the above inequality as

$$E_\sigma\left[\max_{m<k\leq n}|S_k - S_m|^2\right] < \sum_{i=m+1}^{n} d_n^2(i) \quad \text{with} \quad \lim_{n\to\infty}\sum_{i=1}^{n} d_n^2(i) = 3C < \infty. \tag{94}$$

By the so-called Cauchy criterion (see, e.g., Corollary 20.2 of Davidson, 1994) and the Markov inequality, the result (94) implies the existence of a random variable S such that $S_n = n^{1/2}\sum_{i=1}^{n}\eta_n(i) \to S$ almost surely (with respect to the probability measure of σ and Z). Combining this convergence with the condition $m_n^2/n \to 0$, we have

$$m_n\sum_{i=1}^{n}\eta_n(i) = \frac{m_n}{n^{1/2}}S_n \to 0,$$

as desired. Therefore, the limit of an upper bound of $s_n \log E\left[\exp\left\{s_n^{-1}\lambda T_n\right\}\right]$ is given by $\lambda^2\left(1-\rho^4\right)\overline{\sigma^4}$.

References

[1] Aït-Sahalia, Y. and J. Jacod (2009) Testing for jumps in a discretely observed process, *Annals of Statistics*, 37, 184-222.

[2] Andersen, T. G., Bollerslev, T. and F. X. Diebold (2009) Parametric and nonparametric volatility measurement, in Aït-Sahalia, Y. and L. P. Hansen (eds.), *Handbook of Financial Econometrics*, vol. 1 (1), Elsevier.

[3] Andersen, T. G., Bollerslev, T. and D. Dobrev (2007) No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: theory and testable distributional implications, *Journal of Econometrics*, 138, 125-180.

[4] Andersen, T. G., Bollerslev, T., Diebold, F. X. and P. Labys (2003) Modeling and forecasting realized volatility, *Econometrica*, 71, 579-625.

[5] Andersen, T. G., Bollerslev, T. and N. Meddahi (2005) Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities, *Econometrica*, 73, 279-296.

[6] Bandi, F. M. and J. R. Russell (2007) Volatility, in Birge, J. R. and V. Linetsky (eds.), *Handbook of Financial Engineering*, 183-222, Elsevier.

[7] Bandi, F. M. and J. R. Russell (2008) Microstructure noise, realized variance, and optimal sampling, *Review of Economic Studies*, 75, 339-369.

[8] Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J. and N. Shephard (2006a) Limit theorems for bipower variation in financial econometrics, *Econometric Theory*, 22, 677-719.

[9] Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M. and N. Shephard (2006b) A central limit theorem for realised power and bipower variations of continuous semimartingales, in Kabanov, Y. and R. Lispter (eds.), *Stochastic Analysis to Mathematical Finance, Frestchrift for Albert Shiryaev*, 33-68, Springer.

[10] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and N. Shephard (2008) Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica*, 76, 1481-1536.

[11] Barndorff-Nielsen, O. E. and N. Shephard (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, *Journal of the Royal Statistical Society*, B 63, 167-241.

[12] Barndorff-Nielsen, O. E. and N. Shephard (2002) Econometric analysis of realized volatility and its use in estimating volatility models, *Journal of the Royal Statistical Society*, B 64, 253-280.

[13] Barndorff-Nielsen, O. E. and N. Shephard (2004) Power and bipower variation with stochastic volatility and jumps, *Journal of Financial Econometrics*, 2, 1-48.

[14] Barndorff-Nielsen, O. E. and N. Shephard (2006) Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics*, 4, 1-30.

[15] Barndorff-Nielsen, O.E., Shephard, N. and M. Winkel (2006) Limit theorems for multipower variation in the presence of jumps, *Stochastic Processes and their Applications*, 116, 796-806.

[16] Bercu, B., Gamboa, F. and A. Rouault (1997) Large deviations for quadratic forms of stationary Gaussian processes, *Stochastic Processes and their Applications*, 71, 75-90.

[17] Billingsley, P. (1999) *Convergence of Probability Measures*, second edition, Wiley.

[18] Black, F. (1976) Studies on stock market volatility changes, *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.

[19] Bryc, W. and A. Dembo (1997) Large deviations for quadratic functionals of Gaussian processes, *Journal of Theoretical Probability*, 10, 307-322.

[20] Cheung, W. (2001) Generalizations of Hölder's inequalities, *International Journal of Mathematics and Mathematical Sciences*, 26, 7-10.

[21] Christensen, K., Oomen, R. and M. Podokskij (2010) Realised quantile-based estimation of the integrated variance, *Journal of Econometrics*, 159, 74-98.

[22] Comte, F. and E. Renault (1996) Long memory continuous time models, *Journal of Econometrics* 73, 101-149.

[23] Davidson, J. (1994) *Stochastic Limit Theory*, Oxford University Press.

[24] Dembo, A. and O. Zeitouni (1998) *Large Deviation Techniques and Applications*, second edition, Springer.

[25] Ellis, R. S. (1984) Large deviations for a general class of random vectors, *Annals of Probability*, 12, 1-12.

[26] Embrechts, P and M. Maejima (2002) *Selfsimilar Processes*, Princeton University Press.

[27] Feng, J., Forde, M. and J.-P. Fouque (2010) Short-maturity asymptotics for a fast mean-reverting Heston stochastic volatility model, *SIAM Journal on Financial Mathematics*, 1, 126-141.

[28] Finner, H. (1992) A generalization of Hölder's inequality and some probability inequalities, *Annals of Probability*, 20, 1893-1901.

[29] Gärtner, J. (1977) On large deviations from an invariant measure (in Russian), *Teor. Verojatnost. i Primenen*, 22, 27-42.

[30] Gonçalves, S. and N. Meddahi (2009) Bootstraping realized volatility, *Econometrica*, 77, 283-306.

[31] Grama, I. G. (1997) On moderate deviations for martingales, *Annals of Probability*, 25, 152-183.

[32] Grama, I. G. and E. Haeusler (2000) Large deviations for martingales via Cramér's method, *Stochastic Processes and their Applications*, 85, 279-293.

[33] Hansen, P. R. and A. Lunde (2005) Realized variance and market microstructure noise, *Journal of Business and Economic Statistics*, 24, 127-161.

[34] 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, 327-343.

[35] Karatzas, I. and S. E. Shreve (1991) *Brownian Motion and Stochastic Calculus*, Springer.

[36] Lesigne, E. and D. Volný (2001) Large deviations for martingales, *Stochastic Processes and their Applications*, 96, 143-159.

[37] Nelson, D. B. (1990) ARCH models as diffusion approximations, *Journal of Econometrics* 45, 7-39.

[38] Revuz, D. and M. Yor (1999) *Continuous Martingales and Brownian Motion*, third edition, Springer.

[39] Romano, M. and N. Touzi, (1997) Contingent claims and market completeness in a stochastic volatility model, *Mathematical Finance*, 7, 399-410.

[40] Shephard, N. and T. G. Andersen (2009) Stochastic volatility: origins and overview, in Andersen, T. G., Davis, R. A., Kreiß, J. P. and T. Mikosch (eds.), *Handbook of Financial Time Series*, 233-254, Springer.

[41] Tauchen, G. and V. Todorov (2010) Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation, forthcoming in *Annals of Applied Probability.*

[42] Todorov, V. (2008) Econometric analysis of jump-driven stochastic volatility models, forthcoming in *Journal of Econometrics.*

[43] Zhang, L., Mykland, P. M. and Y. Aït-Sahalia (2005a), A tale of two time scales: determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, 100, 1394-1411.

[44] Zhang, L., Mykland, P. M. and Y. Aït-Sahalia (2005b), How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies*, 18, 351-461.

[45] Zhang, L., Mykland, P. M. and Y. Aït-Sahalia (2011) Edgeworth expansions for realized volatility and related estimators, *Journal of Econometrics*, 160, 190-203.