

## Article

# The Convergence Rates of Large Volatility Matrix Estimator Based on Noise, Jumps, and Asynchronization

Erlin Guo <sup>1</sup>, Cuixia Li <sup>1</sup> and Fengqin Tang <sup>2,\*</sup>

<sup>1</sup> School of Mathematics and Statistics, Xuzhou University of Technology, Xuzhou 221018, China; gel@xzit.edu.cn (E.G.); licuixia2002@163.com (C.L.)

<sup>2</sup> School of Mathematics Sciences, Huaibei Normal University, Huaibei 235000, China

\* Correspondence: tfq05@163.com

**Abstract:** At the turn of the 21st century, the wide availability of high-frequency data aroused an increasing demand for better modeling and statistical inference. A challenging problem in statistics and econometrics is the estimation problem of the integrated volatility matrix based on high-frequency data. The existing estimators work well for diffusion processes with micro-structural noise and may get worse when jumps are considered. This paper proposes a novel estimation in the presence of jumps, micro-structural noise, and asynchronization. First, we adopt sub-sampling to synchronize the high-frequency data. Then, we use a two-time scale to realize co-volatility to handle noise. Finally, we employ the threshold parameters to remove the effect of jumps and sparsity in two steps. Both the minimax bound and the convergence rate are discussed in the paper. The estimation procedures of the heavy-tailed data will be solved in the future.

**Keywords:** high-frequency data; asymptotic theory; threshold; jumps; large integrated volatility matrix

**MSC:** 60G51; 60G52; 62G05; 62G35



**Citation:** Guo, E.; Li, C.; Tang, F. The Convergence Rates of Large Volatility Matrix Estimator Based on Noise, Jumps, and Asynchronization. *Mathematics* **2023**, *11*, 1425. <https://doi.org/10.3390/math11061425>

Academic Editor: António Lopes

Received: 13 February 2023

Revised: 12 March 2023

Accepted: 14 March 2023

Published: 15 March 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

At the turn of the 21st century, with the advance of technology in high-frequency trading, high-frequency financial data were recorded at several seconds, even mill- and microseconds. In the United States, tick-by-tick stock transactions can be obtained by the Trade and Quote (TAQ2) database, which includes the New York Stock Exchange. On the other hand, in China, we can obtain one-second records for all stocks from databases of some fund management companies. Such data are widely used for volatility estimation in finance, the environment, and other fields, providing guidance and prediction for financial risk management, environmental monitoring, and other aspects [1–5]. Therefore, the wide availability of high-frequency data arouses increasing demand for better modeling and statistical inference. In financial applications, it often involves dozens or even hundreds of assets, and the corresponding integral volatility matrix is also a high-dimensional problem. When the number of assets is larger than the sample size, the estimation of the integral volatility matrix has been a focus and a challenging problem in finance and statistics recently.

It has long been recognized that market micro-structure noise plays a significant role in the estimation of volatility [6]. The market micro-structure noise is formed by the interaction of transaction cost and transaction friction, which mainly include bid–ask spread, the discreteness of price, etc. The Itô process is used to model the logarithmic price of assets in high-frequency finance, and various non-parametric methods are developed to estimate the integrated (co)volatility for multiple assets based on high-frequency data contaminated with micro-structural noise over a period of time. Such methods include Hayashi and Yoshida (HY) estimators based on overlap intervals [7], multi-scale realized co-volatility (MSRV)

based on previous tick data synchronization [8], a quasi-maximum likelihood estimator (QMLE) based on generalized sampling time [9], realized kernel volatility estimator based on a refresh time scheme [10], and pre-averaging realized volatility [11]. In portfolio risk and hedging of funds management, a large number of assets are often encountered, and one key and challenging problem is estimating the integrated matrix based on high-frequency data. Reference [12] proposed the ARVM estimator under the sparse integrated volatility matrix based on contaminated high-frequency data. Reference [13] employed a large volatility matrix estimator using high-frequency data and established that the convergence rate of the estimator depends on sample size with  $1/6$ -exponent under the general diffusion setup with micro-structure noise in the data. Under sparse conditions, the convergence rate of the kernel and the pre-averaging realized volatility is slower than that of the multi-scale realized volatility, due to the trade-off between positive semi-definiteness and fast convergence rate [14]. An adaptive thresholding estimator of a large volatility matrix with varying entries is developed, and the optimal spectral norm convergence rate of the estimator is shown under sparse conditions [15].

However, in [12–15], the estimations were carried out only under a continuous framework. In fact, in an efficient market environment, the release of major “news” will trigger significant changes in very short time frames, which are called jumps in the pricing process, and the effect of jumps has generally been acknowledged in the high-frequency literature. For a single asset, [16] proposed a non-parametric estimator for the integrated volatility in the presence of jumps and micro-structural noise. Reference [17] developed a statistical test against the necessity of a diffusion component. They used the stock price records of Microsoft (MFST) in 1 November, 1 December, and 11 December 2000 to implement their test, and rejected the existence of the diffusion component. A novel test was proposed to check whether the underlying process of high-frequency data can be modeled using a pure-jump process that was robust to jumps of infinite variation [18]. Reference [19] estimated the integrated volatility in the presence of jumps and endogeneity. Reference [20] developed multipower estimators in the presence of jumps, micro-structural noise, and multiple records of observations. Reference [21] considered a new estimator of the integrated volatility in the presence of jumps and micro-structural noise when the sampling was endogenous. For multiple assets, [22] proposed a threshold estimator to estimate the covariation of two asset prices in the simultaneous presence of jumps and micro-structural noise, but the observed times were synchronous.

These works all concluded that the effect of jumps cannot be appropriately characterized by a continuous model; hence, the estimators proposed in these papers may perform poorly, even if the number of assets is small. When the number of assets becomes larger, the existing studies did not consider the estimation of the integrated volatility matrix in the presence of jumps, micro-structural noise, and nonsynchronous observations.

This paper proposes a thresholding estimator for the integrated volatility matrix which is constructed as follows. First, we synchronize observation time points based on previous tick data. Second, we adopt the threshold method to remove the effect of jumps, such that those increments smaller than the threshold level are included for the calculation of estimator, and those increments larger than the threshold level will be excluded as  $n \rightarrow \infty$ . Finally, we use another threshold to overcome the effect of the sparsity condition. We show that the thresholding estimator is robust to the simultaneous presence of nonsynchronous, jumps, and microstructural noise and establish an asymptotic theory for the proposed estimator when both the assets size and the sample size approach infinity. In this paper, we restrict our study to financial high-frequency data. In addition, our results can be extended to engineering, biomedical, imaging science, and photographic technology [23–29].

The remainder of the paper proceeds as follows. The price model, observed data, and the estimation problem are described in Section 2. Section 3 presents the advised estimator in three steps and the sparse condition. Asymptotic theories are displayed in Section 4, and

proofs are put forward in Section 5. Some conclusions and future research are proposed in Section 6.

## 2. Methodology

### 2.1. Price Model

Let  $\mathbf{X}(t) = (X_1(t), \dots, X_p(t))^T$  be the vector of true log prices of  $p$  assets at time  $t$ . It is well known that  $\mathbf{X}(t)$  is a semimartingale process [30]; thus

$$\mathbf{X}(t) = \mathbf{X}^c(t) + \mathbf{X}^d(t), \tag{1}$$

where

$$d\mathbf{X}^c(t) = \mathbf{b}(t)dt + \sigma(t)^T d\mathbf{B}_t, \quad t \in [0, 1], \tag{2}$$

$$d\mathbf{X}^d(t) = \int_{|x| \leq 1} x(\boldsymbol{\mu} - \boldsymbol{\nu})(dt, dx) + \int_{|x| > 1} x\boldsymbol{\mu}(dt, dx), \tag{3}$$

where  $\mathbf{b}(t) = (b_1(t), \dots, b_p(t))^T$  is a drift vector,  $\mathbf{B}_t = (B_{1t}, \dots, B_{pt})^T$  is standard  $p$ -dimensional Brownian motion, and  $\sigma(t)$  is a càdlàg and locally bounded  $p$ -by- $p$  matrix,  $\boldsymbol{\mu}$  is the jump measure compensated by  $\boldsymbol{\nu}$ , and  $\boldsymbol{\nu}$  has the form  $dt dF_t(x)$ , where  $F_t(dx)$  is a transition measure form  $\Omega \times \mathbb{R}_+$  endowed with the predictable  $\sigma$ -field into  $\mathbb{R}/0$ . The jump activity index is defined as  $\beta =: \{\inf s : \int_{|x| \leq 1} |x|^s F_t(dx) < \infty, \text{ and if } 0 \leq \beta < 1, \text{ we say that } \mathbf{X}^d(t) \text{ has finite variation.}\}$

**Remark 1.** Equation (1) is a rather canonical model in finance theory when the continuous part and discontinuous or jumps part are considered in the model. All powers of  $\sigma$  are locally integrated with respect to the Lebesgue measure, since  $\sigma$  is càdlàg and locally bounded.

### 2.2. Observed Data

Because different assets are traded at distinct times in high-frequency finance, the data for multiple assets often encounter non-synchronization problems, and the true log prices  $\mathbf{X}(t)$  are observed with contamination by the micro-structural noise. On this basis, it is assumed that the observed high-frequency financial data  $Y_i(t_{i,r})$  obey the model

$$Y_i(t_{i,r}) = X_i^c(t_{i,r}) + X_i^d(t_{i,r}) + \epsilon_i(t_{i,r}), \quad i = 1, \dots, p; \quad r = 0, \dots, n_i, \tag{4}$$

where  $t_{i,r}$  denotes the  $r$ -th observation time point for the  $i$ -th asset.

**Assumption 1.** Let  $\epsilon_i(t_{i,r}), i = 1, \dots, p, r = 0, \dots, n_i$ , be independent noises with mean zero, for each fixed  $i$ .  $\epsilon_i(t_{i,r}), r = 0, \dots, n_i$  are i.i.d. random variables with variance  $\eta_{ii}$ , and  $\epsilon_i(\cdot), X_i^c(\cdot)$ , and  $X_i^d(\cdot)$  are independent.

Define the quadratic variation of  $\mathbf{X}(t)$  as

$$[\mathbf{X}^c, \mathbf{X}^c]_1 = \left( \int_0^1 \sum_{k=1}^p \sigma_{ik}(s)\sigma_{kj}(s)ds \right)_{1 \leq i, j \leq p}, \tag{5}$$

and denote the quadratic variation of  $\mathbf{X}(t)$  as the large volatility matrix  $\Gamma$ , i.e.,

$$\Gamma = (\Gamma_{ij})_{1 \leq i, j \leq p} = [\mathbf{X}^c, \mathbf{X}^c]_1, \tag{6}$$

where  $\Gamma_{ij}$  is the  $ij$ -th element of matrix  $\Gamma$ .

The focus of the current paper is to construct a new estimator for the large volatility matrix  $\Gamma$ , and to investigate some asymptotic properties of the proposed estimator in the presence of non-synchronization, micro-structural noise, and jumps.

### 3. The Estimator of the Large-Volatility Matrix

In quantitative finance, capital asset price or return volatility is an important index to measure investment risk. A large number of authors have conducted in-depth studies on financial volatility. Based on the literature, this paper considers the estimation problem of the integral volatility matrix when the asset price process has a jump in a high-dimensional case. It provides a theoretical basis for later statistical inference, asset allocation, risk management, and optimization.

To estimate the large volatility matrix  $\Gamma$ , we need to give a new estimator  $\hat{\Gamma} = (\hat{\Gamma}_{ij}^T(\omega))_{1 \leq i, j \leq p}$  in this section, and the main results, i.e., the asymptotic properties, will be provided in next section. Now, let us describe the estimator in detail.

For a fixed integer  $m$ , we partition the interval  $[0, 1]$  into  $m$  equal sub-intervals, and let  $K = \lfloor n/m \rfloor$ . Take  $\tau_l^k = \frac{l}{m} + \frac{k-1}{n}$ , where  $k = 1, \dots, K; l = 0, 1, \dots, m-1$ , and  $n$  is the average sample size of  $p$  assets and

$$n = \frac{1}{p} \sum_{i=1}^p n_i, \tag{7}$$

is the pre-determined sampling frequency.

For a given  $k, k = 1, \dots, K$  and asset  $i$ , we define the previous-tick times as

$$\tau_{i,l}^k = \max\{t_{i,r} \leq \tau_l^k; r = 1, \dots, n_i\}, \quad l = 0, 1, \dots, m-1. \tag{8}$$

Meanwhile, let

$$\Delta_l = \tau_l^k - \tau_{l-1}^k = \frac{1}{m}, \tag{9}$$

and define  $\Delta_l^k Y_i = Y_i(\tau_{i,l}^k) - Y_i(\tau_{i,l-1}^k)$ .

Thus, under the model (1)–(3), we denote

$$\hat{\Gamma}_{ij}^T = \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} [\Delta_l^k Y_i \Delta_l^k Y_j] \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}}, \tag{10}$$

where  $u_l$  satisfies  $u_l / \Delta_l^{\omega_1} \rightarrow 0, u_l / \Delta_l^{\omega_2} \rightarrow \infty$ , for some  $0 \leq \omega_1 < \omega_2 < 1/6$ .

Since the noise of  $\hat{\Gamma}_{ii}^T$  cannot be easily ignored, we need to redefine  $\hat{\Gamma}_{ii}^T$  as

$$\hat{\Gamma}_{ii}^T = \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} [\Delta_l^k Y_i]^2 \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} - 2m\hat{\eta}_{ii}, \tag{11}$$

where

$$\hat{\eta}_{ii} = \frac{1}{2n_i} \sum_{r=1}^{n_i} [Y_i(t_{i,r}) - Y_i(t_{i,r-1})]^2, \tag{12}$$

is the estimator of  $\eta_{ii}$  [31].

Then, we can redefine  $\tilde{\Gamma}_{ij}^T$  by  $\hat{\Gamma}_{ij}^T$ , for which

$$\tilde{\Gamma}_{ij}^T = \begin{cases} \hat{\Gamma}_{ij}^T, & \text{if } i \neq j, \\ \hat{\Gamma}_{ii}^T - 2m\hat{\eta}_{ii}, & \text{if } i = j. \end{cases} \tag{13}$$

**Remark 2.** First, we use  $\Delta_l$  to sub-sample high-frequency data, and the purpose of sub-sampling is to delete noise. Second, we use the threshold to discard the data larger than  $u_l$ , and the purpose of the threshold is to remove the effect of jumps. Finally, we take their average, and the purpose of averaging is to yield a better convergence rate.

**Remark 3.** After the noise is deleted, the increments from the jump part are equal to or larger than  $\Delta_l^{1/2}$ , while the increments from the diffusion part are still smaller than  $\Delta_l^{1/2}$ . The threshold

level  $u_l$  is provided such that those increments  $(|\Delta_l^k Y_i|)$  smaller than  $u_l$  will be included in the calculation of the integrated volatility matrix, while those increments equal to or larger than  $u_l$  will be gradually excluded.

We denote  $\tilde{\Gamma} = (\tilde{\Gamma}_{ij}^T)_{1 \leq i, j \leq p}$ ; for small  $p$ ,  $\tilde{\Gamma}$  provides a good estimator for  $\Gamma$ . However, when  $p$  is large, even going to infinity as  $n \rightarrow \infty$ ,  $\tilde{\Gamma}$  performs poorly. To estimate a large integrated volatility matrix, one of the key assumptions is that the target matrix of interest is sparse, which is a regularization condition in many studies [12–14,32–35].

**Assumption 2.** Sparsity condition:  $M$  is a positive random variable,  $\pi(p)$  is a deterministic function of  $p$  that grows very slowly in  $p$ , and  $0 \leq \delta < 1$ . Assume that  $\Gamma$  satisfies

$$\sum_{j=1}^p |\Gamma_{ij}|^\delta \leq M\pi(p), \quad i = 1, \dots, p; \quad E[M] \leq C. \tag{14}$$

Under the sparsity condition (14), we regularize  $\tilde{\Gamma}_{ij}^T$  as follows,

$$\hat{\Gamma}_{ij}^T(\omega) = \tilde{\Gamma}_{ij}^T 1(|\tilde{\Gamma}_{ij}^T| \geq \omega), \quad \text{for } i, j = 1, \dots, p, \tag{15}$$

where  $\omega$  is a threshold parameter. Denote  $\hat{\Gamma} = (\hat{\Gamma}_{ij}^T(\omega))_{i, j=1, \dots, p}$ , then the  $(i, j)$ -th element  $\hat{\Gamma}_{ij}^T(\omega)$  of  $\hat{\Gamma}$  is equal to  $\tilde{\Gamma}_{ij}^T$  if its absolute value equals or exceeds  $\omega$  and is zero otherwise, as well as  $\Gamma = (\Gamma_{ij}(\omega))_{1 \leq i, j \leq p}$ .

**Remark 4.** In this part, when both the number of assets and the sample size approach infinity, we provide a new estimator in order to estimate the large dimensional integral volatility matrix, in the presence of jumps, micro-structural noise, and non-synchronization. First, we adopt the sub-sampling to synchronize the high-frequency data. Second, we use the two-time scale realized co-volatility on the off-diagonal elements of  $\Gamma$  but correcting bias on the diagonal elements of  $\Gamma$ , to handle noise. Finally, we employ the threshold parameters to remove the effect of jumps and sparsity in two steps.

#### 4. Asymptotic Properties

In this section, the  $k$ -moment convergence rate and the minimax bound of the suggested estimator are provided, which are the main results of this paper. To provide some theory, some notations are given. Let  $\mathbf{x} = (x_1, \dots, x_p)^T$  denote a vector, and  $\mathbf{U} = (U_{ij})_{p \times p}$  denote a matrix; we can define their  $l_d$ -norms as follows:

$$\|\mathbf{x}\|_d = \left(\sum_{i=1}^p |x_i|^d\right)^{1/d}, \quad \|\mathbf{U}\|_d = \sup\{\|\mathbf{U}\mathbf{x}\|_d, \|\mathbf{x}\|_d = 1\}, \quad d = 1, 2, \infty. \tag{16}$$

then,

$$\|\mathbf{U}\|_1 = \max_{1 \leq j \leq p} \sum_{i=1}^p |U_{ij}|, \quad \|\mathbf{U}\|_\infty = \max_{1 \leq i \leq p} \sum_{j=1}^p |U_{ij}|, \tag{17}$$

and

$$\|\mathbf{U}\|_2^2 \leq \|\mathbf{U}\|_1 \|\mathbf{U}\|_\infty. \tag{18}$$

Furthermore, we need the following assumptions for models (1)–(3).

**Assumption 3.** For some  $\alpha \geq 2$ ,

$$\max_{1 \leq i \leq p} \max_{0 \leq t \leq 1} E[|\sigma_{ii}(t)|^\alpha] < \infty, \quad \max_{1 \leq i \leq p} \max_{0 \leq t \leq 1} E[|b_i(t)|^\alpha] < \infty, \quad \max_{i \leq i \leq p} E[|\epsilon_i(t_{ir})|^{2\alpha}] < \infty. \tag{19}$$

**Assumption 4.** We assume that each of the  $p$  assets has at least one observation between  $\tau_{l-1}^k$  and  $\tau_l^k$ , and

$$C_1 \leq \min_{1 \leq i \leq p} \frac{n_i}{n} \leq \max_{1 \leq i \leq p} \frac{n_i}{n} \leq C_2, \quad \max_{1 \leq i \leq p} \max_{1 \leq l \leq n_i} |t_{i,r} - t_{i,r-1}| = O(n^{-1}), \quad m = o(n). \quad (20)$$

**Theorem 1.** Under models (1)–(3) and Assumptions 1–4, we have, for all  $1 \leq i, j \leq p$ ,

$$E(|\hat{\Gamma}_{ij}^T(\omega) - \Gamma_{ij}|^\alpha) \leq C_n^\alpha, \quad (21)$$

where  $C_n^\alpha = C((Kn^{-1/2})^{-\alpha} + K^{-\alpha/2} + (n/K)^{-\alpha/2} + K^{-\alpha} + n^{-\alpha/2})$  and  $C$  is a generic constant. When  $K = \lfloor n/m \rfloor \sim n^{2/3}$ , the convergence rate  $C_n$  of the estimator is  $n^{-1/6}$ .

**Theorem 2.** Under models (1)–(3) and Assumptions 1–4, if  $h_{n,p}$  is any sequence converging to infinity arbitrarily slow, we have

$$\|\hat{\Gamma} - \Gamma\|_2 \leq \|\hat{\Gamma} - \Gamma\|_\infty = O_P(\pi(p)[C_n p^{2/\alpha} h_{n,p}]^{1-\delta}), \quad (22)$$

where  $C_n \sim n^{-1/6}$  is given in Theorem 1, and  $\omega = C_n p^{2/\alpha} h_{n,p}$ .

**Remark 5.** In order to make  $C_n p^{2/\alpha}$  go to zero,  $p$  needs to grow more slowly than  $n^{\alpha/12}$ . The convergence rate in Theorem 2 is nearly equal to  $\pi(p)[C_n p^{2/\alpha}]^{1-\delta}$ , since  $C_n \sim n^{-\frac{1}{6}}$ .

**5. Proofs**

**Proof of Theorem 1.** Under the above assumptions,  $X_i^d, i = 1, \dots, p$  are of finite variation, i.e.,  $\beta_i < 1$ , we use  $X'$  as the continuous part of  $X$  and  $X''$  as the discontinuous martingale or jumps part, i.e.,

$$X'_{it} = X_0 + \int_0^t b'_{is} ds + \int_0^t \sigma_s dW_s, \quad X''_{it} = X_{it} - X'_{it}, \quad (23)$$

where  $b'_{it} = b_{it} - \int_{|x|>1} F_t(dx)$  and  $X'_i = \sum_{s \leq t} \Delta X_i(s)$ . We also define  $Y'_{it} = X'_{it} + \epsilon_{it}$ , and hence  $Y_{it} = Y'_{it} + X''_{it}$ .

We can decompose  $\tilde{\Gamma}_{ij}^T - \Gamma_{ij}$  as the following: if  $i \neq j$ ,

$$\begin{aligned} \tilde{\Gamma}_{ij}^T - \Gamma_{ij} &= \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y_i \Delta_l^k Y_j) \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} - \Gamma_{ij} \\ &= \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y_i \Delta_l^k Y_j) \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} - \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y'_i \Delta_l^k Y'_j) \\ &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y'_i \Delta_l^k Y'_j) - \Gamma_{ij} \\ &= \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} [\Delta_l^k Y'_i \Delta_l^k Y'_j \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} - \Delta_l^k Y'_i \Delta_l^k Y'_j] \\ &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} \Delta_l^k Y'_i \Delta_l^k X''_j \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} \\ &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} \Delta_l^k X''_i \Delta_l^k Y'_j \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} \\ &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} \Delta_l^k X''_i \Delta_l^k X''_j \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l, |\Delta_l^k Y_j| \leq u_l\}} \\ &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y'_i \Delta_l^k Y'_j) - \Gamma_{ij} \\ &=: \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} [A^1_{ij} + A^2_{ij} + A^3_{ij} + A^4_{ij} + A^5_{ij}]. \end{aligned} \quad (24)$$

From [6], we have  $E|A_{ij}^5|^\alpha \leq C(n^{\alpha/2}K^{-\alpha} + n^{-\alpha/2}K^{\alpha/2})$ , and for  $A_{ij}^s, s = 1, 2, 3, 4$ , we can consider the following disjoint cases of  $|\Delta_l^k Y_i|$  and  $|\Delta_l^k Y_j|$ . Here, we use  $C, l_1, l_2, r_1$ , and  $r_2$  to denote any positive real numbers that may vary from place to place.

- If  $|\Delta_l^k Y_i| \geq u_l/2$  and  $|\Delta_l^k Y_j| \geq u_l/2$ , we have that

$$\begin{aligned}
 |A_{ij}^1| &\leq \frac{C|\Delta_l^k Y_i|^{1+l_1}|\Delta_l^k Y_j|^{1+l_2}}{u_l^{l_1}u_l^{l_2}}, & |A_{ij}^2| &\leq \frac{C|\Delta_l^k Y_i|^{1+l_1}|\Delta_l^k Y_j|^{l_2}|\Delta_l^k X_j''|}{u_l^{l_1}u_l^{l_2}}, \\
 |A_{ij}^3| &\leq \frac{C|\Delta_l^k Y_i|^{l_1}|\Delta_l^k X_i''||\Delta_l^k Y_j|^{1+l_2}}{u_l^{l_1}u_l^{l_2}}, & |A_{ij}^4| &\leq \frac{C|\Delta_l^k Y_i|^{l_1}|\Delta_l^k X_i''||\Delta_l^k Y_j|^{l_2}|\Delta_l^k X_j''|}{u_l^{l_1}u_l^{l_2}}.
 \end{aligned}
 \tag{25}$$

- If  $|\Delta_l^k Y_i| \leq u_l/2$  and  $|\Delta_l^k Y_j| \leq u_l/2$ , we obtain

$$\begin{aligned}
 |A_{ij}^1| &\leq \frac{C|\Delta_l^k Y_i||\Delta_l^k Y_j||\Delta_l^k X_i''|^{r_1}|\Delta_l^k X_j''|^{r_2}}{u_l^{r_1}u_l^{r_2}}, & |A_{ij}^2| &\leq \frac{C|\Delta_l^k Y_i||\Delta_l^k X_i''|^{r_1}|\Delta_l^k X_j''|^{1+r_2}}{u_l^{r_1}u_l^{r_2}}, \\
 |A_{ij}^3| &\leq \frac{C|\Delta_l^k X_i''|^{1+r_1}|\Delta_l^k Y_j||\Delta_l^k X_j''|^{r_2}}{u_l^{r_1}u_l^{r_2}}, & |A_{ij}^4| &\leq \frac{C|\Delta_l^k X_i''|^{1+r_1}|\Delta_l^k X_j''|^{1+r_2}}{u_l^{r_1}u_l^{r_2}}.
 \end{aligned}
 \tag{26}$$

- If  $|\Delta_l^k Y_i| \geq u_l/2$  and  $|\Delta_l^k X_j''| \leq u_l/2$ , we obtain

$$\begin{aligned}
 |A_{ij}^1| &\leq \frac{C|\Delta_l^k Y_i|^{1+l_1}|\Delta_l^k Y_j||\Delta_l^k X_j''|^{r_2}}{u_l^{l_1}u_l^{r_2}}, & |A_{ij}^2| &\leq \frac{C|\Delta_l^k Y_i|^{1+l_1}|\Delta_l^k X_j''|}{u_l^{l_1}}, \\
 |A_{ij}^3| &\leq \frac{C|\Delta_l^k X_i''||\Delta_l^k Y_i|^{l_1}|\Delta_l^k Y_j|}{u_l^{l_1}}, & |A_{ij}^4| &\leq \frac{C|\Delta_l^k X_i''||\Delta_l^k X_j''||\Delta_l^k Y_i|^{l_1}}{u_l^{l_1}}.
 \end{aligned}
 \tag{27}$$

The case of  $|\Delta_l^k Y_i| \leq u_l/2$  and  $|\Delta_l^k X_j''| \geq u_l/2$  is similar to the above.

Next, by Hölder’s and Burkholder’s inequalities, we can estimate  $|\Delta_l^k Y_i|$  and  $|\Delta_l^k X_j''|$  as follows,

$$E(|\Delta_l^k X_i''|^2) \leq C\Delta_l, \quad \text{and} \quad E(|\Delta_l^k Y_i|^s) \leq C_s(\Delta_l)^{s/6}.
 \tag{28}$$

Without loss of generality, we let  $u_l = \Delta_l^{\epsilon+\omega_1}$ , where  $0 < \epsilon < \omega_2 - \omega_1$ . Then, we have

$$\begin{aligned}
 &E\left[\frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (A_{ij}^1 + A_{ij}^2 + A_{ij}^3 + A_{ij}^4 + A_{ij}^5)^\alpha\right] \\
 &\leq CK^{-\alpha}(Km)^{\alpha/2-1} \sum_{k=1}^K \sum_{l=1}^{m-1} [E|A_{ij}^1|^\alpha + E|A_{ij}^2|^\alpha + E|A_{ij}^3|^\alpha + E|A_{ij}^4|^\alpha + E|A_{ij}^5|^\alpha] \\
 &\leq C(K^{-\frac{\alpha}{2}}m^{\frac{\alpha}{2}}\Delta_r^{\alpha[\frac{1}{2}+l_1(\frac{1}{6}-\epsilon-\omega_1)]} + n^{\frac{\alpha}{2}}K^{-\alpha} + n^{-\frac{\alpha}{2}}K^{\frac{\alpha}{2}}) \\
 &\leq C(K^{-\frac{\alpha}{2}}m^{-l_1\alpha[\frac{1}{6}-(\epsilon+\omega_1)]} + n^{\frac{\alpha}{2}}K^{-\alpha} + n^{-\frac{\alpha}{2}}K^{\frac{\alpha}{2}}).
 \end{aligned}
 \tag{29}$$

We have  $l_1\alpha[\frac{1}{6} - (\epsilon + \omega_1)] > 0$ , because  $\omega_1 < \epsilon + \omega_1 < \omega_2 < \frac{1}{4}$ . Thus,

$$E|\tilde{\Gamma}_{ij}^T - \Gamma_{ij}|^\alpha \leq C(n^{\frac{\alpha}{2}}K^{-\alpha} + n^{-\frac{\alpha}{2}}K^{\frac{\alpha}{2}}).
 \tag{30}$$

If  $i = j$ , we have

$$\begin{aligned}
 \tilde{\Gamma}_{ii}^T - 2m\hat{\eta}_{ii} - \Gamma_{ii} &= \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y_i)^2 \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} - 2m\hat{\eta}_{ii} - \Gamma_{ii} \\
 &= \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} ((\Delta_l^k X_i')^2 \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} - (\Delta_l^k X_i')^2) \\
 &\quad + \frac{2}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k X_i') (\Delta_l^k \epsilon_i) \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} \\
 &\quad + \left( \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k \epsilon_i)^2 \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} - 2m\hat{\eta}_{ii} \right) \\
 &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k X_i'')^2 \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} \\
 &\quad + \frac{2}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k Y_i') (\Delta_l^k X_i'') \mathbf{1}_{\{|\Delta_l^k Y_i| \leq u_l\}} \\
 &\quad + \frac{1}{K} \sum_{k=1}^K \sum_{l=1}^{m-1} (\Delta_l^k X_i')^2 - \Gamma_{ii}.
 \end{aligned} \tag{31}$$

Through a long tedious process, we also have

$$E|\tilde{\Gamma}_{ii}^T - 2m\hat{\eta}_{ii} - \Gamma_{ii}|^\alpha \leq C(n^{\frac{\alpha}{2}} K^{-\alpha} + n^{-\frac{\alpha}{2}} K^{\frac{\alpha}{2}}). \tag{32}$$

Now, let  $\frac{K^2}{n} = O(\frac{n}{K})$ , then  $K = \lfloor n/m \rfloor \sim n^{2/3}$ , and the convergence rate  $C_n$  of the estimator is  $n^{-1/6}$ . This completes the proof of Theorem 1.  $\square$

**Proof of Theorem 2.** From the triangle inequality, we have

$$\begin{aligned}
 \|\hat{\Gamma}(\omega) - \Gamma\| &\leq \|\hat{\Gamma}(\omega) - \Gamma(\omega)\|_2 + \|\Gamma(\omega) - \Gamma\|_2 \\
 &\leq \|\hat{\Gamma}(\omega) - \Gamma(\omega)\|_\infty + \|\Gamma(\omega) - \Gamma\|_\infty.
 \end{aligned} \tag{33}$$

Next, Lemma 2 implies

$$\|\hat{\Gamma}(\omega) - \Gamma\|_\infty = \max_{1 \leq i \leq p} \sum_{j=1}^p |\Gamma_{ij}| \mathbf{1}\{|\Gamma_{ij}| \leq \omega\} = O_P(\pi(p)\omega^{1-\delta}). \tag{34}$$

By Lemmas 1 and 2,

$$\begin{aligned}
 &\|\hat{\Gamma}(\omega) - \Gamma(\omega)\|_\infty \\
 \leq &\max_{1 \leq i \leq p} \sum_{j=1}^p |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \mathbf{1}\{|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| \geq \omega\} + \max_{1 \leq i \leq p} \sum_{j=1}^p |\tilde{\Gamma}_{ij}^T| \mathbf{1}\{|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| < \omega\} \\
 &+ \max_{1 \leq i \leq p} \sum_{j=1}^p |\Gamma_{ij}| \mathbf{1}\{|\tilde{\Gamma}_{ij}^T| < \omega, |\Gamma_{ij}| \geq \omega\} \\
 \leq &\max_{1 \leq i, j \leq p} |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \max_{1 \leq i \leq p} \sum_{j=1}^p \mathbf{1}\{|\Gamma_{ij}| \geq \omega\} + \max_{1 \leq i \leq p} \sum_{j=1}^p |\Gamma_{ij}| \mathbf{1}\{|\Gamma_{ij}| < \omega\} \\
 &+ \max_{1 \leq i, j \leq p} |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \max_{1 \leq i \leq p} \sum_{j=1}^p \mathbf{1}\{|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| < \omega\} + \omega \max_{1 \leq i \leq p} \sum_{j=1}^p \mathbf{1}\{|\Gamma_{ij}| \geq \omega\} \\
 = &o_P(\omega) O_P(\pi(p)\omega^{-\delta}) + O_P(\pi(p)\omega^{1-\delta}) + o_P(\omega) O_P(\pi(p)\omega^{-\delta}) + \omega O_P(\pi(p)\omega^{-\delta}) \\
 = &O_P(\pi(p)\omega^{1-\delta}),
 \end{aligned} \tag{35}$$

which immediately shows that  $\|\Gamma(\omega) - \Gamma\|_\infty$  is of order  $\omega^{1-\delta}\pi(p)$ .  $\square$

**Lemma 1.** Under models (1)–(3) and Assumptions 1–4,

$$\max_{1 \leq i, j \leq p} |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| = O_P(C_n p^{2/\alpha}) = o_P(\omega), \tag{36}$$

$$P\left(\max_{1 \leq i \leq p} \sum_{j=1}^p 1\{|\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \geq \omega/2\} > 0\right) = o(1), \tag{37}$$

$$\max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| < \omega) \leq 2\delta M\pi(p)\omega^{-\delta} + o_P(1) = O_P(\pi(p)\omega^{-\delta}), \tag{38}$$

where  $\omega$  is chosen as in Theorem 2.

**Proof of Lemma 1.** Applying the Markov inequality, Theorem 1 and letting  $d = d_1 p^{2/\alpha} e_n$ , we obtain that, as  $n, p \rightarrow \infty$ ,

$$P\left(\max_{1 \leq i, j \leq p} |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| > d\right) \leq \sum_{i, j=1}^p P(|\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| > d) \leq \frac{C p^2 C_n^\alpha}{d^\alpha} = \frac{C}{d_1^\alpha} \rightarrow 0, \tag{39}$$

and then  $d_1 \rightarrow \infty$ . This proves (36).

Since  $h_{n,p} \rightarrow \infty$  as  $n, p \rightarrow \infty$ , using the above inequality,

$$P\left(\max_{1 \leq i \leq p} \sum_{j=1}^p 1\{|\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \geq \omega/2\} > 0\right) \leq P\left(\max_{1 \leq i, j \leq p} |\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \geq \omega/2\right) \leq \frac{2^\alpha C p^2 C_n^\alpha}{\omega^\alpha} = \frac{2^\alpha C}{h_{n,p}^\alpha} \rightarrow 0, \tag{40}$$

which proves (37).

Similarly, the inequality (38) can be obtained by

$$\begin{aligned} & \max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| < \omega) \\ \leq & \max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\tilde{\Gamma}_{ij}^T| \geq \omega, |\Gamma_{ij}| \leq \omega/2) + \max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\tilde{\Gamma}_{ij}^T| \geq \omega, \omega/2 < |\Gamma_{ij}| < \omega) \\ \leq & \max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\tilde{\Gamma}_{ij}^T - \Gamma_{ij}| \geq \omega/2) + \max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\Gamma_{ij}| > \omega/2) \\ \leq & o_P(1) + 2^\delta M\pi(p)\omega^{-\delta} = O_P(\pi(p)\omega^{-\delta}). \end{aligned} \tag{41}$$

$\square$

**Lemma 2 ([12]).** Under models (1)–(3) and Assumptions 1–4, and  $\omega$  is chosen as in Theorem 2. Then, for any fixed  $a > 0$ ,

$$\max_{1 \leq i \leq p} \sum_{j=1}^p |\Gamma_{ij}| 1(|\Gamma_{ij}| \leq a\omega) \leq a^{1-\delta} M\pi(p)\omega^{1-\delta} = O_P(\pi(p)\omega^{1-\delta}), \tag{42}$$

$$\max_{1 \leq i \leq p} \sum_{j=1}^p 1(|\Gamma_{ij}| \geq a\omega) \leq a^{-\delta} M\pi(p)\omega^{-\delta} = O_P(\pi(p)\omega^{-\delta}). \tag{43}$$

### 6. Conclusions

In this work, for cases when both the number of assets and the sample size approach infinity, we provide a new estimator for the large integrated volatility matrix, in the presence of jumps, micro-structural noise, and non-synchronization. First, we adopt the sub-sampling to synchronize the high-frequency data. Second, we use the two-time-scale realized co-volatility on the off-diagonal elements of  $\Gamma$  but correcting the bias on the diagonal elements of  $\Gamma$ , to handle noise. Finally, we employed the threshold parameters to remove the effect of jumps and sparsity in two steps. Both the minimax bound and the convergence rate were investigated. There are still some problems that we are eager to solve in the future research. First, we assume that entries of the integrated volatility matrix

are homogeneous under the sparse condition; however, the volatility of financial assets usually has entries with a very wide range of variability, which motivates us to extend the current work to a general frame. Second, heavy-tailed data are often encountered in financial engineering, which motivates us to develop estimation procedures to solve these issues.

**Author Contributions:** E.G. and C.L. conceived and designed the study; E.G., C.L. and F.T. developed the mathematical model; E.G., C.L. and F.T. conducted the computational implementation of the model and post-processing of the results; E.G., C.L. and F.T. wrote, reviewed, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Erlin Guo OF Jiangsu Province grant number BY2022768, Cuixia Li OF Qing Lan Project of Jiangsu Province and Jiangsu Province grant number BY2022743 and Fengqin Tang OF NSF China grant number 12201235 and Anhui Province grant number 2108085QA14.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the editor, associate editor, and referees.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Andersen, T.G.; Bollerslev, T.; Meddahi, N. Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica* **2005**, *73*, 279–296. [\[CrossRef\]](#)
- Barndorff-Nielsen, O.E.; Shephard, N. Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* **2004**, *72*, 885–925. [\[CrossRef\]](#)
- Todorov, V. Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *J. Econometrics* **2009**, *148*, 131–148. [\[CrossRef\]](#)
- Brogaard, J.A.; Hendershott, T.; Riordan, R. High frequency trading and price discovery. *Rev. Financ. Stud.* **2014**, *27*, 2267–2306. [\[CrossRef\]](#)
- Dai, C.; Lu, K.; Xiu, D. Knowing factors or factor loadings, or neither? Evaluating estimators of large covariance matrices with noisy and asynchronous data. *J. Econometrics* **2019**, *208*, 43–79. [\[CrossRef\]](#)
- Black, F. Noise. *J. Financ.* **1986**, *41*, 529–543. [\[CrossRef\]](#)
- Hayashi, T.; Yoshida, N. On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* **2005**, *11*, 359–379. [\[CrossRef\]](#)
- Zhang, L. Estimating covariation: Epps effect, microstructure noise. *J. Econom.* **2011**, *160*, 33–47. [\[CrossRef\]](#)
- Ait-Sahalia, Y.; Fan, J.; Xiu, D. High-frequency covariance estimates with noisy and asynchronous financial data. *J. Am. Stat. Assoc.* **2010**, *105*, 1504–1517. [\[CrossRef\]](#)
- Barndorff-Nielsen, O.; Hansen, P.; Lunde, A.; Shephard, N. Multivariate realized kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and nonsynchronous trading. *J. Econom.* **2011**, *162*, 149–169. [\[CrossRef\]](#)
- Christensen, K.; Kinnebrock, S.; Podolskij, M. Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *J. Econom.* **2010**, *159*, 116–133. [\[CrossRef\]](#)
- Wang, Y.; Zou, J. Vast volatility matrix estimation for high-frequency financial data. *Ann. Stat.* **2010**, *38*, 943–978. [\[CrossRef\]](#)
- Tao, M.; Wang, Y.; Zhou, H. Optimal sparse volatility matrix estimation for high-dimensional Itô process with measurement error. *Ann. Stat.* **2013**, *41*, 1816–1864. [\[CrossRef\]](#)
- Kim, D.; Wang, Y.; Zou, J. Asymptotic theory for large volatility matrix estimation based on high-frequency financial data. *Stoch. Proc. Appl.* **2016**, *126*, 3527–3577. [\[CrossRef\]](#)
- Kim, D.; Kong, X.; Li, C.; Wang, Y. Adaptive thresholding for large volatility matrix estimation based on high-frequency financial data. *J. Econom.* **2018**, *203*, 69–79. [\[CrossRef\]](#)
- Jing, B.; Liu, Z.; Kong, X. On the estimation of integrated volatility with jumps and microstructure noise. *J. Bus. Econ. Stat.* **2014**, *3*, 457–467. [\[CrossRef\]](#)
- Jing, B.; Kong, X.; Liu, Z. Modeling high-frequency financial data by pure jump processes. *Ann. Stat.* **2012**, *40*, 759–784. [\[CrossRef\]](#)
- Kong, X.; Liu, Z.; Jing, B. Testing for pure-jump processes for high-frequency data. *Ann. Stat.* **2015**, *43*, 847–877. [\[CrossRef\]](#)
- Li, C.; Chen, J.; Liu, Z.; Jing, B. On integrated volatility of Itô semimartingales when sampling times are endogenous. *Commun. Stat.-Theor. Methods* **2014**, *43*, 5263–5275. [\[CrossRef\]](#)
- Liu, Z. Jump-robust estimation of volatility with simultaneous presence of microstructure noise and multiple observations. *Financ. Stoch.* **2017**, *21*, 427–469. [\[CrossRef\]](#)
- Li, C.; Guo, E. Estimation of the integrated volatility using noisy high-frequency data with jumps and endogeneity. *Commun. Stat.-Theor. Methods* **2018**, *3*, 521–531. [\[CrossRef\]](#)

22. Jing, B.; Li, C.; Liu, Z. On estimating the integrated co-volatility using noisy high-frequency data with jumps. *Commun. Stat.-Theor. Methods* **2013**, *42*, 3889–3901. [[CrossRef](#)]
23. Yang, C.; Zhang, J.; Huang, Z. Numerical study on cavitation-vortex-noise correlation mechanism and dynamic mode decomposition of a hydrofoil. *Phys. Fluids* **2022**, *34*, 125105. [[CrossRef](#)]
24. Li, R.; Zhang, H.; Chen, Z.; Yu, N.; Kong, W.; Li, T.; Wang, E.; Wu, X.; Liu, Y. Denoising method of ground-penetrating radar signal based on independent component analysis with multifractal spectrum. *Measurement* **2022**, *192*, 110886. [[CrossRef](#)]
25. Yu, J.; Lu, L.; Chen, Y.; Zhu, Y.; Kong, L. An indirect eavesdropping attack of keystrokes on touch screen through acoustic sensing. *IEEE Trans. Mob. Comput.* **2021**, *20*, 337–351. [[CrossRef](#)]
26. Lu, S.; Yang, B.; Xiao, Y.; Liu, S.; Liu, M.; Yin, L.; Zheng, W. Iterative reconstruction of low-dose CT based on differential sparse. *Biomed. Signal. Proces.* **2023**, *79*, 104204. [[CrossRef](#)]
27. Ye, R.; Liu, P.; Shi, K.; Yan, B. State damping control: a novel simple method of rotor UAV with high performance. *IEEE Access.* **2020**, *8*, 214346–214357. [[CrossRef](#)]
28. Jin, C.; Tsai, F.; Gu, Q.; Wu, B. Does the porter hypothesis work well in the emission trading schema pilot? Exploring moderating effects of institutional settings. *Res. Int. Bus. Financ.* **2022**, *62*, 101732. [[CrossRef](#)]
29. Zhong, T.; Wang, W.; Lu, S.; Dong, X.; Yang, B. RMCHN: A residual modular cascaded heterogeneous network for noise suppression in DAS-VSP records. *IEEE Geosci. Remote Sens.* **2023**, *20*, 7500205. [[CrossRef](#)]
30. Delbaen, F.; Schachermayer, W. A general version of the fundamental theorem of asset pricing. *Math. Ann.* **1994**, *300*, 463–520. [[CrossRef](#)]
31. Zhang, L.; Mykland, P.; Ait-Sahalia, Y. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Am. Stat. Assoc.* **2005**, *100*, 1394–1911. [[CrossRef](#)]
32. Bickel, P.; Levina, E. Covariance regularization by thresholding. *Ann. Stat.* **2008**, *36*, 2577–2604. [[CrossRef](#)] [[PubMed](#)]
33. Lam, C.; Fan, J. Sparsistency and rates of convergence in large covariance matrix estimation. *Ann. Stat.* **2009**, *37*, 4254–4278. [[CrossRef](#)]
34. El Karoui, N. High-dimensionality effects in the markowitz problem and other quadratic programs with linear constraints: Risk underestimation. *Ann. Stat.* **2010**, *38*, 3487–3566. [[CrossRef](#)]
35. Rigollet, P.; Tsybakov, A. Estimation of covariance matrices under sparsity constraints. *Stat. Sin.* **2012**, *22*, 1319–1378.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.