arXiv:2304.04372v1 [stat.ME] 10 Apr 2023

瞬时方差-协方差矩阵的对称半正定傅里叶估计量

Jirô Akahori*、Nien-Lin Liu†、Maria Elvira Mancino‡、Tommaso Mariotti§、Yukie Yasuda*

2023年4月11日

摘要

本文提出了一种即期协方差矩阵估计量，可确保估计结果具有对称半正定性。所提出的估计量基于对 Malliavin 和 Mancino（2009）中的傅里叶协方差估计量进行适当修改，并且在权重核取适当形式时具有一致性。我们通过广泛的数值研究，将该估计量的准确性及其生成半正定协方差矩阵的能力与文献中的其他方法进行比较和评估。模拟研究的结果在多种情形下均得到验证，这些情形考虑了问题的维数、数据的异步性以及多种市场微观结构噪声设定的存在。

1 引言

实证研究指出，考虑资产价格相关性中不同的时间变化非常重要。因此，在过去几年中，多项研究针对如何利用在不同资产之间异步采样的高频数据，高效估计协方差这一问题展开了研究。尽管关于积分协方差估计的文献日益丰富，但关于即期协方差估计的研究仍然较为稀少。Malliavin和Mancino（2002）较早提出了利用异步高频数据处理即期协方差估计的方法。与其他依赖数据预处理以使其同步的估计量不同，例如线性插值、分段常数（前一笔）插值，或Barndorff-Nielsen 等人（2011）提出的刷新时间程序，傅里叶估计量基于积分程序，使用所有可用数据。使用全部数据、避免对数据进行任何初步处理（例如预平均，参见 Aït-Sahalia 和 Jacod（2014年））的可能性，使得在多元情形下可以直接使用不规则采样的收益率，甚至使用异步数据。Hayashi和Yoshida（2005年）提出的估计量通过以下方式规避异步性所造成的不利影响：仅当相应的观测区间存在重叠时，才考虑价格增量乘积的贡献。Christensen等人（2010）提出的预平均全重叠估计量建立在相同思想之上，旨在使Hayashi和Yoshida（2005年）提出的估计量对微观结构噪声具有稳健性。

*立命馆大学
†东京理科大学管理学院，电子邮件：nl-liu@rs.tus.ac.jp
‡佛罗伦萨大学经济与管理系§比萨高等师范学校，电子邮箱：tommaso.mariotti@sns.it

从实用角度来看，选择使用哪些估计量不应仅基于其渐近分布的收敛速率。例如，所估计的协方差矩阵能否保持其对称性和半正定性是一个首要问题。这一性质在多个背景下都具有重要影响，例如近期发展起来的基于高频数据的主成分分析领域（Liu 和 Ngo（2017）、Aït-Sahalia 和 Xiu（2019年）、Chen 等人（2020年））或资产配置框架（例如，参见 Engle 和 Colacito（2006））。尽管一些作者已经针对积分协方差估计量讨论了这一问题（例如，参见 Barndorff-Nielsen 等人（2011）、Mancino 和 Sanfelici（2011）、Park 等人（2016）、Cui 等人（2019年）），但据我们所知，本文提出的即期协方差估计量是第一个能够保证估计结果本身具有半正定性的估计量，而这一问题迄今尚未在文献中得到讨论。例如，在处理现货波动率时，Chen 等人（2020年）会先对估计结果进行积分，再计算协方差矩阵的特征值；而在 Bu 等人（2022年）中，则通过对估计结果应用适当的收缩技术来强制实现半正定性，从而引入了对所估计矩阵的调整。

本文旨在提出一种新颖的、对称且正定的即期协方差估计量，并在模拟环境中分析其有限样本性质。我们的出发点是Malliavin和 Mancino（2009）提出的即期傅里叶估计量。然而，由于 Fejèr 核缺乏对称性，当资产价格在异步网格上被观测时，该估计量也可能无法提供半正定估计。为确保估计结果具有对称性和半正定性，本文引入了傅里叶估计量的一个修正版，并将其称为正定傅里叶估计量（以下简称 PDF 估计量）；我们证明，该估计量确实满足所需性质。

所提出的估计量依赖于两个参数：截断频率 N 和局部化参数 M。我们通过模拟研究来评估如何最优选择这两个参数以最小化误差：针对有效价格过程和加性微观结构噪声的若干不同模型设定，对一组可能的取值进行测试。我们发现，不同波动率过程模型下的最优截断频率 N 保持稳定；在不存在噪声时，应选择 $N = \frac{1}{2}n^{3/4}$，而在存在噪声时， $N = \frac{1}{2}n^{2/3}$ 更为合适，这与其他傅里叶型估计量的情形一致。在这两种情况下，选择 $N < n/2$（即低于奈奎斯特频率）都会降低误差，同时减少估计量所需的计算量。参数 M 似乎不受微观结构噪声的显著影响；其最优取值取决于波动率模型，可能为 $M = N^{4/9}$ 或 $M = N^{1/2}$ 。在本研究分析的情形中，这两种选择之间的差异很小。

此外，还通过专门的模拟实验进一步研究了 N 的敏感性，具体分析了引入异步性与同步采样情形相比所产生的影响。上述对 N 的选择被证实优于奈奎斯特频率，因为当两种资产之间的相关性小于 1 时，在存在异步性的情况下，随着 N 取较大值，绝对偏差和均方误差都会增加；不过，这也表明存在一个 N 取值区间，其表现基本不受影响。

随后，为评估所提出的概率密度函数估计量的有限样本性能，我们比较其准确性以及它能够生成的半正定估计所占比例，并将其与采用 Mykland 等人（2019）的平滑双尺度估计量和 Bibinger 等人（2019）的局部矩估计量得到的结果进行比较；这两种估计量都能够处理

异步观测。在开展这项比较时，我们着重考察可能影响利用高频数据估计方差-协方差矩阵的主要问题。首先，我们考虑不存在市场微观结构噪声的基准情形；其次，我们关注异步程度，考察驱动观测频率的泊松过程具有不同强度时的情况；最后，我们分析市场微观结构噪声的影响，考察取整产生的噪声、独立同分布噪声、自相关噪声、与有效价格过程相关的噪声以及异方差噪声。结果表明，在这项研究中，概率密度函数估计量是唯一能够在 100% 的情况下持续生成半正定估计，同时在均方误差方面保持与竞争方法相当表现的估计量。此外，尽管文献中主要结合主成分分析来讨论维数问题，我们通过比较考察维数增加是否会改变特征值的经验分布，以至于使估计量失去生成半正定矩阵的能力。结果显示，随着所考虑的资产数量增加，竞争估计量生成的半正定矩阵比例会降低。

通过改变分析中采用的模拟模型，所有模拟结果的稳健性均得到验证；具体而言，我们考察了 Heston 随机波动率模型 Heston（1993）、单因子波动率模型、双因子波动率模型（Chernov等（2003））以及粗糙Heston模型（El Euch 和 Rosenbaum（2019）），在每种情况下都得到了相近的结果。

本文其余部分安排如下。第 2 节介绍即期协方差的半正定（PDF）傅里叶估计量，并证明其正定性。第 3 节包含模拟研究，其中对所提出估计量的参数进行敏感性分析，以估计的均方误差为指标；此外，还将所提出的估计量与文献中的其他估计量进行比较，考察其准确性以及生成半正定矩阵的能力。第 4 节作总结。

2 半正定估计量

令 $X=(X^1,\ .....,\ X^d)$ 为一个 d 维伊藤半鞅

$$X_t^j = x_0^j + \int_0^t b_j(s)ds + \sum_{k=1}^{d}\int_0^t \sigma_{jk}(s)dW_s^k, \quad j,k=1,...,d$$

其中 $W=(W^1,...,W^d)$ 是滤过概率空间 $(\Omega,(\mathcal{F}_t)_{t\in[0,T]},P)$ 上的 d 维布朗运动，且 b_j and σ_{jk} 是适应过程。

在本文中，我们关注于 $d\times d$瞬时（现货）协方差矩阵 $V(t)$ 的非参数估计，其元素为

$$V^{j,j'}(t) := \sum_{k=1}^{d}\sigma_{jk}(t)\sigma_{j'k}(t), \quad \text{for} \quad j,j'=1,...,d \quad \text{and} \quad t\in[0,T].$$

我们假设价格是在离散、不规则且异步的时间网格上观测到的。

$$0=t_0^j<t_1^j<,...,<t_{n_j}^j=2\pi \quad \text{for } j=1,...,d,$$

其中，假设 $T=2\pi$ 是为了简化记号而作出的，并不构成限制。在

下文中， ΔX_l^j 表示 $j = 1,...,d$ 和 $l = 1,...,n_j$ 对应的离散收益率 $X_l^j - X_{l-1}^j$。在此设定下，我们提出以下即期协方差估计量。

定义 2.1。 设 $\mathcal{K}$ 是 $\mathbb{Z}$ 的有限子集，$\mathcal{S} \coloneqq \{\mathcal{S}(k) \subset_{finite} \mathbb{Z}^2 \colon k \in \mathcal{K}, (s, s') \in \mathcal{S}(k) \implies s + s' = k\}$，且 c 是定义在 $\mathcal{K}$ 上的复函数；我们将 $V^{j,j'}(t)$ 的估计量定义为：

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t) = \sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}}\sum_{k\in\mathcal{K}} c(k)e^{ikt} \sum_{(s,s')\in\mathcal{S}} e^{-ist_l^j} e^{is't_{l'}^{j'}} \Delta X_l^j \Delta X_{l'}^{j'}. \tag{1}$$

注 2.2。 若取 $\mathcal{K} = \{0, \pm1, \pm2, ..., \pm M\}$，其中 M 为正整数，且 $\mathcal{S}(k) = \{(s, s') | s+s' = k, s = 0, \pm1, \pm2, ..., \pm N\}$，其中 N 为正整数，则

$$c(k) = \frac{1}{2\pi}\left(1 - \frac{|k|}{M+1}\right)\frac{1}{2N+1},$$

我们得到：

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t) = \sum_{k=-M}^{M}\left(1 - \frac{|k|}{M+1}\right)e^{ikt} \sum_{s=-N}^{N}\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}} e^{-ist_l^j} e^{i(k-s)t_{l'}^{j'}} \Delta X_l^j \Delta X_{l'}^{j'}$$

利用定义为 $D_N(x) = \sum_{k=-N}^{N} e^{ikx}$ 和 $F_M(x) = \sum_{k=-M}^{M}\left(1 - \frac{|k|}{M+1}\right)e^{ikx}$ 的狄利克雷核与 *Fejèr* 核，可将该估计量写为：

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t) = \frac{1}{2\pi}\frac{1}{2N+1}\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}} F_M(t - t_l^j) D_N(t_l^j - t_{l'}^{j'}) \Delta X_l^j \Delta X_{l'}^{j'},$$

并且估计量 *(1)* 与 *Malliavin* 和 *Mancino*（*2009*）提出的傅里叶即期协方差估计量一致，其渐近性质已在 *Mancino* 和 *Recchioni*（*2015*）（无噪声时）以及 *Mancino*、*Mariotti*等（*2022*）（有噪声时）中得到研究。然而，尽管 *Mancino* 和 *Sanfelici*（*2011*）证明了相应的积分协方差矩阵估计量具有正性，即期协方差估计量却可能无法生成对称半正定估计，因为 $F_M(t - t_l^j) D_N(t_l^j - t_{l'}^{j'})$ 在 j, j' 中并不对称，从而导致 $\hat{}V_{\mathcal{K},\mathcal{S}}(t)$ 中出现复特征值。

本研究的主要理论结果涉及所提出估计量的半正定性，并在以下定理中给出。

定理 2.3。 假设 $\mathcal{K} = \{0, \pm1, \pm2, ..., \pm2N\}$ 对于某个正整数 N, $c(k)$ 是 $\mathcal{K}$ 上的半正定函数，并且

$$\mathcal{S}(k) = \begin{cases} \{(-N + k + v, N - v) : v = 0, ..., 2N - k\} & 0 \le k \le 2N \\ \{(N + k - v, -N + v) : v = 0, ..., 2N + k\} & -2N \le k < 0. \end{cases}$$

于是，$\hat{}V_{\mathcal{K},\mathcal{S}}(t)$ 在（*1*）中定义的是对称且半正定的。

定理 2.3 的证明见附录 A。

此外，可以发现 $\hat{}V_{\mathcal{K},\mathcal{S}}(t)$ 可以改写为：

$$\hat{V}_N^{j,j'}(t) = \frac{1}{2\pi}\frac{1}{2N+1}\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}}\sum_{u=-N}^{N}\sum_{u'=-N}^{N} c(u-u')e^{iu(t-t_l^j)}e^{-iu'(t-t_{l'}^{j'})}\Delta(X_l^j)\Delta(X_{l'}^{j'}), \qquad (2)$$

对于两个资产 j 和 j' 和 $t \in (0, 2\pi)$，其中 $c(k)$ 仍然是一个半正定函数。

注 2.4。 如 *Akahori* 等人（*2023*）所证明，只要恰当选择函数 $c(k)$，上述半正定傅里叶（*PDF*）估计量就是即期协方差的一致估计量。特别是，高斯核足以保证一致性，而我们对$t \in (0, 2\pi)$ 的选择为：

$$c(k) = e^{-\frac{2\pi^2k^2}{M}}, \quad with \quad M \to \infty \quad as \quad N \to \infty.$$

为强调局部化参数 M 的存在，方程 *(2)* 左侧可写为 $\hat{}V^{j,N}{}_M^{j'},(t)$；关于所提出的估计量对参数 N 和 M 取值的敏感性分析见第 *3.2* 节和第 *3.2.1* 节。

注 2.5。 根据博赫纳定理，对于每个半正定函数 c，都存在一个有界测度 μ 定义在$\mathbf{R}$ 上，使得

$$c(x) = \int_{\mathbf{R}} e^{2\pi iyx}\mu(dy).$$

因此，我们也可以用测度 μ 代替半正定函数 $c(k)$，改写概率密度函数估计量 *(2)*，并得到

$$\hat{V}_{N,M}^{j,j'}(t) = \frac{1}{2\pi(2N+1)}\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}}\int_{\mathbf{R}} D_N(t-t_l^j+y)D_N(t-t_{l'}^{j'}+y)\mu_M(dy)\Delta(X_l^j)\Delta(X_{l'}^{j'}).$$

密度的选择应与备注 *2.4* 中讨论的 c（k）保持一致；有关对 μ 假设的更详细说明，请参见 *Akahori* 等人（*2023* 年）。

3 模拟研究

3.1 模拟设置

本节将介绍一项广泛的数值模拟研究。本研究有两个目的：首先，在 3.2 节中，我们分析估计量对参数 N 和 M 选择的敏感性，并通过不可行的优化探究在不同情景下最优选择的变化程度。其次，在 3.3 节中，我们将所提出的 PDF 估计量与文献中的另外两种替代估计量进行比较，以评估其准确性，以及生成对称且半正定估计的能力。

为增强我们研究结果的稳健性，在上述练习中，我们考虑了许多不同的模拟情景，重点关注高频金融数据的两个组成部分：有效价格和由市场微观结构产生的加性噪声成分，因此˜观测价格 X 为：

$$\tilde{X}_t^j = X_t^j + \eta_t^j, \qquad j = 1, ..., d \qquad (3)$$

其中， η 为噪声成分。

特别地，对于有效价格过程，我们考虑以下设定：

- Heston（1993）提出的Heston随机波动率模型；
- 例如Huang和Tauchen（2005年）所使用的单因子随机波动率模型（SVF1）；
- Chernov等（2003）的双因子随机波动率模型（SVF2）；
- El Euch 和 Rosenbaum（2019）提出的粗糙Heston模型（RH）；

而对于加性微观结构噪声，我们考虑以下几种情况：

- 无噪声；
- 由取整产生的噪声；
- 独立同分布噪声；
- 自相关噪声；
- 与有效价格过程相关的噪声；
- 异方差噪声。

由于在存在噪声的不同情形下，我们分别分析了取整、噪声强度、噪声自相关、噪声相关性和最大噪声强度的 2、4、3、3 和 3 个不同水平，因此在模拟分析中，我们总共研究了 64 种不同情景。

为简化计算，在第 3.2 至 3.3 节中，所有模拟分析均在区间 $[0,1]$ 上进行；因此，根据注释 2.4，概率密度函数估计量为：

$$\hat{V}^{j,j'}_{N,M}(t)=\frac{1}{2N+1}\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}}\sum_{u=-N}^{N}\sum_{u'=-N}^{N}e^{-\frac{2\pi^2(u-u')^2}{M}}e^{iu(t-t^j_l)}e^{-iu'(t-t^{j'}_{l'})}\Delta(X^j_l)\Delta(X^{j'}_{l'}),\qquad(4)$$

除非另有说明，模拟包含 $K=500$ 条日内轨迹，交易日长度设为 6.5 小时，并在间距为 2 秒的等距网格上进行。为在数据中引入异步性，观测取自强度为$\Delta t=10$的泊松过程，即平均每 10 秒产生一次观测的过程。此外，除非另有明确说明，根据 Bibinger 等人（2019），驱动不同资产有效过程的布朗运动之间的相关性固定为纳斯达克成分股已估计实现相关性的中位数，即：

$$\left\langle W^j,W^i\right\rangle=0.312,\qquad j,i=1,...,d,\quad j\neq i.$$

在第3.1.1节和第3.1.2节中，我们定义了用于有效价格过程和市场微观结构噪声的模型。

3.1.1 有效价格过程

Heston模型

Heston（1993）提出的Heston随机波动率模型可能是高频计量经济学文献中使用最广泛的随机波动率模型。其形式为:

$$\begin{cases} dX_t^j & = (\mu - (\sigma_t^j)^2/2)dt + \sigma_t^j dW_t^j \\ d(\sigma_t^j)^2 & = \gamma(\theta - (\sigma_t^j)^2)dt + \nu\sigma_t^j dZ_t^j, \end{cases}$$

其中，$\langle W^j, Z^j \rangle = \lambda$用于考虑杠杆效应，$\langle W^j, Z^i \rangle = 0$用于 $i \neq j$。

参数设定如下:

$$(\mu, \gamma, \theta, \nu, \lambda) = (0.05/252, 5/252, 0.1, 0.5/252, -0.5),$$

这与Zu和Boswijk（2014年）、Mancino和Recchioni（2015年）以及Figueroa-López和Wu（2022年）所采用的设定相同。

因子波动率模型

因子波动率模型在文献中已被长期采用，例如 Huang 和 Tauchen（2005年）。首先，我们考虑以下形式的单因子随机波动率模型（SV1F）:

$$\begin{cases} dX_t^j & = \mu dt + \sigma_t^j dW_t^j \\ \sigma_t^j & = \mathrm{e}^{\beta_0 + \beta_1 \tau_t^j} \\ d\tau_t^j & = \alpha\tau_t^j + dZ_t^j \end{cases}$$

对于 $j = 1, \ldots\ldots, d$，其中 $\langle W^j, \ Z^j \rangle = \lambda$，且 $\left\langle Z^j, \ Z^{j'} \right\rangle = 0$对于 $j \neq j'$。仿真采用以下参数:

$$(\mu, \beta_1, \alpha, \beta_0, \lambda) = (0.03, 0.125, -0.025, \beta_1/(2\alpha), -0.3)$$

这些参数也用于 Zu和Boswijk（2014年）、Mancino 和 Recchioni（2015年）、Figueroa-López和 Wu（2022年）以及 Mancino、Mariotti等（2022年）的研究。

其次，我们考虑 Chernov等（2003）提出的双因子随机波动率模型（SV2F），该模型能够再现更高的波动率的波动值。其形式为:

$$\begin{cases} dX_t^j & = \mu dt + \text{s-exp}[\beta_0 + \beta_1 \tau_t^{j,1} + \beta_2 \tau_t^{j,2}] dW_t^j \\ d\tau_t^{j,1} & = \alpha_1 \tau_t^{j,1} + dZ_t^{j,1} \\ d\tau_t^{j,2} & = \alpha_2 \tau_t^{j,2} + (1 + \beta_v \tau_t^{j,2}) dZ_t^{j,2} \end{cases}$$

with

$$\text{s-exp(x)} = \begin{cases} \exp(x), & \text{if} \quad x \leq x_0 = \log(1.5) \\ \frac{e^{x_0}}{\sqrt{x_0}} \sqrt{x_0 - x_0^2 + x^2}, & \text{otherwise} \end{cases}$$

对于 $j =$ 1, ..., d，其中 $\langle W^j, Z^{j,1} \rangle = \langle W^j, Z^{j,2} \rangle = \lambda$，且 $\left\langle Z^{j,i}, Z^{j',i'} \right\rangle = 0$对于 $j \neq j'$或 $i \neq i'$，

$i, i' = 1,2$. 对于参数，我们选择使用：

$$(\mu, \beta_0, \beta_1, \beta_2, \beta_v, \alpha_1, \alpha_2, \lambda) = (0.03, -1.1, 0.04, 0.3, -0.003, -0.6, 0.25).$$

粗糙波动率

金融计量经济学文献中兴起了一个新的研究方向，探讨并非由标准布朗运动驱动、而是由分数布朗运动驱动的波动率过程动态，其中赫斯特指数为 $H < 0.5$；参见 Alòs 等人（2007）、Gatheral 等人（2018）以及 Euch 和 Rosenbaum（2018）。即使存在粗糙波动率，所提出的概率密度函数估计量仍具有一致性，参见 Akahori 等人（2023）。

粗糙波动率也可以通过随机 Volterra 方程建模，例如 El Euch 和 Rosenbaum（2019年）研究的粗糙Heston模型，我们打算在本分析中采用该模型：

$$\begin{cases} X_t^j & = X_0^j + \int_0^t X_t^j \sigma_s^j dW_s^j \\ (\sigma_t^j)^2 & = (\sigma_0^j)^2 + \int_0^t K(t-s)\left((\theta - \gamma(\sigma_s^j)^2)ds + \nu\sigma_s^j dZ_s^j\right) \end{cases}$$

其中$\left\langle W^j, Z^j \right\rangle = \lambda$和 $K(t) = Ct^{H-\frac{1}{2}}$对应赫斯特指数 $H \in \left(0, \frac{1}{2}\right)$ 以及常数 C。为模拟粗糙Heston模型，我们应用Richard等人（2021）研究的离散时间欧拉型格式，特别参照其中的公式（8）。模型的参数经过设定，以确保在本练习中模拟的波动率过程不会出现负值，具体取值为：

$$(\theta, \gamma, \nu, \lambda, H) = (0.2, 0.3, 0.2, -0.7, 0.1),$$

其中，赫斯特参数的选择由文献中现有的经验证据所驱动；例如参见 Gatheral 等人（2018）。

3.1.2 市场微观结构噪声设定

众所周知（例如，参见 Bandi 和Russell（2008）），高频数据会受到所谓的市场微观结构噪声的污染。具体而言，对观测价格的一种常见表示是将有效价格与噪声成分相加。噪声的来源与金融市场微观结构的某些特征有关，例如买卖价差、取整、策略性交易（例如，参见 Hasbrouck（2007）），并且高频金融计量经济学文献已经提出了关于噪声的若干模型设定。

由取整引起的噪声

在存在取整的情况下，观测价格过程具有以下形式：

$$\tilde{X}_t^j = \log([\exp(X_t^j)/r]r),$$

其中， X 表示有效价格过程。我们考虑两种取整精度，分别对应$r = 1$ 或 5 美分，这是金融市场中最常用的精度。

独立同分布噪声

对噪声最广泛采用的刻画方式，是将其视为一个均值为零、具有给定方差的独立同分布加性成分：

$$\tilde{X}_t^j = X_t^j + \eta_t^j, \tag{5}$$

$$\eta^j \sim i.i.d. \quad E[\eta^j] = 0, \quad E[(\eta^j)^2] = var(\Delta X_{10sec}^j)\sigma_\eta^2,$$

其中， X_{10sec}^j 表示每隔 10 秒对模拟序列进行采样后得到的等间隔序列。

这里，我们为噪声指定高斯分布，并将其方差设为与¯连续两次观测之间的平均时间 Δt成正比，以避免噪声被下采样效应抹去。我们考虑噪声方差的四种取值： $\sigma_\eta^2 = 1$，1。5，2，2。5。

自相关噪声。

在这种情况下，噪声成分中引入了自相关；在保持方程（5）的加性形式的同时，噪声通过如下定义的奥恩斯坦-乌伦贝克过程进行建模：

$$d\eta_t^j = -\theta_\eta \eta_t^j dt + \sigma_\eta dE_t^j,$$

其中， E 是与 W相互独立的标准布朗运动。我们考虑三个不同的自相关水平，分别取 $\theta_\eta = 0.2$，0。3，0。4。 σ_η^2 的设定使其方差水平与上一节第二种情形相同。

与有效价格过程相关的噪声

对前述噪声设定的一种自然推广，是允许驱动噪声成分的布朗运动与驱动有效价格动态的布朗运动之间存在相关性：

$$d\eta_t^j = -\theta_\eta \eta_t^j dt + \sigma_\eta dE_t^j, \qquad \left\langle W^j, E^j \right\rangle = \rho_\eta$$

其中 ρ_η 是一个常数。鉴于文献中已有的经验证据（例如，参见 Hansen 和 Lunde（2006）），我们设定 ρ_η 为负值，考虑三个取值：-0.1、-0.3 和 -0.5；此外，我们考虑的噪声水平与 i.i.d. 情形下噪声信号比为 2 时的水平相同，并假设噪声过程不存在自相关。

异方差噪声

在 MMN 的最后一种表述中，我们允许噪声的方差在该时间区间内随时间变化。具体而言，为了再现 Kalnina 和 Linton（2008）观察到的程式化事实，即交易日开始和结束时的噪声较高，我们遵循 Bu 等人（2022年）的做法，使用以下函数：

$$\sigma_\eta(t) = \bar{\sigma}_\eta \left(\frac{1}{2}(\cos(2\pi t) + 1)0.9 + 0.1 \right)$$

选择 $\bar{\sigma}_\eta = 3$，3。5，4。对于噪声，在这种情况下我们仍采用上文所述的奥恩斯坦-乌伦贝克过程表述，同时考虑自相关以及噪声与有效价格过程之间的依赖关系，其中有 $\theta = 0.3$ 和 $\rho_n = -0.3$。

这一最后的表述允许采用相当一般形式的微观结构噪声，其中包含文献中观察到的程式化事实。在最后一项未报告的实验中，我们还在这一模拟方案中加入了1美分的取整，结果没有发生显著变化。

3.2 频率 N 和 M的选择

本节旨在评估估计量对概率密度函数估计量定义中所出现的频率 N 和 M 选择的敏感性。为在整个时间区间内搜索最优参数，我们评估了每一组参数对应的估计量性能，所依据的是在不可行的优化练习中所考虑的 K 条模拟独立轨迹。在以下所有分析中，现货方差轨迹均在宽度为 20 分钟的规则网格上进行重构。

在本优化研究中，我们令截断频率 N 的取值取决于 n，而定位参数 M 的取值取决于其自身 N，从而构建一个由 (N，M) 组成的网格，其取值为：

$$N = \frac{n^\alpha}{2}, \qquad \alpha = 1, \frac{5}{6}, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}, \frac{1}{3},$$
$$M = N^\beta, \qquad \beta = \frac{5}{6}, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}, \frac{4}{9},$$

其中，N 可能取值的选择依据的是一个已知事实：与所有其他傅里叶型估计量一样，为避免混叠效应，截断频率必须小于奈奎斯特频率，即 $N \leq n/2$。

对于每种情形，在具有 $d = 2$ 的设定下，我们在上述取值网格上考察方差 $\hat{}V^{1,1}$ 和协方差 $\hat{}V^{1,2}$估计中的误差，具体使用积分误差：

$$MISE_j = (K)^{-1} \sum_{k=1}^{K} \int_0^1 (\hat{V}_k^{1,j} - V_k^{1,j})^2 dt, \qquad j = 1, 2$$

并选择使下式最小的参数对

$$0.1 \cdot MISE_1 + 0.9 \cdot MISE_2,$$

其中，对协方差估计赋予更高的权重，因为它是一般方差—协方差矩阵中的主导部分。

表1展示了每种情景下 N 和 M 的最优组合。首先需要注意的是，对于有效价格过程的不同模型，N 的最优值似乎相当稳定，并且始终小于奈奎斯特频率 $n/2$，即使在不存在噪声的情况下也是如此。当数据受到噪声影响更严重时，最优 N 会更小，这与已知的关于积分波动率和现货波动率的其他傅里叶型估计量的结论一致，例如参见 Mancino、Recchioni 和 Sanfelici（2017年）。此外，N的选择虽然对噪声强度较为敏感，但似乎并不取决于噪声的相关结构。至于参数 M，在本次考察的所有情景中，其最优值似乎几乎相同。此外，

对于有效价格的四种模型， $M = N^{4/9}$和 $M = N^{1/2}$之间的差异似乎可以忽略不计，如表2所示；表2报告了选定情景下定义网格上的 $MISE_2$ 。从表2还可以清楚看出，总体而言，该估计量对 N 选择的敏感性远高于对 M选择的敏感性，前者的变化会导致估计误差发生更大的变动；尽管为简便起见，表2仅展示了 Heston 和 SVF2 模型的数值，但其余模型的表现也是类似的。需要强调的是，与 Mancino、Mariotti等（2022年）对原始现货傅里叶估计量的观察一致，通过适当选择截断频率 N，可以消除噪声。事实上， N控制着卷积中所考虑的资产收益率傅里叶系数的数量，从而获得协方差过程的傅里叶系数。众所周知，通过减少频率 N的数量，可以滤除噪声（参见 Mancino 和 Recchioni（2015年））。

	Heston	SVF1	SVF2	RH
/	无噪声			
/	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
r	舍入噪声			
0.01	3/4，4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
0.05	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
σ_η	独立同分布噪声			
1	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
1.5	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
2	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
2.5	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
θ	自相关噪声			
0.2	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3，1/2
0.3	2/3，4/9	2/3，4/9	2/3，1/2	2/3，1/2
0.4	2/3，4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
ρ_η	与 p 相关的噪声			
-0.1	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
-0.3	2/3，4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
-05	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
$\bar{\sigma}_\eta$	异方差噪声			
3	2/3, 4/9	2/3，4/9	2/3, 1/2	2/3, 1/2
3.5	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
4	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2

表*1*：在所考虑的网格中，不同模型下波动率和微观结构噪声对应的最优 α、 β 组合。

注释 3.1。 虽然在本练习和后续分析中，我们着重于针对一对参数 N 和 M，为整个方差—协方差矩阵取得最佳性能，因此赋予协方差估计更高的权重，但值得注意的是，在处理

Heston - 无噪声					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	1.0386e-03	1.0204e-03	1.0069e-03	9.9026e-04	9.8691e-04
$5/6$	2.3748e-04	2.0522e-04	1.7974e-04	1.4534e-04	1.3932e-04
$3/4$	2.3898e-04	1.9593e-04	1.6083e-04	1.1666e-04	1.1096e-04
$2/3$	3.3382e-04	2.8090e-04	2.3916e-04	1.9747e-04	1.9399e-04
$1/2$	7.4551e-04	7.0530e-04	6.8504e-04	6.7519e-04	6.7487e-04
$1/3$	2.0294e-03	2.0282e-03	2.0278e-03	2.0277e-03	2.0277e-03

SVF2 - 无噪声					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	1.4352e-03	1.4189e-03	1.4101e-03	1.4065e-03	1.4081e-03
$5/6$	4.0236e-04	3.7224e-04	3.5338e-04	3.3913e-04	3.4125e-04
$3/4$	3.6879e-04	3.2276e-04	2.9026e-04	2.6256e-04	2.6423e-04
$2/3$	4.2664e-04	3.7789e-04	3.4306e-04	3.2132e-04	3.2378e-04
$1/2$	8.8461e-04	8.5025e-04	8.3790e-04	8.4012e-04	8.4159e-04
$1/3$	2.9639e-03	2.9649e-03	2.9659e-03	2.9668e-03	2.9669e-03

Heston - 独立同分布噪声 , $\sigma_\eta = 2.5$					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	6.6533e-03	5.2534e-03	4.2323e-03	2.9124e-03	2.6191e-03
$5/6$	1.7481e-03	1.3619e-03	1.0611e-03	6.6181e-04	5.9198e-04
$3/4$	7.3901e-04	6.0224e-04	4.9336e-04	3.5747e-04	3.3973e-04
$2/3$	4.7485e-04	3.9982e-04	3.4147e-04	2.8342e-04	2.7856e-04
$1/2$	7.7739e-04	7.3628e-04	7.1565e-04	7.0578e-04	7.0548e-04
$1/3$	2.0670e-03	2.0658e-03	2.0654e-03	2.0653e-03	2.0653e-03

SVF2 - 独立同分布噪声 , $\sigma_\eta = 2.5$					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	9.3093e-03	7.1797e-03	5.5577e-03	3.4186e-03	2.9621e-03
$5/6$	2.8428e-03	2.2971e-03	1.8831e-03	1.3697e-03	1.2876e-03
$3/4$	1.2835e-03	1.0975e-03	9.5587e-04	7.9035e-04	7.7287e-04
$2/3$	6.1650e-04	5.3221e-04	4.6916e-04	4.2024e-04	4.2082e-04
$1/2$	9.1933e-04	8.8559e-04	8.7380e-04	8.7665e-04	8.7821e-04
$1/3$	3.0359e-03	3.0369e-03	3.0380e-03	3.0390e-03	3.0391e-03

表*2*：所选场景下，在所考虑网格上的协方差估计误差。

与仅估计方差（即在单变量情形下）相比，由于这两个量的波动存在差异，最优选择通常不同于协方差估计所得的选择。具体而言，表*3*展示了粗糙 *Heston* 模型的情形：无噪声时，较大的 N 和 M （特别是设为 $\alpha = 1$）可获得最佳结果；而存在噪声时，差异会缩小，最佳选择变为 $\alpha = 2/3$ 和 $\beta = 5/6$。N 的取值与传统现货波动率傅里叶估计量的类似实验结果一致，参见 *Mancino* 和 *Recchioni*（*2015*）；同时，本例也证实了 M 的影响较不显著。下一节将进一步探讨存在异步性时优化方差与协方差之间的差异。

粗糙 Heston 模型 - 无噪声					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	6.4138e-03	6.7283e-03	7.1840e-03	8.0398e-03	8.8810e-03
$5/6$	7.5422e-03	7.8732e-0.3	8.2934e-03	9.3735e-03	9.8061e-03
$3/4$	8.4182e-03	8.7230e-03	9.1001e-03	1.0076e-02	1.0431e-02
$2/3$	9.5777e-03	9.7728e-03	1.0058e-03	1.0887e-02	1.1144e-02
$1/2$	1.3381e-02	1.3525e-02	1.3756e-02	1.4132e-02	1.4190e-02
$1/3$	2.5303e-02	2.5353e-02	2.5384e-02	2.5407e-02	2.5409e-02

粗糙 Heston 模型 - 噪声					
α/β	$5/6$	$3/4$	$2/3$	$1/2$	$4/9$
1	4.6703e$+00$	4.6520e$+00$	4.6384e$+00$	4.6204e$+00$	4.6161e$+00$
$5/6$	4.1665e-01	4.1473e-01	4.1331e-01	0.41143-01	4.1117e-01
$3/4$	6.2516e-02	6.2070e-02	6.1793e-02	6.1843e-02	6.2042e02
$2/3$	1.3675e-02	1.3813e-02	1.4038e-02	1.4408e-02	1.4464e-02
$1/2$	1.4980e-02	1.4994e-02	1.5127e-02	1.5778e-02	1.6012e-02
$1/3$	2.6021e-02	2.6071e-02	2.6102e-02	2.6125e-02	2.6127e-02

表*3*：粗糙 *Heston* 模型在所考虑网格上的方差估计误差。 *el.*

3.2.1 异步性对协方差估计的影响

在本节中，我们将更详细地研究数据中的异步性在多大程度上会影响截断频率 N 的最优选择，因为引入异步性可能会影响瞬时方差估计的效率，正如 Park 等人（2016）所示；并且可能会导致傅里叶型估计量的参数出现不同的最优值，正如 Mancino、Recchioni 和 Sanfelici（2017）针对积分方差的情形所示。本节中，我们比较同步情形与异步情形，以评估其对 N 选择的影响；关于采样方案对估计误差影响的进一步分析，则见第 3.3.2 节。

我们在此采用一个简化的环境，思路与 Mancino、Recchioni 和 Sanfelici

（2017）类似，并根据上一节的结果，在本节中固定 $M = N^{4/9}$。具体而言，我们考虑两种资产，其价格 p^1 和 p^2 分别等于两个标准布朗运动 W^1和 W^2，其中 $\langle W^1,\ W^2\rangle = \rho$。为在此引入观测的异步性，我们只需设定两个不同的观测时间向量：

$$t_i^1 = i/n \quad \text{for} \quad i = 0, ..., n \quad \text{and}$$

$$t_i^2 = i/n + 0.5 \quad \text{for} \quad i = 1, ..., n-1,$$

与 $t_0^2 = 0$ 和 $t_n^2 = 1$。我们设定 $n = 500$，并考虑一组包含10000条模拟价格路径的数据，在这些路径上，我们估计 W^1 与 W^2 之间的协方差，并考虑抽样方案的两种情形：第一种情形中，两种资产在时间网格 t^1上同步观测；第二种情形中，两种资产分别在 t^1 和 t^2 上观测。$\hat{V}^{1,}{}_N{}^2(t)$ 对 $N = 0,...,n$ 以及 $t = 0.5$（即时间窗口的中间位置）进行计算，以评估估计量在改变参数 N 时的表现。分析考虑了 W^1 与 W^2 之间相关性的多个取值，即 $\rho = \{0.2, \pm 0.3, \pm 0.5, \pm 0.7, \pm 1\}$，因此，从这一角度来看，与上一节相比，我们扩大了可能取值的范围；上一节仅考虑了资产之间相关性的单一取值。

图 1–5 展示了这项分析的结果。具体而言，每幅图都绘制了在同步观测和异步观测两种情况下，估计量的相对均方误差和相对偏差随傅里叶系数数量 N 的变化情况；每幅图对应一个相关性取值 ρ。除了考虑两种完全相关的资产这一情形外，不同相关性取值下的结果似乎基本一致。

值得注意的是，尽管在完全相关的情况下，引入异步性似乎几乎没有影响，但当两种资产之间的相关性降低时，情况并非如此。具体而言，我们可以看到，设定 $\rho \neq 1$ 会使 N 取值较大时的均方误差增加（正如预期，在高于奈奎斯特频率的取值下会增加，而在 $n/2$ 左侧的一段区间内也会增加）；而在 N 取值极小时观察到的均方误差上升，似乎是相关性变化所导致的有限样本效应。不过，引入异步性后，均方误差在某个区域内似乎保持稳定，而该区域的起点似乎接近 $N = \frac{1}{2}n^{3/4}$，这正是我们在此前分析中发现的、无噪声情况下的最优值。此外，对于较大的 N，引入异步性似乎会带来显著偏差，因此，选择较小的 N 也能从这一角度提升估计质量。这项分析表明，此前采用的选择似乎会受到数据中异步性的影响，并且对 ρ 的不同选择具有稳健性。

虽然估计量的中心极限定理留待未来研究，但对于不同相关性取值所呈现的相似表现，似乎与渐近分布不依赖于被估计量本身这一情况相符，这通常是高频估计量的特征，参见例如 Aït-Sahalia 和 Jacod（2014年）。

3.3 比较

在评估了我们的估计量对参数选择的敏感性之后，本节将重复第 3.2 节采用的大规模模拟研究，以评估其准确性和

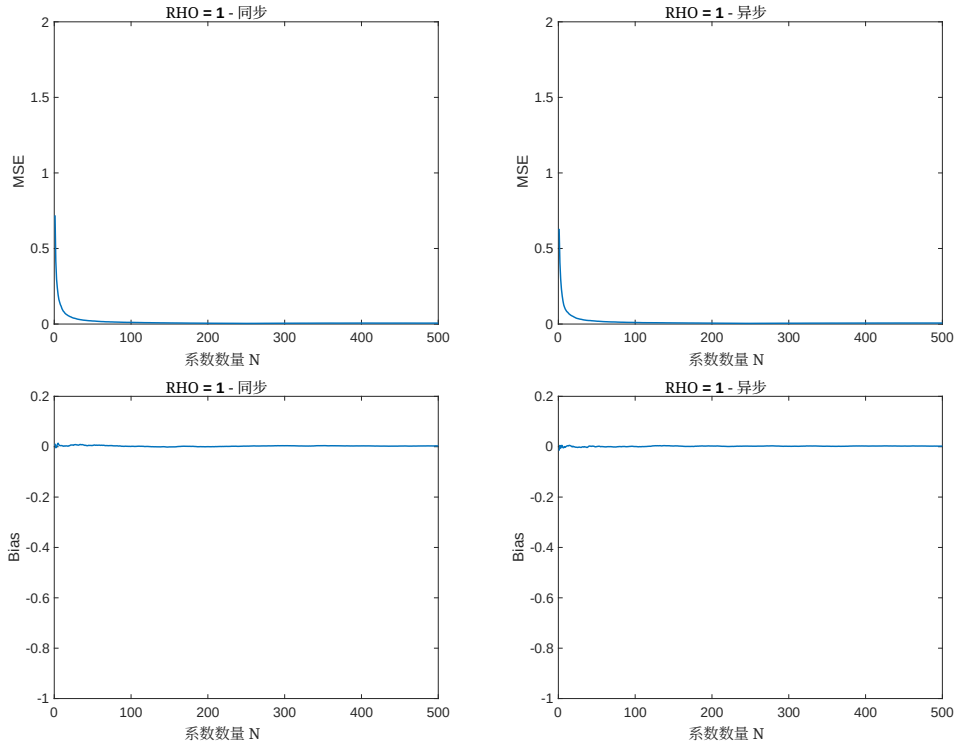

图 *1*： $\rho = 1$ 的均方误差（第一行）和偏差（第二行）。

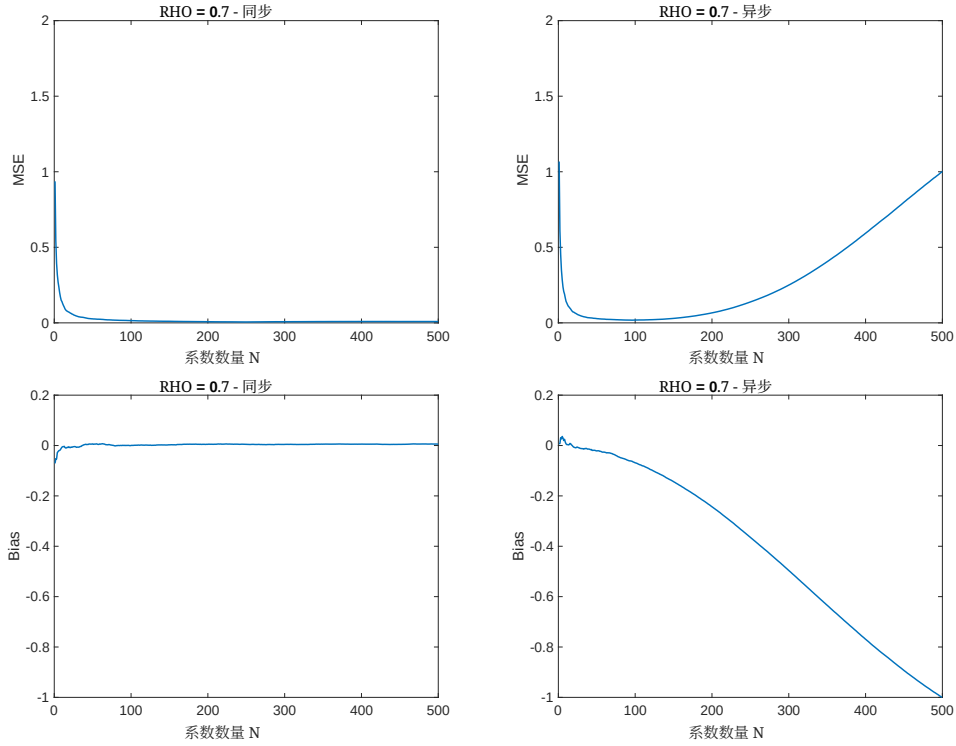

图*2*：均值　　平方误差（第一行）和偏差（第二行）　　*w)* 对于 $\rho = 0.7$。

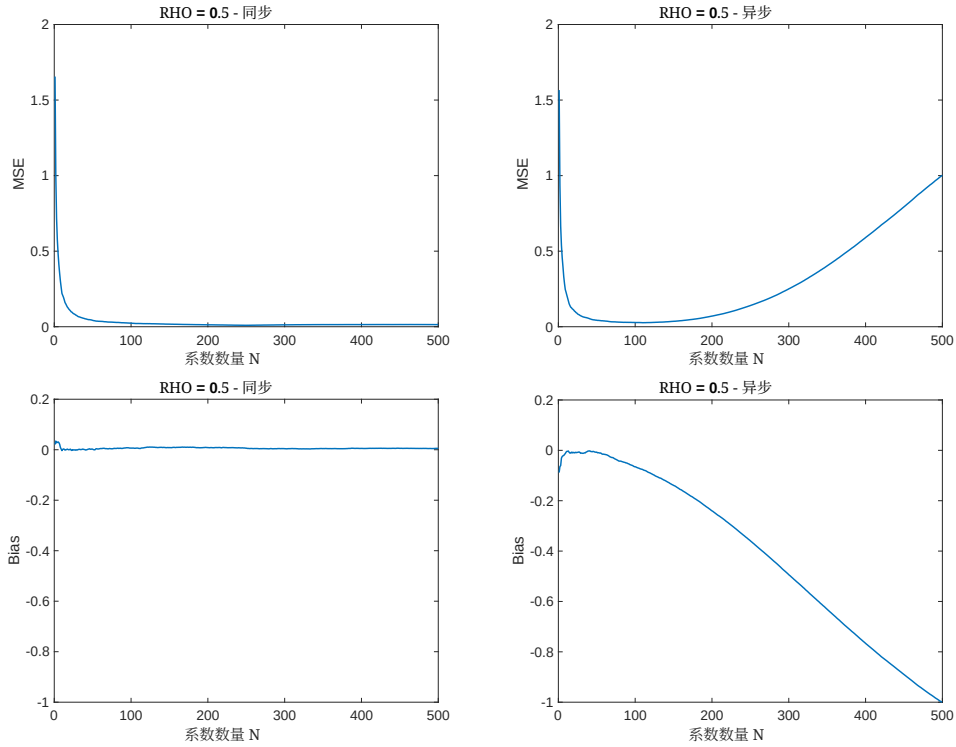

图 3： $\rho = 0.5$ 的均方误差（第一行）和偏差（第二行）。

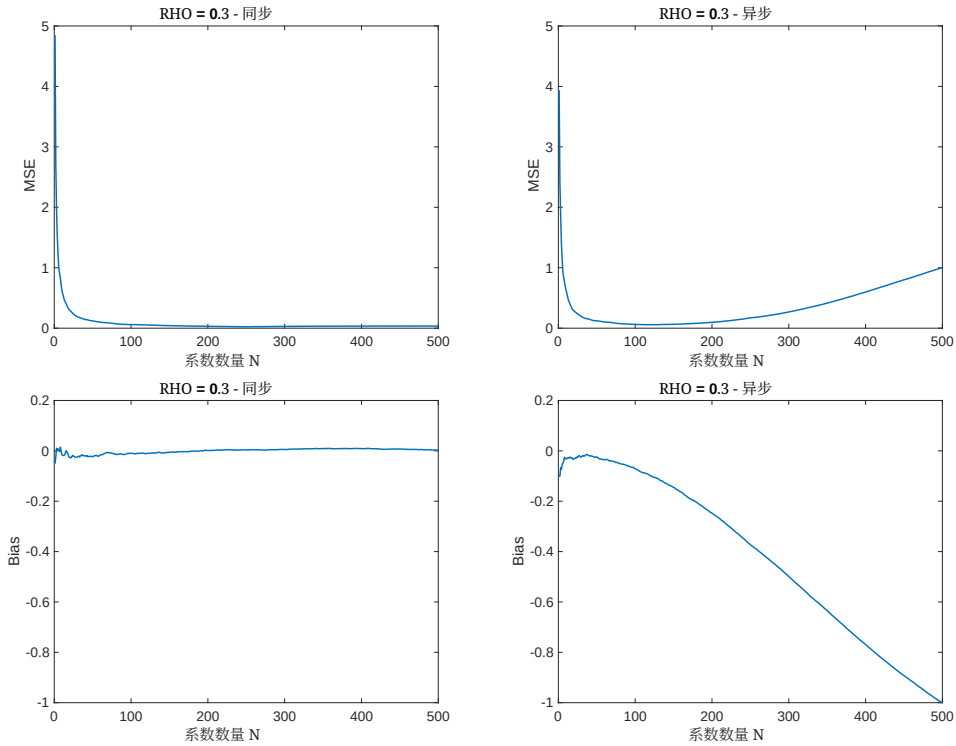

图4：均值　　平方误差（第一行）和偏差（第二行　　）对应于 $\rho = 0.3$。

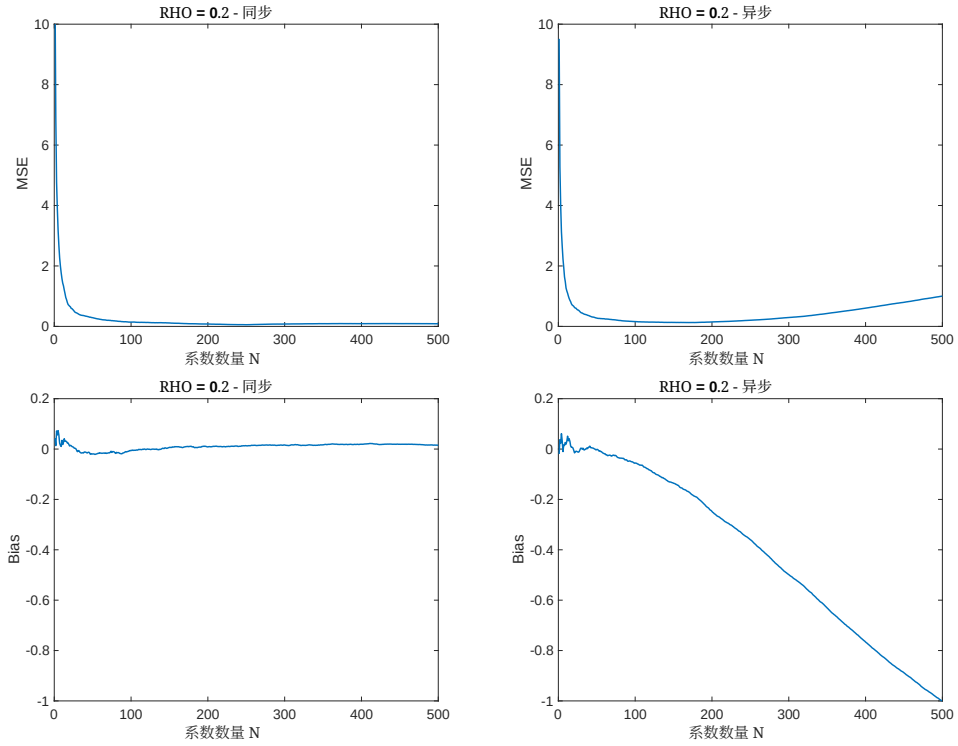

图 5： $\rho = 0.2$ 的均方误差（第一行）和偏差（第二行）。

所提出的概率密度函数估计量与文献中现有的两种竞争估计量相比的能力。聚焦于能够处理异步观测的估计量，我们在此考虑：

- 本文提出的修正傅里叶估计量（PDF）；
- Mykland 等人（2019年）提出的平滑双尺度现货估计量（STS）；
- Bibinger 等人（2019）提出的局部矩法瞬时估计量（LMM）。

Bu等人（2022年）提出的基于核的估计量虽然可以扩展到处理不规则和异步观测，但依赖于经过专门调校的收缩技术来确保估计结果的半正定性，因此在本分析中不予考虑。

分析基于最多 $d = 40$ 项资产进行，并根据均方积分误差（均方积分误差）及其相对形式（均方根积分误差）评估每个估计量的表现，定义如下

$$MISE = (Kd^2)^{-1} \sum_{k=1}^{K} \int_0^1 \sum_{j,i=1}^{d} (\hat{V}_k^{ij}(t) - V_k^{ij}(t))^2 dt,$$

$$RMISE = (Kd^2)^{-1} \sum_{k=1}^{K} \int_0^1 \sum_{j,i=1}^{d} (\hat{V}_k^{ij}(t) - V_k^{ij}(t))^2 / V_k^{ij}(t)^2 dt.$$

未报告的结果表明，采用不同的损失函数，尤其是估计的现货波动率与真实现货波动率之差的弗罗贝尼乌斯范数、欧几里得范数或 1-范数，

矩阵，并不会影响表 4–10 中呈现的排名。最重要的是，在本分析中，我们特别关注各估计量在不同情形下能够生成的对称半正定方差-协方差矩阵的比例。

对于 LMM 估计量，我们采用 Bibinger 等人（2019）数值分析中发现的最优参数；对于 STS 估计量，我们则在 Chen 等人（2020）所用参数附近选取参数，并通过辅助模拟得到的均方误差将其最小化。

3.3.1 无噪声

在第 3.3.1 节和第 3.3.2 节中，我们报告了当有效价格过程遵循 Heston 模型时，三种竞争性估计量在均方积分误差和产生半正定估计的百分比方面的比较结果。在这种设定下，均方根积分误差方面的结果也类似。我们首先考察无噪声数据，并聚焦于高频协方差矩阵估计的两个主要特征：矩阵的维数和观测频率。

众所周知，样本协方差矩阵的谱分布取决于矩阵本身的维数（更准确地说，取决于观测数量与所观测随机向量维数之比）；随机矩阵谱分布的理论可参见 Bai 和 Silverstein（2010）。矩阵维数的影响在应用中也十分重要，例如在主成分分析中（所谓的“维数灾难”及其对特征值估计的影响，参见 Chen 等人（2020））。然而，就本文考察的即期方差估计量而言，估计矩阵的谱密度尚无定论，而且（对于其他估计量而言）估计矩阵甚至无法保证是半正定的。深入研究估计矩阵的谱分布（这或许有助于构建例如基于即期协方差矩阵的有限样本主成分分析框架）超出了本文的研究范围；不过，我们仍有兴趣评估问题维数的变化是否会影响所考察估计量所得半正定估计的比例，因此需要考虑支撑集可能包含负值的谱分布。以下所有实验均通过改变d来探讨这一问题，即逐步增加资产数量。

表 4 展示了未考虑噪声的数据结果，考察的矩阵维数为$d = 2$，5，10，15，20，25，30，40。在这项实验中，当观测到有效价格过程时，对于任意维数的波动率矩阵，我们修改后的傅里叶估计量在均方误差方面表现最佳。随着资产数量增加，平滑双尺度估计量生成正半定估计值的能力似乎有所下降，尤其是在 $d > 20$ 时；另外两种估计量则都能生成 100% 的半正定矩阵。鉴于维数在估计方差-协方差矩阵时发挥着重要作用，并会影响估计值的正定性，在后续分析中始终会考虑维数。为简化起见，以下分析考察的维数限于 $d = 5$，10，15，20。

在无噪声的情形下，我们考虑数据采用不同采样方案的情况。为此，我们研究连续两次观测之间的平均时间增加时性能的变化；具体而言，我们按照泊松过程从模拟轨迹中提取观测，使得平均每 15、30 和 40 秒产生一次观测。表 5 显示，尽管在均方积分误差方面仍优于其他方法，

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=2		d=20	
PDF	1.8453e-04	100%	6.5495e-04	100%
LMM	5.2131e-04	100%	1.6008e-03	100%
STS	8.2259e-04	100%	2.5759e-03	100%
	d=5		d=25	
PDF	5.7567e-04	100%	6.1793e-04	100%
LMM	1.5514e-03	100%	1.5090e-03	100%
STS	2.3805e-03	100%	2.4401e-03	99.20%
	d=10		d=30	
PDF	8.3119e-04	100%	6.8819e-04	100%
LMM	2.0574e-03	100%	1.6779e-03	100%
STS	3.3500e-03	100%	2.7116e-03	98.56%
	d=15		d=40	
PDF	6.9991e-04	100%	6.3139e-04	100%
LMM	1.7120e-03	100%	1.5390e-03	100%
STS	2.7634e-03	100%	2.4887e-03	33.10%

表*4*：当维度 d的 V 增加时，每个估计量生成的半正定矩阵的准确度（均方积分误差）及其百分比。

概率密度函数估计量是唯一能够在 100% 的情况下得到半正定估计的估计量，而 STS 和 LMM估计量都可能以更高的频率失效¯当 Δt 增加时，尽管这种变化对所获得的正估计比例的影响似乎相当小。所有估计量的准确性¯似乎也会随着 Δt取值的增大而下降；这当然与这些估计量的一致性属于渐近性质这一事实相符。

3.3.2 受 MMN 污染的数据

本节运行比较时，考虑第3.1.2 节所述的噪声设定。值得指出的是，LMM估计量包含显式的噪声校正，而STS 估计量依赖于对数据进行预平均；对于所提出的概率密度函数估计量，与原始的现货波动率傅里叶估计量一致，无需操纵数据或校正估计量来处理噪声的存在，只需截断频率 N，如第 3.2 节所示。

表6显示，取整的存在似乎并未显著影响估计量的准确性以及估计结果的半正定性。这种影响可能是由于所采用的不规则采样数据模拟方案造成的，该方案相对于取整后的模拟序列进行了子采样，从而降低了这一噪声源的强度。

表 7 显示，独立同分布噪声，尤其是 σ_η^2 较高时，会对 LMM 和 STS 估计量生成半正定估计值的能力产生负面影响，而且随着维度增加，影响会更加显著

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, $\bar{\Delta}t = 15$		d=5, $\bar{\Delta}t = 30$		d=5, $\bar{\Delta}t = 40$	
PDF	7.7661e-04	100%	1.2284e-03	100%	1.5871e-03	100%
LMM	1.5937e-03	100%	2.0398e-03	100%	2.3025e-03	99.83%
STS	2.5185e-03	100%	3.4587e-03	100%	4.3867e-03	100%
	d=10, $\bar{\Delta}t = 15$		d=10, $\bar{\Delta}t = 30$		d=10, $\bar{\Delta}t = 40$	
PDF	1.1008e-03	100%	1.7661e-03	100%	2.2329e-03	100%
LMM	2.2549e-03	100%	2.7865e-03	100%	3.0841e-03	99.66%
STS	3.4760e-03	100%	4.8142e-03	100%	6.4107e-03	100%
	d=15, $\bar{\Delta}t = 15$		d=15, $\bar{\Delta}t = 30$		d=15, $\bar{\Delta}t = 40$	
PDF	9.0624e-04	100%	1.4832e-03	100%	1.8593e-03	100%
LMM	1.8481e-03	100%	2.2859e-03	100%	2.5646e-03	98.62%
STS	2.8926e-03	100%	4.0185e-03	99.98%	5.3583e-03	99.75 %
	d=20, $\bar{\Delta}t = 15$		d=20, $\bar{\Delta}t = 30$		d=20, $\bar{\Delta}t = 40$	
PDF	8.5236e-04	100%	1.3852e-03	100%	1.8593e-03	100%
LMM	1.7170e-03	100%	2.1356e-03	99.96%	2.5646e-03	98.62%
STS	2.6941e-03	99.98%	3.7669e-03	98.83%	5.3583e-03	99.08 %

表 *5*：平均连续观测间隔$^{-}$发生 Δt 变化时，各估计量生成半正定矩阵的准确率及百分比。

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, r=0.01		d=5, r=0.05	
PDF	5.6886e-04	100%	5.8434e-04	100%
LMM	1.4588e-03	100%	1.4744e-03	100%
STS	2.4015e-03	100%	2.4342e-03	100%
	d=10, r=0.01		d=10, r=0.05	
PDF	8.2010e-04	100%	8.2171e-04	100%
LMM	2.0337e-03	100%	2.0445e-03	100%
STS	3.3250e-03	100%	3.3612e-03	100%
	d=15, r=0.01		d=15, r=0.05	
PDF	6.7300e-04	100%	6.7597e-04	100%
LMM	1.6870e-03	100%	1.6806e-03	100%
STS	2.7879e-03	100%	2.7996e-03	100%
	d=20, r=0.01		d=20, r=0.05	
PDF	6.3197e-04	100%	6.3562e-04	100%
LMM	1.5806e-03	100%	1.5791e-03	100%
STS	2.6104e-03	100%	2.6063e-03	100%

表 *6*：　存在 *1* 或 *5* 美分取整时，各估计量生成半正定矩阵的准确度和比例。

随着估计矩阵规模的扩大，所有估计量的准确性也会随着噪声增大而下降，而在这种情形下，PDF 仍被证实表现最佳。从市场微观结构的角度来看，独立同分布噪声通常与买卖价差的存在有关，例如 Roll（1984）中的建模方式；而买卖价差通常与资产的流动性相关，因此，管理这类噪声的能力可以被视为对非流动性资产的协方差进行准确估计的能力。

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, $\sigma_\eta = 1$		d=5, $\sigma_\eta = 1.5$		d=5, $\sigma_\eta = 2$		d=5, $\sigma_\eta = 2.5$	
PDF	5.8957e-04	100%	1.2411e-03	100%	2.1739e-03	100%	3.5475e-03	100%
LMM	1.4774e-03	100%	2.0615e-03	100%	3.2557e-03	99.91%	5.5029e-03	98.18%
STS	2.4258e-03	100%	3.0744e-03	100%	3.9064e-03	99.95%	5.1760e-03	97.85%
	d=10, $\sigma_\eta = 1$		d=10, $\sigma_\eta = 1.5$		d=10, $\sigma_\eta = 2$		d=10, $\sigma_\eta = 2.5$	
PDF	8.3152e-04	100%	1.7158e-03	100%	2.4388e-03	100%	3.5411e-03	100%
LMM	2.0447e-03	100%	2.6918e-03	100%	3.6557e-03	99.56%	5.5120e-03	91.14%
STS	3.3806e-03	100%	4.3057e-03	99.83%	5.4130e-03	90.31%	7.0991e-03	52.19 %
	d=15, $\sigma_\eta = 1$		d=15, $\sigma_\eta = 1.5$		d=15, $\sigma_\eta = 2$		d=15, $\sigma_\eta = 2.5$	
PDF	6.8032e-04	100%	1.4012e-03	100%	1.8502e-03	100%	2.5615e-03	100%
LMM	1.6911e-03	100%	2.1709e-03	99.98%	2.8183e-03	99.20%	3.9896e-03	80.66%
STS	2.8036e-03	99.87%	3.5990e-03	91.42%	5.8970e-03	34.15%		26.37%
	d=20, $\sigma_\eta = 1$		d=20, $\sigma_\eta = 1.5$		d=20, $\sigma_\eta = 2$		d=20, $\sigma_\eta = 2.5$	
PDF	6.3538e-04	100%	1.3026e-03	100%	1.6698e-03	100%	2.2492e-03	100%
LMM	1.5694e-03	100%	2.0025e-03	99.96%	2.5535e-03	97.44%	3.5224e-03	64.45%
STS	2.6111e-03	97.51%	3.3587e-03	42.48%	4.1644e-03	11.71 %	5.4918e-03	0.50%

表 7：当数据受到独立同分布噪声污染时，各估计量生成的半正定矩阵的准确率和百分比。

表 8 显示，自相关性会显著影响 STS 估计量生成功率谱密度矩阵的能力，尤其是在 θ_η 较低时；LMM估计量也会受到影响，尽管影响较小。 θ_η 较低也会降低所有三种估计量的准确性，但它们的排名保持不变。

表9显示，引入 E 和 W之间的相关性，也就是引入驱动噪声的布朗运动与驱动有效价格动态的布朗运动之间的相关性，似乎并不会特别影响估计量的半正定性。此外，与基准的独立同分布情形相比，所有估计量的准确性似乎都略有提高；这是因为有效价格与噪声之间的负自相关可能会引入与不相关噪声所导致的偏差符号相反的偏差。这种现象通常可以通过波动率特征图观察到，Hansen 和 Lunde（2006）便采用了这种方法。

表10展示了我们所考虑的最后一种噪声设定下的结果，其中包括噪声自协方差、与有效价格的相关性以及随时间变化的噪声方差。结果表明，在这种设定下，LMM 和 STS 估计量能否达到令人满意的半正定估计比例，取决于微观结构成分的强度。我们

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, $\theta_\eta = 0.2$		d=5, $\theta_\eta = 0.3$		d=5, $\theta_\eta = 0.4$	
PDF	5.6169e-03	100%	2.8806e-03	100%	2.1414e-03	100%
LMM	1.0856e-02	100%	4.8943e-03	100%	3.3544e-03	100%
STS	8.5059e-03	97.31%	5.1553e-03	99.08%	4.1302e-03	99.92%
	d=10, $\theta_\eta = 0.2$		d=10, $\theta_\eta = 0.3$		d=10, $\theta_\eta = 0.4$	
PDF	5.1481e-03	100%	3.0019e-03	100%	2.3519e-03	100%
LMM	9.7818e-03	100%	5.0765e-03	99.77%	3.7482e-03	100%
STS	1.0814e-02	60.36%	6.8148e-03	90.08%	5.6242e-03	91.37%
	d=15, $\theta_\eta = 0.2$		d=15, $\theta_\eta = 0.3$		d=15, $\theta_\eta = 0.4$	
PDF	3.5769e-03	100%	2.2202e-03	100%	1.8000e-03	100%
LMM	6.6905e-03	92.90%	3.7483e-03	95.98%	2.8798e-03	96.80%
STS	8.7234e-03	3.70%	5.6586e-03	29.92%	4.6585e-03	35.03%
	d=20, $\theta_\eta = 0.2$		d=20, $\theta_\eta = 0.3$		d=20, $\theta_\eta = 0.4$	
PDF	3.0654e-03	100%	1.9632e-03	100%	1.6249e-03	100%
LMM	5.6948e-03	93.90%	3.3182e-03	92.98%	2.6127e-03	95.48%
STS	8.0435e-03	4.87%	5.2368e-03	9.65%	4.3666e-03	8.23%

表 *8*：数据受自相关噪声污染时，各估计量产生的半正定矩阵的准确性和百分比。

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, $\rho_\eta = -0.1$		d=5, $\rho_\eta = -0.3$		d=5, $\rho_\eta = -0.5$	
PDF	2.0352e-03	100%	1.9671e-03	100%	1.9375e-03	100%
LMM	3.1654e-03	100%	3.0462e-03	100%	2.9484e-03	100%
STS	3.9369e-03	99.81%	3.8525e-03	99.96%	3.8145e-03	99.92%
	d=10, $\rho_\eta = -0.1$		d=10, $\rho_\eta = -0.3$		d=10, $\rho_\eta = -0.5$	
PDF	2.2896e-03	100%	2.3440e-03	100%	2.2369e-03	100%
LMM	3.5607e-03	99.94%	3.4779e-03	99.92%	3.4024e-03	99.98%
STS	5.3662e-03	91.56%	5.2421e-03	94.05%	5.1648e-03	95.49%
	d=15, $\rho_\eta = -0.1$		d=15, $\rho_\eta = -0.3$		d=15, $\rho_\eta = -0.5$	
PDF	1.7608e-03	100%	1.7866e-03	100%	1.7255e-03	100%
LMM	2.7502e-03	98.62%	2.7090e-03	99.44%	2.6483e-03	99.79%
STS	4.4396e-03	37.17%	4.3377e-03	44.13%	4.2908e-03	53.57%
	d=20, $\rho_\eta = -0.1$		d=20, $\rho_\eta = -0.3$		d=20, $\rho_\eta = -0.5$	
PDF	1.6065e-03	100%	1.6253e-03	100%	1.5582e-03	100%
LMM	2.5180e-03	98.35%	2.4725e-03	98.89%	2.3988e-03	95.84%
STS	4.1353e-03	0.82%	4.0389e-03	2.92%	3.9787e-03	5.43%

表*9*： 当数据受到与有效价格过程相关的噪声污染时，各估计量的准确度及其生成半正定矩阵的百分比。

再次确认概率密度函数估计量能够生成具有所需性质且估计误差相对较低的方差-协方差矩阵。

估计量	MISE	% 功率谱密度	MISE	% 功率谱密度	MISE	% 功率谱密度
	d=5, $\bar{\sigma}_\eta = 3$		d=5, $\bar{\sigma}_\eta = 3.5$		d=5, $\bar{\sigma}_\eta = 4$	
PDF	4.7998e-03	100%	7.8667e-03	100%	1.3212e-02	100%
LMM	6.8314e-03	100%	1.1376e-02	98.81%	1.8293e-02	93.86%
STS	8.6164e-03	84.19%	1.3869e-02	71.84%	2.0190e-02	61.96%
	d=10, $\bar{\sigma}_\eta = 3$		d=10, $\bar{\sigma}_\eta = 3.5$		d=10, $\bar{\sigma}_\eta = 4$	
PDF	4.5443e-03	100%	6.9249e-03	100%	1.1111e-02	100%
LMM	6.7069e-03	99.89%	1.0364e-02	95.89%	1.6497e-02	83.48%
STS	1.1635e-02	52.86%	1.7614e-02	36.65%	2.6044e-02	30.51%
	d=15, $\bar{\sigma}_\eta = 3$		d=15, $\bar{\sigma}_\eta = 3.5$		d=15, $\bar{\sigma}_\eta = 4$	
PDF	3.2065e-03	100%	4.7542e-03	100%	7.2763e-03	100%
LMM	4.8670e-03	99.37%	7.2946e-03	89.26%	1.1172e-02	68.90%
STS	9.4853e-03	35.25%	1.4449e-02	24.56%	2.1268e-02	19.98%
	d=20, $\bar{\sigma}_\eta = 3$		d=20, $\bar{\sigma}_\eta = 3.5$		d=20, $\bar{\sigma}_\eta = 4$	
PDF	2.7842e-03	100%	4.0233e-03	100%	6.0704e-03	100%
LMM	4.2737e-03	98.14%	6.2574e-03	80.82%	9.4823e-03	54.03%
STS	8.7501e-03	25.69%	1.3279e-02	16.33%	1.9663e-02	11.38%

表10 ：数据受到异方差噪声过程污染时，各估计量生成的半正定矩阵的准确率和百分比。

3.4 替代波动率模型

前几节展示了我们针对采用 Heston 模型模拟有效价格过程所得结果的比较。尽管分析中使用的模拟模型发生变化时，三种估计量产生的误差可能有所不同；特别是，采用 SVF2 或粗糙 Heston 模型时，PDF 估计量与 LMM 估计量之间的均方积分误差差距会缩小，但总体结果基本得到验证：PDF 估计量仍是唯一能够持续产生半正定估计的估计量，而且在几乎所有情形下，其均方误差表现最佳。表11列出了在不存在微观结构噪声的情况下，采用替代波动率模型所得半正定估计的比例。可以看到，在这项研究中，改用 SV1F、SV2F 或粗糙 Heston 模型，似乎并不会显著影响各估计量生成正定矩阵的能力；而且 LMM 估计量生成半正定矩阵的能力与 PDF 估计量相同。附录B报告了更全面的替代模型结果，其中展示了在 i.i.d.、自相关和异方差噪声情形下半正定估计的比例。最后值得指出的是，从均方根积分误差来看，PDF 估计量始终表现最佳；而且改用替代模型后，在大多数情形下，按均方积分误差衡量的排名也保持不变。

估计量	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=2			d=20		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	99.86%	99.86%	100%
	d=5			d=25		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	98.45%	97.20%	96.54%
	d=10			d=30		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	94.25%	92.80	29.72%
	d=15			d=40		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	34.04%	32.12%	27.71%

表*11*：当有效价格过程由替代模型生成时，各估计量生成半正定矩阵的百分比。

4 结论

本文提出了 Malliavin 和 Mancino（2009）经典即期协方差傅里叶估计量的一种改进版本，以克服难以获得对称且半正定的即期方差-协方差矩阵估计这一问题。已有研究证明该估计量具有一致性（Akahori 等人（2023）），本文则证明了其半正定性。我们开展了数值研究，以评估在多种设定下参数 N 和 M的最优选择。最优参数组合似乎相当稳定；按照傅里叶估计量这一类方法的通常规律，当数据受到严重噪声污染时，应减小 N 。根据这些发现，我们开展了全面的模拟研究，以评估该估计量的准确性及其实际生成半正定估计的能力。我们将结果与文献中两种尚未证明具有半正定性的替代估计量进行比较，发现所提出的 PDF 估计量在均方误差方面通常优于竞争方法，表明无需牺牲估计准确性，也能使估计得到的即期协方差矩阵具备理想性质。此外，通过恰当选择截断频率 N，该估计量被证明能够适应多种市场微观结构噪声设定。采用替代数据生成过程后，我们的结果仍然稳健。

附录A 定理2.3的证明

设 a_j 对于 $j = 1$，2，3 是定义在 $\mathbb{Z}$上的任意函数，根据 K 和 S(k) 的定义，我们注意到：

$$\sum_{k\in\mathcal{K}}\sum_{(s,s')\in\mathcal{S}(k)} a_1(k)a_2(s)a_3(s')$$
$$=\sum_{k=0}^{2N}\sum_{v=0}^{2N-k} a_1(k)a_2(-N+k+v)a_3(N-v)+\sum_{k=-2N}^{-1}\sum_{v=0}^{2N+k} a_1(k)a_2(N+k-v)a_3(-N+v)=:A+B.$$

对于第一项，我们有：

$$\begin{aligned}A&=\sum_{k=0}^{2N}\sum_{u=k-N}^{N} a_1(k)a_2(k-u)a_3(u)\\&=\sum_{u=-N}^{N}\sum_{k=0}^{u+N} a_1(k)a_2(k-u)a_3(u)\\&=\sum_{u=-N}^{N}\sum_{u'=-N}^{N} a_1(u+u')a_2(u')a_3(u),\end{aligned}$$

其中，第一行中我们设定了 $u=N-v$，第二行中改变了求和运算的顺序，并代入了 $u'=k-u$。类似地，按照 $\sum_{u=0}^{-1}=0$ 的约定，对于第二项有：

$$\begin{aligned}B&=\sum_{k=-2N}^{-1}\sum_{u=-N}^{N+k} a_1(k)a_2(k-u)a_3(u)\\&=\sum_{u=-N}^{N}\sum_{k=u-N}^{-1} a_1(k)a_2(k-u)a_3(u)\\&=\sum_{u=-N}^{N}\sum_{u'=-N}^{-u-1} a_1(u+u')a_2(u')a_3(u).\end{aligned}$$

因此我们可以看到

$$\sum_{k\in\mathcal{K}}\sum_{(s,s')\in\mathcal{S}(k)} a_1(k)a_2(s)a_3(s')=\sum_{u=-N}^{N}\sum_{u'=-N}^{N} a_1(u+u')a_2(u')a_3(u).$$

当 $a_1(k)=c(k)e^{ikt}$，$a_2(s)=e^{-ist_l^{j'}}$，且 $a_3(s')=e^{-ist_l^{j}}$ 时，利用变量替换 $u-\to -u'$，可得：

$$\hat{V}_N^{j,j'}(t)=\sum_{l=1}^{n_j}\sum_{l'=1}^{n_{j'}}\sum_{u=-N}^{N}\sum_{u'=-N}^{N} c(u-u')e^{iu(t-t_l^j)}e^{-iu'(t-t_{l'}^{j'})}\Delta(X_l^j)\Delta(X_{l'}^{j'}).$$

然后，对于 $x\in\mathbf{C}^d$

$$\sum_{j,j'} \hat{V}_N^{j,j'}(t) x_j \overline{x_{j'}}$$
$$= \sum_{u=-N}^{N} \sum_{u=-N}^{N} c(u-u') \left(\sum_{j=1}^{d} x_j \sum_{l=1}^{n_j} e^{iu(t-t_l^j)} \Delta X_l^j \right) \left(\sum_{j'=1}^{d} x_j \sum_{l'=1}^{n_{j'}} e^{-iu(t-t_{l'}^{j'})} \Delta X_{l'}^{j'} \right)$$
$$= \sum_{u=-N}^{N} \sum_{u'=-N}^{N} c(u-u') f(u) \overline{f(u')} \geq 0.$$

其中 $f(u) := \sum_{j=1}^{d} x_j \sum_{l=1}^{n_j} e^{iu(t-t_l^j)} \Delta X_l^j$。证明完毕。

附录B 替代模型比较的其他结果

估计量	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\sigma_\eta = 1$			d=5, $\sigma_\eta = 1.5$			d=5, $\sigma_\eta = 2$			d=5, $\sigma_\eta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%	100%	99.85%	99.68%	97.28%	93.98%	94.98%
STS	100%	100%	100%	100%	100%	100%	99.77%	99.77%	99.79%	97.95%	97.33%	97.70%
	d=10, $\sigma_\eta = 1$			d=10, $\sigma_\eta = 1.5$			d=10, $\sigma_\eta = 2$			d=10, $\sigma_\eta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	100%	100%	99.70%	99.25%	95.26	90.03%	85.48%	87.11%
STS	100%	100%	100%	99.60%	99.69%	99.54%	90.91%	89.16%	91.10	52.28%	49.62%	54.59%
	d=15, $\sigma_\eta = 1$			d=15, $\sigma_\eta = 1.5$			d=15, $\sigma_\eta = 2$			d=15, $\sigma_\eta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	99.80%	99.80%	99.25%	97.85%	98.02%	79.28%	75.16%	77.98%
STS	100%	99.90%	99.96%	90.52%	88.66%	90.81%	37.21%	29.88%	39.24%	23.72%	15.64%	18.25%
	d=20, $\sigma_\eta = 1$			d=20, $\sigma_\eta = 1.5$			d=20, $\sigma_\eta = 2$			d=20, $\sigma_\eta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	98.27%	98.82%	97.95%	95.23%	95.88%	65.91%	63.84%	65.95%
STS	97.37%	96.68%	97.14%	42.42%	49.39%	44.44%	28.56%	22.64%	25.84%	1.80%	0.28%	0.02%

表 *12*：当有效价格过程由替代模型生成且存在独立同分布噪声时，各估计量生成半正定矩阵的百分比。

估计量	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\theta_\eta = 0.2$			d=5, $\theta_\eta = 0.3$			d=5, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	99.96%	100%	100%	100%	100%	100%	100%	100%
STS	97.39%	97.20%	97.83%	99.21%	99.14%	99.83%	99.77%	99.25%	99.73%
	d=10, $\theta_\eta = 0.2$			d=10, $\theta_\eta = 0.3$			d=10, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.85%	98.20%	100%	98.16%	96.90%	100%	100%	100%	100%
STS	49.23%	47.87%	53.55%	73.25%	71.10%	73.25%	90.39%	88.34%	89.65%
	d=15, $\theta_\eta = 0.2$			d=15, $\theta_\eta = 0.3$			d=15, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.12%	97.90%	96.55%	99.85%	99.73%	98.16%	99.88%	95.36%	97.22%
STS	10.22%	10.63%	9.36%	11.32%	15.38%	14.64%	21.32%	20.44%	20.89%
	d=20, $\theta_\eta = 0.2$			d=20, $\theta_\eta = 0.3$			d=20, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	98.52%	97.50%	96.55%	98.63%	99.20%	98.25%	99.85%	96.17%	96.82%
STS	1.83%	1.32%	3.74%	4.57%	3.22%	4.72%	8.94%	8.55%	9.33%

表 *13*：当有效价格过程由替代模型产生且存在自相关噪声时，各估计量生成的半正定矩阵百分比。

估计量	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\bar{\sigma}_\eta = 3$			d=5, $\bar{\sigma}_\eta = 3.5$			d=5, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	99.96%	99.96%	99.25%	98.95%	99.18%	92.79%	91.65%	91.31%
STS	89.60%	82.73%	84.52%	72.26%	72.97%	74.06%	61.17%	60.53%	63.01%
	d=10, $\bar{\sigma}_\eta = 3$			d=10, $\bar{\sigma}_\eta = 3.5$			d=10, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.85%	99.73%	99.77%	95.78%	95.74%	95.61%	79.97%	80.03%	81.27%
STS	45.15%	44.82%	45.95%	36.34%	36.11%	36.38%	30.33%	29.74%	30.45%
	d=15, $\bar{\sigma}_\eta = 3$			d=15, $\bar{\sigma}_\eta = 3.5$			d=15, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.12%	98.02%	99.25%	90.68%	89.58%	91.22%	67.04%	66.54%	68.96%
STS	30.58%	29.88%	30.81%	23.92%	23.39%	23.66%	19.40%	19.32%	18.94%
	d=20, $\bar{\sigma}_\eta = 3$			d=20, $\bar{\sigma}_\eta = 3.5$			d=20, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	97.58%	97.09%	97.58%	83.06%	80.64%	83.44%	52.92%	51.88%	55.24%
STS	21.37%	21.37%	21.30%	15.77%	15.66%	15.31%	11.09%	11.05%	10.99%

表*14*：当有效价格过程由替代模型生成且存在异方差噪声时，各估计量生成的半正定矩阵所占百分比。

参考文献

Aït-Sahalia，Yacine和 JeanJacod（2014年）。高频金融计量经济学。普林斯顿大学出版社。

Aït-Sahalia，Yacine 和 Dacheng Xiu（2019年）。“高频数据的主成分分析”。载于：《美国统计学会会刊》114.525，第 287–303 页。Akahori，Jirô、ReikaKambara 和Gô Yûki（2023）。“概率密度函数估计量的一致性”。载于：*ArXiv* 预印本，即将出版。Alòs，Elisa、Jorge ALeón和JosepVives（2007）。“具有随机波动率的跳扩散模型中隐含波动率的短期行为”。载于：《金融与随机过程》4.11，第 571–589 页。Bai，Zhidong 和 JackW Silverstein（2010）。大维随机矩阵的谱分析。第 20 卷。Springer。Bandi，Federico M和JeffreyRRussell（2008）。“微观结构噪声、已实现方差与最优采样”。载于：《经济研究评论》75.2，第 339–369 页。

Barndorff-Nielsen，Ole E、Peter Reinhard Hansen、AsgerLunde 和Neil Shephard（2011）。“多元已实现核：在存在噪声和非同步交易的情况下，对股票价格协变异进行一致的半正定估计”。载于：《计量经济学杂志》162.2，第 149–169 页。

Bibinger, Markus, Nikolaus Hautsch,Peter Malec, 和Markus Reiss（2019年）。“资产价格现货协变性的估计——统计理论与经验证据”。载于：《商业与经济统计期刊》37.3，第419–435页。Bu,Ruijun, Degui Li, OliverLinton,和Hanchao Wang（2022年）。“高频金融数据的大型现货波动率矩阵的非参数估计”。载于。Chen, Dachuan,PerAMykland,和Lan Zhang（2020年）。“桥下的五个巨魔：使用异步和含噪高频数据的主成分分析”。载于：《美国统计学会会刊》115.532，第1960–1977页。Chernov, Mikhail, A Ronald Gallant, EricGhysels,和GeorgeTauchen（2003年）。“股票价格动态的替代模型”。载于：《计量经济学杂志》116.1-2，第225–257页。Christensen, Kim, Silja Kinnebrock,和 MarkPodolskij（2010年）。“含噪扩散模型中非同步数据的事后协方差矩阵预平均估计量”。载于：《计量经济学杂志》159.1，第116–133页。Cui, Liyuan, Yongmiao Hong,YingxingLi,和JunhuiWang（2019年）。“通过低秩与稀疏矩阵分解对高频数据进行大维正定协方差估计”。载于：可见于 *SSRN 3414910*。El Euch, Omar 和 MathieuRosenbaum（2019年）。“粗糙 Heston 模型的特征函数”。载于：《数学金融》29.1，第3–38页。Engle,Robert和 Riccardo Colacito（2006年）。“检验和评估用于资产配置的动态相关性”。载于：《商业与经济统计期刊》24.2，第238–253页。Euch,OmarEl 和 Mathieu Rosenbaum（2018年）。“粗糙 Heston 模型中的完美对冲”。载于：《应用概率年刊》28.6，第3813–3856页。Figueroa-López,JoséE和 BeiWu（2022年）。“使用预平均方法在微观结构噪声下进行现货波动率的核估计”。载于：《计量经济理论》，第1–50页。Gatheral, Jim, Thibault Jaisson,和Mathieu Rosenbaum（2018年）。“波动率是粗糙的”。载于：《量化金融》18.6，第933–949页。Hansen, Peter R和 AsgerLunde（2006年）。“已实现方差与市场微观结构噪声”。载于：《商业与经济统计期刊》24.2，第127–161页。Hasbrouck, Joel（2007年）。《实证市场微观结构：证券交易的制度、经济学与计量经济学》。牛津大学出版社。Hayashi,Takaki 和 NakahiroYoshida（2005年）。“关于非同步观测扩散过程的协方差估计”。载于：《伯努利》11.2，第359–379页。Heston,StevenL（1993年）。“具有随机波动率的期权的闭式解及其在债券和货币期权中的应用”。载于：《金融研究评论》6.2，第327–343页。Huang,Xin和 GeorgeTauchen（2005年）。“跳跃对总价格方差的相对贡献”。载于：《金融计量经济学杂志》3.4，第456–499页。Kalnina, Ilze和Oliver Linton（2008年）。“在存在内生性和日内测量误差的情况下对二次变差进行一致估计”。载于：《计量经济学杂志》147.1，第47–59页。Liu,Nien-Lin和Hoang-LongNgo（2017年）。“现货交叉波动率矩阵特征值的近似：面向主成分分析”。载于：《日本工业与应用数学杂志》34.3，第747–761页。Malliavin,Paul 和MariaElvira Mancino（2002年）。“用于测量多元波动率的傅里叶级数方法”。载于：《金融与随机过程》6.1，第49–61页。

Malliavin,Paul 和 MariaElviraMancino（2009）。《一种用于多元波动率非参数估计的傅里叶变换方法》。载于：《统计年鉴》37.4，第1983–2010页。Mancino, Maria Elvira、TommasoMariotti 和GiacomoToscano（2022年）。《存在微观结构噪声时傅里叶现货波动率估计量的渐近正态性》。载于：*arXiv*预印本 *arXiv:2209.08967*。Mancino, Maria Elvira和MariaCristinaRecchioni（2015年）。《傅里叶现货波动率估计量：流动性高频数据和非流动性高频数据下的渐近正态性与效率》。载于：《*PLOS ONE*》10.9，e0139041。Mancino, Maria Elvira、Maria Cristina Recchioni 和 SimonaSanfelici（2017）。傅里叶–*Malliavin*波动率估计：理论与实践。Springer。Mancino,Maria Elvira 和Simona Sanfelici（2011）。《存在异步交易和微观结构噪声时通过傅里叶方法估计协方差》。载于：《金融计量经济学杂志》9.2，第367–408页。Mykland, Per A、LanZhang和 Dachuan Chen（2019）。《双尺度估计的代数及S-TSRV：对采样时间具有稳健性的高频估计》。载于：《计量经济学杂志》208.1，第101–119页。Park, Sujin、Seok YoungHong 和Oliver Linton（2016）。《对受加性测量误差污染的异步观测高频股票收益的二次协方差矩阵进行估计》。载于：《计量经济学杂志》191.2，第325–347页。Richard,Alexandre、Xiaolu Tan和 Fan Yang（2021）。《关于粗糙Heston模型的离散时间模拟》。载于：*arXiv*预印本 *arXiv:2107.07835*。Roll,Richard（1984）。《有效市场中有效买卖价差的一种简单隐含测度》。载于：《金融学杂志》39.4，第1127–1139页。Zu, Yang 和 HPeter Boswijk（2014年）。《利用高频金融数据估计现货波动率》。载于：《计量经济学杂志》181.2，第117–135页。