Modeling and Forecasting Realized Volatility with Multivariate Fractional Brownian Motion[*]

Markus Bibinger[a], Jun Yu[b], Chen Zhang[b]

[a]*Faculty of Mathematics and Computer Science, Institute of Mathematics, University of Würzburg*
[b]*Faculty of Business Administration, University of Macau*

Abstract

A multivariate fractional Brownian motion (mfBm) with component-wise Hurst exponents is used to model and forecast realized volatility. We investigate the interplay between correlation coefficients and Hurst exponents and propose a novel estimation method for all model parameters, establishing consistency and asymptotic normality of the estimators. Additionally, we develop a time-reversibility test, which is typically not rejected by real volatility data. When the data-generating process is a time-reversible mfBm, we derive optimal forecasting formulae and analyze their properties. A key insight is that an mfBm with different Hurst exponents and non-zero correlations can reduce forecasting errors compared to a one-dimensional model. Consistent with optimal forecasting theory, out-of-sample forecasts using the time-reversible mfBm show improvements over univariate fBm, particularly when the estimated Hurst exponents differ significantly. Empirical results demonstrate that mfBm-based forecasts outperform the (vector) HAR model.

Keywords: Forecasting, Hurst exponent, multivariate fractional Brownian motion, realized volatility, rough volatility

MSC classification: 62M10, 60G18, 62P05

1. Introduction

Modeling and forecasting financial asset volatility play a crucial role in numerous areas of finance. In the univariate context, several stylized facts about volatility dynamics—particularly long memory—have been well established since the seminal work of Ding et al. (1993). Subsequent research has introduced multiple generations of long-memory volatility models and their approximations, including contributions by Baillie et al. (1996); Harvey (2007); Comte and Renault (1996); Corsi (2009). Andersen and Bollerslev (1997) fur-

---

ther elucidated the mechanisms generating long memory in volatility. A recent advancement is the rough fractional stochastic volatility (RFSV) model of Gatheral et al. (2018), which employs fractional Brownian motion (fBm) with a Hurst exponent $H < 0.5$. This model has been linked to long-memory ARFIMA processes (Shi and Yu, 2023; Li et al., 2024; Wang and Yu, 2023) and has demonstrated superior forecasting performance, as evidenced by Gatheral et al. (2018) and Wang et al. (2024).

It is well-documented that financial volatilities exhibit strong co-movement across assets and markets. This stylized fact has spurred extensive research on multivariate volatility modeling, including surveys of multivariate GARCH and stochastic volatility models (Bauwens et al., 2006; Asai et al., 2006). These frameworks effectively capture not only individual asset volatilities but also cross-asset correlations, which are often even more critical for applications such as portfolio optimization and risk diversification. At the same time, they enhance the accuracy of volatility forecasts for individual assets. In contrast, the existing literature on fractional models predominantly concentrates on univariate approaches, neglecting the correlation structure and probably resulting in inefficiencies.

This paper proposes modeling and forecasting realized volatility using a multivariate fBm (mfBm) with component-wise Hurst exponents. The model extends the univariate fractional Brownian motion (fBm) framework, building on foundational work by Amblard et al. (2013) and Coeurjolly et al. (2013), which itself derives from the theoretical contributions of Lavancier et al. (2009, 2010); Didier and Pipiras (2011). It retains the appealing properties of univariate fBm—which has seen growing application across various domains—while addressing its limitations in capturing cross-sectional dynamics. The inclusion of component-specific Hurst exponents introduces significant mathematical complexity. This leads to a fundamental theoretical question that motivates our current investigation:

Can two real-valued processes exhibit markedly different autocorrelation structures and

persistence patterns while maintaining strong contemporaneous correlation? $\qquad$ (**Q**)

The mfBm framework allows for distinct Hurst exponents across components while capturing their interdependencies through correlation parameters. Each component possesses its own scale parameter, and cross-component covariances incorporate both correlation and asymmetry parameters—the latter capturing lead-lag effects between series. As illustrated in Figure 1, these features provide preliminary evidence addressing question (**Q**), with a com-
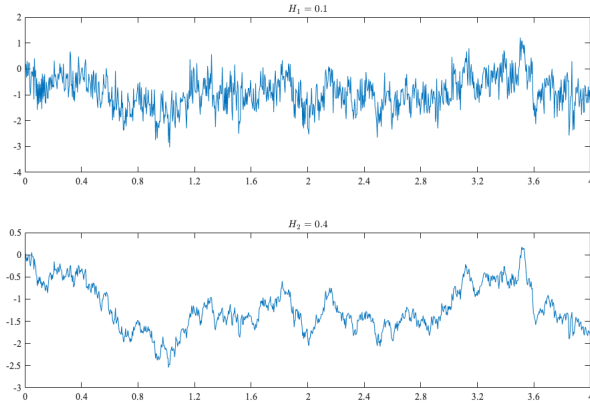
Figure 1: Simulated sample paths of fBm with $H_1 = 0.1$ (top), and $H_2 = 0.4$ (bottom), which are correlated with $\rho = 0.8$ and zero asymmetry parameter.

prehensive analysis developed in the following section. Given the well-documented strong correlation among realized volatilities (Ding et al., 2024), we investigate the forecasting implications of multivariate rough fractional volatility modeling.

We propose a novel moment-based estimation method for all model parameters. The Hurst exponents and scale parameters are estimated component-wise, while correlation and asymmetry parameters are estimated pairwise. Under mild regularity conditions, we establish the consistency of these estimators. The new estimators are compared to the existing estimators proposed by Amblard and Coeurjolly (2011). We demonstrate that our approach yields substantially improved estimation of asymmetry parameters while maintaining comparable performance for other parameters. Furthermore, we derive the asymptotic distribution of our estimators, enabling formal statistical inference. For forecasting applications, we focus on the time-reversible specification of mfBm (with zero asymmetry parameters), as our proposed hypothesis test fails to reject this null hypothesis for most assets at the 1% significance level.

Our theoretical analysis of time-reversible mfBm for realized volatility forecasting yields several key insights. First, forecasting improvements emerge when components exhibit heterogeneous Hurst exponents. Second, stronger cross-asset correlations further enhance forecast accuracy. Contrary to traditional wisdom that correlation enhances forecasting accuracy, the benefits of time-reversible mfBm depend critically on its Hurst exponents. We implement this framework to model multivariate log realized volatility series while producing univariate volatility forecasts. By partitioning the out-of-sample period into distinct temporal windows, we identify significant efficiency gains from mfBm precisely when estimated

Hurst exponents diverge across assets. Conversely, when Hurst exponents are similar, mfBm forecasts converge to their univariate fBm counterparts. These empirical findings robustly confirm our theoretical predictions about the conditions under which multivariate modeling provides forecasting advantages.

We empirically examine whether incorporating additional information through multivariate modeling necessarily improves forecasting efficiency by comparing the heterogeneous autoregressive (HAR) and vector HAR models, where the univariate HAR serves as our benchmark. While theory suggests that correctly specified multivariate models should outperform their univariate counterparts by leveraging additional information, our analysis reveals two key findings: (1) Contrary to theoretical expectations, the vector HAR model does not consistently deliver superior forecasting accuracy despite its richer information set; and (2) When utilizing identical information sets, mfBm consistently generates more precise forecasts than HAR-type models. These results demonstrate that mfBm's mathematical structure provides distinct advantages in effectively extracting forecasting value from multivariate information.

The remainder of this paper is organized as follows. Section 2 introduces the model and explains the covariance structure and probabilistic properties of mfBm. Section 3 develops our novel parameter estimators for mfBm and the associated asymptotic theory and proposes a test for time-reversibility. Section 4 develops a formula for the optimal forecast with time-reversible mfBm. Section 5 conducts a Monte Carlo study to examine the performance of the statistical methods and the forecast formula. Section 6 conducts an empirical study where mfBm is fitted to log realized volatility sequences and forecasting performance from alternative models is compared. Section 7 concludes. Section 8 collects proofs of the theorems in the paper. An appendix provides details about HAR and vector HAR models, and additional empirical results. Throughout the paper, we denote $\overset{p}{\to}$ convergence in probability, $\overset{d}{\to}$ convergence in distribution, $\sim$ asymptotic equivalence of real-valued sequences, $\overset{d}{=}$ equality in distribution and log natural logarithm.

2. Model, Correlation Structure and Properties

We begin with a review of univariate fBm and its key properties. A univariate fBm $(B_t^H)_{t \in \mathbb{R}}$ is a continuous Gaussian process with almost surely continuous sample paths satisfying: (1) Zero-mean: $\mathbb{E}[B_t^H] = 0$ $\forall t$; (2) Covariance structure: $\mathbb{Cov}(B_{t+h}^H, B_t^H) =$

$w(t, h, H)$ for $t, h \in \mathbb{R}$, where

$$w(t, h, H) = \frac{1}{2}\left(|t + h|^{2H} + |t|^{2H} - |h|^{2H}\right). \tag{1}$$

The process exhibits two fundamental characteristics: (1) Self-similarity: for any scaling factor $\lambda > 0$, $\left(B_{\lambda t}^H\right)_{t \in \mathbb{R}} \overset{d}{=} \left(\lambda^H B_t^H\right)_{t \in \mathbb{R}}$; (2) Stationary increments: for any $s, t \in \mathbb{R}$, $B_t^H - B_s^H \sim \mathcal{N}(0, |t - s|^{2H})$. $H \in (0, 1)$ governs both the covariance structure and self-similarity properties. Three important properties emerge. (1) When $H = 0.5$, the process reduces to standard Brownian motion; (2) When $H \neq 0.5$, fBm is neither a (semi-)martingale nor Markovian; (3) The sample paths exhibit Hölder continuity of order γ for any $\gamma < H$. For discrete observations at intervals $\Delta > 0$, the increment process $(\Delta_i B^H)_{1 \leq i \leq n}$ defined by $\Delta_i B^H = B_{i\Delta}^H - B_{(i-1)\Delta}^H$ is a stationary Gaussian process whose autocovariance function is

$$\mathbb{C}\text{ov}\left(\Delta_i B^H, \Delta_j B^H\right) = \frac{\Delta^{2H}}{2}\left((i - j + 1)^{2H} + (i - j - 1)^{2H} - 2(i - j)^{2H}\right), \text{ for any } i \geq j,$$
$$\sim \Delta^{2H} H(2H - 1)(i - j)^{2H-2}, \text{ for large } i - j. \tag{2}$$

Equation (2) shows that for $H \in (0.5, 1)$, $(\Delta_i B^H)_{1 \leq i \leq n}$ is serially dependent with positive autocorrelations and its autocovariances are not absolutely summable. That is, it is a long-memory process for $H > 0.5$. In contrast, if $H \in (0, 0.5)$, $(\Delta_i B^H)_{1 \leq i \leq n}$ has negative autocorrelations and

$$\sum_{i-j=-\infty}^{+\infty} \mathbb{C}\text{ov}\left(\Delta_i B^H, \Delta_j B^H\right) = 0,$$

such that the spectral density is zero at the origin. In this case, $(\Delta_i B^H)_{1 \leq i \leq n}$ is an antipersistent, or short-memory, process.

The mfBm is an extension of the real-valued univariate fBm to a multivariate process with similar properties and such that its components are fBms, with possibly different Hurst exponents, which are not required to be independent.[2] The mfBm is defined by Amblard et al. (2013) and its covariance structure can be introduced in a pairwise manner. For simplicity, we focus on a bivariate setting in the sequel. The bivariate fBm $(B_t) =$

---

[2]It should not be confused with the *multifractional Brownian motion* which is a univariate process with time-varying Hurst function.

$\left((B_t^{(1)}, B_t^{(2)})^\top\right)_{t\in\mathbb{R}}$ is constructed as a Gaussian process satisfying self-similarity:

$$\left((B_{\lambda t}^{(1)}, B_{\lambda t}^{(2)})^\top\right)_{t\in\mathbb{R}} \overset{d}{=} \left((\lambda^{H_1} B_t^{(1)}, \lambda^{H_2} B_t^{(2)})^\top\right)_{t\in\mathbb{R}}, \quad \text{for } \lambda > 0, \tag{3}$$

where $(H_1, H_2)^\top \in (0,1)^2$. Each component is a fBm with Hurst exponent $H_i \in (0,1)$ and scaling parameter $\sigma_i > 0$. Their covariance functions are

$$\mathbb{Cov}\left(B_t^{(i)}, B_s^{(i)}\right) = \frac{\sigma_i^2}{2}\left(|s|^{2H_i} + |t|^{2H_i} - |s-t|^{2H_i}\right) = \sigma_i^2\, w(t, s-t, H_i), \ i = 1, 2,$$

where $\sigma_i^2 = \mathbb{Var}(B_1^{(i)})$ and $w(\cdot,\cdot,\cdot)$ defined in (1). The cross-covariances of $B_s^{(1)}$ and $B_t^{(2)}$ given in (Amblard et al., 2013, Prop. 3) are more complex and divided into two cases: The general case with $H_1 + H_2 \neq 1$ and the specific case with $H_1 + H_2 = 1$.

1. If $H_1 + H_2 \neq 1$, there exists $(\rho, \eta_{1,2}) \in [-1, 1] \times \mathbb{R}$, such that

$$\mathbb{Cov}\left(B_s^{(1)}, B_t^{(2)}\right) = \frac{\sigma_1\sigma_2}{2}\left((\rho + \eta_{1,2}\,\text{sign}(s))\,|s|^{H_1+H_2} + (\rho - \eta_{1,2}\,\text{sign}(t))\,|t|^{H_1+H_2}\right.$$
$$\left. -(\rho - \eta_{1,2}\,\text{sign}(t-s))\,|t-s|^{H_1+H_2}\right). \tag{4}$$

2. If $H_1 + H_2 = 1$, there exists $(\tilde{\rho}, \tilde{\eta}_{1,2}) \in [-1, 1] \times \mathbb{R}$, such that

$$\mathbb{Cov}\left(B_s^{(1)}, B_t^{(2)}\right) = \frac{\sigma_1\sigma_2}{2}\left(\tilde{\rho}\,(|s|+|t|-|s-t|)+\tilde{\eta}_{1,2}(t\log|t|-s\log|s|-(t-s)\log|t-s|)\right).$$

The existence of the process can be established through a spectral integral representation (Amblard et al., 2013, Sec. 3), provided the parameters are chosen such that the covariance is well-defined (i.e., positive semi-definite). As the notation suggests, $\rho = \text{corr}(B_1^{(1)}, B_1^{(2)})$ is a correlation parameter. The second parameter carries a subscript to reflect its asymmetry: $\eta_{1,2} = -\eta_{2,1}$ and $\tilde{\eta}_{1,2} = -\tilde{\eta}_{2,1}$. This asymmetry arises when swapping the roles of the two components, i.e., transitioning from $\mathbb{Cov}(B_s^{(1)}, B_t^{(2)})$ to $\mathbb{Cov}(B_t^{(1)}, B_s^{(2)})$. These asymmetry parameters are defined by

$$\eta_{1,2} = \frac{\mathbb{E}[B_1^{(1)} B_{-1}^{(2)}] - \mathbb{E}[B_1^{(2)} B_{-1}^{(1)}]}{\sigma_1\sigma_2(2 - 2^{H_1+H_2})}, \ \text{if } H_1 + H_2 \neq 1,$$

$$\tilde{\eta}_{1,2} = \frac{\mathbb{E}[B_1^{(1)} B_{-1}^{(2)}] - \mathbb{E}[B_1^{(2)} B_{-1}^{(1)}]}{\sigma_1\sigma_2 2\log(2)}, \ \text{if } H_1 + H_2 = 1.$$

A key observation is that the cross-covariance function depends on the Hurst exponents only through their sum. While the general covariance structure is complex and potentially

difficult to interpret, an important special case is the symmetric scenario where $B_t \overset{d}{=} B_{-t}$, which is particularly relevant for financial applications. This defines the bivariate time-reversible fBm. In this simplified case, when asymmetry parameters vanish, the spectral density becomes real, and the covariance turns symmetric. That is, $\mathbb{C}\mathrm{ov}(B_s^{(1)}, B_t^{(2)}) = \mathbb{C}\mathrm{ov}(B_t^{(1)}, B_s^{(2)})$. This is the process we consider in the forecasting application. It has the following cross-covariance function

$$\mathbb{C}\mathrm{ov}\left(B_s^{(1)}, B_t^{(2)}\right) = \frac{\rho \sigma_1 \sigma_2}{2}\left(|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right) = \rho \sigma_1 \sigma_2 \, w(t, s-t, H), \qquad (5)$$

where $H = (H_1 + H_2)/2$ denotes the cross Hurst exponent. This process possesses a covariance structure where the cross-covariances mirror the form of univariate fBm autocovariances, incorporating both a correlation coefficient ρ and the cross Hurst exponent H. Importantly, the value of H determines the nature of interdependence. $H < 1/2$ results in a short-range interdependent process, while $H > 1/2$ leads to long-range interdependency.

A multivariate process with asymmetric covariances—distinct from the well-known asymmetry in return distributions—can induce lead-lag effects between components. Specifically, one component influences another with a delayed reaction. Although this phenomenon is relevant in certain applications (Wackernagel, 2003, p. 147), it appears less critical for volatility time series. For realized volatilities, estimated asymmetry parameters are typically close to zero. In Section 3, we develop a hypothesis test for time-reversibility. Empirical results show that the null hypothesis cannot be rejected in most cases. Consequently, we focus on the time-reversible mfBm in our forecasting exercise.

To ensure the covariance kernel of mfBm remains positive semi-definite, certain parameter restrictions must hold. In this section, we examine these constraints for the bivariate time-reversible fBm case. The more general setting, which incorporates asymmetry, leads to greater complexity and less intuitive parameter coherence; we refer to (Amblard et al., 2013, Fig. 1) for illustrative examples. In the bivariate case, the correlation is bounded above by a maximum value

$$\rho_{max}(H_1, H_2) = \frac{\sqrt{\sin(\pi H_1) \cdot \sin(\pi H_2) \cdot \Gamma(2H_1 + 1) \cdot \Gamma(2H_2 + 1)}}{\sin(\pi H) \cdot \Gamma(2H + 1)}.$$

For correlations ρ, with $|\rho| \le \rho_{max}$, time-reversible fBm exists; see (Amblard et al., 2013, page 10). Due to the symmetry between positive and negative correlations, i.e., $\rho_{min} =$
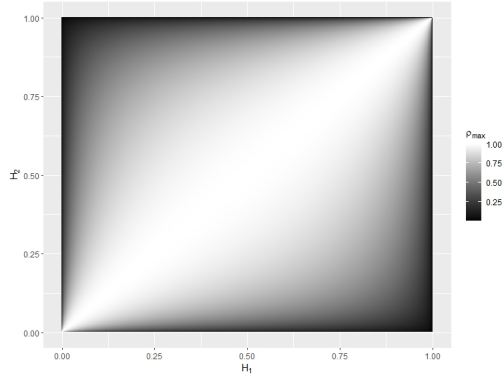
Figure 2: Maximal possible (absolute) correlation ρ_{max} of the bivariate time-reversible fBm with Hurst exponents H_1 and H_2.

$-\rho_{max}$, we only display the maximal absolute correlation ρ_{max} in Figure 2.

The plot resembles a heatmap, showing that smaller differences $|H_1 - H_2|$ permit higher correlations between components. Additionally, correlations tend to be larger when H_1, H_2, H are closer to $1/2$. In the special case where $H = H_1 = H_2$—which we term unifractional—there are no restrictions on $\rho \in [-1, 1]$. Here, the covariance structure simplifies, as the same Hurst exponent H governs all autocovariances, both marginal and cross-component. Only in this scenario can mfBm be represented as a linear transformation of a vector of independent fBms, a property we exploit in certain proofs.

However, when Hurst exponents differ significantly, the constraints tighten. For example, if $H_1 = 0.2$, $H_2 = 0.8$, the maximal correlation drops to $\rho_{max}(0.2, 0.8) \approx 0.662$. This answers our initial question ($\mathbf{Q}$): while not all combinations are feasible, the mfBm framework still accommodates such cases. Remarkably, even when the components exhibit starkly contrasting properties—one being rough and short-range dependent ($H_1 = 0.2$), the other smooth and long-range dependent ($H_1 = 0.8$)—a substantial correlation (up to $\sim 2/3$) is achievable. For more extreme disparities, such as $H_1 = 0.1$ and $H_2 = 0.9$, the maximum correlation further declines to $\rho_{max}(0.1, 0.9) \approx 0.383$. Sample paths illustrating such a behavior, with $H_1 = 0.1$, $H_2 = 0.4$, $\rho = 0.8$, are shown in Figure 1.

In Figure 3, we examine three fixed correlation values: $\rho = 0.5$, $\rho = 0.75$, and $\rho = 0.9$, shown in separate panels. The gray region in each panel represents the set of Hurst exponent pairs (H_1, H_2), for which bivariate time-reversible fBm exists. These admissible regions exhibit symmetry about the diagonal, a property that holds for all $\rho \in [-1, 1]$. Notably, the diagonal itself always lies entirely within the admissible parameter space. The left panel demonstrates that $\rho = 0.5$ is achievable for nearly all (H_1, H_2) combinations, with exclusion limited to cases where $|H_1 - H_2|$ approaches 1. The middle panel shows that $\rho = 0.75$
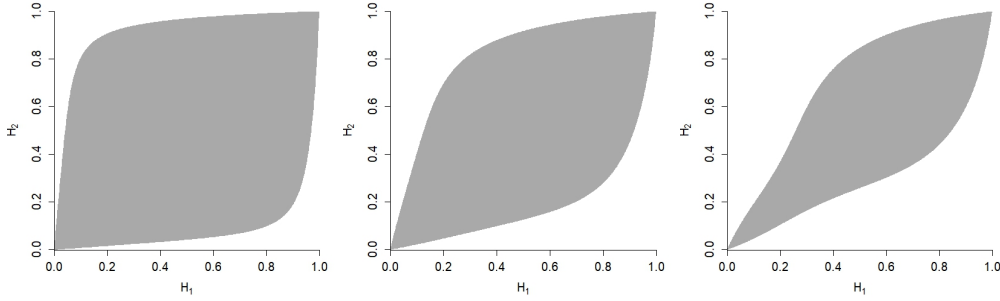
Figure 3: Parameter values for which bivariate time-reversible fBm exists (gray), and combinations that are not possible (white). The left-hand side is for $\rho = 0.5$, the middle for $\rho = 0.75$, and the right-hand side for $\rho = 0.9$.

remains feasible for most parameter pairs, though the admissible region begins to narrow. This restriction becomes particularly evident in the right panel ($\rho = 0.9$), where the gray area contracts significantly, indicating stronger constraints on permissible Hurst exponent combinations.

3. Statistics for Multivariate Fractional Brownian Motion

3.1. Observation model

Fractional time series models have demonstrated strong performance in volatility fore-casting applications, as evidenced by (Andersen et al., 2003; Chiriac and Voev, 2011) and related literature. Recent work, building on the framework established by Gatheral et al. (2018), has expanded this line of research to include fractional continuous-time models with Hurst exponents $H < 1/2$. These approaches, grounded in fBm, have shown particular promise for modeling and forecasting log-volatility processes.

Traditionally, fBm and its increment process—fractional Gaussian noise—were primar-ily employed to model long memory phenomena. Given the well-documented presence of long memory in volatility processes, Comte and Renault (1998) proposed using fBm with Hurst exponent $H > 1/2$ capturing this persistence property. However, recent empirical studies have consistently found that when modeling volatility using fBm, estimated Hurst exponents typically fall below $1/2$ (Gatheral et al., 2018; Wang et al., 2023; Bolko et al., 2023; Fukasawa et al., 2022; Bibinger and Sonntag, 2025). This is evident even from sim-ple autocorrelation plots of the data, which typically show negative autocorrelations in the increments. Therefore, the rough fractional stochastic volatility model from Gatheral et al. (2018) has become increasingly popular and has been utilized already for options pricing (Bayer et al., 2016; Garnier and Sølna, 2018), variance swaps (Bayer et al., 2016), portfolio

choice (Fouque and Hu, 2018), designing trading strategies (Glasserman and He, 2020), and dynamic hedging (Euch and Rosenbaum, 2018).

Akin to the approach of fractional time series in Andersen et al. (2003), Wang et al. (2023) employs discrete observations from fBm (or generalized fractional processes) to model log realized volatility—an observable quantity. In contrast, recent theoretical work has focused on latent spot volatility derived from discrete price recordings, providing evidence of roughness in spot volatility (Chong et al., 2024; Bolko et al., 2023; Fukasawa et al., 2022). Since our objective, like in Andersen et al. (2003); Wang et al. (2023), is to forecast realized volatility, we adopt a modeling framework for this observable measure. Specifically, we treat available time series as discrete observations of fractional processes, where $B_t^{(i)}$ is the log realized volatility of asset i at time t. This approach ensures that any empirical findings we establish pertain strictly to realized volatility, not spot volatility.

Consistent with the existing literature that employs univariate fBm for modeling and forecasting log realized volatilities, we adopt a statistical observation model based on equispaced high-frequency observations over a fixed time interval, such as $[0, 1]$. While the univariate framework provides valuable insights, extending the statistical inference methods to the multivariate case offers greater practical utility. We therefore consider observations specified by

$$\left(B_{j\Delta} \right) = \left(B_{j\Delta}^{(1)}, B_{j\Delta}^{(2)}, ..., B_{j\Delta}^{(d)} \right)^{\top}, \ 0 \leq j \leq n, \ \text{with} \ \Delta = 1/n \,,$$

where d denotes the dimension of $B_{j\Delta}$ and asymptotic results are established under the asymptotic scheme where $\Delta \to 0$, $n \to \infty$.[3]

3.2. Existing estimation method

The first and only available parameter estimation method for mfBm was introduced in Amblard and Coeurjolly (2011). Let $B_{k\Delta}^m$ denote the time series obtained by applying the

---

[3]The same high-frequency asymptotics have been employed in existing literature (Gatheral et al., 2018; Wang et al., 2023). However, our theory and proofs for mfBm leverage self-similarity to extend beyond high-frequency settings—for instance, to low-frequency data over long time spans (i.e., $T \to \infty$ with Δ fixed where $\Delta = T/n$). This long-span asymptotic scheme does not universally apply to univariate fractional models, as many rely on high-frequency assumptions to handle drift terms. In contrast, our results depend solely on self-similarity and the framework of Arcones (1994) for sums of transforms of Gaussian time series, rather than triangular arrays.

m-th dilated version of a filter to the d-dimensional $B_{k\Delta}$, where m is an integer and[4]

$$a_t^m = \begin{cases} a_{t/m}, & \text{if } t \in m\mathbb{Z}, \\ \\ 0, & \text{if } t \notin m\mathbb{Z} \end{cases},$$

with $\mathbb{Z} = \{0, 1, 2, ...\}$. To be more specific, the i-th component of $B_{k\Delta}^m$, denoted by $B_{k\Delta}^{(i),m}$, is

$$B_{k\Delta}^{(i),m} = \sum_{t \in \mathbb{Z}} a_t^m B_{(k-t)\Delta}^{(i)}.$$

The estimation principle for the case $H_i + H_j \neq 1$ in Amblard and Coeurjolly (2011) relies for a given $m \geq 1$ on the theoretical covariance

$$\begin{aligned} \mathbb{COV}_{ij}^m(h\Delta) :=& \mathbb{COV}(B_{k\Delta}^{(i),m}, B_{(k+h)\Delta}^{(j),m}) \\ =& -\frac{\sigma_i \sigma_j}{2} \sum_{t,l \in \mathbb{Z}} a_t a_l \left(\rho_{ij} - \eta_{ij} \text{sign}(h + m(t - l)) \right) |(h + m(t - l))\Delta|^{H_i + H_j}, \end{aligned}$$

for all $i, j = 1, \ldots, d$, and its empirical estimator

$$C_{ij}^m(h\Delta) = \frac{1}{n - ml - h} \sum_{k=ml+1}^{n-h} B_{k\Delta}^{(i),m} B_{(k+h)\Delta}^{(j),m}.$$

Subsequently, the differences between the theoretical covariances and their empirical counterparts are

$$\begin{aligned} \epsilon_{v_i}^m =& \log C_{ii}^m(0) - \log \mathbb{COV}_{ii}^m(0), \\ \epsilon_{c_{ij}}^m =& \log |C_{ij}^m(0)| - \log |\mathbb{COV}_{ij}^m(0)|, \\ \epsilon_{d_{ij}}^m =& \log \left(0.5 |C_{ij}^m(ml\Delta) - C_{ji}^m(ml\Delta)| \right) - \log \left(0.5 |\mathbb{COV}_{ij}^m(ml\Delta) - \mathbb{COV}_{ij}^m(ml\Delta)| \right). \end{aligned}$$

Amblard and Coeurjolly (2011) proposed to estimate the Hurst exponents, scaling, correlation and asymmetry parameters by minimizing the following weighted mean squared error

---

[4]a is in the set of filters,

$$\mathcal{A}_{l,q} = \left\{ (a_t)_{t \in \mathbb{Z}} : a_t = 0, \forall t \in \mathbb{Z}^- \cup \{l+1, ..., +\infty\} \text{ and } \sum_{t \in \mathbb{Z}} t^l a_t = 0, \forall l = 0, ..., q-1 \right\},$$

where $\mathbb{Z}^- = \{..., -2, -1, 0\}$, l and q be two positive integers. Typical examples are the difference filter and its compositions, Daubechies wavelet filters, and any known wavelet filter with compact support and a sufficient number of vanishing moments.

over all values of m taken from a discrete set $\mathcal{M}$:

$$\sum_{m \in \mathcal{M}} \left(w_v \sum_{i=1}^{d} (\epsilon_{v_i}^m)^2 + w_c \sum_{i=1, j>i}^{d} (\epsilon_{c_{ij}}^m)^2 + w_{dd} \sum_{i=1, j>i}^{d} (\epsilon_{d_{ij}}^m)^2 \right).$$

Amblard and Coeurjolly (2011) state that in a typical setting one can choose the weights $w_c = w_{dd} = 0$, which is what we adopt in the present paper. Moreover, Amblard and Coeurjolly (2011) prove that the estimator is consistent regardless of the choice of weights. They also provide the condition under which the estimator is asymptotically normally distributed. However, the complexity of the asymptotic variance expression poses significant challenges for constructing asymptotic confidence intervals in practice.

3.3. Our estimation method

Like the method of Amblard and Coeurjolly (2011), our new estimation method is also based on moments. The estimators of H_j and σ_j^2 are based on the marginal distributions of mfBm, while correlation and asymmetry parameters are estimated pairwise. Recall the notation for increments $(\Delta_k B^{(j)})_{1 \le k \le n}$, defined by $\Delta_k B^{(j)} = B_{k\Delta}^{(j)} - B_{(k-1)\Delta}^{(j)}$, which we use in the sequel with the components $B^{(j)}, 1 \le j \le d$.

For a simple notation, consider the estimation of a bivariate fBm with components $B^{(1)}$, and $B^{(2)}$, and with Hurst exponents H_1, H_2, correlation ρ, variance parameters σ_1^2, σ_2^2, and asymmetry parameter $\eta_{1,2}$. The Hurst exponent is component-wise estimated with

$$\hat{H}_j = \frac{1}{2 \log(2)} \log \left(\frac{\sum_{k=1}^{n-1} \left(\Delta_{k,2} B^{(j)} \right)^2}{\sum_{k=1}^{n} \left(\Delta_k B^{(j)} \right)^2} \right), \ j = 1, 2, \tag{6}$$

with the short notation $\Delta_{k,2} B^{(j)} = B_{(k+1)\Delta}^{(j)} - B_{(k-1)\Delta}^{(j)}, j = 1, 2$, for lag-2-increments. Then, the variance parameters can be estimated component-wise with

$$\hat{\sigma}_j^2 = \frac{\sum_{k=1}^{n} \left(\Delta_k B^{(j)} \right)^2}{n \Delta^{2\hat{H}_j}}, \ j = 1, 2. \tag{7}$$

To estimate ρ and $\eta_{1,2}$, as in Amblard and Coeurjolly (2011), we restrict to the case that $H_1 + H_2 \ne 1$, which is relevant for modeling realized volatility, as prior studies (e.g., Gatheral et al. (2018); Wang et al. (2023)) and our own estimates of roughness (reported in subsequent sections) consistently find Hurst exponents below 0.5. Our proposed estimators

of ρ and $\eta_{1,2}$ are

$$\hat{\rho} = \frac{\sum_{k=1}^{n} \Delta_k B^{(1)} \cdot \Delta_k B^{(2)}}{\sqrt{\sum_{k=1}^{n} \left(\Delta_k B^{(1)}\right)^2 \cdot \sum_{k=1}^{n} \left(\Delta_k B^{(2)}\right)^2}}, \tag{8a}$$

$$\hat{\eta}_{1,2} = \frac{\sum_{k=1}^{n-1} \left(\Delta_{k+1} B^{(2)} \cdot \Delta_k B^{(1)} - \Delta_{k+1} B^{(1)} \cdot \Delta_k B^{(2)}\right)}{\sqrt{\sum_{k=1}^{n-1} \left(\Delta_{k,2} B^{(1)}\right)^2 \sum_{k=1}^{n-1} \left(\Delta_{k,2} B^{(2)}\right)^2} - 2\sqrt{\sum_{k=1}^{n} \left(\Delta_k B^{(1)}\right)^2 \sum_{k=1}^{n} \left(\Delta_k B^{(2)}\right)^2}}. \tag{8b}$$

The estimators constitute proper statistics that depend solely on observed increments. In the construction of $\hat{\rho}$, the identical normalizing factors in both the numerator and denominator cancel out in the ratio. For $\hat{\eta}_{1,2}$, the covariance structure introduces an additional factor $2^{H_1+H_2} - 2$ in the numerator's expectation. This explains the inclusion of lag-2-increments in the denominator—their specific form ensures cancellation of this factor.

3.4. Asymptotic properties of estimators and a test for time-reversibility

The proposed estimators (6, 7, 8a, 8b) are consistent for general bivariate fBm under mild regularity conditions. Tables 8 and 9 report the estimated parameters from our dataset. The Hurst exponents are all below 0.5, the correlations of increments across components are approximately 0.3, and the asymmetry parameters are close to zero. The presence of nonzero correlations supports the adoption of a multivariate modeling framework, while the near-zero asymmetry parameters justify a focus on the time-reversible case. We derive the asymptotic distribution for the bivariate time-reversible fBm and develop a formal hypothesis test for $H_0 : \eta_{1,2} = 0$. The test results provide strong statistical evidence in favor of the time-reversible modeling assumption. Our asymptotic theory require the additional restriction $H < 3/4$, a condition satisfied by empirical estimates of realized volatility roughness, as discussed earlier.

Theorem 3.1. *For a general bivariate fBm satisfying* $\max(H_1, H_2) < 3/4$, $H_1 + H_2 \neq 1$, *the component-wise estimators obey the following asymptotic property: For each j, as $n \to \infty$,*

$$\hat{H}_j \xrightarrow{p} H_j, \quad \sqrt{n}\big(\hat{H}_j - H_j\big) \xrightarrow{d} \mathcal{N}\big(0, AVAR_{\hat{H}_j}\big), \tag{9}$$

where

$$AVAR_{\hat{H}_j} = \frac{1}{4\log^2(2)} \left(4 + \sum_{r=1}^{\infty} \left(|r+1|^{2H_j} + |r-1|^{2H_j} - 2|r|^{2H_j}\right)^2\right)$$

$$+ 2^{-4H_j} \sum_{r=1}^{\infty} \left(|r+2|^{2H_j} + |r-2|^{2H_j} - 2|r|^{2H_j} \right)^2$$

$$- 2^{1-2H_j} \sum_{r=1}^{\infty} \left(|r+1|^{2H_j} + |r-2|^{2H_j} - |r|^{2H_j} - |r-1|^{2H_j} \right)^2 \bigg).$$

Under the same conditions, for each j, as $n \to \infty$,

$$\hat{\sigma}_j^2 \xrightarrow{p} \sigma_j^2, \quad \frac{\sqrt{n}}{\log(n)} \left(\hat{\sigma}_j^2 - \sigma_j^2 \right) \xrightarrow{d} \mathcal{N} \left(0, AVAR_{\hat{H}_j} \cdot 4\sigma_j^4 \right). \tag{10}$$

Remark 3.1. *The component-wise estimators $\hat{H}_j$ and $\hat{\sigma}_j$, for $j = 1, 2$, are unaffected by asymmetry. Their estimation results remain invariant regardless of whether the underlying model is a general bivariate fBm or a bivariate time-reversible fBm.*

The estimator (6) achieves consistency with the optimal parametric convergence rate of $\sqrt{n}$ when $H_j < 3/4$. This condition mirrors the requirement for limit theory of normalized realized variances in univariate fBm (see, e.g., Nourdin et al. (2010)). While general asymptotic theory for univariate fBm estimators based on first-order increments exists (e.g., Coeurjolly (2001)), we provide a complete proof for (9) to derive explicit variance formulas for our standard estimator (6). Alternative estimators using second-order increments, similar in form to (6), can be constructed to satisfy central limit theorems for all $H \in (0,1)$ (Kubilius, 2020). Such second-order increment statistics have been extensively studied (e.g., Bégyn (2007), Wang et al. (2023), Bibinger and Sonntag (2025)). However, our estimator (6) offers two key advantages: (1) a simpler asymptotic variance structure, and (2) uniformly smaller variance across $H \in (0, 3/4)$ compared to second-order increment estimators. These properties make (6) particularly suitable for rough volatility series with small Hurst exponents. The convergence rate in (10), including its logarithmic factor, is also optimal (Brouste and Fukasawa, 2018, Section 4.1). Notably, the relationship between asymptotic variances in (9) and (10) parallels that of second-order increment estimators in Theorem 4.1 of Wang et al. (2023).

Theorem 3.2. *For a general bivariate fBm with $\max(H_1, H_2) < 3/4$, $H_1 + H_2 \neq 1$, as $n \to \infty$, we have*

$$\hat{\eta}_{1,2} \xrightarrow{p} \eta_{1,2}, \quad \hat{\eta}_{1,2} - \eta_{1,2} = \mathcal{O}_{\mathbb{P}} \left(n^{-1/2} \right).$$

In the time-reversible case, it satisfies a central limit theorem with an asymptotic variance $AVAR_{\hat{\eta}_{1,2}}(H_1, H_2, \rho)$ given by (26) in Section 8.

Based on the developed asymptotic theory for $\hat{\eta}_{1,2}$, in the following corollary, we propose a test for time-reversibility.

Corollary 3.1. *Given a bivariate fBm with* $\max(H_1, H_2) < 3/4$, $H_1 + H_2 \neq 1$, *the test rejects the null hypothesis* $\eta_{1,2} = 0$, *in favor of the alternative hypothesis that* $\eta_{1,2} \neq 0$, *when*

$$\sqrt{n} \frac{|\hat{\eta}_{1,2}|}{\sqrt{AVAR_{\hat{\eta}_{1,2}}(\hat{H}_1, \hat{H}_2, \hat{\rho})}} > \Phi^{-1}(1 - \alpha/2), \tag{11}$$

where $\alpha \in (0,1)$, *and* Φ *is the cumulative distribution function of the standard normal. This test attains the asymptotic size* α *and the asymptotic power* 1.

Remark 3.2. *There are several advantages in our estimators relative to those proposed by Amblard and Coeurjolly (2011). First and foremost, our estimators have closed-form expressions, facilitating the computation. Second, our estimators have simpler asymptotic variance structures that are more straightforward to estimate. Third, our asymmetry parameter estimator shows superior performance compared to that of Amblard and Coeurjolly (2011), as demonstrated in Table 3. This improvement is notable given Amblard and Coeurjolly (2011)'s own observation that estimating* $\eta_{1,2}$ *was "very difficult to estimate, at least with the method adopted here". Our approach successfully overcomes this estimation challenge.*

Theorem 3.3. *For a bivariate time-reversible fBm with* $\max(H_1, H_2) < 3/4$, $|\rho| < 1$, $H_1 + H_2 \neq 1$, *as* $n \to \infty$, *we have*

$$n^{-1} \Delta^{-H_1 - H_2} \sum_{k=1}^{n} \Delta_k B^{(1)} \cdot \Delta_k B^{(2)} \xrightarrow{p} \varsigma = \rho \sigma_1 \sigma_2,$$

$$\sqrt{n} \left(n^{-1} \Delta^{-H_1 - H_2} \sum_{k=1}^{n} \Delta_k B^{(1)} \cdot \Delta_k B^{(2)} - \varsigma \right) \xrightarrow{d} \mathcal{N} \left(0, \sigma_1^2 \sigma_2^2 \left(1 + \rho^2 + \rho^2 \upsilon_1 + \upsilon_2 \right) \right), \tag{12}$$

where

$$\upsilon_1 = \upsilon_1(H_1, H_2) = \frac{1}{2} \sum_{r=1}^{\infty} \left(|r+1|^{H_1 + H_2} + |r-1|^{H_1 + H_2} - 2|r|^{H_1 + H_2} \right)^2, \tag{13a}$$

$$\upsilon_2 = \frac{1}{2} \sum_{r=1}^{\infty} \left(|r+1|^{2H_1} + |r-1|^{2H_1} - 2|r|^{2H_1} \right) \left(|r+1|^{2H_2} + |r-1|^{2H_2} - 2|r|^{2H_2} \right). \tag{13b}$$

Under the same set of assumptions, the estimator $\hat{\rho}$ *defined in (8a) satisfies, as* $n \to \infty$,

$$\hat{\rho} \xrightarrow{p} \rho, \ \sqrt{n}(\hat{\rho} - \rho) \xrightarrow{d} \mathcal{N} \left(0, AVAR_{\hat{\rho}} \right), \tag{14}$$

15

where

$$AVAR_{\hat{\rho}} = (1 - \rho^2)^2 + \rho^2\Big((1 + \rho^2)v_1(H_1, H_2) + \tfrac{v_1(H_1, H_1)}{2} + \tfrac{v_1(H_2, H_2)}{2} - v_3(H_1, H_2)$$
$$-v_3(H_2, H_1)\Big) + v_2,$$

$$v_3(H_1, H_2) = \sum_{r=1}^{\infty} \big(|r+1|^{2H_1} + |r-1|^{2H_1} - 2|r|^{2H_1}\big)\big(|r+1|^{H_1+H_2} + |r-1|^{H_1+H_2} - 2|r|^{H_1+H_2}\big).$$

Remark 3.3. $\hat{\rho} - \rho = \mathcal{O}_{\mathbb{P}}\big(n^{-1/2}\big)$ *will generalize to a general bivariate fBm. A non-zero asymmetry parameter will just influence the asymptotic variance, which we compute for the time-reversible case of interest.*

Remark 3.4. *When we set all series coefficients (v_1, v_2, v_3) in Theorem 3.3 to zero, the resulting asymptotic variances correspond exactly to those of a multivariate i.i.d. model—or equivalently, a multivariate Brownian motion. Specifically: (1) For the empirical covariance in (12), the asymptotic variance reduces to $(1 + \rho^2)\sigma_1^2\sigma_2^2$; (2) For the empirical correlation in (14), the asymptotic variance becomes $(1 - \rho^2)^2$.*

4. Optimal Forecast

Given the empirical evidence showing only weak asymmetry in multivariate volatility dynamics, we recommend employing time-reversible mfBm for modeling log realized volatility. In what follows, we examine the forecasting advantages offered by this modeling approach.

The conditional expectation represents the optimal forecast in that it minimizes the mean squared forecast error (MSFE). In practice, almost always discrete observations over multiple periods are available. Consequently, the optimal forecast of the mfBm is obtained through conditional expectation based on the multivariate normal distribution[5]. Consider one has equispaced discrete observations $\mathcal{X}_{t,d} = \{B_\Delta, \ldots, B_{t\Delta}\}$ from a d-dimensional mfBm $B = (B^{(1)}, \ldots, B^{(d)})^\top$, where Δ is the sampling interval and t is the number of periods where the mfBm is observed. In the volatility literature, the length of a year is conventionally

---

[5]When continuous records are available, there exist formulas for generating the optimal forecast. For example, in an idealized continuous-time setting where a univariate fBm is observed over the infinite past $(-\infty, t]$, (Nuzman and Poor, 2000, Eq. (34)) derive the conditional expectation for the process at future time $t + h$. However, this result cannot be directly applied in practice due to two key limitations: (1) real-world observations are available only at discrete time points, and (2) historical data spans a finite, rather than infinite, time window. To address these constraints, Gatheral et al. (2018) propose necessary modifications to the original formula. As comprehensively reviewed in Wang et al. (2024), the method proposed by Gatheral et al. (2018) performs worse than the conditional expectation based on the multivariate normal distribution.

normalized to 1. Hence, for $\Delta = 1/12, 1/52, 1/252$, $\mathcal{X}_{t,d}$ represents monthly, weekly, and daily observations accordingly. We stack the data $\mathcal{X}_{t,d}$ into a column vector, such that

$$\mathcal{X}_{t,d} = (B_\Delta^{(1)}, \ldots, B_\Delta^{(d)}, \ldots, B_{t\Delta}^{(1)}, \ldots, B_{t\Delta}^{(d)})^\top.$$

The optimal h-step-ahead forecast of $B_{(t+h)\Delta}^{(j)}$, for any $j \in \{1, \ldots, d\}$, is again the conditional mean and has a closed-form expression of the form

$$\hat{B}_{(t+h)\Delta|(1:t)\Delta}^{(j)} = \mathbb{E}\left[B_{(t+h)\Delta}^{(j)} | \mathcal{X}_{t,d}\right] = \left(\gamma_{t,h}^j\right)^\top \Sigma_{t,d}^{-1} \mathcal{X}_{t,d}, \tag{15}$$

where $\Sigma_{t,d}$ denotes the covariance matrix of $\mathcal{X}_{t,d}$, whose elements are readily obtained from (5) and $\gamma_{t,h}^j$ is a vector that contains the covariance of $B_{(t+h)\Delta}^{(j)}$ with each entry of $\mathcal{X}_{t,d}$, which are also readily obtained from (5).

Since the explicit formula (15) is generally complex, the forecasting advantages of this modeling approach over the univariate framework are not immediately apparent. To develop key insights into forecasting within the mfBm model, we begin with simplified scenarios—specifically, the optimal forecast based on a single observation: (1) single-period forecasts in bivariate settings, and (2) their multivariate extensions, for which explicit expressions for the weights in (15) are available. The general case, involving multiple observations of discretized paths, will be treated later. This approach is inspired by known results for univariate fBm, where optimal forecasts assign maximal weight to the most recent observation and nearly all weight to only a few recent data points—a pattern demonstrated in Wang et al. (2024).

4.1. Optimal forecast with observations in one period

In this subsection, we first assume that a single observation in a bivariate setting is the only available data. For simplicity, we ignore the sampling frequency Δ and denote the observation as $(B_t^{(1)}, B_t^{(2)})$. Recall that $H = (H_1 + H_2)/2$.

Proposition 4.1. *Consider a bivariate time-reversible fBm $(B_t^{(1)}, B_t^{(2)})^\top$. The optimal h-step-ahead forecast of $B_{t+h}^{(1)}$, conditional on $B_t^{(1)}, B_t^{(2)}$, is given by*

$$\hat{B}_{t+h|t}^{(1)} = w_{t+h|t}^{11} \cdot B_t^{(1)} + w_{t+h|t}^{12} \cdot B_t^{(2)}, \tag{16}$$

for any $\rho \in (-1, 1)$, with weight functions

$$w^{11}_{t+h|t} = \frac{1}{1-\rho^2}\Big(\frac{w(t,h,H_1)}{t^{2H_1}} - \rho^2 \frac{w(t,h,H)}{t^{2H}}\Big),$$

$$w^{12}_{t+h|t} = \frac{\rho}{1-\rho^2}\frac{\sigma_1}{\sigma_2}\Big(\frac{w(t,h,H)}{t^{2H_2}} - \frac{w(t,h,H_1)}{t^{2H}}\Big),$$

where $w(t,h,H)$ is defined in (1). Its MSFE is

$$\mathbb{E}\Big[\big(\hat{B}^{(1)}_{t+h|t} - B^{(1)}_{t+h}\big)^2\Big] = \sigma_1^2(t+h)^{2H_1} - \frac{\sigma_1^2}{1-\rho^2}\frac{(w(t,h,H_1))^2}{t^{2H_1}}$$
$$+ \frac{\sigma_1^2\rho^2}{1-\rho^2}\Big(\frac{2w(t,h,H_1)w(t,h,H)}{t^{2H}} - \frac{(w(t,h,H))^2}{t^{2H_2}}\Big).$$

The roles of both components are symmetric and can be interchanged. We now examine important implications arising from Proposition 4.1. Incorporating $B^{(2)}_t$ consistently improves forecasting accuracy by reducing the MSFE, or at worst maintains equivalent performance. This improvement becomes evident when we express the term $(1-\rho^2)^{-1}$ as $1 + \rho^2/(1-\rho^2)$ and represent the reduction in forecasting error relative to the univariate case as a squared quantity. Regarding the role of each parameter, the MSFE of $\hat{B}^{(1)}_{t+h|t}$ does not depend on σ_2; only the Hurst exponents and correlation coefficients are relevant. We highlight several key special cases related to these parameters.

- If $H_1 = H_2 = H$, then $w^{12}_{t+h|t} = 0$ for any ρ. Thus, in the unifractional case, the optimal forecast coincides with the optimal forecast conditional on $B^{(1)}_t$ only. That is, we do not attain any gains from observing the second component $B^{(2)}_t$. The MSFE is $\sigma_1^2((t+h)^{2H_1} - (w(t,h,H_1))^2 t^{-2H_1})$, and $w^{11}_{t+h|t} = w(t,h,H_1)t^{-2H_1}$.

- If $\rho = 0$, then $w^{12}_{t+h|t} = 0$ for any H_2. Hence, the optimal forecast coincides with the optimal forecast conditional on $B^{(1)}_t$ only.

Contrary to traditional wisdom that correlation enhances forecasting accuracy, the benefits of time-reversible mfBm depend critically on its Hurst exponents. Additionally, for arbitrary parameter values, the weights satisfy the following limiting behavior as the forecast horizon approaches zero: $w^{11}_{t+h|t} \to 1$ and $w^{12}_{t+h|t} \to 0$.

Figure 4 illustrates the role of the Hurst exponents in determining the weights by fixing the moderate and empirically relevant coefficients at $\rho = 1/2$ and $H_1 = 0.4$, while allowing H_2 to vary. The other parameters are set to $\sigma_1 = \sigma_2 = 1$, $t = 1$, and $h = 1$. Although in this case the weight function is drawn over the whole support $H_2 \in (0,1)$,
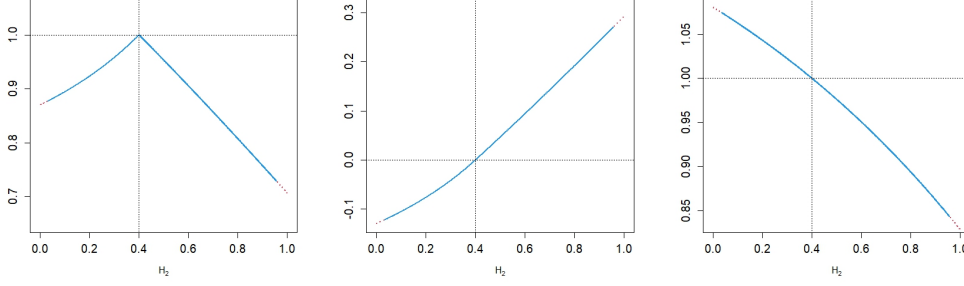
Figure 4: Relative weights $w_{2|1}^{11}/(|w_{2|1}^{11}| + |w_{2|1}^{12}|)$ (left) and $w_{2|1}^{12}/(|w_{2|1}^{11}| + |w_{2|1}^{12}|)$ (middle) and $w_{2|1}^{11}t^{2H_1}/w(1,1,H_1)$ (right), for parameters $\rho = 1/2$, $\sigma_1 = \sigma_2 = 1$ and $H_1 = 0.4$ as functions of H_2.

some parameter configurations, especially when H_2 takes values close to the boundaries, are not possible. Over these sub-intervals the function is drawn with a dotted red line. The left panel presents the relative weight when the first component is forecasted, computed as $w_{2|1}^{11}/(|w_{2|1}^{11}| + |w_{2|1}^{12}|)$. The middle panel shows the relative weight on the other component, computed as $w_{2|1}^{12}/(|w_{2|1}^{11}| + |w_{2|1}^{12}|)$. The right panel plots $w_{2|1}^{11}t^{2H_1}/w(1,1,H_1)$, which is the ratio of the values of the first weight when $\rho = 0.5$ and those when $\rho = 0$, the latter coinciding with the factor of the univariate forecast based on observing $B_t^{(1)}$ only.

Several observations emerge. First, in general, the weights on the second component are nonzero, indicating that relying solely on the first component's own past observation leads to inefficiency in forecasting. Second, as predicted by our theory, when $H_2 = H_1 = 0.4$, all the weight is placed on $B_t^{(1)}$, and $B_t^{(2)}$ is not used in the forecast, despite the moderate correlation coefficient $\rho = 1/2$, as shown in all three panels. Third, the sign of the weight depends on the relative magnitude of H_2 and H_1. The middle panel shows that the relative weight is a strictly increasing function: it equals zero when $H_2 = 0.4$, becomes negative when $H_2 < 0.4$, and approaches 0.3 as H_2 nears 1. Fourth, the right panel shows that $w_{2|1}^{11}t^{2H_1}/w(1,1,H_1)$ is strictly decreasing, implying that the univariate forecast shifts from underweighting to overweighting its own past observation as H_2 increases.

The left panel of Figure 5 displays the relative MSFE for forecasting $B_2^{(1)}$ based on $(B_1^{(1)}, B_1^{(2)})$, relative to the univariate benchmark that uses only $B_1^{(1)}$, corresponding to the parameter settings in Figure 4. The values are hence normalized so that the maximum equals 1. In all cases where $H_1 \neq H_2$, we observe efficiency gains (values below 1) relative to the univariate forecast, with the improvements being meaningful. If we increase the correlation coefficient to $\rho = 0.9$, as shown in the right panel, the forecast error reduction becomes more pronounced, highlighting the role of correlation in enhancing bivariate forecast.

19
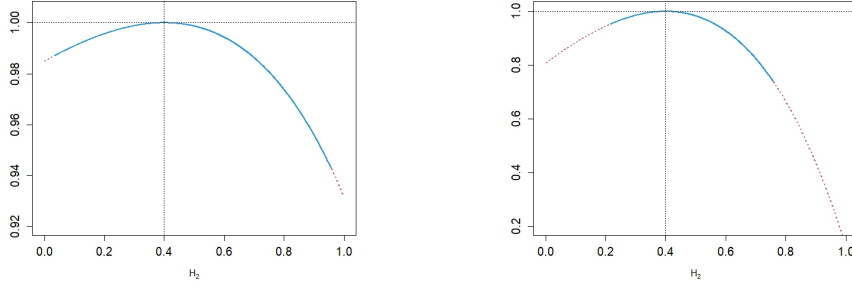
Figure 5: Relative MSFEs as functions of H_2 for $t = 1$, $h = 1$, $\sigma_1 = \sigma_2 = 1$, with $\rho = 0.5$ and $H_1 = 0.4$ (left), with $\rho = 0.9$ and $H_1 = 0.4$ (right).

Propositions 4.2 and 4.3 extend our bivariate results to the general d-dimensional setting. A key finding that persists in this broader context is the absence of forecasting gains from multivariate modeling when all Hurst exponents are equal.

Proposition 4.2. *Consider a time-reversible mfBm* $(B_t^{(1)}, \ldots, B_t^{(d)})^\top$*, and denote* $\Sigma = (\Sigma_{ij})_{1 \leq i,j \leq d} \in \mathbb{R}^{d \times d}$ *the positive definite covariance matrix of* $(B_1^{(1)}, \ldots, B_1^{(d)})^\top$*, with inverse* $\Sigma^{-1} = ((\Sigma^{-1})_{ij})_{1 \leq i,j \leq d}$*. The optimal h-step-ahead forecast of* $B_{t+h}^{(1)}$*, given* $B_t^{(1)}, \ldots, B_t^{(d)}$*, is given by*

$$\hat{B}_{t+h|t}^{(1)} = \sum_{j=1}^{d} w_{t+h|t}^{1j} \cdot B_t^{(j)}, \tag{17}$$

where

$$w_{t+h|t}^{1j} = \sum_{k=1}^{d} \Sigma_{1k}(\Sigma^{-1})_{kj}\Big(\frac{w\big(t, h, (H_1 + H_k)/2\big)}{t^{H_j + H_k}}\Big), 1 \leq j \leq d.$$

A key remaining question concerns how higher dimensions ($d > 2$) may further reduce forecasting errors when (1) correlations are non-zero and (2) Hurst exponents differ across components. The following specific result provides valuable insight through its tractable closed-form expressions, revealing how forecast accuracy improves with increasing dimensionality.

Proposition 4.3. *Consider mfBm with dimension $d > 2$, with $\sigma_j^2 = 1$ for all j, and all correlations equal to $\rho > 0$. Assume H_1 is the Hurst exponent of $(B_t^{(1)})$, while $(B_t^{(j)})$ has Hurst exponent $H_j = H \neq H_1$, for all $2 \leq j \leq d$.[6] The optimal forecast for $B_{t+h}^{(1)}$, given*

---

[6]We slightly abuse the notation here as H is not the average of all Hurst exponents.

20
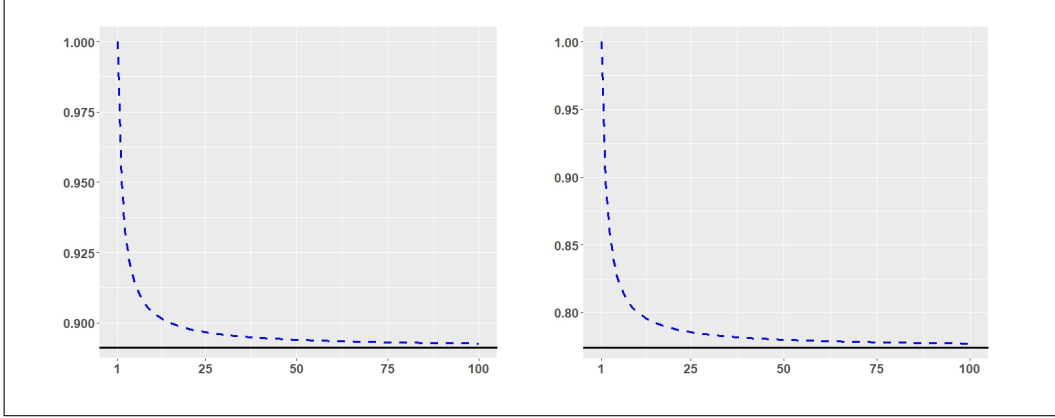
Figure 6: Ratios of MSFE in Prop. 4.3, with $\rho = 0.8$, $H_1 = 0.4$, $H = 0.1$, and that in the univariate model, for $t = 1, h = 1$ (left) and $t = 10, h = 1$ (right), as a function of dimension d. The horizontal reference lines show the limits as $d \to \infty$.

$(B_t^{(1)}, B_t^{(2)}, \ldots, B_t^{(d)})^\top$, *has the following expression for MSFE*

$$\mathbb{E}\left[\left(\hat{B}_{t+h|t}^{(1)} - B_{t+h}^{(1)}\right)^2\right] = (t+h)^{2H_1} - \frac{1}{1 + (d-2)\rho - (d-1)\rho^2}\left((1 + (d-2)\rho)\frac{(w(t,h,H_1))^2}{t^{2H_1}}\right.$$
$$\left. - 2(d-1)\rho^2\frac{w(t,h,H_1)w(t,h,H)}{t^{2H}} + (d-1)\rho^2\frac{(w(t,h,H))^2}{t^{2H_2}}\right).$$

In particular, a strictly positive limit exists as $d \to \infty$.

Since the equal correlation condition requires $\rho \in (-(d-1)^{-1}, 1)$, we restrict our analysis to $\rho > 0$ in Proposition 4.3, which is empirically relevant. Figure 6 illustrates the results by plotting the relative MSFEs for $t = 1$ and $t = 10$, $h = 1$, $\rho = 0.8$, $H_1 = 0.1$ and $H = 0.4$. Here, the relative MSFE is normalized by the minimal univariate MSFE and evaluated across dimensions $d = 1, \ldots, 100$. The horizontal lines indicate the asymptotic limits as $d \to \infty$.

Compared to the bivariate case, incorporating additional components significantly reduces the forecast error. However, the risk does not vanish as d increases. In fact, for the scenario in Figure 6, the MSFE improvement becomes almost negligible beyond $d \geq 20$. The limiting relative MSFE as $d \to \infty$ settles around 89% of the minimal univariate case for $t = 1$. Setting $t = 10$, yields a smaller limit of approximately 77.4% (even for much lager t, the limit further decreases only slightly). Thus, Proposition 4.3 presents a mixed outcome for forecasting efficiency: while substantial gains are achievable by increasing dimensionality, these gains plateau early, and the forecasting risk remains bounded away from zero.

4.2. Optimal forecast with all historical observations

Proposition 4.4 establishes that, for a bivariate time-reversible fBm with identical Hurst exponents $H_1 = H_2 = H$, observing the historical path of the second component $B_t^{(2)}$ provides no additional forecasting benefit. This result extends naturally to the d-dimensional case in Proposition 4.5, provided all Hurst exponents coincide.

Table 1: Bias, standard error (asymptotic standard errors in parentheses) and RMSE of our method: $H_1 = 0.1$, $H_2 = 0.4$, $\sigma_1^2 = 1$, $\sigma_2^2 = 1$, $\rho = 0$ and $\eta_{1,2} = 0$ (denoted η in the table).

		H_1	H_2	σ_1^2	σ_2^2	ρ	η
	Bias	0.0013	-0.0012	0.0567	0.0222	0.0015	0.0003
$n = 500, \Delta = 1/52$	Std	0.0441	0.0356	0.3449	0.2934	0.0482	0.1192
		(0.0431)	(0.0351)	(0.3404)	(0.2774)	(0.0472)	(0.1137)
	RMSE	0.0441	0.0357	0.3496	0.2943	0.0483	0.1192
	Bias	-0.0009	0.0001	0.0144	0.0156	-0.0005	-0.0041
$n = 1000, \Delta = 1/52$	Std	0.0293	0.0248	0.2246	0.1977	0.0334	0.0788
		(0.0305)	(0.0248)	(0.2407)	(0.1962)	(0.0334)	(0.0804)
	RMSE	0.0293	0.0248	0.2250	0.1984	0.0334	0.0789
	Bias	-0.0014	-0.0025	0.0837	0.0445	-0.0011	-0.0017
$n = 500, \Delta = 1/250$	Std	0.0430	0.0353	0.5132	0.4075	0.0480	0.1109
		(0.0431)	(0.0351)	(0.4756)	(0.3876)	(0.0472)	(0.1137)
	RMSE	0.0430	0.0354	0.5203	0.4101	0.0480	0.1110
	Bias	-0.0005	0.0001	0.0399	0.0366	-0.0015	0.0022
$n = 1000, \Delta = 1/250$	Std	0.0301	0.0246	0.3285	0.2800	0.0333	0.0759
		(0.0305)	(0.0248)	(0.3363)	(0.2741)	(0.0334)	(0.0804)
	RMSE	0.0301	0.0246	0.3309	0.2824	0.0334	0.0760

Proposition 4.4. *Consider a bivariate time-reversible fBm with a positive definite covariance and $H_1 = H_2 = H$. The optimal forecast for $B_{t+h}^{(1)}$, given t discrete observations $(B_\Delta^{(1)}, B_\Delta^{(2)}), \dots, (B_{t\Delta}^{(1)}, B_{t\Delta}^{(2)})$, does not depend on the component of $(B^{(2)})$ and coincides with the optimal forecast based on univariate fBm.*

The proof of Proposition 4.4 benefits from a more explicit and transparent representation of the conditional expectation in the bivariate case compared to the general d-dimensional setting. Nevertheless, the overarching result in Proposition 4.5 and its proof strategy follow a similar approach.

Proposition 4.5. *Consider a time-reversible mfBm with a positive definite covariance and all component-wise Hurst exponents equal. The optimal forecast for $B_{t+h}^{(1)}$, given t discrete observations $(B_{\Delta}^{(1)}, \ldots, B_{\Delta}^{(d)}), \ldots, (B_{t\Delta}^{(1)}, \ldots, B_{t\Delta}^{(d)})$, does only depend on the first component $(B^{(1)})$ and coincides with the optimal forecast based on univariate fBm.*

5. Monte Carlo Study

5.1. Finite-sample performance of proposed estimators

We first examine the finite-sample performance of the proposed estimators. Since our estimators are derived either component-wise or pairwise, we consider the bivariate time-reversible fBm without loss of generality. Two parameter settings in the rough region are considered: $\{H_1, H_2, \sigma_1^2, \sigma_2^2, \rho_{12}\} = \{0.1, 0.4, 1, 1, 0\}, \{0.1, 0.4, 1, 1, 0.4\}$. We set the sample size to $n = 500, 1000$, $\Delta = 1/52, 1/250$, and conduct 1000 replications.

Tables 1-2 report the bias, the standard deviation, and the root mean squared error (RMSE) of the estimators across replications, with the standard error implied by our asymptotic theory in parentheses. The proposed method performs well and the empirical standard deviation is accurately predicted by the asymptotic theory.

We next evaluate the finite-sample performance of our proposed estimators in comparison with those from Amblard and Coeurjolly (2011). As our estimators are consistent for general mfBm, we can evaluate their performance when $\eta_{1,2} \neq 0$. We focus on the bivariate case, as Amblard and Coeurjolly (2011) noted that estimating $\eta_{1,2}$ was "very difficult" with their method. If their estimator for $\eta_{1,2}$ performs poorly in the bivariate case, it is unlikely to improve in higher dimensions, whereas our estimators are dimension-independent. Four parameter configurations are considered: $\{H_1, H_2, \sigma_1^2, \sigma_2^2, \rho_{12}, \eta_{1,2}\} = \{0.1, 0.4, 1, 1, 0, 0\}$, $\{0.1, 0.4, 1, 1, 0, 0.5\}$, $\{0.1, 0.4, 1, 1, 0.4, 0\}$, $\{0.1, 0.4, 1, 1, 0.4, 0.5\}$. The sample size is set to $n = 500, 1000$, and the sample interval Δ to $1/250$. For clarity, we denote our estimators as BYZ and those from Amblard and Coeurjolly (2011) as AC. The AC estimators are implemented using Coeurjolly's `SimEstFBM.R` package with its default configuration.

Table 2: Bias, standard error (asymptotic standard errors in parentheses) and RMSE of our method: $H_1 = 0.1$, $H_2 = 0.4$, $\sigma_1^2 = 1$, $\sigma_2^2 = 1$, $\rho = 0.4$ and $\eta_{1,2} = 0$ (denoted η in the table).

		H_1	H_2	σ_1^2	σ_2^2	ρ	η
	Bias	-0.0029	-0.0011	0.0214	0.0206	0.0006	0.0024
$n = 500, \Delta = 1/52$	Std	0.0420	0.0341	0.3285	0.2738	0.0402	0.1019
		(0.0431)	(0.0351)	(0.3404)	(0.2774)	(0.0394)	(0.1036)
	RMSE	0.0421	0.0341	0.3294	0.2747	0.0402	0.1019
	Bias	-0.0007	-0.0011	0.0157	0.0084	-0.0006	-0.0014
$n = 1000, \Delta = 1/52$	Std	0.0300	0.0250	0.2225	0.1990	0.0284	0.0746
		(0.0305)	(0.0248)	(0.2407)	(0.1962)	(0.0279)	(0.0733)
	RMSE	0.0300	0.0250	0.2230	0.1991	0.0284	0.0746
	Bias	-0.0008	-0.0005	0.0868	0.0726	0.0014	0.0041
$n = 500, \Delta = 1/250$	Std	0.0421	0.0353	0.4956	0.4242	0.0388	0.1080
		(0.0431)	(0.0351)	(0.4756)	(0.3877)	(0.0394)	(0.1036)
	RMSE	0.0421	0.0354	0.5034	0.4306	0.0389	0.1081
	Bias	-0.0032	0.0003	0.0149	0.0387	-0.0013	0.0000
$n = 1000, \Delta = 1/250$	Std	0.0306	0.0250	0.3275	0.2889	0.0280	0.0744
		(0.0305)	(0.0248)	(0.3363)	(0.2741)	(0.0279)	(0.0733)
	RMSE	0.0308	0.0250	0.3279	0.2915	0.0280	0.0744

Table 3 presents the bias, standard deviation, and RMSE of both estimators across 1000 replications. Our estimator for ρ shows slightly inferior performance compared to that of Amblard and Coeurjolly (2011), while our estimator for $\eta_{1,2}$ demonstrates substantially improved accuracy, particularly when $\eta_{1,2} = 0$. For instance, when $\rho = 0.4$, $\eta_{1,2} = 0$, and $n = 500$, the RMSE ratio for ρ is 100.3%, indicating nearly identical performance, whereas the ratio for $\eta_{1,2}$ is 34.8%, clearly favoring our estimator. Real-data results echo this scenario where $\eta_{1,2} = 0$, providing further support for the practical relevance of our estimation approach.

Notably, when $\rho = 0$, an interesting contrast emerges: while the estimator from Amblard and Coeurjolly (2011) exhibits smaller variance, our estimator achieves significantly reduced bias. For both parameters (ρ and $\eta_{1,2}$), we find that when their true values are near zero—a scenario explicitly excluded in the construction of Amblard and Coeurjolly (2011)'s estimators—their estimators exhibit substantial bias. Although their correlation estimator maintains reasonable RMSE performance, their asymmetry estimates suffer from severe bias. Overall, our estimators deliver smaller biases and markedly improved statistical

Table 3: Bias, standard error (asymptotic standard errors in parentheses), and RMSE of 2 methods for different combinations of ρ and $\eta_{1,2}$ (denoted η in the table), with $H_1 = 0.1$, $H_2 = 0.4$, $\sigma_1^2 = 1$ and $\sigma_2^2 = 1$.

| | | $n = 500$ | | | | $n = 1000$ | | | |
| | | BYZ | | AC | | BYZ | | AC | |
		ρ	η	ρ	η	ρ	η	ρ	η
	Bias	-0.0006	-0.0009	0.0383	0.3096	0.0027	0.0037	0.0274	0.2090
$\rho = 0$	Std	0.0473	0.1156	0.0231	0.1487	0.0342	0.0812	0.0164	0.0997
$\eta = 0$	RMSE	0.0473	0.1157	0.0447	0.3435	0.0343	0.0813	0.0319	0.2316
	Bias	-0.0022	0.0005	-0.0026	0.2843	-0.0012	-0.0018	-0.0014	0.1936
$\rho = 0.4$	Std	0.0381	0.1093	0.0380	0.1407	0.0276	0.0735	0.0269	0.0950
$\eta = 0$	RMSE	0.0382	0.1093	0.0381	0.3173	0.0276	0.0735	0.0270	0.2157
	Bias	-0.0010	0.0040	0.0361	-0.0415	0.0009	0.0050	0.0255	-0.0770
$\rho = 0$	Std	0.0474	0.1296	0.0228	0.2042	0.0329	0.0961	0.0151	0.1617
$\eta = 0.5$	RMSE	0.0474	0.1298	0.0427	0.2085	0.0329	0.0962	0.0297	0.1791
	Bias	0.0017	0.0051	-0.0011	-0.0573	0.0006	0.0037	-0.0013	-0.0852
$\rho = 0.4$	Std	0.0372	0.1279	0.0358	0.1835	0.0269	0.0874	0.0253	0.1468
$\eta = 0.5$	RMSE	0.0372	0.1280	0.0358	0.1924	0.0269	0.0875	0.0253	0.1697

Table 4: Size and power of the test for $\eta_{1,2}$ (denoted η in the table).

	$\eta = 0$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.3$	$\eta = 0.4$	$\eta = 0.5$	$\eta = 0.6$	$\eta = 0.65$
				1% significance level				
$n = 500$	0.0120	0.0616	0.2608	0.6214	0.8852	0.9856	0.9986	0.9998
$n = 1000$	0.0122	0.1074	0.5556	0.9230	0.9978	1.0000	1.0000	1.0000
				5% significance level				
$n = 500$	0.0582	0.1610	0.4806	0.8126	0.9634	0.9972	1.0000	1.0000
$n = 1000$	0.0526	0.2742	0.7736	0.9772	0.9996	1.0000	1.0000	1.0000

inference for $\eta_{1,2}$ compared to Amblard and Coeurjolly (2011).

5.2. Size and power of the test

To examine the size and power of the proposed test, we consider the following parameter configurations, which are empirically relevant: $H_1 = 0.1$, $H_2 = 0.4$, $\rho = 0.4$, $\Delta = 1/250$, $\sigma_1^2 = 1$, and $\sigma_2^2 = 1$. The true value for $\eta_{1,2}$ varies over the interval $[0, 0.65]$. The sample sizes are set to 500 and 1000, and the number of replications is 5000. The test has good size when $\eta_{1,2} = 0$ with the nominal sizes close to the real sizes in finite sample. Moreover, testing at both 1% and 5% significance levels, we find the procedure maintains good finite-sample properties (Table 4). Consistent with theoretical expectations, the test's power grows monotonically as the true $\eta_{1,2}$ value deviates further from zero and gets larger for a larger sample size.

5.3. Finite-sample performance of optimal forecasts

This subsection examines potential improvements in out-of-sample forecasting accuracy through mfBm modeling. While Section 4.1 established theoretical foundations through Propositions 4.1-4.3 under simplified conditions, we now conduct numerical simulations to

assess forecasting performance in empirically realistic settings. We design three Monte Carlo experiments. In all experiments, we generate forecasts based on the first 500 observations ($n = 500$) sampled at frequency $\Delta = 1/250$. Through 10,000 Monte Carlo replications, we compute the root MSFE (RMSFE). Each replication simulates 505 observations to accommodate forecast horizons $h \in \{1, \ldots, 5\}$. To isolate the pure forecasting gains from estimation effects, we assume that Σ and H are known in these simulations.

The first Monte Carlo experiment is designed to investigate how forecast gains vary with inter-component correlation. Using simulated data from a bivariate time-reversible fBm with Hurst exponents $H_1 = 0.1$, $H_2 = 0.4$, we evaluate three distinct correlation structures for Σ (i.e., $\rho = 0, 0.4, 0.8$):

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}.$$

Table 5: RMSFEs of h-day-ahead forecasts for $B^{(i)}_{(t+h)\Delta}$, with and without taking the other component into account. The theoretical RMSFEs are in the brackets.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$
	$i=1$					$i=2$				
	$Corr = 0$									
fBm&bfBm	.4845	.5084	.5237	.5389	.5501	.1093	.1441	.1683	.1881	.2040
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.1085)	(.1430)	(.1681)	(.1886)	(.2061)
	$Corr = 0.4$									
fBm	.4832	.5119	.5239	.5370	.5551	.1096	.1433	.1689	.1899	.2080
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.1085)	(.1430)	(.1681)	(.1886)	(.2061)
bfBm	.4794	.5080	.5203	.5337	.5517	.1088	.1424	.1680	.1888	.2066
	(.4756)	(.5035)	(.5213)	(.5348)	(.5456)	(.1075)	(.1417)	(.1666)	(.1869)	(.2043)
	$Corr = 0.8$									
fBm	.4839	.5044	.5220	.5370	.5492	.1085	.1420	.1683	.1883	.2067
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.1085)	(.1430)	(.1681)	(.1886)	(.2061)
bfBm	.4286	.4502	.4698	.4830	.4940	.0946	.1224	.1441	.1593	.1734
	(.4246)	(.4526)	(.4700)	(.4827)	(.4927)	(.0953)	(.1242)	(.1443)	(.1602)	(.1734)

The results, presented in Table 5, reveal two key findings. First, incorporating the second (correlated) component in bivariate fBm (bfBm) yields significant improvements over univariate fBm forecasting. Second, these gains increase monotonically with the absolute correlation strength. As expected, forecasts coincide when $\rho = 0$ due to component inde-

pendence. For $\rho = 0.4$, we observe moderate efficiency gains, while $\rho = 0.8$ demonstrates the potential for substantial RMSFE reduction (up to 15.4% in our simulations).

The second experiment is designed to examine how gains depend on the differences between Hurst exponents. We fix $\sigma_1^2 = \sigma_2^2 = 1$, $\rho = 0.4$, and allow $(H_1, H_2) = (0.1, 0.1)$, $(0.1, 0.2)$ and $(0.1, 0.4)$. The RMSFE is computed from 10,000 replications. Results are given in Table 6. Consistent with our theoretical insights, gains become more pronounced for increasing differences between the idiosyncratic Hurst exponents.

The third experiment is designed to examine gains when including more components of mfBm. Assume there are three components with $H_1 = 0.1$, $H_2 = H_3 = 0.4$, and

$$
\Sigma_1 = \begin{bmatrix} 1 & 0.4 & 0.4 \\ 0.4 & 1 & 0 \\ 0.4 & 0 & 1 \end{bmatrix}.
$$

Table 6: RMSFEs of h-day-ahead forecasts for $B_{(t+h)\Delta}^{(i)}$, with and without taking the second component into account. The theoretical RMSFEs are in the brackets.

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$
			$i=1$					$i=2$		
					$H_1 = 0.1, H_2 = 0.1$					
fBm&bfBm	.4754	.5103	.5348	.5396	.5533	.4847	.5143	.5321	.5356	.5469
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)
					$H_1 = 0.1, H_2 = 0.2$					
fBm	.4818	.5098	.5226	.5375	.5495	.3019	.3425	.3670	.3904	.4059
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.2999)	(.3411)	(.3682)	(.3890)	(.4061)
bfBm	.4810	.5093	.5217	.5370	.5492	.3013	.3420	.3665	.3900	.4055
	(.4795)	(0.5071)	(.5249)	(.5382)	(.5490)	(.2995)	(.3407)	(.3679)	(.3887)	(.4058)
					$H_1 = 0.1, H_2 = 0.4$					
fBm	.4832	.5119	.5239	.5370	.5551	.1096	.1433	.1689	.1899	.2080
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.1085)	(.1430)	(.1681)	(.1886)	(.2061)
bfBm	.4794	.5080	.5203	.5337	.5517	.1088	.1424	.1680	.1888	.2066
	(.4756)	(.5035)	(.5213)	(.5348)	(.5456)	(.1075)	(.1417)	(.1666)	(.1869)	(.2043)

We evaluate three forecasting approaches: univariate fBm, bfBm, and three-dimensional mfBm (mfBm3). The forecasting results are reported in Table 7. As expected from the theoretical insights, including more correlated components further increases the forecasting performance. By further incorporating an additional component that is correlated only with

the first component, with $H_4 = 0.4$, such that

$$\Sigma_2 = \begin{bmatrix} 1 & w_{0.4}^\top \\ w_{0.4} & I_3 \end{bmatrix}, \text{ with } I_3 \text{ the } (3 \times 3) \text{ identity matrix and } w_{0.4} = (0.4, 0.4, 0.4)^\top,$$

the forecasting error will be reduced accordingly.

Table 7: RMSFEs of h-day-ahead forecasts for $B^{(1)}_{(n+h)\Delta}$, with and without taking other components into account. The theoretical RMSFEs are in the brackets.

h	1	2	3	4	5	1	2	3	4	5
			$\Sigma = \Sigma_1$					$\Sigma = \Sigma_2$		
fBm	.4784	.5025	.5213	.5361	.5458	.4833	.5081	.5293	.5359	.5463
	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)	(.4802)	(.5077)	(.5254)	(.5387)	(.5495)
bfBm	.4740	.4982	.5177	.5328	.5427	.4783	.5040	.5248	.5318	.5421
	(.4756)	(.5035)	(.5213)	(.5348)	(.5456)	(.4756)	(.5035)	(.5213)	(.5348)	(.5456)
mfBm3	.4656	.4919	.5108	.5266	.5352	.4697	.4970	.5176	.5249	.5351
	(.4686)	(.4969)	(.5150)	(.5286)	(.5396)	(.4686)	(.4969)	(.5150)	(.5286)	(.5396)
mfBm4						.4586	.4852	.5053	.5136	.5225
						(.4563)	(.4851)	(.5035)	(.5173)	(.5284)

In sum, we draw the following conclusions about efficiency gains from mfBm based on the analytical and simulation results:

1. If all Hurst exponents are identical, the forecasts based on fBm and mfBm coincide. Hence, there is no efficiency gain.

2. If all correlations are zero, i.e., Σ is diagonal, the forecasts based on fBm and mfBm coincide. Hence, there is no efficiency gain.

3. In a scenario that is not covered by the previous two cases, the forecast efficiency improves compared to the univariate model. Larger differences between Hurst exponents lead to larger efficiency gains. Larger absolute correlations between components lead to larger efficiency gains.

4. Including more correlated components yields further improvements of the forecasting performance.

6. Empirical Study

Volatility modeling and forecasting have been subjects of significant interest for decades. Recent advances in continuous-time finance highlight the strong performance of rough fBm and rough fOU processes in volatility forecasting. In this study, we investigate the practical implications of using rough mfBm and its associated theoretical framework. Our analysis focuses on daily realized volatilities of the Dow Jones 30 (DJ30) constituents from March 28, 2013, to August 21, 2021—a span of eight years—using data provided by Risk Lab.[7]

Table 8: Estimates of H and ρ for 7 Dow Jones 30 stocks between March 28, 2013, and August 21, 2021.

Ticker	$\hat{H}$	Correlation estimates						
AAPL	0.28	1.00						
ALD	0.19	0.39	1.00					
AMGN	0.21	0.37	0.31	1.00				
AXP	0.22	0.38	0.39	0.30	1.00			
BA	0.25	0.38	0.36	0.29	0.36	1.00		
BEL	0.18	0.34	0.33	0.27	0.37	0.27	1.00	
CAT	0.21	0.37	0.38	0.25	0.41	0.32	0.30	1.00

Table 9: Estimates of the asymmetry parameter based on (8b) and test outcomes at the 1% significance level based on Corollary 3.1 for 7 Dow Jones 30 stocks between March 28, 2013, and August 21, 2021.

	AAPL	ALD	AMGN	AXP	BA	BEL	CAT
ALD	0.17						
	Reject						
AMGN	0.02	-0.12					
	Not reject	Not reject					
AXP	0.09	-0.18	0.05				
	Not reject	Reject	Not reject				
BA	0.11	-0.10	-0.01	-0.02			
	Not reject	Not reject	Not reject	Not reject			
BEL	0.17	-0.06	-0.01	-0.02	-0.05		
	Reject	Not reject	Not reject	Not reject	Not reject		
CAT	0.08	-0.05	0.03	0.00	0.09	-0.02	
	Not reject	Not reject	Not reject	Not reject	Not reject	Not reject	

---

[7]See Dacheng Xiu's website: https://dachxiu.chicagobooth.edu/#risklab.

Table 10: RMSFE alternative models for h-day-ahead forecasts of RV during 3 forecast periods

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=10$	$h=15$	$h=20$
	March 20, 2015 to July 30, 2021							
fBm	0.0592	0.0698	0.0764	0.0816	0.0848	0.0964	0.1048	0.1119
bfBm	0.0589	0.0694	0.0758	0.0809	0.0841	0.0952	0.1034	0.1103
mfBm3	0.0590	0.0694	0.0758	0.0809	0.0840	0.0950	0.1031	0.1099
mfBm4	0.0589	0.0694	0.0758	0.0808	0.0840	0.0950	0.1030	0.1099
mfBm5	0.0589	0.0694	0.0758	0.0809	0.0840	0.0952	0.1031	0.1100
	March 20, 2015 to April 11, 2017							
fBm	0.0514	0.0621	0.0674	0.0715	0.0719	0.0806	0.0863	0.0941
bfBm	0.0504	0.0608	0.0655	0.0691	0.0693	0.0764	0.0811	0.0881
mfBm3	0.0504	0.0606	0.0654	0.0689	0.0690	0.0758	0.0803	0.0874
mfBm4	0.0503	0.0606	0.0654	0.0687	0.0687	0.0759	0.0798	0.0869
mfBm5	0.0503	0.0607	0.0655	0.0688	0.0688	0.0761	0.0800	0.0870
	April 12, 2017 to July 30 2021							
fBm	0.0628	0.0734	0.0805	0.0862	0.0906	0.1034	0.1129	0.1199
bfBm	0.0628	0.0734	0.0805	0.0862	0.0906	0.1033	0.1129	0.1198
mfBm3	0.0628	0.0734	0.0805	0.0862	0.0907	0.1032	0.1128	0.1196
mfBm4	0.0628	0.0734	0.0805	0.0863	0.0907	0.1033	0.1128	0.1197
mfBm5	0.0628	0.0734	0.0805	0.0863	0.0907	0.1034	0.1130	0.1199

6.1. Estimation results

Under the assumption that a single log realized volatility series follows fBm, we use the method-of-moment defined in (6) to estimate the Hurst exponents. To model multiple log realized volatilities jointly, we assume the data-generating process follows the time-reversible mfBm, with parameters estimated using the methodology developed in Section 3. For brevity, we present the point estimates of H and ρ's for the first seven Dow Jones 30 stocks in alphabetical order in Table 8.[8] We also report the point estimates of the asymmetry parameter and the test outcomes at the 1% level in Table 9. Table 8 suggests that there are significant differences among the estimated Hurst exponents and that the estimated correlation parameters are not close to zero. According to our theory, mfBm should offer forecasting improvements over univariate fBm.

---

[8]We report more empirical results for 20 Dow Jones 30 stocks in Tables B.12 and B.13 in the appendix.

6.2. Forecasting results

For brevity, we focus on forecasting the realized volatility of AAPL[9], examining whether incorporating additional cross-asset information enhances predictive accuracy. Using a two-year rolling window approach, we compare forecasts derived from AAPL's historical data alone—as is standard in the literature—with those augmented by data from four additional stocks (ALD, AMGN, AXP, and BA), selected alphabetically. We evaluate five model specifications in terms of RMSFE, progressively expanding the asset set: fBm (AAPL), bfBm (AAPL and ALD), mfBm3 (AAPL, ALD and AMGN), mfBm4 (AAPL, ALD, AMGN and AXP) and mfBm5 (AAPL, ALD, AMGN, AXP and BA). The forecast period is between March 20, 2015 and July 30 2021.

The top panel in Table 10 reports RMSFEs for five competing models for the full forecast period. It is clear that bfBm outperforms fBm and, in some cases mfBm processes outperform bfBm, consistent with the prediction of our theory.

Figure 7 displays the sequences of the rolling window estimates of the Hurst exponents for each log realized volatility series. It reveals a key pattern: after 2017, the estimated Hurst exponents for AAPL and other assets converge, with their differences approaching zero. This observation motivate us to split the forecast period into two distinct time intervals: from March 20, 2015 to April 11, 2017, from April 12, 2017 to July 30, 2021. As the differences between the estimated Hurst exponents are large in the first forecast period, we expect substantial gains in using mfBm over fBm. Whereas, the estimated Hurst exponents converge in the second forecast period, we expect no gains in using mfBm over fBm.

To validate our theory, we report RMSFEs in the middle and bottom panels of Table 10 for the two forecast periods, respectively. In the first period, mfBm demonstrates superior forecasting performance compared to fBm. Conversely, in the second period, the forecasts generated by mfBm and fBm become nearly identical. This empirical evidence strongly supports our analytical conclusions.

6.3. More information necessarily leads to efficiency gains?

Our analysis naturally prompts two fundamental questions: First, what justifies our use of the mfBm framework over conventional multivariate time series models? Second, does

---

[9]We also obtain similar results using both log realized volatility and realized volatility, which is consistent with the theoretical predictions. Since realized volatility is of greater interest, we focus on reporting the results based on realized volatility.

incorporating additional information (e.g., moving from univariate to multivariate specifications) necessarily yield efficiency gains in forecasting? These questions address both the methodological choice and the practical implications of model complexity.

Figure 7: Rolling window estimates of H for log realized volatilities

(a) AAPL (red line) and ALD (blue line)   (b) AAPL (red line) and AMGN (blue line)
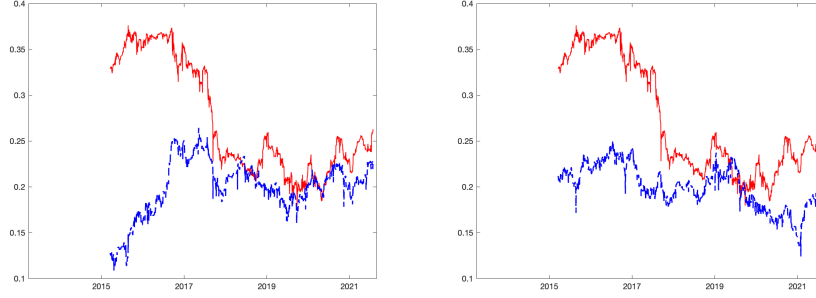
(c) AAPL (red line) and AXP (blue line)   (d) AAPL (red line) and BA (blue line)
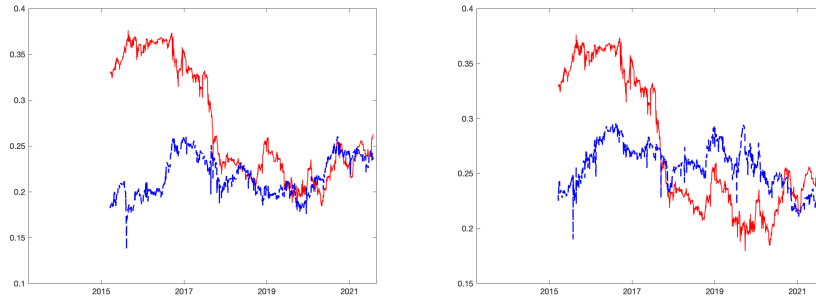
Given the HAR model's established role as a benchmark in volatility forecasting, we systematically compare the performance of univariate HAR and vector HAR (VHAR) specifications to evaluate whether increased information incorporation necessarily improves forecast accuracy.[10] The VHAR model can be viewed as a multivariate approximation of a long-memory vector autoregressive fractionally integrated moving average (VARFIMA) process (Chiriac and Voev, 2011), in a manner similar to the HAR model. To do so, we forecast AAPL's realized volatility using the same set of information sources (AAPL, ALD, AMGN, AXP, and BA) across the same five progressively expanded model specifications as before. Instead of expanding the specification from fBm to mfBm, we now expand the specification from HAR to VHAR. To maintain parsimony, we assume a diagonal covariance matrix

---

[10]The detailed specifications of HAR and VHAR models can be found in the appendix.

structure for the error term ϵ_t, enabling computationally efficient row-by-row estimation and forecasting. Table 11 reports RMSFE for HAR and four VHAR models.

Contrary to conventional expectations, our analysis reveals that the multivariate vector HAR model underperforms its univariate counterpart. By examining Table 11 and comparing it with Table 10, we draw the following conclusions: First, unlike the mfBm framework, which benefits from incorporating additional information, the vector HAR model fails to improve forecasting accuracy. Second, when using the same information, the mfBm generally delivers more accurate forecasts than HAR-type models. Given the strong performance of fBm-based models, the mfBm emerges as a powerful tool for leveraging extra information to enhance predictions. These findings underscore a key advantage of the mfBm framework: its unique capacity to effectively utilize additional information where traditional multivariate approaches fall short.

Table 11: RMSFE of HAR-type models for h-day-ahead forecasts of RV between March 20, 2015 and July 30 2021

	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=10$	$h=15$	$h=20$
HAR	0.0597	0.0695	0.0760	0.0813	0.0862	0.0995	0.1015	0.1044
VHAR2	0.0603	0.0702	0.0779	0.0849	0.0916	0.1093	0.1137	0.1182
VHAR3	0.0607	0.0706	0.0779	0.0844	0.0905	0.1063	0.1129	0.1208
VHAR4	0.0607	0.0712	0.0791	0.0862	0.0928	0.1071	0.1172	0.1193
VHAR5	0.0614	0.0717	0.0798	0.0868	0.0936	0.1061	0.1131	0.1212

7. Conclusion

This paper introduces the first multivariate rough fractional volatility model, proposing a time-reversible multivariate fBm to jointly model a panel of realized volatilities. We develop the moment-based estimation method and investigate its asymptotic properties. We establish theoretical results on optimal forecasts, demonstrating that the multivariate model outperforms its univariate counterpart when correlations are non-zero and Hurst exponents differ across components. Empirically, the model delivers significant efficiency gains in volatility forecasting compared to both univariate and benchmark approaches.

During the preparation of this paper, we became aware of an independent study by Dugo et al. (2024), which constructs a multivariate fractional Ornstein-Uhlenbeck (mfOU) process. While their work primarily explores probabilistic properties rather than feasible estimation and forecasting, they also motivate their model using the concept of rough volatility.

Our study opens several promising avenues for future research: (1) Asymmetric forecasting theory—developing a general theoretical framework for forecasting with asymmetry, which could broaden the model's applicability across different financial contexts. (2) mfOU modelling and forecasting—investigating the modelling and forecasting performance of mfOU processes and assessing their potential practical advantages over existing approaches. (3) Efficient estimation methods—exploring improved estimation techniques, such as maximum likelihood estimation for mfBm, and quantifying the expected efficiency gains relative to current methods. (4) Structural break analysis—extending the model to incorporate change-point detection, as suggested by the regime-switching behavior observed in AAPL's Hurst exponent, to better identify and interpret structural shifts in volatility dynamics.

8. Proofs

8.1. Proofs on asymptotic statistics

8.1.1. Proof of Theorem 3.1

For components $(p, q) \in \{1, \ldots, d\}^2$, the covariance of the kth and jth increments, $(j, k) \in \{1, \ldots, n\}^2$, of the time-reversible mfBm yields

$$\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)}, \Delta_j B^{(q)}\big) = \rho_{p,q}\sigma_p\sigma_q\Delta^{H_p+H_q}\gamma_{p,q}(k-j)\,, \tag{18}$$

where the dependence on time, or the lag $l \in \mathbb{Z}$, is via the function

$$\gamma_{p,q}(l) = \frac{1}{2}\Big(|l+1|^{H_p+H_q} + |l-1|^{H_p+H_q} - 2|l|^{H_p+H_q}\Big). \tag{19}$$

The function is symmetric in the sense that $\gamma_{p,q}(l) = \gamma_{p,q}(-l)$. For the correlation parameters, we use the notation $\rho_{p,p} = 1$, for $p = q$. For $p = q$, we further write shortly γ_p instead of $\gamma_{p,p}$. It holds $\gamma_{p,q}(0) = 1$, for any p, q. Having Gaussian processes, a crucial ingredient to compute variances and covariances of our statistics, is obtained from the Isserlis' theorem (Isserlis (1918)):

$$\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)}\Delta_j B^{(q)}, \Delta_i B^{(r)}\Delta_l B^{(s)}\big) = \mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)}, \Delta_i B^{(r)}\big)\mathbb{C}\mathrm{ov}\big(\Delta_j B^{(q)}, \Delta_l B^{(s)}\big) \tag{20}$$
$$+ \mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)}, \Delta_l B^{(s)}\big)\mathbb{C}\mathrm{ov}\big(\Delta_j B^{(q)}, \Delta_i B^{(r)}\big),$$

for all $(p, q, r, s) \in \{1, \ldots, d\}^4$ and $(j, k, l, i) \in \{1, \ldots, n\}^4$. We obtain the following variances of suitably normalized realized volatilities:

$$
\begin{aligned}
\mathbb{Var}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}} \sum_{k=1}^{n} \big(\Delta_k B^{(p)}\big)^2\Big) &= \frac{\Delta^{-4H_p}}{n} \sum_{1 \leq k, l \leq n} \mathbb{Cov}\big(\big(\Delta_k B^{(p)}\big)^2, \big(\Delta_l B^{(p)}\big)^2\big) \\
&\stackrel{(20)}{=} \frac{\Delta^{-4H_p}}{n} \sum_{1 \leq k, l \leq n} 2\mathbb{Cov}^2\big(\Delta_k B^{(p)}, \Delta_l B^{(p)}\big) \\
&\stackrel{(18)}{=} \frac{2\sigma_p^4}{n} \sum_{1 \leq k, l \leq n} \gamma_p^2(k - l) \\
&= \frac{2\sigma_p^4}{n} \Big(n\gamma_p^2(0) + 2\sum_{r=1}^{n}(n - r)\gamma_p^2(r)\Big) \\
&= 2\sigma_p^4\Big(1 + 2\sum_{r=1}^{n} \frac{n - r}{n}\gamma_p^2(r)\Big).
\end{aligned}
$$

With dominated convergence, we obtain that

$$
\begin{aligned}
\lim_{n \to \infty} \mathbb{Var}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}} \sum_{k=1}^{n} \big(\Delta_k B^{(p)}\big)^2\Big) &= 2\sigma_p^4\Big(1 + 2\sum_{r=1}^{\infty} \gamma_p^2(r)\Big) \\
&\stackrel{(19)}{=} \sigma_p^4\Big(2 + \sum_{r=1}^{\infty} \big(|r + 1|^{2H_p} + |r - 1|^{2H_p} - 2|r|^{2H_p}\big)^2\Big).
\end{aligned}
$$

Applying twice the mean value theorem for differentiation with the functions $x \mapsto x^{2H_p}$, and $x \mapsto x^{2H_p - 1}$, we find that

$$
\begin{aligned}
(r + 1)^{2H_p} + (r - 1)^{2H_p} - 2(r)^{2H_p} &= 2H_p\big(\xi_r^{2H_p - 1} - \xi_{r-1}^{2H_p - 1}\big) \\
&= 2H_p(2H_p - 1)\zeta_r^{2H_p - 2},
\end{aligned}
$$

with some $\xi_r \in (r, r + 1), r \in \mathbb{N}$, and $\zeta_r \in (\xi_{r-1}, \xi_r) \subset (r - 1, r + 1)$. This suffices to see (based on a majorant and minorant) that the series $\sum_{r=1}^{\infty} \gamma_p^2(r)$ converges if and only if $H_p \in (0, 3/4)$. Similar arguments apply to upcoming series expressions in the proofs. Univariate central limit theorems

$$
\sqrt{n}\Big(\frac{\Delta^{-2H_p}}{n} \sum_{k=1}^{n} \big(\Delta_k B^{(p)}\big)^2 - \sigma_p^2\Big) \stackrel{d}{\to} \mathcal{N}\Big(0, \sigma_p^4\Big(2 + \sum_{r=1}^{\infty} \big(|r + 1|^{2H_p} + |r - 1|^{2H_p} - 2|r|^{2H_p}\big)^2\Big)\Big)
$$

are known to hold in this case by an application of the celebrated limit theorem by Breuer and Major (1983) with the Hermite polynomial $x \mapsto x^2 - 1$ of Hermite rank 2. For its rank, it suffices to see that $\mathbb{E}[(X^2 - 1)X] = 0$ and $\mathbb{E}[(X^2 - 1)^2] = 2 \neq 0$, for a normalized centred

normal random variable X.

Next, consider lag-2-increments, which equal (telescoping) sums of increments:

$$\Delta_{k,2}B^{(p)} = B^{(p)}_{(k+1)\Delta} - B^{(p)}_{(k-1)\Delta} = \Delta_k B^{(p)} + \Delta_{k+1}B^{(p)}, \ k \in \{1, \ldots, n-1\}.$$

It holds that

$$\mathbb{C}\text{ov}(\Delta_{k,2}B^{(p)}, \Delta_{l,2}B^{(p)}) = \mathbb{C}\text{ov}(\Delta_k B^{(p)} + \Delta_{k+1}B^{(p)}, \Delta_l B^{(p)} + \Delta_{l+1}B^{(p)})$$
$$= \Delta^{2H_p}\sigma_p^2\big(2\gamma_p(k-l) + \gamma_p(k+1-l) + \gamma_p(k-l-1)\big)$$

by linearity. Note the relation $2^{2H_p} = (2\gamma_p(1) + 2)$ for the variances. This relation and applying the Isserlis formula directly to lag-2-increments, what is possible, yields a simple computation of the variance

$$\mathbb{V}\text{ar}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n-1}\big(\Delta_{k,2}B^{(p)}\big)^2\Big) = \frac{\Delta^{-4H_p}}{n}\sum_{1\leq k,l\leq n-1}2\mathbb{C}\text{ov}^2(\Delta_{k,2}B^{(p)}, \Delta_{l,2}B^{(p)})$$
$$= \frac{2\sigma_p^4}{n}\Big((n-1)(2+2\gamma_p(1))^2 + 2\sum_{r=1}^{n-2}(n-1-r)\big(2\gamma_p(r) + \gamma_p(r+1) + \gamma_p(r-1)\big)^2\Big)$$
$$= \sigma_p^4\Big(2\cdot 2^{4H_p}\frac{n-1}{n} + 4\sum_{r=1}^{n-2}\frac{n-1-r}{n}\big(2\gamma_p(r) + \gamma_p(r+1) + \gamma_p(r-1)\big)^2\Big).$$

With dominated convergence and elementary transformations, we obtain that

$$\lim_{n\to\infty}\mathbb{V}\text{ar}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n-1}\big(\Delta_{k,2}B^{(p)}\big)^2\Big) = \sigma_p^4\Big(2\cdot 2^{4H_p} + 4\sum_{r=1}^{\infty}\big(2\gamma_p(r) + \gamma_p(r+1) + \gamma_p(r-1)\big)^2\Big)$$
$$\overset{(19)}{=} \sigma_p^4\Big(2\cdot 2^{4H_p} + \sum_{r=1}^{\infty}\big(|r+2|^{2H_p} + |r-2|^{2H_p} - 2|r|^{2H_p}\big)^2\Big).$$

Since we sum over all k, and do not only use half of the observations, this is not obvious from the above asymptotics of the realized volatility. A central limit theorem follows analogously as for the realized volatility. For the covariance of both statistics, we use that

$$\mathbb{C}\text{ov}\big(\Delta_k B^{(p)}, \Delta_{l,2}B^{(p)}\big) = \mathbb{C}\text{ov}\big(\Delta_k B^{(p)}, \Delta_l B^{(p)} + \Delta_{l+1}B^{(p)}\big)$$
$$= \Delta^{2H_p}\sigma_p^2\big(\gamma_p(k-l) + \gamma_p(k-l-1)\big), \ 1\leq k\leq n, 1\leq l\leq n-1.$$

We derive that

$$\mathbb{Cov}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2, \frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{l=1}^{n-1}\big(\Delta_{l,2}B^{(p)}\big)^2\Big) = \frac{\Delta^{-4H_p}}{n}\sum_{\substack{1\le k\le n \\ 1\le l\le n-1}} 2\mathbb{Cov}^2\big(\Delta_k B^{(p)}, \Delta_{l,2}B^{(p)}\big)$$

$$= \frac{2\sigma_p^4}{n}\sum_{\substack{1\le k\le n \\ 1\le l\le n-1}}\big(\gamma_p(k-l) + \gamma_p(k-l-1)\big)^2$$

$$= \frac{2\sigma_p^4}{n}\Big(2(n-1)\big(\gamma_p(0)+\gamma_p(1)\big)^2 + \sum_{r=1}^{n-2}(n-r-1)\big(\gamma_p(r)+\gamma_p(r+1)\big)^2 + \sum_{r=2}^{n-1}(n-r)\big(\gamma_p(r)+\gamma_p(r-1)\big)^2\Big)$$

$$= 4\sigma_p^4\sum_{r=1}^{n-1}\frac{n-r}{n}\big(\gamma_p(r)+\gamma_p(r-1)\big)^2,$$

such that with dominated convergence

$$\lim_{n\to\infty}\mathbb{Cov}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2, \frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{l=1}^{n-1}\big(\Delta_{l,2}B^{(p)}\big)^2\Big) = 4\sigma_p^4\sum_{r=1}^{\infty}\big(\gamma_p(r)+\gamma_p(r-1)\big)^2$$

$$\overset{(19)}{=} \sigma_p^4\sum_{r=1}^{\infty}\Big(|r+1|^{2H_p} + |r-2|^{2H_p} - |r|^{2H_p} - |r-1|^{2H_p}\Big)^2.$$

If we show for scalar products $\langle\,\cdot\,,\,\cdot\,\rangle$, with any $\alpha\in\mathbb{R}^2$, that

$$\sqrt{n}\Big\langle\alpha, \Big(\frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}(\Delta_k B^{(p)})^2 - \sigma_p^2, \frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n-1}(\Delta_{k,2}B^{(p)})^2 - 2^{2H_p}\sigma_p^2\Big)^{\top}\Big\rangle \overset{d}{\to} \langle\alpha, Z\rangle, \quad (21)$$

with the bivariate normal

$$Z \sim \mathcal{N}\Big(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \mathrm{AVAR}_{11} & \mathrm{AVAR}_{12} \\ \mathrm{AVAR}_{12} & \mathrm{AVAR}_{22} \end{pmatrix}\Big),$$

where

$$\mathrm{AVAR}_{11} = \sigma_p^4\Big(2 + \sum_{r=1}^{\infty}\Big(|r+1|^{2H_p} + |r-1|^{2H_p} - 2|r|^{2H_p}\Big)^2\Big), \tag{22a}$$

$$\mathrm{AVAR}_{12} = \sigma_p^4\sum_{r=1}^{\infty}\Big(|r+1|^{2H_p} + |r-2|^{2H_p} - |r|^{2H_p} - |r-1|^{2H_p}\Big)^2, \tag{22b}$$

$$\mathrm{AVAR}_{22} = \sigma_p^4\Big(2\cdot 2^{4H_p} + \sum_{r=1}^{\infty}\Big(|r+2|^{2H_p} + |r-2|^{2H_p} - 2|r|^{2H_p}\Big)^2\Big), \tag{22c}$$

the Cramér-Wold device implies the bivariate central limit theorem

$$\sqrt{n}\left(\begin{array}{c} \frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}(\Delta_k B^{(p)})^2 - \sigma_p^2 \\ \frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n-1}(\Delta_{k,2} B^{(p)})^2 - 2^{2H_p}\sigma_p^2 \end{array}\right) \xrightarrow{d} Z\,. \qquad (23)$$

By self-similarity $B_{k\Delta}^{(p)} \stackrel{d}{=} \Delta^{H_p} B_k^{(p)}$, such that (instead of a triangular array) we consider the $\mathbb{R}^2$-valued stationary, mean-zero Gaussian sequence

$$(X_1, \ldots, X_{n-1}), \text{ with } X_j = \left(X_j^{(1)}, X_j^{(2)}\right)^\top = \left(\sigma_p(B_{j+1}^{(p)} - B_j^{(p)}), \sigma_p(B_{j+1}^{(p)} - B_{j-1}^{(p)})\right)^\top,$$

and for $\alpha = (\alpha_1, \alpha_2)^\top$ the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \alpha_1 x^2 + \alpha_2 y^2$. Since Hermite polynomials $H_r, r \in \mathbb{N}$, satisfy $\mathbb{E}[H_{r_1}(X_j^{(1)})H_{r_2}(X_j^{(2)})] = \delta_{r_1,r_2} r_1! c_{12}$, with (covariance) constant c_{12}, i.e., zero for $r_1 \neq r_2$, and non-zero for $r_1 = r_2$, f has Hermite rank 2 in the sense of Eq. (2.2) in Arcones (1994); see (27) also. Neglecting the first increment in the first component, which is asymptotically negligible, the limit theorem from Arcones (1994) thus readily yields (21) and we conclude (23). A (more direct) proof based on moment generating functions can be worked out similar to Theorem 4 of Kubilius (2020). We have that

$$\hat{H}_p = \frac{1}{2\log(2)}\log\left(\frac{T}{S}\right), \text{ with } S = \frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}(\Delta_k B^{(p)})^2\,, \text{ and } T = \frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n-1}(\Delta_{k,2} B^{(p)})^2\,.$$

From (23), we can therefore derive a central limit theorem for $\hat{H}_p$ based on the multivariate Δ-method applied to the function $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x,y) = \log(y/x)/(2\log(2))$. Since $g\left(\sigma_p^2, 2^{2H_p}\sigma_p^2\right) = H_p$, the Δ-method yields

$$\sqrt{n}(\hat{H}_p - H_p) \xrightarrow{d} \mathcal{N}\left(0, \left(\nabla g\left(\sigma_p^2, 2^{2H_p}\sigma_p^2\right)\right)^\top \left(\begin{array}{cc} \text{AVAR}_{11} & \text{AVAR}_{12} \\ \text{AVAR}_{12} & \text{AVAR}_{22} \end{array}\right) \nabla g\left(\sigma_p^2, 2^{2H_p}\sigma_p^2\right)\right),$$

with $\left(\nabla g(x,y)\right)^\top = (-x^{-1}, y^{-1})^\top/(2\log(2))$, such that

$$\nabla g\left(\sigma_p^2, 2^{2H_p}\sigma_p^2\right))^\top \left(\begin{array}{cc} \text{AVAR}_{11} & \text{AVAR}_{12} \\ \text{AVAR}_{12} & \text{AVAR}_{22} \end{array}\right) \nabla g\left(\sigma_p^2, 2^{2H_p}\sigma_p^2\right)$$

$$= \frac{1}{4\log^2(2)\sigma_p^4}\left(\text{AVAR}_{11} - 2 \cdot 2^{-2H_p}\text{AVAR}_{12} + 2^{-4H_p} \cdot \text{AVAR}_{22}\right).$$

We conclude with (22a)-(22c) for $H_p \in (0, 3/4)$ that

$$\sqrt{n}(\hat{H}_p - H_p) \xrightarrow{d} \mathcal{N}(0, \mathrm{AVAR}_{\hat{H}_p}), \tag{24}$$

with asymptotic variance as in (9)

$$
\begin{aligned}
\mathrm{AVAR}_{\hat{H}_p} = \frac{1}{4 \log^2(2)} \Big(& 4 + \sum_{r=1}^{\infty} \Big(|r+1|^{2H_p} + |r-1|^{2H_p} - 2|r|^{2H_p} \Big)^2 \\
& + 2^{-4H_p} \sum_{r=1}^{\infty} \Big(|r+2|^{2H_p} + |r-2|^{2H_p} - 2|r|^{2H_p} \Big)^2 \\
& - 2 \cdot 2^{-2H_p} \sum_{r=1}^{\infty} \Big(|r+1|^{2H_p} + |r-2|^{2H_p} - |r|^{2H_p} - |r-1|^{2H_p} \Big)^2 \Big).
\end{aligned}
$$

This asymptotic variance depends on H_p, but not on σ_p^2. Based on the expansion

$$\Delta^{2(H_p - \hat{H}_p)} = \exp\big(2(\hat{H}_p - H_p)\log(\Delta^{-1})\big) = 1 + 2(\hat{H}_p - H_p)\log(n) + \mathcal{O}_{\mathbb{P}}\big(n^{-1}\log^2(n)\big),$$

we obtain that

$$
\begin{aligned}
\frac{\sqrt{n}}{\log(n)}\big(\hat{\sigma}_p^2 - \sigma_p^2\big) &= \frac{\sqrt{n}}{\log(n)}\Big(\frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2 \Delta^{2(H_p - \hat{H}_p)} - \sigma_p^2\Big) \\
&= \frac{\sqrt{n}}{\log(n)}\Big(\frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2 - \sigma_p^2\Big) + 2\sqrt{n}\big(\hat{H}_p - H_p\big)\frac{\Delta^{-2H_p}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2 + \mathcal{O}_{\mathbb{P}}\big(n^{-1/2}\log^2(n)\big) \\
&= 2\sqrt{n}\big(\hat{H}_p - H_p\big)\sigma_p^2 + \mathcal{O}_{\mathbb{P}}(1) \xrightarrow{d} \mathcal{N}\big(0, 4\sigma_p^4 \mathrm{AVAR}_{\hat{H}_p}\big),
\end{aligned}
$$

by Slutsky's lemma and (24). This proves (10) and completes the proof of Theorem 3.1.

8.1.2. Proof of Theorem 3.2

For the general, bivariate fBm with asymmetry parameter $\eta_{1,2} = -\eta_{2,1}$, the cross-covariances of increments generalize from (18) to

$$\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(1)}, \Delta_j B^{(2)}\big) = \big(\rho_{1,2} + \eta_{1,2}\,\mathrm{sign}\,(k-j)\big)\sigma_1\sigma_2\Delta^{H_1 + H_2}\gamma_{1,2}(k-j), \tag{25}$$

for $(j, k) \in \{1, \ldots, n\}^2$, and with $\gamma_{1,2}$ from (19). Covariances of increments of one component are not affected by the asymmetry. This implies

$$\mathbb{E}\big[\Delta_{k+1} B^{(1)} \Delta_k B^{(2)} - \Delta_k B^{(1)} \Delta_{k+1} B^{(2)}\big] = \mathbb{E}\big[\Delta_{k+1} B^{(1)} \Delta_k B^{(2)}\big] - \mathbb{E}\big[\Delta_k B^{(1)} \Delta_{k+1} B^{(2)}\big]$$

$$\overset{(25)}{=} \sigma_1\sigma_2\Delta^{H_1+H_2}\gamma_{1,2}(1)\big((\rho_{1,2}+\eta_{1,2}) - (\rho_{1,2}+\eta_{2,1})\big)$$

$$= 2\eta_{1,2}\sigma_1\sigma_2\Delta^{H_1+H_2}\gamma_{1,2}(1)\,,$$

with $2\gamma_{1,2}(1) = 2^{H_1+H_2}-2$. With a simple bound as $\big(\rho_{1,2}+\eta_{1,2}\,\mathrm{sign}\,(k-j)\big)^2 \le 2(\rho_{1,2}^2+\eta_{1,2}^2)$, the above analysis for the covariation readily yields that

$$\frac{\Delta^{-H_1-H_2}}{n}\sum_{k=1}^{n-1}\Big(\Delta_{k+1}B^{(2)}\cdot\Delta_k B^{(1)}-\Delta_{k+1}B^{(1)}\cdot\Delta_k B^{(2)}\Big)-\eta_{1,2}\sigma_1\sigma_2(2^{H_1+H_2}-2) = \mathcal{O}_{\mathbb{P}}(n^{-1/2})\,.$$

The asymptotics of rescaled realized variances, with increments and lag-2-increments, i.e., the above central limit theorems, imply that

$$\frac{\Delta^{-2H_j}}{n}\sum_{k=1}^{n-1}\big(\Delta_{k,2}B^{(j)}\big)^2 - 2^{2H_j}\sigma_j^2 = \mathcal{O}_{\mathbb{P}}(n^{-1/2}),\ j=1,2,$$

$$\frac{\Delta^{-2H_j}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(j)}\big)^2 - \sigma_j^2 = \mathcal{O}_{\mathbb{P}}(n^{-1/2}),\ j=1,2,$$

such that elementary computations yield that

$$\hat{\eta}_{1,2} - \eta_{1,2} = \mathcal{O}_{\mathbb{P}}(n^{-1/2})\,.$$

Since we aim to construct an asymptotic test for $\eta_{1,2} = 0$ vs. $\eta_{1,2} \ne 0$, we work out the central limit theorem for $\hat{\eta}_{1,2}$ under the null hypothesis that $\eta_{1,2} = 0$. Therefore, the following derivation of the asymptotic variance may restrict to covariances (18) instead of (25), what simplifies some terms. We deduce that

$$\mathbb{V}\mathrm{ar}\bigg(\frac{\Delta^{-H_1-H_2}}{\sqrt{n}}\sum_{k=1}^{n-1}\Big(\Delta_{k+1}B^{(2)}\cdot\Delta_k B^{(1)}-\Delta_{k+1}B^{(1)}\cdot\Delta_k B^{(2)}\Big)\bigg)$$

$$= \frac{\Delta^{-2(H_1+H_2)}}{n}\sum_{1\le i,j\le n-1}\mathbb{C}\mathrm{ov}\Big(\Delta_{i+1}B^{(2)}\Delta_i B^{(1)}-\Delta_{i+1}B^{(1)}\Delta_i B^{(2)},\Delta_{j+1}B^{(2)}\Delta_j B^{(1)}-\Delta_{j+1}B^{(1)}\Delta_j B^{(2)}\Big)$$

$$= \frac{\sigma_1^2\sigma_2^2}{n}\sum_{1\le i,j\le n-1}\Big(2\big(\gamma_1(i-j)\gamma_2(i-j)+\rho^2\gamma_{1,2}(i+1-j)\gamma_{1,2}(i-j-1)\big)$$

$$- \gamma_1(i+1-j)\gamma_2(i-j-1) - \gamma_1(i-j-1)\gamma_2(i+1-j) - 2\rho^2\gamma_{1,2}^2(i-j)\Big)$$

$$= \sigma_1^2\sigma_2^2\Big(\frac{n-1}{n}\big(2 - 2\gamma_1(1)\gamma_2(1) - 2\rho^2\big(1-\gamma_{1,2}^2(1)\big)\big)$$

$$+ 2\sum_{r=1}^{n-1}\frac{n-1-r}{n}\big(2\gamma_1(r)\gamma_2(r) - \gamma_1(r+1)\gamma_2(r-1) - \gamma_1(r-1)\gamma_2(r+1)\big)$$

$$-4\rho^2 \sum_{r=1}^{n-1} \frac{n-1-r}{n}\big(\gamma_{1,2}^2(r) - \gamma_{1,2}(r+1)\gamma_{1,2}(r-1)\big)\bigg).$$

With dominated convergence, we conclude under the null, $\eta_{1,2} = 0$, the asymptotic variance

$$\lim_{n\to\infty} n \cdot \mathbb{Var}\big(\hat{\eta}_{1,2}\big) = (2^{H_1+H_2} - 2)^{-2}\Big(2\big(1 - \gamma_1(1)\gamma_2(1)\big) + 2\rho^2\big(\gamma_{1,2}^2(1) - 1\big)$$

$$- \rho^2 \sum_{r=1}^{\infty} \Big(\big(|r+1|^{H_1+H_2} + |r-1|^{H_1+H_2} - 2|r|^{H_1+H_2}\big)^2$$

$$- \big(|r+2|^{H_1+H_2} + |r|^{H_1+H_2} - 2|r+1|^{H_1+H_2}\big)\big(|r|^{H_1+H_2} + |r-2|^{H_1+H_2} - 2|r-1|^{H_1+H_2}\big)\Big)$$

$$+ \frac{1}{2} \sum_{r=1}^{\infty} \Big(2\big(|r+1|^{2H_1} + |r-1|^{2H_1} - 2|r|^{2H_1}\big)\big(|r+1|^{2H_2} + |r-1|^{2H_2} - 2|r|^{2H_2}\big)$$

$$- \big(|r+2|^{2H_1} + |r|^{2H_1} - 2|r+1|^{2H_1}\big)\big(|r|^{2H_2} + |r-2|^{2H_2} - 2|r-1|^{2H_2}\big)$$

$$- \big(|r+2|^{2H_2} + |r|^{2H_2} - 2|r+1|^{2H_2}\big)\big(|r|^{2H_1} + |r-2|^{2H_1} - 2|r-1|^{2H_1}\big)\Big)\Big)$$

$$= \text{AVAR}_{\hat{\eta}_{1,2}}. \tag{26}$$

For the proof of Theorem 3.2, we use the asymptotic variance of $\hat{\eta}_{1,2}$ from (26) and show a central limit theorem. By self-similarity (3), it holds that $\big(B_{k\Delta}^{(1)}, B_{k\Delta}^{(2)}\big)^\top \overset{d}{=} \big(\Delta^{H_1} B_k^{(1)}, \Delta^{H_2} B_k^{(2)}\big)^\top$. To prove asymptotic normality of the numerator in (8b), multiplied with $\Delta^{-H_1-H_2}/\sigma_1\sigma_2$, we apply Theorem 4 from Arcones (1994) to prove asymptotic normality of $\sum_{k=1}^{n-1} f(X_k)$, with the $\mathbb{R}^4$-valued, stationary, mean-zero Gaussian sequence $(X_1, \ldots, X_{n-1})$, with

$$X_j = \big(X_j^{(1)}, X_j^{(2)}, X_j^{(3)}, X_j^{(4)}\big)^\top = \big((B_j^{(1)} - B_{j-1}^{(1)}), (B_j^{(2)} - B_{j-1}^{(2)}), (B_{j+1}^{(1)} - B_j^{(1)}), (B_{j+1}^{(2)} - B_j^{(2)})\big)^\top,$$

where the entries are already standardized with variances equal to one, and the function

$$f : \mathbb{R}^4 \to \mathbb{R}, f(x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}) = x^{(1)} \cdot x^{(4)} - x^{(2)} \cdot x^{(3)}.$$

It suffices to show that the Hermite rank of f with respect to X_1 is 2 to conclude with Theorem 4 from Arcones (1994). The Hermite rank is defined by

$$\text{rank}(f) := \inf\Big\{\tau \geq 1 \big| \sum_{j=1}^{4} l_j = \tau \text{ and } \mathbb{E}\big[(f(X_1) - \mathbb{E}[f(X_1)]) \prod_{k=1}^{4} H_{l_k}(X_1^{(k)})\big] \neq 0\Big\}, \tag{27}$$

41

with $H_l, l \geq 0$, the Hermite polynomials. We have $\mathrm{rank}(f) > 1$, since with $H_1(x) = x$ and $H_0 = 1$, it suffices to show

$$\mathbb{E}\big[\big(f(X_1) - \mathbb{E}[f(X_1)]\big)X_1^{(k)}\big] = 0, \ k = 1, 2, 3, 4.$$

For all $k = 1, 2, 3, 4$, the proof is similar, we consider without loss of generality $k = 1$:

$$\mathbb{E}\big[\big(f(X_1) - \mathbb{E}[f(X_1)]\big)X_1^{(1)}\big] = \mathbb{E}\big[\big(X_1^{(1)} \cdot X_1^{(4)} - X_1^{(2)} \cdot X_1^{(3)} - \mathbb{E}[f(X_1)]\big)X_1^{(1)}\big]$$
$$= \mathbb{E}\big[(X_1^{(1)})^2 \cdot X_1^{(4)}\big] - \mathbb{E}\big[X_1^{(1)} \cdot X_1^{(2)} \cdot X_1^{(3)}\big] = 0,$$

where the last identity follows since $-X_1 \overset{d}{=} X_1$, tantamount with the (trivial) odd case in Isserlis' theorem. To show $\mathrm{rank}(f) = 2$, consider

$$\mathbb{E}\big[\big(f(X_1) - \mathbb{E}[f(X_1)]\big)X_1^{(1)}X_1^{(4)}\big]$$
$$= \mathbb{E}\big[(X_1^{(1)} \cdot X_1^{(4)})^2\big] - \mathbb{E}\big[X_1^{(1)} \cdot X_1^{(2)} \cdot X_1^{(3)} \cdot X_1^{(4)}\big] - \big(\mathbb{E}\big[X_1^{(1)} \cdot X_1^{(4)}\big]\big)^2 + \mathbb{E}\big[X_1^{(1)} \cdot X_1^{(4)}\big]\mathbb{E}\big[X_1^{(2)} \cdot X_1^{(3)}\big]$$
$$\overset{(20)}{=} 1 + \big(\mathbb{E}\big[X_1^{(1)} \cdot X_1^{(4)}\big]\big)^2 - \mathbb{E}\big[X_1^{(1)} \cdot X_1^{(3)}\big]\mathbb{E}\big[X_1^{(2)} \cdot X_1^{(4)}\big] - \mathbb{E}\big[X_1^{(1)} \cdot X_1^{(2)}\big]\mathbb{E}\big[X_1^{(3)} \cdot X_1^{(4)}\big]$$
$$= 1 + \rho^2 \gamma_{1,2}^2(1) - \rho^2 - \gamma_1(1)\gamma_2(1).$$

We make a case differentiation. If $\rho = 0$, the above yields $1 - \gamma_1(1)\gamma_2(1) > 0$. If $H_1 = H_2$, we have that $\gamma_1(1) = \gamma_2(1) = \gamma_{1,2}(1) = \gamma(1)$, what yields $(1 - \rho^2)(1 - \gamma^2(1)) > 0$. For the other case, it suffices to consider with $H_2(x) = x^2 - 1$ the expectation

$$\mathbb{E}\big[\big(f(X_1) - \mathbb{E}[f(X_1)]\big)\big((X_1^{(1)})^2 - 1\big)\big]$$
$$= \mathbb{E}\big[(X_1^{(1)})^3 \cdot X_1^{(4)} - X_1^{(1)} \cdot X_1^{(4)} + X_1^{(2)} \cdot X_1^{(3)} - (X_1^{(1)})^2 \cdot X_1^{(2)} \cdot X_1^{(3)}\big]$$
$$= 3\mathbb{E}\big[(X_1^{(1)})^2\big]\mathbb{E}\big[X_1^{(1)}X_1^{(4)}\big] - \mathbb{E}\big[X_1^{(1)}X_1^{(4)}\big]$$
$$\quad + \mathbb{E}\big[X_1^{(2)}X_1^{(3)}\big] - 2\mathbb{E}\big[X_1^{(1)}X_1^{(2)}\big]\mathbb{E}\big[X_1^{(1)}X_1^{(3)}\big] - \mathbb{E}\big[(X_1^{(1)})^2\big]\mathbb{E}\big[X_1^{(2)}X_1^{(3)}\big]$$
$$= 2\mathbb{E}\big[X_1^{(1)}X_1^{(4)}\big] - 2\mathbb{E}\big[X_1^{(1)}X_1^{(2)}\big]\mathbb{E}\big[X_1^{(1)}X_1^{(3)}\big]$$
$$= 2\rho\big(\gamma_{1,2}(1) - \gamma_1(1)\big),$$

what is non-zero if $\rho \neq 0$, and $H_1 \neq H_2$. The central limit theorem for $\hat{\eta}_{1,2}$, under the null hypothesis $\eta_{1,2} = 0$, is hence implied by Slutsky's lemma. Since the map $(\rho, H_1, H_2) \mapsto \mathrm{AVAR}_{\hat{\eta}_{1,2}}$ is continuous, what can be seen by dominated convergence, Corollary 3.1 is hence implied by the convergence of the normalized realized volatilities in the denominator, con-

tinuous mapping for convergence in probability and Slutsky's lemma.

8.1.3. Proof of Theorem 3.3

Consider the variances of suitably normalized realized covariances

$$\mathbb{V}\mathrm{ar}\Big(\frac{\Delta^{-H_p-H_q}}{\sqrt{n}}\sum_{k=1}^{n}\Delta_k B^{(p)}\Delta_k B^{(q)}\Big) = \frac{\Delta^{-2(H_p+H_q)}}{n}\sum_{1\le k,l\le n}\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)}\Delta_k B^{(q)},\Delta_l B^{(p)}\Delta_l B^{(q)}\big)$$

$$\stackrel{(20)}{=}\frac{\Delta^{-2(H_p+H_q)}}{n}\sum_{1\le k,l\le n}\Big(\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)},\Delta_l B^{(p)}\big)\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(q)},\Delta_l B^{(q)}\big)$$

$$+\,\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(p)},\Delta_l B^{(q)}\big)\mathbb{C}\mathrm{ov}\big(\Delta_k B^{(q)},\Delta_l B^{(p)}\big)\Big)$$

$$\stackrel{(18)}{=}\frac{1}{n}\sum_{1\le k,l\le n}\Big(\sigma_p^2\sigma_q^2\gamma_p(k-l)\gamma_q(k-l)+\rho_{p,q}^2\sigma_p^2\sigma_q^2\gamma_{p,q}^2(k-l)\Big)$$

$$=\sigma_p^2\sigma_q^2\Big(\gamma_p(0)\gamma_q(0)+2\sum_{r=1}^{n}\frac{n-r}{n}\gamma_p(r)\gamma_q(r)+\rho_{p,q}^2\big(\gamma_{p,q}^2(0)+2\sum_{r=1}^{n}\frac{n-r}{n}\gamma_{p,q}^2(r)\big)\Big),$$

such that with dominated convergence, we obtain that

$$\lim_{n\to\infty}\mathbb{V}\mathrm{ar}\Big(\frac{\Delta^{-H_p-H_q}}{\sqrt{n}}\sum_{k=1}^{n}\Delta_k B^{(p)}\Delta_k B^{(q)}\Big) = \sigma_p^2\sigma_q^2(1+\rho_{p,q}^2)+2\sigma_p^2\sigma_q^2\sum_{r=1}^{\infty}\gamma_p(r)\gamma_q(r)$$

$$+\,2\rho_{p,q}^2\sigma_p^2\sigma_q^2\sum_{r=1}^{\infty}\gamma_{p,q}^2(r)\,.$$

The Hermite rank of $G:\mathbb{R}^2\to\mathbb{R}, G(x,y)=x\cdot y$, defined in general in Eq. (2.2) of Arcones (1994) as in (27), is 2. This is checked using that for $(X,Y)^\top$ jointly centred normal we have $\mathbb{E}[X^2Y^2]\ne 0$, while $\mathbb{E}[X^2Y]=0$. Therefore, the central limit theorem

$$\sqrt{n}\Big(\frac{\Delta^{-(H_p+H_q)}}{n}\sum_{k=1}^{n}\Delta_k B^{(p)}\Delta_k B^{(q)}-\rho_{p,q}\sigma_p\sigma_q\Big)\stackrel{d}{\to}\mathcal{N}\Big(0,\upsilon\Big)$$

is implied by Arcones (1994), where

$$\upsilon=\sigma_p^2\sigma_q^2(1+\rho_{p,q}^2)+\frac{\sigma_p^2\sigma_q^2}{2}\Big(\rho_{p,q}^2\sum_{r=1}^{\infty}\big(|r+1|^{H_p+H_q}+|r-1|^{H_p+H_q}-2|r|^{H_p+H_q}\big)^2$$

$$+\sum_{r=1}^{\infty}\big(|r+1|^{2H_p}+|r-1|^{2H_p}-2|r|^{2H_p}\big)\big(|r+1|^{2H_q}+|r-1|^{2H_q}-2|r|^{2H_q}\big)\Big)\,.$$

Based on similar computations, we deduce the asymptotic covariances

$$\lim_{n\to\infty}\mathbb{C}\mathrm{ov}\Big(\frac{\Delta^{-H_p-H_q}}{\sqrt{n}}\sum_{k=1}^{n}\Delta_k B^{(p)}\Delta_k B^{(q)},\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2\Big)=2\sigma_p^3\sigma_q\rho\Big(1+2\sum_{r=1}^{\infty}\gamma_p(r)\gamma_{p,q}(r)\Big)$$

between normalized realized variance and covariance, and between realized variances:

$$\lim_{n\to\infty} \mathbb{C}\text{ov}\Big(\frac{\Delta^{-2H_p}}{\sqrt{n}}\sum_{k=1}^{n}\big(\Delta_k B^{(p)}\big)^2, \frac{\Delta^{-2H_q}}{\sqrt{n}}\sum_{k=1}^{n}\big(\Delta_k B^{(q)}\big)^2\Big) = 2\rho^2\sigma_p^2\sigma_q^2\Big(1+2\sum_{r=1}^{\infty}\gamma_{p,q}^2(r)\Big).$$

Based on Arcones (1994) and Cramér-Wold, analogously as above, we obtain a multivariate central limit theorem

$$\sqrt{n}\left(\begin{array}{c}\frac{\Delta^{-2H_1}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(1)}\big)^2 - \sigma_1^2 \\[2mm] \frac{\Delta^{-2H_2}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(2)}\big)^2 - \sigma_2^2 \\[2mm] \frac{\Delta^{-(H_1+H_2)}}{n}\sum_{k=1}^{n}\Delta_k B^{(1)}\Delta_k B^{(2)} - \rho\sigma_1\sigma_2\end{array}\right) \xrightarrow{d} \mathcal{N}\big(0, \text{AV}\big), \qquad (28)$$

with the (3×3) asymptotic variance-covariance matrix AV, which contains the above given limiting variances and covariances. Since

$$\hat{\rho} = \frac{\frac{\Delta^{-(H_1+H_2)}}{n}\sum_{k=1}^{n}\Delta_k B^{(1)}\Delta_k B^{(2)}}{\sqrt{\frac{\Delta^{-2H_1}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(1)}\big)^2\,\frac{\Delta^{-2H_2}}{n}\sum_{k=1}^{n}\big(\Delta_k B^{(2)}\big)^2}},$$

we derive (14) from (28) with the multivariate Δ-method. With the function $h : (x, y, z)^\top \mapsto z/\sqrt{xy}$, $\nabla h(x, y, z) = \big(-yz/(2(xy)^{3/2}), -xz/(2(xy)^{3/2}), (xy)^{-1/2}\big)^\top$, we obtain

$$\begin{aligned}\text{AVAR}_{\hat{\rho}} &= \big(\nabla h(\sigma_1^2, \sigma_2^2, \rho\sigma_1\sigma_2)\big)^\top \cdot \text{AV} \cdot \nabla h(\sigma_1^2, \sigma_2^2, \rho\sigma_1\sigma_2) \\[2mm] &= \frac{\rho^2}{4\sigma_1^4}\,\text{AV}_{11} + \frac{\rho^2}{4\sigma_2^4}\,\text{AV}_{22} + \frac{\text{AV}_{33}}{\sigma_1^2\sigma_2^2} + \frac{\rho^2\,\text{AV}_{12}}{2\sigma_1^2\sigma_2^2} - \frac{\rho\,\text{AV}_{13}}{\sigma_1^3\sigma_2} - \frac{\rho\,\text{AV}_{23}}{\sigma_1\sigma_2^3}.\end{aligned}$$

Inserting the above limiting variances and covariances and simplifying the terms yields (14). This finishes the proof of Theorem 3.3.

8.2. Proofs of theoretical results on forecasting

8.2.1. Proof of Proposition 4.1

Consider a multivariate normally distributed random vector

$$X = \left(\begin{array}{c}Y \\ Z\end{array}\right) \sim \mathcal{N}\left(\left(\begin{array}{c}\mu_Y \\ \mu_Z\end{array}\right), \left(\begin{array}{cc}\Sigma_Y & \Sigma_{Y,Z} \\ \Sigma_{Y,Z}^\top & \Sigma_Z\end{array}\right)\right) = \mathcal{N}(\mu, \Sigma),$$

which is $\mathbb{R}^d$-valued with sub-vectors Y taking values in $\mathbb{R}^p$ and Z taking values in $\mathbb{R}^{d-p}$, $p, (d-p) \in \mathbb{N}$, with marginals $Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$, $Z \sim \mathcal{N}(\mu_Z, \Sigma_Z)$, where $\mu_Y \in \mathbb{R}^p$, $\mu_Z \in \mathbb{R}^{d-p}$,

$\Sigma_Y \in \mathbb{R}^{p \times p}$, $\Sigma_Z \in \mathbb{R}^{(d-p) \times (d-p)}$ and $\Sigma_{Y,Z} \in \mathbb{R}^{p \times (d-p)}$. In this case, it is well known[11] that the conditional distribution of Y given $Z = z$ is multivariate normal:

$$(Y \mid Z = z) \sim \mathcal{N}\big(\mu_Y + \Sigma_{Y,Z}\Sigma_Z^{-1}(z - \mu_Z), \Sigma_Y - \Sigma_{Y,Z}\Sigma_Z^{-1}\Sigma_{Y,Z}^\top\big). \tag{29}$$

Having a Gaussian process, we apply this general result with $d = 3$, $p = 1$, to $Y = B_{t+h}^{(1)}$, with $\mu_Y = 0$, $\Sigma_Y = \sigma_1^2(t+h)^{2H_1}$, and $Z = (B_t^{(1)}, B_t^{(2)})^\top$, with $\mu_Z = (0,0)^\top$, and

$$\Sigma_Z = \begin{pmatrix} \sigma_1^2 t^{2H_1} & \rho\sigma_1\sigma_2 t^{2H} \\ \rho\sigma_1\sigma_2 t^{2H} & \sigma_2^2 t^{2H_2} \end{pmatrix},$$

with H the cross Hurst exponent from (5). The determinant

$$\det(\Sigma_Z) = \sigma_1^2 \sigma_2^2 (1 - \rho^2) t^{4H}$$

depends on (H_1, H_2) only via H. We obtain that

$$\Sigma_Z^{-1} = \begin{pmatrix} \big(\sigma_1^2(1-\rho^2)t^{2H_1}\big)^{-1} & -\rho\big(\sigma_1\sigma_2(1-\rho^2)t^{2H}\big)^{-1} \\ -\rho\big(\sigma_1\sigma_2(1-\rho^2)t^{2H}\big)^{-1} & \big(\sigma_2^2(1-\rho^2)t^{2H_2}\big)^{-1} \end{pmatrix},$$

such that with $\Sigma_{Y,Z} = (\sigma_1^2 w(t,h,H_1), \rho\sigma_1\sigma_2 w(t,h,H))$, we obtain

$$\hat{B}_{t+h|t}^{(1)} = \Sigma_{Y,Z}\Sigma_Z^{-1}z$$

$$= (\sigma_1^2 w(t,h,H_1), \rho\sigma_1\sigma_2 w(t,h,H)) \cdot \begin{pmatrix} \frac{B_t^{(1)}}{\sigma_1^2(1-\rho^2)t^{2H_1}} - \frac{\rho B_t^{(2)}}{\sigma_1\sigma_2(1-\rho^2)t^{2H}} \\ \frac{B_t^{(2)}}{\sigma_2^2(1-\rho^2)t^{2H_2}} - \frac{\rho B_t^{(1)}}{\sigma_1\sigma_2(1-\rho^2)t^{2H}} \end{pmatrix}.$$

Computing this product proves (16). The MSFE equals the conditional variance:

$$\mathbb{E}\left[\big(\hat{B}_{t+h|t}^{(1)} - B_{t+h}^{(1)}\big)^2\right] = \Sigma_Y - \Sigma_{Y,Z}\Sigma_Z^{-1}\Sigma_{Y,Z}^\top$$

$$= \sigma_1^2\left((t+h)^{2H_1} - \frac{1}{1-\rho^2} \cdot \left(\frac{(w(t,h,H_1))^2}{t^{2H_1}} + \rho^2\left(\frac{2w(t,h,H_1)w(t,h,H)}{t^{2H}} - \frac{(w(t,h,H))^2}{t^{2H_2}}\right)\right)\right). \quad \square$$

---

[11] See, for instance, (Eaton, 1983, Proposition 3.13).

8.2.2. Proof of Proposition 4.2

With similar notation as in the proof of Proposition 4.1, we apply (29) with $Y = B_{t+h}^{(1)}$, and $Z = (B_t^{(1)}, \ldots, B_t^{(d)})^\top$. The main step is to see that the $(d \times d)$ inverse of the covariance matrix Σ_Z has entries

$$(\Sigma_Z^{-1})_{ij} = (\Sigma^{-1})_{ij}\, t^{-(H_i + H_j)}\,, \quad 1 \leq i, j \leq d\,.$$

First, $\det(\Sigma_Z) = \det(\Sigma) \cdot t^{2 \sum_{k=1}^d H_k}$ is verified by induction. For $d = 3$, this follows readily with the rule of Sarrus. Then, a Laplace expansion, e.g., with respect to the first row, yields the induction step. Hence, the determinant depends in general only on the sum of all Hurst exponents. This yields for the inverse for $1 \leq i, j \leq d$ that

$$\begin{aligned}
\left(\Sigma_Z^{-1}\right)_{ij} &= \frac{1}{\det(\Sigma_Z)}\, \mathrm{adj}(\Sigma_Z) = \frac{(-1)^{i+j}}{\det(\Sigma_Z)} \det\left((\Sigma_Z)_{-j,-i}\right) \\
&= \frac{(-1)^{i+j}}{\det(\Sigma)} t^{-2 \sum_{k=1}^d H_k} \det\left((\Sigma)_{-j,-i}\right) t^{\sum_{k=1}^d H_k (1 + \mathbb{1}(k \notin \{H_i, H_j\}))} \\
&= (\Sigma^{-1})_{ij}\, t^{-(H_i + H_j)}\,,
\end{aligned}$$

where we write $(\Sigma_Z)_{-j,-i}$ for the $((d-1) \times (d-1))$ matrix obtained from Σ_Z by deleting the jth row and ith column, and $(\Sigma)_{-i,-j}$ analogously. We obtain the result from

$$\hat{B}_{t+h|t}^{(1)} = \Sigma_{Y,Z} \Sigma_Z^{-1} \cdot (B_t^{(1)}, \ldots, B_t^{(d)})^\top\,,$$

with $\Sigma_{Y,Z} = (\Sigma_{11} w(t, h, H_1), \ldots, \Sigma_{1d} w(t, h, (H_1 + H_d)/2))$. $\qquad \square$

8.2.3. Proof of Proposition 4.3

We use analogous notation as in the previous proofs. Based on similar linear algebra as above, we can prove inductively that Σ^{-1} equals

$$\frac{1}{1 + (d-2)\rho - (d-1)\rho^2} \begin{pmatrix}
1 + (d-2)\rho & -\rho & -\rho & \cdots & -\rho \\
-\rho & 1 + (d-2)\rho & -\rho & \cdots & -\rho \\
-\rho & \cdots & \ddots & \cdots & -\rho \\
-\rho & \cdots & \cdots & 1 + (d-2)\rho & -\rho \\
-\rho & -\rho & \cdots & -\rho & 1 + (d-2)\rho
\end{pmatrix}.$$

Computing

$$\big(w(t,h,H_1),\rho w(t,h,(H_1+H)/2),\ldots,\rho w(t,h,(H_1+H)/2)\big)\Sigma_Z^{-1}\begin{pmatrix} w(t,h,H_1) \\ \rho w(t,h,(H_1+H)/2) \\ \vdots \\ \rho w(t,h,(H_1+H)/2) \end{pmatrix}$$

and re-using that $\big(\Sigma_Z^{-1}\big)_{ij}=(\Sigma^{-1})_{ij}\,t^{-(H_i+H_j)}$ yields the result. $\qquad\square$

8.2.4. Proof of Proposition 4.4

Since for $H=H_1=H_2$, it holds that

$$\mathbb{E}\big[B_sB_t^\top\big]=w(t,s-t,H)\cdot\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

the bivariate time-reversible fBm satisfies in this case the equality in distribution $B_t \overset{d}{=} A\cdot(W_t^{(1)},W_t^{(2)})^\top$, where $W_t^{(1)}$ and $W_t^{(2)}$ are two independent univariate fBms with Hurst exponents H, and

$$A=\begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}\text{ with }AA^\top=\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Since the covariance function uniquely determines the distribution of a centered Gaussian process, the equality is implied by that of the covariance functions. Let us point out that for $H_1\neq H_2$, it is impossible to illustrate (B_t) as a linear transformation of independent fBms, unless all correlations between components of (B_t) directly vanish.

The main step of the proof is now to show that for any $c,d,x,y\in\mathbb{R}$, we have that

$$(\sigma_1^2\,c,\rho\sigma_1\sigma_2\,c)A^{-\top}\begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix}A^{-1}\begin{pmatrix} x \\ y \end{pmatrix}=c\cdot d\cdot x. \tag{30}$$

We use the standard notation $A^{-\top}$ for $(A^{-1})^\top=(A^\top)^{-1}$. Since the inverse

$$A^{-1}=\begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix},$$

47

of the triangular matrix is easily determined, we obtain that

$$(\sigma_1^2 c, \rho\sigma_1\sigma_2 c)A^{-\top}\begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix}A^{-1}\begin{pmatrix} x \\ y \end{pmatrix} =$$

$$= (\sigma_1^2 c, \rho\sigma_1\sigma_2 c)\begin{pmatrix} \frac{1}{\sigma_1} & -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} \\ 0 & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix}\begin{pmatrix} \frac{d}{\sigma_1} & 0 \\ -\frac{d\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{d}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= (\sigma_1^2 c, \rho\sigma_1\sigma_2 c)\begin{pmatrix} \left(\frac{d}{\sigma_1^2} + \frac{d\rho^2}{\sigma_1^2(1-\rho^2)}\right)x - \frac{d\rho}{\sigma_1\sigma_2(1-\rho^2)}y \\ -\frac{d\rho}{\sigma_1\sigma_2(1-\rho^2)}x + \frac{d}{\sigma_2^2(1-\rho^2)}y \end{pmatrix}$$

$$= cd\left(1 + \frac{\rho^2}{1-\rho^2}\right)x - \frac{cd\rho\sigma_1}{\sigma_2(1-\rho^2)}y - \frac{cd\rho^2}{1-\rho^2}x + \frac{cd\rho\sigma_1}{\sigma_2(1-\rho^2)}y$$

$$= cdx\,.$$

Consider a forecast based on observations at two time points t_1 and t_2, being more general than equidistant ones. The covariance matrix $\mathfrak{W}$ of $(W_{t_1}^{(1)}, W_{t_1}^{(2)}, W_{t_2}^{(1)}, W_{t_2}^{(2)})^\top$ contains due to the independence of the components four (2×2)-diagonal blocks:

$$\mathfrak{W} = \begin{pmatrix} (t_1)^{2H} & 0 & w(t_1, t_2, H) & 0 \\ 0 & (t_1)^{2H} & 0 & w(t_1, t_2, H) \\ w(t_1, t_2, H) & 0 & (t_2)^{2H} & 0 \\ 0 & w(t_1, t_2, H) & 0 & (t_2)^{2H} \end{pmatrix},$$

and its inverse $\mathfrak{W}^{-1}$ is of similar block-diagonal form. This and its explicit entries are determined in Lemma 1, but the explicit entries are not required to conclude. Based on the linear transformation

$$(B_{t_1}^{(1)}, B_{t_1}^{(2)}, B_{t_2}^{(1)}, B_{t_2}^{(2)})^\top \stackrel{d}{=} \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}(W_{t_1}^{(1)}, W_{t_1}^{(2)}, W_{t_2}^{(1)}, W_{t_2}^{(2)})^\top$$

the covariance matrix of $(B_{t_1}^{(1)}, B_{t_1}^{(2)}, B_{t_2}^{(1)}, B_{t_2}^{(2)})^\top$ is given by

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}\mathfrak{W}\begin{pmatrix} A^\top & 0 \\ 0 & A^\top \end{pmatrix}.$$

Writing $\mathfrak{W}^{-1}$ with three (2×2)-diagonal blocks D_1, D_2 and D_{12}, the inverse covariance

matrix of $(B_{t_1}^{(1)}, B_{t_1}^{(2)}, B_{t_2}^{(1)}, B_{t_2}^{(2)})^\top$ yields

$$
\begin{pmatrix} A^{-\top} & 0 \\ 0 & A^{-\top} \end{pmatrix} \begin{pmatrix} D_1 & D_{12} \\ D_{12} & D_2 \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ 0 & A^{-1} \end{pmatrix} = \begin{pmatrix} A^{-\top} D_1 A^{-1} & A^{-\top} D_{12} A^{-1} \\ A^{-\top} D_{12} A^{-1} & A^{-\top} D_2 A^{-1} \end{pmatrix}.
$$

Computing the optimal forecast based on the conditional expectation (29), (30) applies to all four blocks in the same way, such that the forecast does not contain the other component. The same (2×2)-diagonal block structure readily generalizes to arbitrarily many observation times, such that an algebraic induction step only requires an iterated application of (30). The induction hence yields the general result. $\qquad\square$

To prove Proposition 4.5, we require the following algebraic lemma on inverses of block matrices of a specific form.

Lemma 1. *A positive definite, symmetric matrix $\mathfrak{W} \in \mathbb{R}^{nd \times nd}$ of the form*

$$
\mathfrak{W} = \begin{pmatrix} w_{1,1} I_d & w_{1,2} I_d & \dots & w_{1,n} I_d \\ \vdots & w_{2,2} I_d & \dots & \vdots \\ \vdots & & \ddots & \\ w_{1,n} I_d & w_{2,n} I_d & \dots & w_{n,n} I_d \end{pmatrix}, \ w_{i,j} \in \mathbb{R}, 1 \le i, j \le n,
$$

where I_d denotes the $(d \times d)$ identity matrix, has an inverse of the same form, i.e. there exist $\tilde{w}_{i,j} \in \mathbb{R}, 1 \le i, j \le n$, such that

$$
\mathfrak{W}^{-1} = \begin{pmatrix} \tilde{w}_{1,1} I_d & \tilde{w}_{1,2} I_d & \dots & \tilde{w}_{1,n} I_d \\ \vdots & \tilde{w}_{2,2} I_d & \dots & \vdots \\ \vdots & & \ddots & \\ \tilde{w}_{1,n} I_d & \tilde{w}_{2,n} I_d & \dots & \tilde{w}_{n,n} I_d \end{pmatrix}.
$$

We prove the lemma by algebraic induction. For $n = 2$, it is simple to verify by matrix multiplication that

$$
\begin{pmatrix} w_{1,1} I_d & w_{1,2} I_d \\ w_{1,2} I_d & w_{2,2} I_d \end{pmatrix} = \frac{1}{w_{1,1} w_{2,2} - w_{1,2}^2} \begin{pmatrix} w_{2,2} I_d & -w_{1,2} I_d \\ -w_{1,2} I_d & w_{1,1} I_d \end{pmatrix},
$$

what serves as the induction basis and yields an explicit formula for the one-period setting.

If a positive definite, symmetric $\mathfrak{W} \in \mathbb{R}^{nd \times nd}$ is of the above form, write $\mathfrak{W} \in \mathrm{BD}(nd)$. Assume as induction hypothesis that for $\mathfrak{W}_{n-1} \in \mathrm{BD}((n-1)d)$, $n \geq 3$ arbitrary, it holds that $\mathfrak{W}_{n-1}^{-1} \in \mathrm{BD}((n-1)d)$. To prepare the induction step, write for $\mathfrak{W}_n \in \mathrm{BD}(nd)$:

$$\mathfrak{W}_n = \begin{pmatrix} \mathfrak{W}_{n-1} & C \\ C^\top & w_{n,n} I_d \end{pmatrix},$$

with $w_{n,n} \in \mathbb{R}$, $C^\top = (w_{1,n} I_d, \ldots, w_{n-1,n} I_d) \in \mathbb{R}^{d \times (n-1)d}$. A general block inversion formula, see Sec. 9.3.1 in Petersen et al. (2012), yields

$$\mathfrak{W}_n^{-1} = \begin{pmatrix} (\mathfrak{W}_{n-1} - C w_{n,n}^{-1} I_d C^\top)^{-1} & -\mathfrak{W}_{n-1}^{-1} C (w_{n,n} I_d - C^\top \mathfrak{W}_{n-1}^{-1} C)^{-1} \\ -(w_{n,n} I_d - C^\top \mathfrak{W}_{n-1}^{-1} C)^{-1} C^\top \mathfrak{W}_{n-1}^{-1} & (w_{n,n} I_d - C^\top \mathfrak{W}_{n-1}^{-1} C)^{-1} \end{pmatrix}.$$

Since C and $C^\top$ consist of $(d \times d)$-diagonal blocks in that I_d is multiplied with real scalars, it follows that $C^\top \mathfrak{W}_{n-1}^{-1} C$ is a $(d \times d)$-diagonal matrix of the same form, such that $w_{n,n} I_d - C^\top \mathfrak{W}_{n-1}^{-1} C$ and its inverse are as well of this form. Since

$$CC^\top = \begin{pmatrix} w_{1,n}^2 I_d & \cdots & w_{1,n} w_{n-1,n} I_d \\ & \ddots & \\ w_{1,n} w_{n-1,n} I_d & \cdots & w_{n-1,n}^2 I_d \end{pmatrix} \in \mathrm{BD}((n-1)d),$$

$\mathfrak{W}_{n-1} - w_{n,n}^{-1} CC^\top$, and its inverse, are as well in $\mathrm{BD}((n-1)d)$. C containing $(d \times d)$-diagonal blocks in that I_d is multiplied with real scalars, and since $\mathfrak{W}_{n-1}^{-1} \in \mathrm{BD}((n-1)d)$, $\mathfrak{W}_{n-1}^{-1} C (w_{n,n} I_d - C^\top \mathfrak{W}_{n-1}^{-1} C)^{-1}$ preserves this structure, such that we conclude the induction step $\mathfrak{W}_n^{-1} \in \mathrm{BD}((n-1)d)$.

8.2.5. Proof of Proposition 4.5

In the general, d-dimensional setting with a mfBm $B_t = (B_t^{(1)}, \ldots, B_t^{(d)})^\top$, with all Hurst exponents being $H \in (0,1)$, it holds that

$$\mathbb{E}[B_s B_t^\top] = w(t, s-t, H) \cdot \Sigma,$$

with the static, positive definite variance-covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ of $(B_1^{(1)}, \ldots, B_1^{(d)})^\top$. By the spectral theorem, there is a diagonal matrix $\Lambda \in \mathbb{R}^{d \times d}$, with strictly positive entries on the diagonal, and an orthogonal matrix $O \in \mathbb{R}^{d \times d}$, such that $\Sigma = O \Lambda O^\top$ and we set

$\Sigma^{1/2} = O\Lambda^{1/2}O^\top$, where $\Lambda^{1/2}$ is obtained from Λ taking square roots entry-wise. Similar as in the proof of Proposition 4.4, (5) yields an illustration

$$B_t \overset{d}{=} \Sigma^{1/2} \cdot W_t,$$

with $W_t = (W_t^{(1)}, \ldots, W_t^{(d)})^\top$ being a vector of d independent fBms with equal Hurst exponents H. Considering t observations in a vector $(B_\Delta^{(1)}, \ldots, B_\Delta^{(d)}, \ldots, B_{t\Delta}^{(1)}, \ldots, B_{t\Delta}^{(d)})^\top \in \mathbb{R}^{td}$, this vector has the variance-covariance matrix $\mathfrak{C} \in \mathbb{R}^{td \times td}$ of the form

$$\mathfrak{C} = BD_{\Sigma^{1/2}} \cdot \begin{pmatrix} \Delta^{2H}I_d & w(\Delta, 2\Delta, H) & \ldots & w(\Delta, t\Delta, H) \\ w(\Delta, 2\Delta, H) & (2\Delta)^{2H}I_d & \ldots & \vdots \\ \vdots & & \ddots & \vdots \\ w(\Delta, t\Delta, H) & \ldots & \ldots & (t\Delta)^{2H}I_d \end{pmatrix} \cdot BD_{\Sigma^{1/2}},$$

where I_d denotes the $(d \times d)$ identity matrix and $BD_{\Sigma^{1/2}}$ the block-diagonal matrix

$$BD_{\Sigma^{1/2}} = \begin{pmatrix} \Sigma^{1/2} & 0 & \ldots & 0 \\ 0 & \Sigma^{1/2} & 0 \ldots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \Sigma^{1/2} \end{pmatrix}.$$

The inverse $\mathfrak{C}^{-1} \in \mathbb{R}^{td \times td}$, entering the conditional expectation of $B_{t\Delta+h}^{(1)}$, $h > 0$, given the t observations at times $j\Delta, 1 \leq j \leq t$, is with Lemma 1 hence of the form

$$\mathfrak{C}^{-1} = \begin{pmatrix} \Sigma^{-1}c_{1,1} & \Sigma^{-1}c_{1,2} & \ldots & \Sigma^{-1}c_{1,t} \\ \Sigma^{-1}c_{1,2} & \Sigma^{-1}c_{2,2} & \ldots & \vdots \\ \vdots & & \ddots & \\ \Sigma^{-1}c_{1,t} & \ldots & \ldots & \Sigma^{-1}c_{t,t} \end{pmatrix},$$

with real scalars $c_{i,j}$, $1 \leq i, j \leq t$, which depend on Δ, H and the observation times, but whose explicit values are not important. With algebraic induction as in the proof of

Proposition 4.4, it suffices to conclude block-wise that

$$c \cdot \Sigma_{1\bullet} \cdot \Sigma^{-1} \cdot d \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} = c \cdot d \cdot x_1 \, , \tag{31}$$

for arbitrary $c, d \in \mathbb{R}$, with $\Sigma_{1\bullet} \in \mathbb{R}^{1 \times d}$ being the first row of Σ, entering the conditional expectation of $B_{t\Delta+h}^{(1)}$ given $(B_\Delta^{(1)}, \ldots, B_\Delta^{(d)}, \ldots, B_{t\Delta}^{(1)}, \ldots, B_{t\Delta}^{(d)})^\top$. (31) is readily obtained, since

$$\Sigma \cdot \Sigma^{-1} = I_d = \begin{pmatrix} \Sigma_{1\bullet} \cdot \Sigma^{-1} \\ \vdots \\ \vdots \end{pmatrix} \, .$$

Similar as in the proof of Proposition 4.4, the main ingredient is that $\mathfrak{C}^{-1}$ only contains block matrices of the form $\Sigma^{-1} \cdot d$, $d \in \mathbb{R}$. The equidistant observation scheme, $t_j = j\Delta, 1 \leq j \leq t$, is not important for this proof which generalizes to arbitrary observation times. $\qquad\square$

References

Amblard, P.-O. and Coeurjolly, J.-F. (2011). Identification of the multivariate fractional Brownian motion. *IEEE Transactions on Signal Processing*, 59(11):5152–5168.

Amblard, P.-O., Coeurjolly, J.-F., Lavancier, F., and Philippe, A. (2013). Basic properties of the multivariate fractional Brownian motion. In *Séminaires et congrès*, volume 28, pages 65–87.

Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance*, 52(3):975–1005.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Arcones, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability*, 22(4):2242 – 2274.

Asai, M., McAleer, M., and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews*, 25(2-3):145–175.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.

Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). Multivariate garch models: A survey. *Journal of Applied Econometrics*, 21(1):79–109.

Bayer, C., Friz, P., and Gatheral, J. (2016). Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904.

Bégyn, A. (2007). Functional limit theorems for generalized quadratic variations of Gaussian processes. *Stochastic Processes and their Applications*, 117(12):1848–1869.

Bibinger, M. and Sonntag, M. (2025). Testing for jumps in processes with integral fractional part and jump-robust inference on the Hurst exponent. *Econometrics and Statistics, forthcoming.*

Bolko, A. E., Christensen, K., Pakkanen, M. S., and Veliyev, B. (2023). A GMM approach to estimate the roughness of stochastic volatility. *Journal of Econometrics*, 235(2):745–778.

Breuer, P. and Major, P. (1983). Central limit theorems for non-linear functionals of Gaussian fields. *Journal of Multivariate Analysis*, 13(3):425–441.

Brouste, A. and Fukasawa, M. (2018). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *The Annals of Statistics*, 46(5):2045–2061.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26(6):922–947.

Chong, C. H., Hoffmann, M., Liu, Y., Rosenbaum, M., and Szymanski, G. (2024). Statistical inference for rough volatility: Central limit theorems. *The Annals of Applied Probability*, 34(3):2600–2649.

Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes*, 4:199–227.

Coeurjolly, J.-F., Amblard, P.-O., and Achard, S. (2013). Wavelet analysis of the multivariate fractional Brownian motion. *ESAIM: Probability and Statistics*, 17:592–604.

Comte, F. and Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics*, 73(1):101–149.

Comte, F. and Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance*, 8(4):291–323.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Didier, G. and Pipiras, V. (2011). Integral representations and properties of operator fractional Brownian motions. *Bernoulli*, pages 1–33.

Ding, Y., Engle, R. F., Li, Y., and Zheng, X. (2024). Multiplicative factor modeling for volatility. *Journal of Econometrics*, 249:105959.

Ding, Z., Granger, C. W., and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106.

Dugo, R., Giorgio, G., and Pigato, P. (2024). The multivariate fractional Ornstein-Uhlenbeck process. *arXiv preprint arXiv:2408.03051*.

Eaton, M. L. (1983). *Multivariate statistics: A vector space approach*, volume 512. Wiley New York.

Euch, O. E. and Rosenbaum, M. (2018). Perfect hedging in rough Heston models. *The Annals of Applied Probability*, 28(6):3813–3856.

Fouque, J.-P. and Hu, R. (2018). Optimal portfolio under fast mean-reverting fractional stochastic environment. *SIAM Journal on Financial Mathematics*, 9(2):564–601.

Fukasawa, M., Takabatake, T., and Westphal, R. (2022). Consistent estimation for fractional stochastic volatility model under high-frequency asymptotics. *Mathematical Finance*, 32(4):1086–1132.

Garnier, J. and Sølna, K. (2018). Option pricing under fast-varying and rough stochastic volatility. *Annals of Finance*, 14(4):489–516.

Gatheral, J., Jaisson, T., and Rosenbaum, M. (2018). Volatility is rough. *Quantitative Finance*, 18(6):933–949.

Glasserman, P. and He, P. (2020). Buy rough, sell smooth. *Quantitative Finance*, 20(3):363–378.

Harvey, A. C. (2007). Long memory in stochastic volatility. In Knight, J. and Satchell, S., editors, *Forecasting Volatility in the Financial Markets*, Quantitative Finance, pages 351–363. Butterworth-Heinemann, Oxford, third edition edition.

Isserlis, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1-2):134–139.

Kubilius, K. (2020). CLT for quadratic variation of Gaussian processes and its application to the estimation of the Orey index. *Statistics & Probability Letters*, 165:108845.

Lavancier, F., Philippe, A., and Surgailis, D. (2009). Covariance function of vector self-similar processes. *Statistics & Probability Letters*, 79(23):2415–2421.

Lavancier, F., Philippe, A., and Surgailis, D. (2010). A two-sample test for comparison of long memory parameters. *Journal of Multivariate Analysis*, 101(9):2118–2136.

Li, J., Phillips, P. C. B., Shi, S., and Yu, J. (2024). Weak identification of long memory with implications for volatility modelling. *Review of Financial Studies, forthcoming.*

Nourdin, I., Nualart, D., and Tudor, C. A. (2010). Central and non-central limit theorems for weighted power variations of fractional Brownian motion. *Annales de l'IHP Probabilités et Statistiques*, 46(4):1055–1079.

Nuzman, C. J. and Poor, H. V. (2000). Linear estimation of self-similar processes via Lamperti's transformation. *Journal of Applied Probability*, 37(2):429–452.

Petersen, K. B., Pedersen, M. S., et al. (2012). The matrix cookbook. *Technical University of Denmark*, 7(15):510. `https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf`.

Shi, S. and Yu, J. (2023). Volatility puzzle: Long memory or antipersistency. *Management Science*, 69(7):3861–3883.

Wackernagel, H. (2003). *Multivariate geostatistics: An introduction with applications*. Springer Science & Business Media.

Wang, X., Xiao, W., and Yu, J. (2023). Modeling and forecasting realized volatility with the fractional Ornstein–Uhlenbeck process. *Journal of Econometrics*, 232(2):389–415.

Wang, X. and Yu, J. (2023). Latent local-to-unit root models. *Econometric Reviews*, 42(7):586–611.

Wang, X., Yu, J., and Zhang, C. (2024). On the optimal forecast with the fractional Brownian motion. *Quantitative Finance*, 24(2):337–346.

Appendix A. HAR and Vector HAR

Appendix A.1. HAR model

The HAR model proposed by Corsi (2009) has a simple linear specification:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t|t-4} + \beta_3 RV_{t|t-21} + \epsilon_t, \tag{A.1}$$

where RV_t is the daily realized volatility and $RV_{t-j|t-k} = \frac{1}{k-j+1} \sum_{i=j}^{k} RV_{t-i}$ with $j \leq k$, and ϵ_t is the disturbance term. Based on this definition, $RV_{t|t-4}$ and $RV_{t|t-21}$ represent the weekly and monthly average realized volatility, respectively. The HAR model can be estimated using the ordinary least squares (OLS) method.

Appendix A.2. Vector HAR model

The HAR model only handles a univariate time series. If we extend it to a multivariate setting and incorporate information from other time series for forecasting, a natural approach is to stack the realized volatilities and formulate a vector autoregressive model as follows (referred to as the Vector HAR model or VHAR for short):

$$\mathbf{RV}_{t+h} = \beta_\mathbf{0} + \mathbf{\Phi}_1 \mathbf{RV}_t + \mathbf{\Phi}_2 \mathbf{RV}_{t|t-4} + \mathbf{\Phi}_3 \mathbf{RV}_{t|t-21} + \epsilon_\mathbf{t},$$

where

$$\mathbf{RV}_t = \begin{pmatrix} RV_t^1 \\ RV_t^2 \\ ... \\ RV_t^J \end{pmatrix}, \; \mathbf{RV}_{t|t-4} = \begin{pmatrix} RV_{t|t-4}^1 \\ RV_{t|t-4}^2 \\ ... \\ RV_{t|t-4}^J \end{pmatrix}, \; \mathbf{RV}_{t|t-21} = \begin{pmatrix} RV_{t|t-21}^1 \\ RV_{t|t-21}^2 \\ ... \\ RV_{t|t-21}^J \end{pmatrix}, \; \mathbf{C} = \begin{pmatrix} \beta_0^1 \\ \beta_0^2 \\ ... \\ \beta_0^J \end{pmatrix}, \; \epsilon_t = \begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \\ ... \\ \epsilon_t^J \end{pmatrix},$$

J represents the number of realized volatilities and the coefficient matrices are defined as $(i = 1, 2, 3)$:

$$
\mathbf{\Phi}_i = \begin{pmatrix} \phi_{11,i} & \phi_{12,i} & ... & \phi_{1J,i} \\ \phi_{21,i} & \phi_{22,i} & ... & \phi_{2J,i} \\ ... & ... & ... & ... \\ \phi_{J1,i} & \phi_{J2,i} & ... & \phi_{JJ,i} \end{pmatrix},
$$

Assuming that the covariance matrix of the error term, ϵ_t, is a diagonal matrix allows estimation and forecasting to be performed row by row. To forecast the first element in $\mathbf{RV}_t$, denoted as RV_t^1, the regression model can be expressed as follows and estimated using OLS:

$$
\begin{aligned}
RV_{t+h}^1 =& \beta_0^1 + \phi_{11,1} RV_t^1 + \phi_{11,2} RV_{t|t-4}^1 + \phi_{11,3} RV_{t|t-21}^2 \\
& + \sum_{j=2}^{J} \left\{ \phi_{1(j+1),1} RV_t^j + \phi_{1(j+1),2} RV_{t|t-4}^j + \phi_{1(j+1),3} RV_{t|t-21}^j \right\} + \epsilon_t^1.
\end{aligned}
$$

The additional terms in the vector HAR model, compared to the standard HAR model,

$$
\sum_{j=2}^{J} \left\{ \phi_{1(j+1),1} RV_t^j + \phi_{1(j+1),2} RV_{t|t-4}^j + \phi_{1(j+1),3} RV_{t|t-21}^j \right\}
$$

represent information from other time series.

Appendix B. Parameter estimates for 20 Dow Jones 30 stocks

Table B.12 reports the point estimates of H and ρ for 20 Dow Jones 30 stocks in alphabetical order. We also report the point estimates of η in Table B.13 for these stocks. Table B.12 suggests that there are differences among the estimated Hurst exponents and that the estimated correlation parameters are always positive and far away from zero. We expect mfBm should offer forecasting improvements over univariate fBm. Table B.13 indicates that the asymmetry parameters are close to zero, indicating the usefulness of time-reversible mfBm.

	H	AAPL	ALD	AMGN	AXP	BA	BEL	CAT	CHV	CRM	CSCO	DIS	GS	HD	IBM	INTC	JNJ	JPM	KO	MCD	MMM
AAPL	0.2821	1																			
ALD	0.1932	0.39	1																		
AMGN	0.2066	0.37	0.31	1																	
AXP	0.2216	0.38	0.39	0.30	1																
BA	0.2481	0.38	0.36	0.29	0.36	1															
BEL	0.1777	0.34	0.33	0.27	0.36	0.27	1														
CAT	0.2050	0.37	0.38	0.24	0.40	0.32	0.30	1													
CHV	0.2196	0.41	0.40	0.31	0.44	0.34	0.39	0.41	1												
CRM	0.2530	0.40	0.31	0.27	0.35	0.28	0.28	0.31	0.33	1											
CSCO	0.2336	0.46	0.39	0.35	0.44	0.37	0.35	0.38	0.46	0.38	1										
DIS	0.2135	0.35	0.32	0.31	0.37	0.30	0.33	0.30	0.39	0.30	0.35	1									
GS	0.2616	0.46	0.41	0.32	0.48	0.39	0.36	0.42	0.45	0.38	0.46	0.42	1								
HD	0.2237	0.41	0.40	0.31	0.36	0.33	0.37	0.34	0.40	0.30	0.41	0.36	0.45	1							
IBM	0.2349	0.39	0.36	0.28	0.35	0.33	0.35	0.36	0.38	0.33	0.40	0.35	0.42	0.35	1						
INTC	0.2247	0.45	0.35	0.32	0.38	0.32	0.32	0.36	0.41	0.35	0.42	0.31	0.40	0.37	0.35	1					
JNJ	0.1363	0.36	0.36	0.32	0.36	0.34	0.39	0.29	0.37	0.31	0.37	0.33	0.45	0.38	0.31	0.31	1				
JPM	0.2326	0.52	0.46	0.35	0.53	0.45	0.45	0.48	0.52	0.39	0.54	0.44	0.68	0.54	0.42	0.45	0.49	1			
KO	0.2122	0.38	0.37	0.29	0.37	0.35	0.39	0.35	0.42	0.32	0.40	0.36	0.44	0.41	0.36	0.36	0.37	0.50	1		
MCD	0.2172	0.35	0.31	0.30	0.32	0.26	0.32	0.32	0.37	0.25	0.33	0.31	0.37	0.39	0.31	0.31	0.31	0.39	0.34	1	
MMM	0.2135	0.35	0.44	0.28	0.40	0.32	0.36	0.44	0.42	0.31	0.39	0.36	0.45	0.39	0.35	0.32	0.41	0.46	0.39	0.34	1

Table B.12: Estimates of H and ρ for 20 Dow Jones 30 stocks

	AAPL	ALD	AMGN	AXP	BA	BEL	CAT	CHV	CRM	CSCO	DIS	GS	HD	IBM	INTC	JNJ	JPM	KO	MCD
ALD	0.17																		
AMGN	0.02	-0.12																	
AXP	0.09	-0.18	0.05																
BA	0.11	-0.10	-0.01	-0.02															
BEL	0.17	-0.06	-0.01	-0.02	-0.05														
CAT	0.08	-0.05	0.03	0.00	0.09	-0.02													
CHV	0.13	-0.08	0.10	-0.03	0.01	-0.04	0.01												
CRM	0.09	-0.02	-0.01	-0.05	0.02	-0.05	0.00	0.00											
CSCO	0.21	0.03	0.07	0.06	0.07	-0.02	-0.02	0.04	0.05										
DIS	0.13	-0.09	-0.05	-0.10	-0.07	-0.05	-0.06	-0.14	-0.02	-0.11									
GS	0.11	-0.11	0.02	0.01	0.06	0.02	-0.04	-0.02	0.07	-0.12	0.14								
HD	-0.02	-0.13	-0.06	-0.11	-0.03	-0.13	-0.12	-0.12	-0.08	-0.13	-0.01	-0.06							
IBM	0.02	-0.12	-0.03	-0.11	-0.03	-0.08	-0.00	0.00	-0.02	-0.12	0.04	0.02	0.00						
INTC	0.15	-0.03	0.15	-0.05	0.03	-0.02	-0.01	-0.03	-0.01	0.01	0.14	0.07	0.19	0.05					
JNJ	0.07	-0.07	-0.03	-0.05	-0.00	-0.05	-0.03	-0.07	-0.05	-0.13	-0.05	-0.09	0.03	-0.11	-0.18				
JPM	0.08	-0.09	0.03	0.04	0.09	-0.02	-0.07	0.02	0.01	-0.10	0.08	-0.03	0.04	-0.01	-0.00	0.09			
KO	0.11	-0.04	0.01	-0.06	-0.01	0.00	-0.06	-0.01	-0.01	-0.06	-0.01	-0.04	0.06	0.09	-0.04	0.09	0.01		
MCD	0.13	-0.10	-0.03	0.03	-0.01	-0.03	-0.09	-0.02	-0.02	0.06	-0.01	0.00	0.06	-0.03	-0.10	-0.03	0.04	-0.04	
MMM	0.04	-0.08	-0.04	-0.03	-0.01	-0.11	0.10	-0.01	0.03	-0.15	0.11	0.00	0.08	0.03	-0.14	0.08	0.02	-0.08	0.01

Table B.13: Estimates of $\eta_{1,2}$ for 20 Dow Jones 30 stocks