SFB 649 Discussion Paper 2011-086

Spectral estimation of covolatility from noisy observations using local weights

Markus Bibinger*
Markus Reiß*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

Spectral estimation of covolatility from noisy observations using local weights

Markus Bibinger[1] & Markus Reiß[1]

Institute of Mathematics, Humboldt-Universität zu Berlin

ABSTRACT. We propose localized spectral estimators for the quadratic covariation and the spot covolatility of diffusion processes which are observed discretely with additive observation noise. The eligibility of this approach to lead to an appropriate estimation for time-varying volatilities stems from an asymptotic equivalence of the underlying statistical model to a white noise model with correlation and volatility processes being constant over small intervals. The asymptotic equivalence of the continuous-time and the discrete-time experiments are proved by a construction with linear interpolation in one direction and local means for the other. The new estimator outperforms earlier nonparametric approaches in the considered model. We investigate its finite sample size characteristics in simulations and draw a comparison between the various proposed methods.

Key words: asymptotic equivalence, covariation, integrated covolatility, microstructure noise, spectral adaptive estimation

MSC Classification: 62M10, 62G05, 62G20, 91B84

JEL Classification: C14, C32, C58, G10

1 Introduction

The estimation of the quadratic (co-)variation of semimartingales is of large interest in statistics and financial econometrics. Especially, statistical models taking market microstructure frictions into account have attracted a lot of attention in recent years. Inspired by empirical studies of the characteristics of high-frequency financial data, a prominent approach is to describe asset prices as a superposition of a discretely sampled semimartingale with an independent additive noise component.
The finding that observations of a Brownian motion with noise on a discretely arranged grid possesses the LAN-property in Le Cam's sense with the rate $n^{-1/4}$ by Gloter & Jacod (2001), instead of the usual $n^{-1/2}$ rate in the absence of noise, has provided the optimal rate and a parametric efficiency bound for the asymptotic variance as a benchmark for this estimation problem. Interestingly, the nuisance quantity, namely the noise level, can be estimated with the usual faster rate in this model in contrast to the parameter of interest. This is caused by observation errors with non-decreasing variances perturbing diffusion increments of order $n^{-1/2}$. These features carry over to the estimation problem of covariation in a multidimensional setting as has been shown in Bibinger (2011a).
The key role of quantifying integrated (co-)volatilities in portfolio optimization and risk management has stimulated an increasing interest in estimation methods for these models starting with Aït-Sahalia *et al.* (2005) and Zhang *et al.* (2005). Subsequently three nonparametric approaches for integrated volatility estimation have been suggested, the multi-scale realized volatility by Zhang (2006), a pre-average strategy by Jacod *et al.* (2009) and the realized kernels from Barndorff-Nielsen *et al.* (2008). All estimators are based on quadratic forms of the observations and depend on a globally chosen tuning parameter. For that reason, when ignoring the treatment of end-effects, all three share a similar asymptotic behavior. They attain the optimal rate, but cannot be asymptotically efficient for time-varying volatility functions. Still, several robustness results to more realistic models incorporating non-i. i. d. noise and stochastic volatilities with leverage have been established and make these approaches quite attractive.

[1]Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 'Ökonomisches Risiko', Humboldt-Universität zu Berlin, is gratefully acknowledged.

An alternative approach for the estimation of the quadratic variation arising in Aït-Sahalia *et al.* (2005) from the parametric point of view is based on the MLE for this model. It turned out in Xiu (2010) that the MLE for integrated volatility can cope with a nonparametric volatility specification. This quasi-maximum likelihood estimator (QMLE) also attains the optimal rate. Asymptotic efficiency, however, is achieved only in the parametric setup with constant volatility. In Reiß (2011) an asymptotically efficient estimator based on spectral theory and localized MLEs for asymptotically shrinking blocks has been constructed. The idea stems from an asymptotic equivalence result in the spirit of Grama & Nussbaum (2002) pertaining the underlying nonparametric setting and a piecewise constant local parametric approximation. In Curci & Corsi (2011) a related estimation strategy using a discrete sine transform approach is considered and tested in an application study.
Ongoing progress in this research area has recently led to estimation approaches for the integrated covolatility in multidimensional models. The above-mentioned methods carry over to a multidimensional setting. Rate-optimal estimators, which also cope with asynchronous observations, have been established by Christensen *et al.* (2010), Aït-Sahalia *et al.* (2010) and Bibinger (2011a), while Barndorff-Nielsen *et al.* (2011) focusses on positive-definite (co)volatility matrix estimators.
The motivation and contribution of the article at hand is twofold. First, we step forward towards a deeper understanding of the statistical properties of covariation estimation from noisy discretely observed diffusions. In particular, we prove that observing two correlated diffusion processes with noise at synchronous times is asymptotically equivalent (in the sense of Le Cam's equivalence of statistical experiments) to observations in a related continuous time white noise model. The procedure is completely explicit and thus allows to transfer estimators and tests from one model to the other with the same asymptotic properties. In particular, for bounded loss functions asymptotic efficiency results are the same in both model sequences. The white noise model itself is asymptotically equivalent to a piecewise constructed parametric model. That result is an extension of the one-dimensional findings in Reiß (2011) and gives rise to our local spectral approach. The second contribution are our nonparametric spectral estimators of covolatility (SPECV) for both, the integrated covolatility (i.e. covariation) and the spot (i.e. instantaneous) covolatility. The estimators are based on certain empirical bivariate Fourier coefficients on each block in time which in the piecewise parametric white noise model are just independent Gaussian vectors in $\mathbb{R}^2$ with volatilities and covolatilities appearing in the covariance structures. This very simple structure allows a straight-forward analysis and often reduces the estimation variance compared to the previously suggested methods. This is corroborated by simulation results which show good finite-sample properties.
The article is arranged in three upcoming sections and an appendix comprising the technical proofs. Section 2 is devoted to the underlying statistical experiments and the asymptotic equivalence results. In Section 3, we develop the SPECV, spectral estimator of covolatility, and investigate its mathematical properties. A discussion and simulation study is provided in Section 4, where the SPECV of integrated covolatility is compared to concurrent nonparametric approaches. Owing to its local spectral construction principle, the new approach outperforms earlier methods if the correlation or volatility processes vary in time.

2 Asymptotic equivalence of the discrete regression-type and the continuous white noise experiment

Consider the statistical experiment in which a two-dimensional discrete time process $\tilde{\mathbf{Z}}$ defined by

$$(\mathcal{E}_0) \qquad \tilde{\mathbf{Z}}_{t_i^n} = \mathbf{Z}_{t_i^n} + \varepsilon_i,\ 0 \le i \le n \text{ with } \mathbf{Z}_t = \mathbf{Z}_0 + \int_0^t \Sigma_s^{1/2}\, d\mathbf{B}_s,\ t \in [0,1]$$

is observed, where $\mathbf{B}$ is a two-dimensional standard Brownian motion and

$$\Sigma_t = \begin{pmatrix} (\sigma_t^X)^2 & \rho_t \sigma_t^X \sigma_t^Y \\ \rho_t \sigma_t^X \sigma_t^Y & (\sigma_t^Y)^2 \end{pmatrix}$$

the (spot) volatility matrix. The signal part of $\tilde{\mathbf{Z}} = (\tilde{X}, \tilde{Y})^\top$ denoted $\mathbf{Z} = (X, Y)^\top$, which is called efficient price process in finance, is independent of the observation noise $\varepsilon = (\varepsilon^X, \varepsilon^Y)^\top$. The observation

errors (ε_i) are i. i. d. centred normal with covariance matrix

$$\mathbf{H} = \begin{pmatrix} \eta_X^2 & \eta_{XY} \\ \eta_{XY} & \eta_Y^2 \end{pmatrix} .$$

We consider time-varying volatility matrices Σ belonging to a Hölder ball of order $\alpha \in (0,1]$ and radius $R > 0$, i.e. $\Sigma \in C^\alpha(R)$ with

$$C^\alpha(R) = \{f \in C^\alpha([0,1], \mathbb{R}^{2\times 2}) | \|f\|_{C^\alpha} \leq R\} \text{ where } \|f\|_{C^\alpha} := \|f\|_\infty + \sup_{x\neq y} \frac{\|f(x) - f(y)\|}{|x-y|^\alpha} .$$

We denote the spectral norm in $\mathbb{R}^{2\times 2}$ always by $\|\cdot\|$ and define $\|f\|_\infty := \sup_{t\in[0,1]} \|f(t)\|$.
In $(\mathcal{E}_0)$ we allow for a non-equidistant synchronous observation scheme $(t_i^n)_{0\leq i\leq n}$, but we will have to impose that the sampling can be transferred to an equidistant scheme by a quantile transformation independent of n.

Assumption 1. *Suppose that there exists a differentiable distribution function $F : [0,1] \to [0,1]$ with $F(0) = 0$, $F(1) = 1$ and $F' > 0$, such that the observation times in $(\mathcal{E}_0)$ are generated by $t_i^n = F^{-1}(i/n)$, $i = 0, \ldots, n$.*

Note that we only consider deterministic designs of observation times. Under random sampling schemes the estimators should have similar properties, but the mathematical analysis is much harder.
We use a similar notation for the white noise experiment

$$(\mathcal{E}_1) \qquad d\tilde{\mathbf{Z}}_t = \mathbf{Z}_t\, dt + n^{-1/2}\mathbf{H}^{1/2}\, d\mathbf{W}_t\,, t \in [0,1]\,,$$

with the covariance matrix $\mathbf{H}$ of ε, $\mathbf{Z}_t = \mathbf{Z}_0 + \int_0^t \Sigma_s^{1/2}\, d\mathbf{B}_s$ and a standard two-dimensional Brownian motion $\mathbf{W}$ independent of $\mathbf{B}$.
In the following, we shall prove the results for an equidistant setting $t_i^n = i/n, i = 0, \ldots, n$. This is founded on the connection between a sampling scheme based on a quantile transformation of the equidistant grid and an equidistantly observed process with transformed volatility matrix by the identity in law

$$\mathbf{Z}_u^F := \mathbf{Z}_{F^{-1}(u)} = \int_0^u (\Sigma_s^F)^{1/2}\, d\mathbf{B}_s \text{ with } \Sigma_s^F = \Sigma_{F^{-1}(s)}(F^{-1})'(s),$$

which follows directly from the identity for covariance functions of these Gaussian processes via Itô isometry. Hence, upcoming results can be generalized for all F satisfying Assumption 1, replacing everywhere $\mathbf{Z}$ by $\mathbf{Z}^F$, t_i^n by i/n and Σ by Σ^F. Yet, the ease in dealing with transformations in the white noise model even gives another useful representation for non-equidistant design. Experiment $(\mathcal{E}_1)$ in terms of observing $\mathbf{Z}^F$ in noise is equivalent to observing

$$d\tilde{\mathbf{Z}}_{F(t)} = \mathbf{Z}_t F'(t)\, dt + n^{-1/2}\mathbf{H}^{1/2} F'(t)^{1/2}\, d\mathbf{W}_t\,, t \in [0,1],$$

see below for the exact notion of Le Cam equivalence which can be easily verified here by the identity of likelihood processes. Dividing by $F'(t)$ yields further equivalence with observing

$$d\overline{\mathbf{Z}}_t = \mathbf{Z}_t\, dt + (nF'(t))^{-1/2}\mathbf{H}^{1/2}\, d\mathbf{W}_t\,, t \in [0,1]. \tag{1}$$

As we shall establish next, experiments $(\mathcal{E}_0)$ and $(\mathcal{E}_1)$ will be asymptotically equivalent for $n \to \infty$ and the formulation (1) has a very intuitive meaning: the local noise level at t is proportional to $(nF'(t))^{-1/2}$, one over the square root of the local sample size $nF'(t)$.

Definition 1. *Let $\mathcal{E}_0(n, \alpha, R, \underline{\Sigma})$ with $n \in \mathbb{N}, \alpha \in (0,1], R, \underline{\Sigma} \geq 0$ be the statistical experiment generated by observations from $(\mathcal{E}_0)$ with $t_i^n = i/n$. The unknown parameter Σ in $(\mathcal{E}_0)$ belongs to the class $C^\alpha(R)$ and satisfies $\Sigma_t \geq \underline{\Sigma} E_2$ for all $t \in [0,1]$ with the identity matrix $E_2 \in \mathbb{R}^{2\times 2}$, i.e. the smallest eigenvalues of Σ_t are larger than $\underline{\Sigma}$.*

Analogously, let $\mathcal{E}_1(n, \alpha, R, \underline{\Sigma})$ be the statistical experiment generated by observing $(\mathcal{E}_1)$ with the same parameter class for Σ.

For the following results, we will briefly recall the notion of asymptotic equivalence, Le Cam deficiency and Le Cam distance. We refer interested readers to Le Cam & Yang (2000) for more information on the underlying theory. For statistical experiments $\mathcal{E}_0 = \left(\mathcal{X}_0, \mathcal{F}_0, \{\mathbb{P}_\theta^0 \,|\theta \in \Theta\}\right)$ and $\mathcal{E}_1 = \left(\mathcal{X}_1, \mathcal{F}_1, \{\mathbb{P}_\theta^1 \,|\theta \in \Theta\}\right)$ with the same parameter set Θ defined on (possibly) different Polish spaces, their Le Cam deficiency is defined by

$$\delta\left(\mathcal{E}_0, \mathcal{E}_1\right) = \inf_K \sup_{\theta\in\Theta} \|K\mathbb{P}_\theta^0 - \mathbb{P}_\theta^1\|_{\mathrm{TV}}\,,$$

where the infimum is taken over all Markov kernels (or randomisations) K from $(\mathcal{X}_0, \mathcal{F}_0)$ to $(\mathcal{X}_1, \mathcal{F}_1)$. The Le Cam distance is defined by

$$\Delta\left(\mathcal{E}_0, \mathcal{E}_1\right) = \max\left(\delta\left(\mathcal{E}_1, \mathcal{E}_0\right), \delta\left(\mathcal{E}_0, \mathcal{E}_1\right)\right).$$

If

$$\lim_{n\to\infty} \Delta\left(\mathcal{E}_0^n, \mathcal{E}_1^n\right) = 0$$

holds for sequences of experiments $\left(\mathcal{E}_0^n\right)_n$ and $\left(\mathcal{E}_1^n\right)_n$, then these sequences are called asymptotically equivalent.
The construction of the Markov kernel K will be explicit in all the proofs given in this article in terms of data transformations and randomisations.

Theorem 1. *The statistical experiments $\mathcal{E}_0(n, \alpha, R, \underline{\Sigma})$ and $\mathcal{E}_1(n, \alpha, R, \underline{\Sigma})$ are for any $\alpha > 0$ and $n \to \infty$ asymptotically equivalent. More precisely, the Le Cam distance is of order*

$$\Delta\left(\mathcal{E}_0, \mathcal{E}_1\right) = \mathcal{O}\left(Rn^{-(\alpha\wedge 1/2)}\underline{\mathbf{H}}^{-1}\right), \tag{2}$$

where $\underline{\mathbf{H}}$ denotes the smallest eigenvalue of $\mathbf{H}$.

We explicitly state how asymptotic terms hinge on $\underline{\mathbf{H}}$, since this is of interest when considering noise levels decreasing with n. A concise proof of this theorem is given in the appendix. The strategy of proof follows the same principle as for the one-dimensional setting in Reiß (2011). For the proof that $(\mathcal{E}_0)$ is at least as informative as $(\mathcal{E}_1)$, we construct a continuous time observation by linear interpolation. The interpolated process $\hat{\mathbf{Z}}$ is a centred Gaussian process on $[0,1]$. The associated covariance operator $\hat{C}$ on $L^2\left([0,1], \mathbb{R}^2\right)$ is such that the difference $(\bar{C} - \hat{C})$, where $\bar{C}$ is the covariance operator in a white noise model comprising the interpolated signal term, is positive (semi-) definite. For this reason observations from such a white noise model can be generated by adding an independent Gaussian noise component to $\hat{\mathbf{Z}}$. Now a process $\bar{\mathbf{Z}}$ from this white noise model and $\tilde{\mathbf{Z}}$ in $(\mathcal{E}_1)$ can be defined on the same probability space and it suffices to show that the total variation distance of the laws converges uniformly over Σ to zero. This is accomplished by bounding the squared Hellinger distance. For the proof of the intuitive converse, that $(\mathcal{E}_1)$ is at least as informative as $(\mathcal{E}_0)$, we take means symmetrically around the points (i/n), $1 \le i \le (n-1)$ from $(\mathcal{E}_1)$ and verify that the Hellinger distance between the processes generated in this manner and $\tilde{\mathbf{Z}}$ from $(\mathcal{E}_0)$ tends to zero.
An important setting in which the volatility processes follow again semimartingales is covered by Theorem 1 for the case that $\mathbf{Z}$ remains conditionally Gaussian.

Definition 2. *Write $\lfloor t\rfloor_h = \lfloor t/h\rfloor h$ for $h > 0$, assume $h^{-1}, nh \in \mathbb{N}$ and let $\mathbf{Z}_t^h = \mathbf{Z}_0 + \int_0^t \Sigma_{\lfloor s\rfloor_h}^{1/2}\, d\mathbf{B}_s$ with a two-dimensional standard Brownian motion $\mathbf{B}$. Let Σ_t belong to $C^\alpha(R)$ and satisfy $\Sigma_t \ge \underline{\Sigma} E_2$. Define the process*

$$d\tilde{\mathbf{Z}}_t = \mathbf{Z}_t^h\, dt + n^{-1/2}\mathbf{H}^{1/2}\, d\mathbf{W}_t,\ t \in [0,1], \tag{$\mathcal{E}_2$}$$

where $\mathbf{W}$ is a standard Brownian motion independent of $\mathbf{B}$. The statistical model generated by the observations from $(\mathcal{E}_2)$ is denoted by $\mathcal{E}_2(n, h, \alpha, R, \underline{\Sigma})$.

In experiment $(\mathcal{E}_2)$ we thus observe a process with a volatility matrix which is constant on each block $[kh,(k+1)h)$, $k=0,1,\dots,h^{-1}-1$. It is intuitive that for small block sizes h and sufficient Hölder regularity α this piecewise constant approximation is sufficiently close to render the approximation error statistically negligible. This is made precise in the following theorem.

Theorem 2. *Assume* $h^\alpha = \mathcal{O}\left(n^{-1/4}\right)$ *for* $1/2 < \alpha \leq 1$ *and* $\underline{\Sigma} > 0$. *Then the statistical experiments* $\mathcal{E}_1(n,\alpha,R,\underline{\Sigma})$ *and* $\mathcal{E}_2(n,h,\alpha,R,\underline{\Sigma})$ *are asymptotically equivalent:*

$$\Delta\left(\mathcal{E}_1,\mathcal{E}_2\right) = \mathcal{O}\left(Rh^\alpha \underline{\Sigma}^{-3/4}\underline{\mathbf{H}}^{-1/4}n^{1/4}\right)\,. \tag{3}$$

The asymptotic equivalence results lead to a new approach for the covariation estimation problem. Following the idea for the scalar case from Reiß (2011), we consider an orthonormal system in $L^2([0,1])$ of specific cosine functions with support on the blocks $[kh,(k+1)h]$ and frequencies of order $j \geq 1$. Their antiderivatives are sine functions on the same support and will also play a crucial role. We set

$$\varphi_{jk}(t) = \sqrt{\frac{2}{h}}\cos\left(j\pi h^{-1}\left(t-kh\right)\right)\mathbb{1}_{[kh,(k+1)h]}(t), \quad j \geq 1\,, k = 0,\dots,h^{-1}-1\,, \tag{4a}$$

$$\Phi_{jk}(t) = \left(\sqrt{2h}\,n\sin\left(\frac{j\pi}{2nh}\right)\right)^{-1}\sin\left(j\pi h^{-1}(t-kh)\right)\mathbb{1}_{[kh,(k+1)h]}(t),\; j \geq 1\,, k = 0,\dots,h^{-1}-1\,. \tag{4b}$$

Differently from Reiß (2011), we appropriately renormalize the antiderivatives (4b) to be equipped for the discrete analysis. The functions (4a) and (4b), evaluated on the grid given by the observation times, provide spectral weights for local blockwise averages. By virtue of the transformation for general observation schemes discussed above, we may for ease of exposition consider the equidistant grid:

$$\tilde{x}_{jk} = \sum_{l=1}^{n}\left(\tilde{X}_{\frac{l}{n}} - \tilde{X}_{\frac{l-1}{n}}\right)\Phi_{jk}\left(\frac{l}{n}\right), \tag{5a}$$

$$\tilde{y}_{jk} = \sum_{l=1}^{n}\left(\tilde{Y}_{\frac{l}{n}} - \tilde{Y}_{\frac{l-1}{n}}\right)\Phi_{jk}\left(\frac{l}{n}\right). \tag{5b}$$

Since $\Phi_{j(h^{-1}-1)}(1) = 0$, the last addend is zero for all blocks k. We stress that by the indicator functions in (4a) and (4b) and since $\Phi_{jk}(kh) = \Phi_{jk}((k+1)h) = 0$, the sums in (5a) and (5b) only extend over $l = k\cdot nh+1,\dots,(k+1)\cdot nh-1$. Therefore, families $(\tilde{x}_{jk},\tilde{y}_{jk})_j$ are uncorrelated and thus by Gaussianity independent for different blocks k. Besides the independence between blocks, we additionally benefit from the orthogonality of each family of functions associated with a specific period or frequency. The orthogonality relations $\int\varphi_{jk}\varphi_{ik} = 0$ and $\int\Phi_{jk}\Phi_{ik} = 0\;\forall i\neq j$ in $L^2([0,1])$ will remain valid for the discretized versions and the corresponding sums when $i,j\in\{1,\dots,nh\}$. For the purpose of explicitly analyzing the discrete terms, we introduce the notion of empirical scalar products:

$$\langle f,g\rangle_n := \frac{1}{n}\sum_{l=1}^{n} f\left(\frac{l}{n}\right)g\left(\frac{l}{n}\right) \text{ and } \|f\|_n^2 := \frac{1}{n}\sum_{l=1}^{n} f^2\left(\frac{l}{n}\right) = \langle f,f\rangle_n\,, \tag{6a}$$

$$[f,g]_n := \frac{1}{n}\sum_{l=1}^{n} f\left(\frac{l-\frac{1}{2}}{n}\right)g\left(\frac{l-\frac{1}{2}}{n}\right), \text{ for } f,g:[0,1]\to\mathbb{R}\,. \tag{6b}$$

By abuse of notation for a vector $Z=(Z_1,\dots,Z_n)$ and $f:[0,1]\to\mathbb{R}$, we will also write

$$\langle Z,f\rangle_n := \frac{1}{n}\sum_{l=1}^{n} Z_l\, f\left(\frac{l}{n}\right) \text{ and } [Z,f]_n := \frac{1}{n}\sum_{l=1}^{n} Z_l\, f\left(\frac{l-\frac{1}{2}}{n}\right)\,. \tag{6c}$$

For two vectors Z and $\tilde{Z}$ it is convenient to introduce the notation

$$\langle Z, \tilde{Z} \rangle_{nh;k} := \frac{1}{n} \sum_{l=1}^{n} Z_l \tilde{Z}_l \mathbb{1}_{[(kh,(k+1)h]} \left(\frac{l}{n} \right) = \frac{1}{n} \sum_{i=0}^{nh} Z_{knh+i} \tilde{Z}_{knh+i} \,. \tag{6d}$$

The following identity is a main ingredient in the construction of the estimator and for its error analysis below.

Proposition 1. *For the blockwise weighted sums $\tilde{x}_{jk}, \tilde{y}_{jk}, j \in \{1, \dots, nh\}, k \in \{0, \dots, h^{-1} - 1\}$, the following summation by parts formula holds true:*

$$\begin{aligned} \tilde{y}_{jk} &= \sum_{l=1}^{n} \Delta \tilde{Y}_l \, \Phi_{jk} \left(\frac{l}{n} \right) = - \sum_{l=0}^{n-1} \tilde{Y}_{\frac{l}{n}} \, \varphi_{jk} \left(\frac{l + \frac{1}{2}}{n} \right) \frac{1}{n} \\ &= \langle n \Delta Y, \Phi_{jk} \rangle_n - \left[\varepsilon^Y, \varphi_{jk} \right]_n \,, \end{aligned} \tag{7}$$

and for $\tilde{x}_{jk}$ analogously, where Δ denotes the backward difference operator $\Delta \tilde{Y}_l := \tilde{Y}_{\frac{l}{n}} - \tilde{Y}_{\frac{l-1}{n}}$ and $\Delta Y = \left(\Delta \tilde{Y}_1, \dots, \Delta \tilde{Y}_n \right)$. Moreover, we have the following orthogonality identities:

$$\left[\varphi_{jk}, \varphi_{rk} \right]_n = \delta_{jr} \,, \; j, r \in \{1, \dots, nh\} \,, k = 0, \dots, h^{-1} - 1 \,, \tag{8a}$$

$$\langle \Phi_{jk}, \Phi_{rk} \rangle_n = \| \Phi_{jk} \|_n^2 \, \delta_{jr} \,, \; j, r \in \{1, \dots, nh\} \,, k = 0, \dots, h^{-1} - 1 \,, \tag{8b}$$

where δ_{jr} is Kronecker's delta. The empirical norm

$$\| \Phi_{jk} \|_n^2 = \left(4n^2 \sin^2 \left(j\pi / (2nh) \right) \right)^{-1} , \; k \in \{0, \dots, h^{-1} - 1\}, \tag{9}$$

does not depend on the block k and appears in our estimator in the next section.

The two representations of the blockwise sums in (7) are very useful when disentangling the estimation error emerging from the two independent error sources: discretization and observation noise. In particular, we use the left-hand side which involves the increments of the processes only when considering the signal parts X and Y. For the analysis of cross terms and the pure noise parts the right-hand side of (7) permits a significant simplification. In the next section, we use these ideas and the insight into the structure of the estimation problem to construct a new estimation approach for the quadratic covariation and the spot covolatility of diffusion processes based on the original model $(\mathcal{E}_0)$. The final estimator for the quadratic covariation appears as a linear combination of the products of the local spectral averages $\tilde{x}_{jk} \tilde{y}_{jk}$ over all j and k combined with a bias correction. We will benefit from the asymptotic equivalence results for the mathematical analysis of our estimator by the following conclusion that we can straiten the analysis to the statistical experiment

$$(\mathcal{E}_3) \qquad \tilde{\mathbf{Z}}^h_{t_i^n} = \mathbf{Z}^h_{t_i^n} + \varepsilon_i, \; 0 \le i \le n \text{ with } \mathbf{Z}^h_t = \mathbf{Z}_0 + \int_0^t \Sigma^{1/2}_{\lfloor s \rfloor_h} \, d\mathbf{B}_s, \; t \in [0,1] \,,$$

where we have noisy discrete observations with the volatility matrix being constant on blocks.

Proposition 2. *For $nh \in \mathbb{N}$, $\alpha, R > 0$ and $\underline{\Sigma} \ge 0$ the statistical experiments $\mathcal{E}_2(n, h, \alpha, R, \underline{\Sigma})$ and $\mathcal{E}_3(n, h, \alpha, R, \underline{\Sigma})$ with $t_i^n = i/n$ are asymptotically equivalent:*

$$\Delta \left(\mathcal{E}_2, \mathcal{E}_3 \right) = \mathcal{O} \left(R \underline{\mathbf{H}}^{-1} n^{-1/2} \right) \,. \tag{10}$$

Consequently, observing $\tilde{\mathbf{Z}}$ in $(\mathcal{E}_0)$ is asymptotically equivalent to observations of $\tilde{\mathbf{Z}}^h$ from $(\mathcal{E}_3)$. Note that for constant Σ_{kh} on each block, the Φ_{jk} have the same structure as the eigenvectors of the covariance matrix associated with the vector of the $(nh - 1)$ observed increments on the block. The local weighted sums (5a) and (5b) on each block hence constitute the corresponding Karhunen-Loève expansion. We refer to Bibinger (2011a) and for the one-dimensional case to Gloter & Jacod (2001) and Curci & Corsi (2011) for the explicit computation of the eigenvalues and eigenvectors.

3 Local spectral estimation of covolatility

In the sequel, we always assume $h^\alpha = o\left(n^{-1/4}\right)$, Assumption 1 on the sampling scheme and that the volatility matrix belongs to $C^\alpha(R)$ for some $\alpha > 1/2$ and $R > 0$ and is bounded from below by a positive constant. By virtue of Proposition 2, we can then work within the simpler model $(\mathcal{E}_3)$. We present all results for the equidistant design $t_i^n = i/n$, noting again that the general case follows by substituting $\tilde{\mathbf{Z}}$ by $\tilde{\mathbf{Z}}^F$, Σ by Σ^F etc. Interestingly, integrated volatility is even invariant under this transformation:

$$\int_0^1 \Sigma_u^F \, du = \int_0^1 \Sigma_{F^{-1}(u)} (F^{-1})'(u) \, du = \int_0^1 \Sigma_t \, dt.$$

For estimation purposes this means that we can just neglect the design in the implementation. The invariance property, however, does not hold for powers of Σ or for polynomials in σ^X, σ^Y of degree different from two such that the asymptotic variance will significantly depend on the design function F.

On each of the independent blocks, we have observations (5a) and (5b) with

$$(\tilde{x}_{jk}, \tilde{y}_{jk}) \sim \mathbf{N}\left(\mathbf{0}, \begin{pmatrix} \eta_X^2/n + \|\Phi_{jk}\|_n^2 (\sigma_{kh}^X)^2 & \eta_{XY}/n + \|\Phi_{jk}\|_n^2 \rho_{kh} \sigma_{kh}^X \sigma_{kh}^Y \\ \eta_{XY}/n + \|\Phi_{jk}\|_n^2 \rho_{kh} \sigma_{kh}^X \sigma_{kh}^Y & \eta_Y^2/n + \|\Phi_{jk}\|_n^2 (\sigma_{kh}^Y)^2 \end{pmatrix}\right),$$

independently for all j, k, what can be proved by a standard calculation. We will postpone a detailed computation of estimation errors to the Appendix B. For each j, k fixed, the empirical covariance yields a natural estimator of the spot covolatility $\rho_{kh} \sigma_{kh}^X \sigma_{kh}^Y$ on each block provided we correct the bias by subtracting η_{XY}/n.

Remark 1. *In the following we assume for the ease of exposition that η_{XY} is known. Yet we can estimate η_{XY} from the observations with faster rate $\sqrt{n}$ by*

$$\widehat{\eta_{XY}} = \frac{1}{2n} \sum_{l=1}^{n} \left(\tilde{Y}_{\frac{l}{n}} - \tilde{Y}_{\frac{l-1}{n}}\right) \left(\tilde{X}_{\frac{l}{n}} - \tilde{X}_{\frac{l-1}{n}}\right) \tag{11}$$

or as well by $-n^{-1} \sum_l (\tilde{Y}_{l/n} - \tilde{Y}_{(l-1)/n})(\tilde{X}_{(l+1)/n} - \tilde{X}_{l/n})$. For the first estimator $\sqrt{n}$-consistency and a central limit theorem can be proved in the spirit of Zhang et al. (2005) for its one-dimensional counterpart $1/(2n) \sum_l (\tilde{X}_{l/n} - \tilde{X}_{(l-1)/n})^2$. The second estimator and its one-dimensional analogue $-n^{-1} \sum_l (\tilde{X}_{l/n} - \tilde{X}_{(l-1)/n})(\tilde{X}_{(l+1)/n} - \tilde{X}_{l/n})$ have a slightly bigger variance but the benefit of no finite sample bias due to the quadratic (co-)variation of the signal part.

By using just the lowest frequency $j = 1$ in each block, we obtain a simple rate-optimal estimator of integrated covolatility when summing over all blocks $[kh, (k+1)h]$ multiplied by the block length h:

$$\widehat{IC}^{(SPECV, j=1)} = h \sum_{k=0}^{h^{-1}-1} \|\Phi_{1k}\|_n^{-2} \left(\tilde{x}_{1k} \tilde{y}_{1k} - \eta_{XY}/n\right). \tag{12}$$

By independence between the blocks, its variance is of order $\mathcal{O}(h^{-3}(\eta_X^2/n + h^2)(\eta_Y^2/n + h^2))$. For fixed noise levels $\eta_X, \eta_Y, \eta_{XY}$ the rate-optimal choice $h \sim n^{-1/2}$ thus yields a variance of order $\mathcal{O}(n^{-1/2})$ (note that for $\alpha > 1/2$ and $h \sim n^{-1/2}$ the condition $h^\alpha = o(n^{-1/4})$ always holds).

It is possible to obtain a pointwise estimator of the spot covolatility $SCV_t := \rho_t \sigma_t^X \sigma_t^Y$ by the average of the spectral estimators over a set $\mathcal{K}_t$ of K adjacent blocks containing t:

$$\widehat{SCV}_t = K^{-1} \sum_{k \in \mathcal{K}_t} \|\Phi_{1k}\|_n^{-2} \left(\tilde{x}_{1k} \tilde{y}_{1k} - \eta_{XY}/n\right). \tag{13}$$

Since the observation times in $\mathcal{K}_t$ have at most distance Kh to t, the approximation error bound for the α-Hölder continuous function Σ yields a squared bias of order $\mathcal{O}((Kh)^{2\alpha})$. The variance is $\mathcal{O}(K^{-1})$ for $h \gtrsim n^{-1/2}$, and we obtain for the rate-optimal choices $h \sim n^{-1/2}$, $K \sim n^{\alpha/(2\alpha+1)}$ a root mean

squared error of order $\mathcal{O}(n^{-\alpha/(4\alpha+2)})$. Standard nonparametric techniques based on Gaussian measure concentration then even give the same rate times a log-factor in n for uniform loss in t, i.e.

$$\mathbb{E}\left[\sup_{t\in[0,1]}|\widehat{SCV}_t - SCV_t|\right] = \mathcal{O}\left((n/\log n)^{-\alpha/(4\alpha+2)}\right).$$

For estimation of the integrated covolatility we are not content with rate-optimality, but we also want to minimize the asymptotic variance. By independence we gain in efficiency by using on each block a convex combination of the estimators over all frequencies j. In order to estimate the integrated covolatility, we then just sum these estimators over all blocks. We end up with the following spectral estimation approach with local weights w_{jk}, satisfying $\sum_j w_{jk} = 1$:

$$\widehat{IC}_{w,n}^{(SPECV)} = \sum_{k=0}^{h^{-1}-1} h\sum_{j=1}^{nh} w_{jk}\|\Phi_{jk}\|_n^{-2}\left(\tilde{x}_{jk}\tilde{y}_{jk} - \eta_{XY}/n\right). \tag{14}$$

The optimal weights (minimizing the variance) depend on the unknown spot volatility matrix. As will be shown in the proof of Theorem 3, they are given by $w_{jk}^{oracle} = w_j(\Sigma_{kh})$ with

$$w_j(\Sigma) = \frac{\left(\|\Phi_{jk}\|_n^{-4}\,\frac{\eta_X^2\eta_Y^2+\eta_{XY}^2}{n^2} + (1+\rho^2)(\sigma^X\sigma^Y)^2 + \|\Phi_{jk}\|_n^{-2}\left((\sigma^X)^2\frac{\eta_Y^2}{n} + (\sigma^Y)^2\frac{\eta_X^2}{n}\right)\right)^{-1}}{\sum_{r=1}^{nh}\left(\|\Phi_{rk}\|_n^{-4}\,\frac{\eta_X^2\eta_Y^2+\eta_{XY}^2}{n^2} + (1+\rho^2)(\sigma^X\sigma^Y)^2 + \|\Phi_{rk}\|_n^{-2}\left((\sigma^X)^2\frac{\eta_Y^2}{n} + (\sigma^Y)^2\frac{\eta_X^2}{n}\right)\right)^{-1}}. \tag{15}$$

They give rise to the oracle version of our spectral estimator of covolatility (SPECV)

$$\widehat{IC}_{oracle,n}^{(SPECV)} = \sum_{k=0}^{h^{-1}-1} h\sum_{j=1}^{nh} w_j(\Sigma_{kh})\|\Phi_{jk}\|_n^{-2}\left(\tilde{x}_{jk}\tilde{y}_{jk} - \eta_{XY}/n\right). \tag{16}$$

Using adequate consistent pilot estimates, we obtain a feasible estimator which is asymptotically as efficient as the oracle estimator. Besides (13) we need the corresponding estimators for the spot volatilities $(\sigma_t^X)^2, (\sigma_t^Y)^2$:

$$\widehat{(\sigma_t^X)^2} = K^{-1}\sum_{k\in\mathcal{K}_t}\|\Phi_{1k}\|_n^{-2}\left(\tilde{x}_{1k}^2 - \eta_X^2/n\right), \quad \widehat{(\sigma_t^Y)^2} = K^{-1}\sum_{k\in\mathcal{K}_t}\|\Phi_{1k}\|_n^{-2}\left(\tilde{y}_{1k}^2 - \eta_Y^2/n\right), \tag{17}$$

which also satisfy

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left(\left|\widehat{(\sigma_t^X)^2} - (\sigma_t^X)^2\right| + \left|\widehat{(\sigma_t^Y)^2} - (\sigma_Y^X)^2\right|\right)\right] = \mathcal{O}\left((n/\log n)^{-\alpha/(4\alpha+2)}\right)$$

for $h \sim n^{-1/2}$, $K \sim n^{\alpha/(2\alpha+1)}$. In particular, all estimators are uniformly (in t) consistent provided the sample size tends to zero. By using just a negligible fraction of the data with sample size $m_n \to \infty$, $m_n = \mathrm{o}(n)$, we dispose of a uniformly consistent estimator $\widehat{\Sigma}_{t,n}$ of Σ_t which is independent from the SPECV estimator when the latter is based on the remaining $n - m_n = n(1 - \mathrm{o}(1))$ observations. This gives a concrete construction for the pilot estimator used in the following main theorem.

Theorem 3. *We observe from model $(\mathcal{E}_0)$ with $\Sigma \in C^\alpha(R), R > 0, \alpha > 1/2$ and $\underline{\Sigma} > 0$. Choose $h \sim n^{-1/2}\log(n)$. The resulting adaptive spectral estimator of covolatility (SPECV) for the integrated covolatility is*

$$\widehat{IC}_n^{(SPECV)} = \sum_{k=0}^{h^{-1}-1} h\sum_{j=1}^{nh} w_j\left(\widehat{\Sigma}_{kh,n}\right)\|\Phi_{jk}\|_n^{-2}\left(\tilde{x}_{jk}\tilde{y}_{jk} - \eta_{XY}/n\right) \tag{18}$$

with a pilot estimator $\widehat{\Sigma}_{t,n}$ of the spot covolatility matrix Σ_t inserted into the oracle weight formula (15)*. If $\widehat{\Sigma}_{t,n}$ is uniformly in $t \in [0,1]$ consistent and independent of the data used in $(\tilde{x}_{jk}, \tilde{y}_{jk})_{jk}$, then both the adaptive and the oracle SPECV estimator satisfy the same central limit theorem:*

$$n^{1/4}\left(\widehat{IC}^{(SPECV)}_{oracle,n} - \int_0^1 \rho_t \sigma_t^X \sigma_t^Y \, dt\right) \rightsquigarrow \mathbf{N}\left(0, (\eta_X^2\eta_Y^2 + \eta_{XY}^2)^{1/4} \int_0^1 \mathfrak{v}_s \, ds\right), \tag{19}$$

$$n^{1/4}\left(\widehat{IC}^{(SPECV)}_{n} - \int_0^1 \rho_t \sigma_t^X \sigma_t^Y \, dt\right) \rightsquigarrow \mathbf{N}\left(0, (\eta_X^2\eta_Y^2 + \eta_{XY}^2)^{1/4} \int_0^1 \mathfrak{v}_s \, ds\right), \tag{20}$$

with

$$\mathfrak{v}_t = \sqrt{2(A_t^2 - B_t)B_t}\left(\sqrt{A_t + \sqrt{A_t^2 - B_t}} - \operatorname{sgn}\left(A_t^2 - B_t\right)\sqrt{A_t - \sqrt{A_t^2 - B_t}}\right)^{-1} \tag{21}$$

and $A_t = 1/\sqrt{\eta_X^2\eta_Y^2 + \eta_{XY}^2}\left(\eta_Y^2\left(\sigma_t^X\right)^2 + \eta_X^2\left(\sigma_t^Y\right)^2\right)$ and $B_t = 4\left(\sigma_t^X\sigma_t^Y\right)^2(1+\rho_t^2)$.

The independence of the pilot estimator from the data used in the main estimator is assumed for technical reasons. It is believed that the result continues to hold without this assumption, which is also confirmed in simulations. In the Appendix B we learn that high spectral frequencies have decreasing weights and exceeding some threshold will asymptotically not contribute to the estimation. For practicable and tractable application of the SPECV it suffices to sum up frequencies in (18) only up to a spectral cut-off $J_n \ll nh$. We refer to Reiß (2011) for more information on the cut-off.

We give a complete overview on the estimation of the (co)volatility matrix here by recalling the according univariate estimator for the integrated volatilities:

$$\widehat{IV}^{(SPEV)}_n = \sum_{k=0}^{h^{-1}-1} h \sum_{j=1}^{nh} w_j^X\left(\hat{\sigma}^X_{kh}\right)\|\Phi_{jk}\|_n^{-2}\left(\tilde{x}_{jk}^2 - \eta_X^2/n\right), \tag{22}$$

which we call SPEV, with the oracle weights

$$w_j^X\left(\sigma^X\right) = \frac{\left(\|\Phi_{jk}\|_n^{-2}(\eta_X^2/n) + (\sigma^X)^2\right)^{-2}}{\sum_{l=1}^{nh}\left(\|\Phi_{lk}\|_n^{-2}(\eta_X^2/n) + (\sigma^X)^2\right)^{-2}}$$

and analogously for $\tilde{Y}$.

In general, the noise levels $\eta_X, \eta_Y, \eta_{XY}$ are unknown, but they can be estimated with faster rate $\sqrt{n}$ as mentioned above. A result with preestimated error covariance matrix $\widehat{\mathbf{H}}$ can be derived as for the preestimated Σ_{kh} above. Furthermore, it is of high practical interest to study how our covolatility estimator behaves under vanishing microstructure noise level, i.e. in the case $\mathbf{H} = 0$. In that case the oracle weights are all equal $w_{jk} = 1/(nh)$ and on each block we estimate the block covolatility by the sum

$$(nh)^{-1}\sum_{j=1}^{nh}\|\Phi_{jk}\|_n^{-2}\langle \Delta X, n\Phi_{jk}\rangle_n\langle \Delta Y, n\Phi_{jk}\rangle_n$$

of discrete Fourier coefficients with respect to $(\Phi_{jk})_{1\le j\le nh}$. By Parseval identity this sum is equal to $n\langle \Delta X, \Delta Y\rangle_{nh;k}$. In conclusion, in the case $\mathbf{H} = 0$ and for oracle weights our SPECV estimator reduces to the realized covolatility, which is the natural estimator in this situation.

Let us finally mention that the pilot estimators (17) and the estimator (22) slightly differ from the one in Reiß (2011) because we use the accurate Φ_{jk} for the discrete setup and their empirical norms defined above.

4 Discussion and simulations

As mentioned before, previously proposed nonparametric approaches have in common that they are quadratic forms of the observation vectors and when choosing corresponding weights or weight functions translate into each other and show accordant asymptotic properties. Nevertheless, each method is motivated

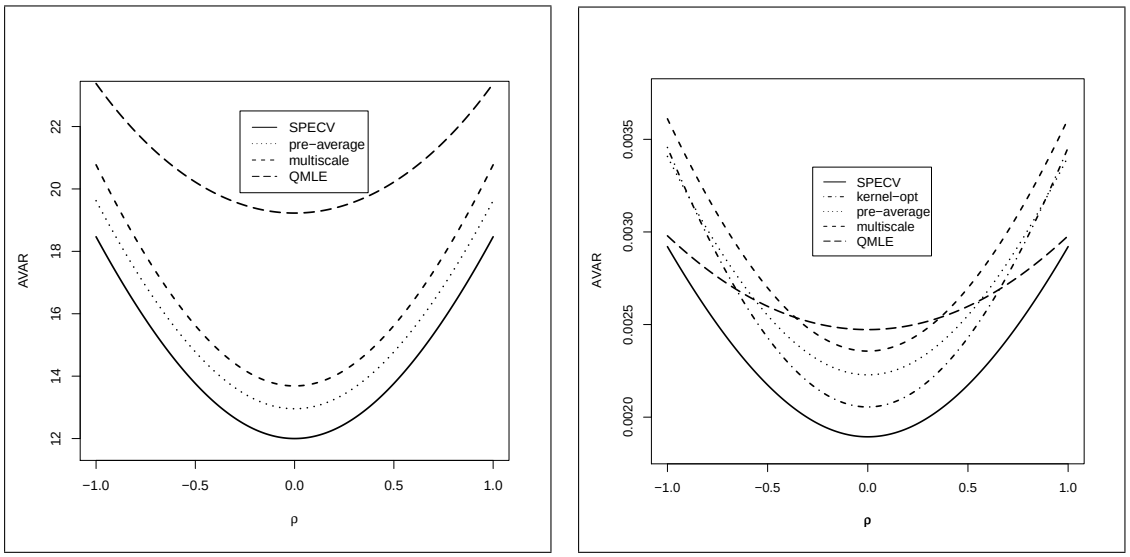

Figure 1: Asymptotic variances of estimators for the covolatility in specific scalar case (left) and with time varying volatilities (right).

from a slightly different point of view. The first two-scales realized volatility (TSRV) approach by Zhang *et al.* (2005) for the integrated volatility has been grounded on a subsampling method and a bias correction. Disregarding the bias correction, the subsampling estimator is the mean of lower frequent and hence less noise-sensitive realized volatilities. Zhang (2006) has extended this procedure to a linear combination using different time-scales (MSRV). The kernel approach by Barndorff-Nielsen *et al.* (2008) can be viewed as a linear combination of empirical autocovariances. Finally, the pre-average principle by Jacod *et al.* (2009) pursuant to its name incorporates (pre-) averaged weighted observations on blocks. The latter is closest to our methodology, but using Haar functions instead of (4b).

For all three the trade-off between the error due to noise and discretization is handled by choosing a global tuning parameter $c\sqrt{n}$, where c is a constant, minimizing the MSE to order $n^{-1/2}$. Thus, the optimal convergence rate is attained. If we neglect in support of these methods the possible asymptotic influence of end effects, they have an asymptotic variance structure $\mathfrak{N}c^{-3} + \mathfrak{D}c + \mathfrak{C}c^{-1}$, where the signal part $\mathfrak{D}$ depends in our notation on Σ, the noise part $\mathfrak{N}$ on $\mathbf{H}$ and the cross term $\mathfrak{C}$ on both. Minimization leads to $c = \left(\left(-\mathfrak{C} + \sqrt{\mathfrak{C}^2 + 12\mathfrak{N}\mathfrak{D}}\right)/6\mathfrak{N}\right)^{-1/2}$. The oracle solution is proportional to η^{-1} for equal noise variances η^2 of $\tilde{X}$ and $\tilde{Y}$. Interestingly, Barndorff-Nielsen *et al.* (2008) have succeeded in the univariate case with constant volatility in approximately attaining the lower bound from Gloter & Jacod (2001) by a clever selection method for their bandwidth and weights and also a feasible version with Tukey-Hanning kernels comes very close to that bound. Essentially, the main difference to our proposed approach is that we do not need to fix a tuning parameter and weights globally – but are able to adapt weights locally dependent on the observations only on each particular block.

We content ourselves with the findings in an idealized statistical model which gives insight into the fundamental structure of the estimation problem. Note, that an i. i. d. assumption on the noise and Hölder-continuity conditions on the volatility processes are customary in the strand of literature on nonparametric estimation methods. In our opinion, it is convenient to look at methods derived from a simple model and inspect the effect of misspecification on them. In the microstructure noise setup, we might first think of a diffusion with constant parameters. Xiu (2010) has taken a path in this vein with reviving the classical MLE in this framework and proving its robustness to a typical nonparametric setup. A local parametric approach is more flexible and increases in general the performance. More surprising than the accordance of asymptotic properties for the aforementioned three nonparametric methods, is that Xiu (2010) reports that the Quasi-MLE approach is in this sense asymptotically equivalent to the kernel approach as well. This is not the case for our SPEV/SPECV approach what underlines the originality of our local spectral

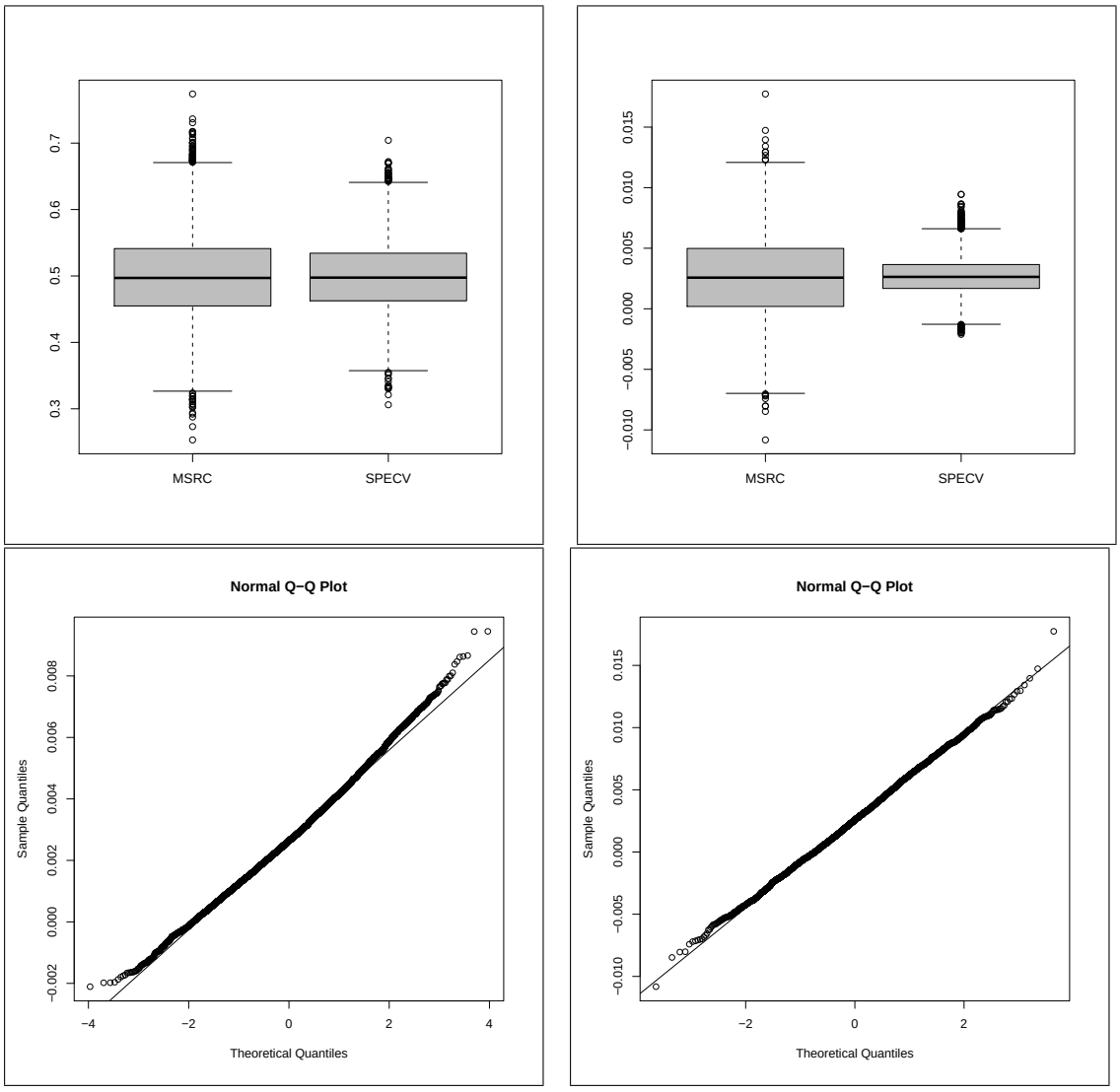

Figure 2: Boxplots for constant (left) and time varying (right) spot correlation and volatilities and normal QQ-Plots for MSRC and SPECV estimates in the time varying setting.

estimation method. Extensions of the theory that investigate the properties of the SPEV/SPECV in more general models, e.g. incorporating stochastic volatility and non-Gaussian errors, remain an open task for further research. For the moment the simple structure of SPEV/SPECV makes us confident that it can be robust to much more general model specifications.

Let us give concrete examples to compare the asymptotic variances of our SPECV and the other methods. For the simple parametric setting with constant $\sigma^X = 2$ and $\sigma^Y = 1$ and $\eta := \eta_X = \eta_Y = 1$, in Figure 1 we depict the asymptotic variances for $\rho \in (-1, 1)$ of the SPECV from Theorem 3, of the multi-scale realized covariance as deduced in Bibinger (2011b), of the pre-average estimator as given in Christensen *et al.* (2010) and of the QMLE from Aït-Sahalia *et al.* (2010), all with an optimal oracle tuning parameter selected as described above. The asymptotic variances are proportional to η, so that Figure 1 rescaled by η is meaningful for arbitrary noise levels. The SPECV has the smallest asymptotic variance and the QMLE the largest in this particular setup, all to the same optimal rate of convergence. The kernel method according to Barndorff-Nielsen *et al.* (2011) is not included, since the multivariate version has a non-optimal $n^{1/5}$-rate by oversmoothing to the benefit of positive semi-definiteness. We stress that we in-

tentionally have picked unequal constant volatilities here to the disadvantage of the QMLE which relies on the polarization identity. For equal volatilities $\sigma^X = \sigma^Y$ in our model $X + Y$ and $X - Y$ are independent and the polarized QMLE (also a polarized SPEV concurrent to the SPECV) will not suffer a disadvantage by polarization. From the comparison to the Fisher information in Bibinger (2011a), we can learn that the QMLE and the SPECV exhibit asymptotic efficiency for $\rho = 0$ in this setting. The approach presented in Barndorff-Nielsen *et al.* (2008) to derive asymptotic efficiency in the one-dimensional scalar case can easily be extended to a bivariate synchronous setting and renders a rate-optimal approach, asymptotically efficient for $\rho = 0$ as well. Yet none of these estimators is asymptotically efficient on the whole parameter space $(\rho, \sigma^X, \sigma^Y) \in ((0,1), \mathbb{R}_+, \mathbb{R}_+)$ and it is beyond the scope of this article to finish the quest for a globally asymptotically efficient estimator that outperforms the concurrent methods in each case.
Here we aim at providing with SPECV a method that performs well for time varying functions by local adaptivity and thus focus on that setting in the following. For this purpose we compare asymptotic variances in the same spirit for $\rho \in (-1, 1)$ and

$$\begin{aligned}\sigma_t^X &= 0.1 - 0.08 \cdot \sin(\pi t),\ t \in [0,1]\,,\\ \sigma_t^Y &= 0.15 - 0.07 \cdot \sin((6/7) \cdot \pi t),\ t \in [0,1]\,,\end{aligned}$$

which will as well be considered for the simulation part below. We add the theoretical asymptotic variance of a simple extension of the optimal kernel estimator for integrated volatility from Barndorff-Nielsen *et al.* (2008), which can be approximated by Tukey-Hanning kernels. This approach features the smallest asymptotic variance in a wide domain of ρ among the compared non-locally adaptive methods. Even so, the SPECV clearly comes below this benchmark. The right display of Figure 1 shows that the gains of SPECV compared to the previously proposed methods are much more distinctively than in the scalar case. After this theoretical comparison and the conclusion that the SPECV is preferable, especially in the general nonparametric setting, we shed light on the finite sample size behaviour of our approach in a Monte Carlo study.

In the first simulation, we compare the SPECV with the multiscale realized covariance (MSRC), both with an oracle choice of weights and tuning parameter, respectively. First, we implement a simple parametric model with $n = 30000$ equidistant observations of $\tilde{X}$ and $\tilde{Y}$, where $\sigma^X = \sigma^Y = 1, \rho = 1/2$ and noise levels $\eta_X = \eta_Y = 0.1$. The implemented MSRC as given in Bibinger (2011a) is for synchronous observations a direct extension of the MSRV by Zhang (2006) and translates asymptotically to the kernel estimator with a cubic kernel. It is known to have a good finite sample size behavior. We implement the SPECV with an adequate heuristic choice $h = 1/30$ such that $nh = 1000$.
The empirical distribution of the estimates from 10 000 MC iterations are visualized in a boxplot in Figure 2. The SPECV estimates have an empirical variance of $0.49 \cdot \sqrt{n}$ and the MSRC of $0.71 \cdot \sqrt{n}$. The empirical finding is that in this setting the SPECV is closer to its theoretical asymptotic variance of about 0.46 than the MSRC to its theoretical value of 0.52.
Our main focus will be the non-scalar case. For an example of deterministic time-varying functions, set

$$\begin{aligned}\sigma_t^X &= 0.1 - 0.08 \cdot \sin(\pi t),\ t \in [0,1]\,,\\ \sigma_t^Y &= 0.15 - 0.07 \cdot \sin((6/7) \cdot \pi t),\ t \in [0,1]\,,\\ \rho_t &= 0.5 + 0.01 \cdot \sin(\pi t),\ t \in [0,1]\,,\end{aligned}$$

where the volatilities are higher at the beginning and end of the observed interval and the correlation is only slowly varying, which mimics the basic realistic features. We keep the noise levels $\eta_X = \eta_Y = 0.1$ fixed and rather high compared to the signal part. The known integrated covolatility equals 0.00269 here. Since the noise level is high and dominates the signal part, the frequencies chosen according to the above given selection rule for the MSRC estimator become large (over 1 000) and the computing time increases for these kind of nonparametric estimators. As can be seen in the right boxplot of Figure 2, the SPECV outperforms the MSRC for non-constant volatilities and correlation more clearly. This confirms that the spectral local technique is more adequate to capture the effect of time-varying volatilities by local adaptation, not only theoretically but significantly in the finite sample case. The QQ-Plots in Figure 2 inspect the normal approximation for the two estimators from this Monte Carlo study in the time varying case.

Estimator	**RMSE**
$\widehat{IV}_n^{(SPEV)}$ for $\int_0^1 \left(\sigma_t^X\right)^2 \, dt$	0.0072
$\widehat{IV}_n^{(SPEV)}$ for $\int_0^1 \left(\sigma_t^Y\right)^2 \, dt$	0.0086
$\widehat{IC}_n^{(SPECV)}$ for $\int_0^1 \rho_t \sigma_t^X \sigma_t^X \, dt$	0.0034
$\widehat{IC}_{oracle}^{(MSRC)}$ for $\int_0^1 \rho_t \sigma_t^X \sigma_t^X \, dt$	0.0035
$\widehat{IC}_{oracle,n}^{(SPECV)}$ for $\int_0^1 \rho_t \sigma_t^X \sigma_t^X \, dt$	0.0015

Table 3: Comparison of root mean squared errors of MC (co-)volatility estimates.

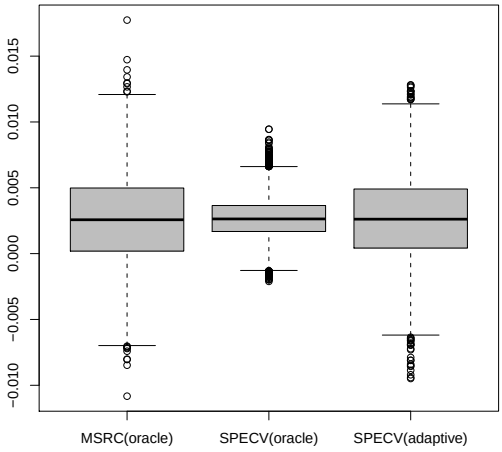

Figure 4: Boxplot of 10000 MC iterations of the oracle MSRC and oracle/adaptive SPECV.

We conclude the simulation study with an implementation of the adaptive SPEV/SPECV. We use pilot estimators (13) and (17) for Σ at times $l \cdot nh$, $l = 0, \ldots, 30$, with $K = 30$ adjacent blocks.
The 10 000 MC estimates of the adaptive SPECV are illustrated in Figure 4. Table 3 summarizes the root mean squared errors of all three adaptive SPEV/SPECV estimators and the oracle SPECV and the oracle MSRC. The performance of the adaptive version of SPECV can not keep up with the oracle version, but in our simulation it is still slightly better than the oracle MSRC. For an adaptive MSRC the root mean squared error will clearly become larger and we refer to Bibinger (2011b) for the method and simulation results pertaining this point.

A Appendix: Proofs of asymptotic equivalence

Proof of Theorem 1

We start with the constructive proof that $(\mathcal{E}_0)$ is at least as informative as $(\mathcal{E}_1)$. We use the linear B-splines

$$b_i(t) = \mathbb{1}_{[\frac{i-1}{n}\,,\,\frac{i+1}{n}]}(t) \min\left(1 + n\left(t - \frac{i}{n}\right)\,,\, 1 - n\left(t - \frac{i}{n}\right)\right),$$

i. e. $\operatorname{supp} b_i = [(i-1)/n, (i+1)/n]$, $b_i(i/n) = 1$, and b_i linear on $[(i-1)/n, i/n]$ and $[i/n, (i+1)/n]$. Consider the centred Gaussian process $\hat{\mathbf{Z}}$ defined by

$$\hat{\mathbf{Z}}_t = \sum_{i=1}^n \tilde{\mathbf{Z}}_i b_i(t) = \sum_{i=1}^n \mathbf{Z}_{\frac{i}{n}} b_i(t) + \sum_{i=1}^n \varepsilon_i b_i(t)\,. \tag{23}$$

The covariance function of $\hat{\mathbf{Z}}$ is

$$\mathbb{E}\left[\hat{\mathbf{Z}}_t \hat{\mathbf{Z}}_s^\top\right] = \sum_{i,j=1}^n \mathbf{A}\left(\frac{i \wedge j}{n}\right) b_i(t) b_j(s) + \mathbf{H} \sum_{i=1}^n b_i(t) b_i(s)$$

with

$$\mathbf{A}(t) := \int_0^t \Sigma_s\, ds = \begin{pmatrix} \int_0^t (\sigma_s^X)^2\, ds & \int_0^t \rho_s \sigma_s^X \sigma_s^Y\, ds \\ \int_0^t \rho_s \sigma_s^X \sigma_s^Y\, ds & \int_0^t (\sigma_s^Y)^2\, ds \end{pmatrix} \text{ and } \mathbf{H} = \begin{pmatrix} \eta_X^2 & \eta_{XY} \\ \eta_{XY} & \eta_Y^2 \end{pmatrix}.$$

For any $\mathbf{f} = (f_X, f_Y)^\top \in L^2\left([0,1], \mathbb{R}^2\right)$, we have

$$\begin{aligned}
&\mathbb{E}\left[\langle \mathbf{f}, \hat{\mathbf{Z}}\rangle^2\right] = \mathbb{E}\left[\left(\langle f_X, \hat{X}\rangle + \langle f_Y, \hat{Y}\rangle\right)^2\right] \\
&= \mathbb{E}\left[\langle f_X, \hat{X}\rangle^2\right] + \mathbb{E}\left[\langle f_Y, \hat{Y}\rangle^2\right] + 2\,\mathbb{E}\left[\langle f_X, \hat{X}\rangle \langle f_Y, \hat{Y}\rangle\right] \\
&= \sum_{i,j=1}^n A_{11}\left(\frac{i \wedge j}{n}\right) \langle f_X, b_i\rangle \langle f_X, b_j\rangle + \sum_{i=1}^n \eta_X^2 \langle f_X, b_i\rangle^2 \\
&+ \sum_{i,j=1}^n A_{22}\left(\frac{i \wedge j}{n}\right) \langle f_Y, b_i\rangle \langle f_Y, b_j\rangle + \sum_{i=1}^n \eta_Y^2 \langle f_Y, b_i\rangle^2 \\
&+ 2 \sum_{i,j=1}^n A_{12}\left(\frac{i \wedge j}{n}\right) \langle f_X, b_i\rangle \langle f_Y, b_j\rangle + 2 \sum_{i=1}^n \eta_{XY} \langle f_X, b_i\rangle \langle f_Y, b_i\rangle\,.
\end{aligned}$$

The sum of the three terms induced by the observation noise is bounded from above by $n^{-1}(\eta_X^2 \|f_X\|^2 + \eta_Y^2 \|f_Y\|^2 + 2\,\eta_{XY} \langle f_X, f_Y\rangle)$, since

$$\begin{aligned}
&\eta_X^2 \sum_{i=1}^n \langle f_X, b_i\rangle^2 + \eta_Y^2 \sum_{i=1}^n \langle f_Y, b_i\rangle^2 + 2\eta_{XY} \sum_{i=1}^n \langle f_X, b_i\rangle \langle f_Y, b_i\rangle \\
&= \left(\frac{1}{2} + \frac{\eta_{XY}}{2\eta_X \eta_Y}\right) \sum_{i=1}^n \langle \eta_X f_X + \eta_Y f_Y, b_i\rangle^2 + \left(\frac{1}{2} - \frac{\eta_{XY}}{2\eta_X \eta_Y}\right) \sum_{i=1}^n \langle \eta_X f_X - \eta_Y f_Y, b_i\rangle^2 \\
&\le \left(\frac{1}{2} + \frac{\eta_{XY}}{2\eta_X \eta_Y}\right) n^{-1} \|\eta_X f_X + \eta_Y f_Y\|^2 + \left(\frac{1}{2} - \frac{\eta_{XY}}{2\eta_X \eta_Y}\right) n^{-1} \|\eta_X f_X - \eta_Y f_Y\|^2 \\
&= n^{-1}\left(\eta_X^2 \|f_X\|^2 + \eta_Y^2 \|f_Y\|^2 + 2\,\eta_{XY} \langle f_X, f_Y\rangle\right).
\end{aligned}$$

For the upper bound we have used that $\int_0^1 n b_i(t)\, dt = 1$ implies $\langle f_X, n b_i\rangle^2 \le \langle f_X^2, n b_i\rangle$ by Jensen's inequality and $\sum_i b_i \le 1$ and analogously for the other terms. Now observe that $\mathbb{E}\left[\langle f, \mathbf{H}\, d\mathbf{W}\rangle\right] =$

$\mathbb{E}\left[\int f_t^\top \mathbf{H}\, d\mathbf{W}_t\right] = (\eta_X^2\|f_X\|^2 + \eta_Y^2\|f_Y\|^2 + 2\,\eta_{XY}\langle f_X, f_Y\rangle)$.
As a consequence, observations from $\bar{\mathbf{Z}}$ defined by

$$d\bar{\mathbf{Z}} = \sum_{i=1}^n \mathbf{Z}_{\frac{i}{n}} b_i(t)\, dt + \frac{1}{\sqrt{n}}\mathbf{H}^{1/2}\, d\mathbf{W}_t \tag{24}$$

with a two-dimensional standard Brownian motion $\mathbf{W}$ can be generated from $(\mathcal{E}_0)$ by adding additional $\mathbf{N}\left(0, \bar{C} - \hat{C}\right)$-noise, where $\hat{C} : L^2 \to L^2$ is the covariance operator of $\hat{Z}$ and the covariance operator $\bar{C} : L^2 \to L^2$ associated with (24) is given by

$$\bar{C}\mathbf{f}(t) = \sum_{i,j=1}^n \mathbf{A}\left(\frac{i\wedge j}{n}\right)\langle f, b_j\rangle b_i(t) + n^{-1}\mathbf{H}\mathbf{f}(t)\,, \mathbf{f} \in L^2\left([0,1], \mathbb{R}^2\right)\,.$$

Let C be the covariance operator

$$C\mathbf{f}(t) = \int_0^1 \left(\int_0^{t\wedge u} \mathbf{A}(s)\, ds\right)\mathbf{f}(u)\, du + n^{-1}\mathbf{H}\mathbf{f}(t)$$

from $(\mathcal{E}_1)$. In the following $\underline{\mathbf{H}}$ denotes the smallest eigenvalue of $\mathbf{H}$ as in Section 2. In the extension of the findings for the one-dimensional case, which has been treated in Section **A.2** in Reiß (2011), we make use of the convenient upper bound for the squared Hellinger distance between two normal measures by the squared Hilbert-Schmidt norm denoted $\|\cdot\|_{\mathrm{HS}}$. For a concise introduction on Hellinger distances between Gaussian measures and the Hilbert-Schmidt norm we refer to Section **A.1** in Reiß (2011).
The asymptotic equivalence of observing $\bar{\mathbf{Z}}$ and $\tilde{\mathbf{Z}}$ in $(\mathcal{E}_1)$ is ensured by the Hellinger distance bound

$$\begin{aligned}
\mathrm{H}^2\left(\mathcal{L}\left(\bar{\mathbf{Z}}\right), \mathcal{L}\left(\tilde{\mathbf{Z}}\right)\right) &\le 2\,\|C^{-1/2}\left(\bar{C} - C\right)C^{-1/2}\|_{\mathrm{HS}}^2 \\
&\le 2\,\underline{\mathbf{H}}^{-2}n^2\int_0^1\int_0^1 \left\|\mathbf{A}(t\wedge s) - \sum_{i,j=1}^n \mathbf{A}\left(\frac{i\wedge j}{n}\right) b_i(t)b_j(s)\right\|^2 dt\, ds \\
&= \mathcal{O}\left(\underline{\mathbf{H}}^{-2}R^2 n^{-(2\alpha\wedge 1)}\right) = o(1) \text{ for } \alpha > 0\,.
\end{aligned}$$

Note that we have estimated the L^2-distance between $\mathbf{A}(t\wedge s)$ and its coordinate-wise linear interpolation by $\mathcal{O}(n^{-1-\alpha})$ using a standard approximation result based on the fact that the function $(t,s) \mapsto \mathbf{A}(t\wedge s)$ lies in the class $C^{1+\alpha}$ away from the diagonal $\{t = s\}$ due to $\mathbf{A}'(t) = \Sigma_t \in C^\alpha$ and is Lipschitz at the diagonal (on the $n-1$ squares $[(i-1)/n, (i+1)/n]$ the pointwise bound $\mathcal{O}(n^{-1})$ only contributes $(n-1)\mathcal{O}(n^{-2}) = \mathcal{O}(n^{-1})$ to the squared L^2-distance).

The proof that $(\mathcal{E}_1)$ is at least as informative as $(\mathcal{E}_0)$ is obtained by a similar estimate and a generalization of the construction technique from the one-dimensional setting. For this purpose, set

$$\begin{aligned}
\mathbf{Z}'_i = n\int_{(2i-1)/2n}^{(2i+1)/2n} d\tilde{\mathbf{Z}}_t = n\int_{(2i-1)/2n}^{(2i+1)/2n} \mathbf{Z}_t\, dt + \varepsilon_i,\ 1 \le i \le (n-1), \\
\mathbf{Z}'_n = 2\,n\int_{(2n-1)/2n}^1 d\tilde{\mathbf{Z}}_t = 2\,n\int_{(2n-1)/2n}^1 \mathbf{Z}_t\, dt + \varepsilon_n,
\end{aligned}$$

with

$$\varepsilon_i = \sqrt{n}\int_{(2i-1)/2n}^{(2i+1)/2n} \mathbf{H}^{1/2}\, d\mathbf{W}_t \sim \mathbf{N}(0, \mathbf{H})\,.$$

The estimate that

$$\begin{aligned}
&\mathrm{H}^2\left(\mathcal{L}\left(\mathbf{Z}'_1, \ldots, \mathbf{Z}'_n\right), \mathcal{L}\left(\tilde{\mathbf{Z}}_1, \ldots, \tilde{\mathbf{Z}}_n\right)\right) \le 2\,\|\tilde{C}^{-1/2}\left(C' - \tilde{C}\right)\tilde{C}^{-1/2}\|_{\mathrm{HS}}^2 \\
&\le 2\,\underline{\mathbf{H}}^{-2}\|C' - \tilde{C}\|_{\mathrm{HS}}^2 = \mathcal{O}\left(\underline{\mathbf{H}}^{-2}R^2 n^{-2\alpha}\right)
\end{aligned}$$

establishes the result. Altogether, the Le Cam distance between the experiments $(\mathcal{E}_0)$ and $(\mathcal{E}_1)$ is of order $\mathcal{O}\left(\underline{\mathbf{H}}^{-1}Rn^{-\alpha}\right)$. Assuming that $\mathbf{A}$ is $(1+\alpha)$-Hölder continuous (α-Hölder regularity of the covolatility and volatilities), the asymptotic equivalence of the statistical experiments with discretely observed noisy diffusions and the continuous time white noise model is deduced. □

Proof of Theorem 2

The proof affiliates to the one-dimensional result and its proof in Section **A.3** of Reiß (2011). It is shown that the Hilbert-Schmidt norm of the difference between the experiment $(\mathcal{E}_1)$ and the one where Σ is evaluated at times $\lfloor t \rfloor_h := \min\{k\,h | k\,h \le t\}$, $1 \le k \le h^{-1} - 1$ tends to zero.
In the two-dimensional setting, we have a Hölder bound

$$\|\Sigma_t - \Sigma_{\lfloor t \rfloor_h}\|_\infty \le R h^\alpha\,,\ t \in [0,1]\,.$$

Denote C_Σ the covariance operator associated with the experiment $(\mathcal{E}_1)$ with volatility matrix Σ. For $f \in L^2\left([0,1], \mathbb{R}^2\right)$ let $F : [0,1] \to \mathbb{R}^2$ be the corresponding antiderivative with $F(1) = (0,0)^\top$. The difference of the two covariance operators of experiments with Σ and Σ^h where $\Sigma^h_t := \Sigma_{\lfloor t \rfloor_h}$, respectively, pertains only the signal part:

$$\langle (C_\Sigma - C_{\Sigma^h}) f, f\rangle = \int_0^1 F_t^\top \left(\Sigma_t - \Sigma^h_t\right) F_t\, dt \le \|\Sigma - \Sigma^h\|_\infty \langle \mathfrak{C} f, f\rangle$$

by partial integration, where $\mathfrak{C}$ denotes the covariance operator of a standard two-dimensional Brownian motion. We end up with the following upper bound for the Hilbert-Schmidt norm:

$$\begin{aligned}
\|C_\Sigma^{-1/2}(C_{\Sigma^h} - C_\Sigma)C_\Sigma^{-1/2}\|_{\text{HS}} &\le \|\Sigma - \Sigma^h\|_\infty \|C_\Sigma^{-1/2}\,\mathfrak{C}\,C_\Sigma^{-1/2}\|_{\text{HS}} \\
&\le \|\Sigma - \Sigma^h\|_\infty \left\| \left(\mathfrak{C}\underline{\Sigma} + \underline{\mathbf{H}} n^{-1}\,\mathrm{id}\right)^{-1/2} \mathfrak{C} \left(\mathfrak{C}\underline{\Sigma} + \underline{\mathbf{H}} n^{-1}\,\mathrm{id}\right)^{-1/2} \right\|_{\text{HS}} \\
&\le R h^\alpha \|G(\mathfrak{C})\|_{\text{HS}}\,.
\end{aligned}$$

The function $G(z) = z\left(z\underline{\Sigma} + \underline{\mathbf{H}} n^{-1}\right)^{-1}$ is applied to $\mathfrak{C}$ employing functional calculus.
The operator $\mathfrak{C}$ has the same spectral values as the covariance operator of a one-dimensional standard Brownian motion with double multiplicity. Hence, the result is derived directly from the spectral analysis for the one-dimensional case in Reiß (2011). □

A.1 Proof of Proposition 2

The proof follows exactly along the lines of proof for Theorem 1. The only difference is the bound on the $L^2([0,1]^2)$-distance between the functions $\mathbf{A}(t \wedge s)$ and $\sum_{i,j} \mathbf{A}((i \wedge j)/n) b_i(t) b_j(s)$. Since Σ is block-wise constant, $\mathbf{A}$ is linear on each interval $[(i-1)/n, i/n]$. By the linear interpolation property the two functions coincide on each square $[(i-1)/n, i/n] \times [(j-1)/n, j/n]$ for $i \ne j$. For the n squares where $i = j$, the Lipschitz property of $(t,s) \mapsto \mathbf{A}(t \wedge s)$ yields a total L^2-distance of order $n^{-3/2}$ (cf. again proof of Theorem 1) and the bound on the Le Cam distance follows.

B Appendix: Asymptotics of the local spectral (co-)volatility estimator

We start with the following standard formula for a bivariate normal distribution which will be used implicitly several times.

Lemma 1. *For a Gaussian random vector*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathbf{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}\right)$$

it holds true that

$$\mathbb{V}\mathrm{ar}(X^2Y^2) = (1+\rho^2)\sigma_X^2\sigma_Y^2 \ \textit{ and } \ \mathbb{V}\mathrm{ar}(X^2) = 2\sigma_X^4,\ \mathbb{V}\mathrm{ar}(Y^2) = 2\sigma_Y^4\,. \tag{25}$$

Proof. The nature of the Gaussian distribution allows us to write $Y = \rho(\sigma_Y/\sigma_X)X + \sqrt{1-\rho^2}\sigma_Y Z$ where $Z \sim \mathbf{N}(0,1)$ is independent of X. Since $\mathbb{E}\left[X^4\right] = 3\sigma_X^4$ and

$$\begin{aligned}\mathbb{E}\left[X^2Y^2\right] = \mathbb{E}\left[X^2\mathbb{E}\left[Y^2|X^2\right]\right] &= \mathbb{E}\left[X^2\left(\rho^2(\sigma_Y^2/\sigma_X^2)X^2 + (1-\rho^2)\sigma_Y^2\mathbb{E}\left[Z^2\right]\right)\right] \\ &= \rho^2(\sigma_Y^2/\sigma_X^2)\mathbb{E}\left[X^4\right] + (1-\rho^2)\mathbb{E}\left[X^2\right]\sigma_Y^2 = (1+2\rho^2)\sigma_X^2\sigma_Y^2\,,\end{aligned}$$

we directly conclude the statement of the Lemma. □

The elementary identities (25) and (7) are central tools in the error analysis for the SPECV-estimator. The latter is proved in the following.

Proof of Proposition 1

Equation (7) is basically an application of the discrete summation by parts analogue to the integration by parts formula, also called Abel transformation.
The elementary identity $\sin(x+h) - \sin x = 2\cos(x+h/2)\sin(h/2)$ yields:

$$\begin{aligned}\sum_{l=1}^{n}\Delta\tilde{Y}_l\Phi_{jk}\left(\frac{l}{n}\right) &= -\sum_{l=1}^{n-1}\tilde{Y}_{\frac{l}{n}}\left(\Phi_{jk}\left(\frac{l+1}{n}\right) - \Phi_{jk}\left(\frac{l}{n}\right)\right) + \tilde{Y}_1\Phi_{jk}(1) - \tilde{Y}_0\Phi_{jk}\left(n^{-1}\right) \\ &= -\sum_{l=0}^{n-1}\tilde{Y}_{\frac{l}{n}}\left(\Phi_{jk}\left(\frac{l+1}{n}\right) - \Phi_{jk}\left(\frac{l}{n}\right)\right) \\ &= -\sum_{l=0}^{n-1}\tilde{Y}_{\frac{l}{n}}\;\varphi_{jk}\left(\frac{l+\frac{1}{2}}{n}\right)\frac{1}{n}\,.\end{aligned}$$

The boundary terms vanish due to $\Phi_{jk}(0) = \Phi_{jk}(1) = 0$. Further simple relations for trigonometric functions reveal the orthogonality properties (8a) and (8b). Without loss of generality, consider the first block $k = 0$:

$$\begin{aligned}\left[\varphi_{j0},\varphi_{r0}\right]_n &= \frac{1}{n}\sum_{l=0}^{nh-1}\frac{2}{h}\cos\left(j\pi h^{-1}n^{-1}(l+1/2)\right)\cos\left(r\pi h^{-1}n^{-1}(l+1/2)\right) \\ &= \frac{1}{n}\sum_{l=0}^{nh-1}h^{-1}\left(\cos\left((j+r)\pi h^{-1}n^{-1}(l+1/2)\right) + \cos\left((j-r)\pi h^{-1}n^{-1}(l+1/2)\right)\right) \\ &= \delta_{jr}\,.\end{aligned}$$

The last equality holds since for arbitrary $m \in \mathbb{N}$:

$$\begin{aligned}\sum_{l=0}^{nh-1}\cos\left(\frac{m\pi}{hn}\left(l+\frac{1}{2}\right)\right) &= \sum_{l=0}^{\lfloor\frac{nh-1}{2}\rfloor}\sin\left(\pi\left(\frac{2l+1}{2}\frac{m}{hn}+\frac{1}{2}\right)\right) + \sum_{l=0}^{\lfloor\frac{nh-1}{2}\rfloor}\sin\left(\pi\left(\frac{1}{2}+m-\frac{2l+1}{2}\frac{m}{hn}\right)\right) \\ &= \sum_{l=0}^{\lfloor\frac{nh-1}{2}\rfloor}\sin\left(\pi\left(\frac{2l+1}{2}\frac{m}{hn}+\frac{1}{2}\right)\right) - \sum_{l=0}^{\lfloor\frac{nh-1}{2}\rfloor}\sin\left(\pi\left(\frac{1}{2}+\frac{2l+1}{2}\frac{m}{hn}\right)\right) = 0\,.\end{aligned}$$

Analogously we deduce that

$$\begin{aligned}\langle\Phi_{j0},\Phi_{r0}\rangle_n &= \frac{1}{n}\sum_{l=1}^{nh}\frac{\sin\left(j\pi h^{-1}n^{-1}l\right)\sin\left(r\pi h^{-1}n^{-1}l\right)}{2hn^2\sin\left(\frac{j\pi}{2nh}\right)\sin\left(\frac{r\pi}{2nh}\right)} \\ &= \frac{1}{n}\sum_{l=1}^{nh}\frac{\cos\left((j-r)\pi h^{-1}n^{-1}l\right) - \cos\left((j+r)\pi h^{-1}n^{-1}l\right)}{4hn^2\sin\left(\frac{j\pi}{2nh}\right)\sin\left(\frac{r\pi}{2nh}\right)} \\ &= \delta_{jr}\left(4n^2\sin^2\left(j\pi/(2nh)\right)\right)^{-1} = \delta_{jr}\|\Phi_{j0}\|_n^2\,.\end{aligned}$$

We conclude that the families of functions $(\varphi_{jk}), (\Phi_{jk})$ are orthogonal systems with respect to $[\cdot,\cdot]_n$ and $\langle\cdot,\cdot\rangle_n$, respectively. □

Proof of Theorem 3

Though we have exploited the well-known distribution characteristics for blockwise averages $(\tilde{x}_{jk}, \tilde{y}_{jk})$ directly in our simple Gaussian model in order to motivate the spectral estimator, for a better transparency and clarity, we give a detailed analysis for the asymptotic expectation and variance here. By Proposition 2 we can equivalently work with model $(\mathcal{E}_3)$, where the observations are generated by a blockwise constant spot volatility $\Sigma_{\lfloor t\rfloor}$.
Consider at first $\widehat{IC}^{(SPECV)}_{oracle,n}$ with known spot volatility matrix and correlation. We drop the superscript and subscripts in the following. The estimator is (asymptotically in $(\mathcal{E}_0)$) unbiased since

$$\begin{aligned}\mathbb{E}\left[\widehat{IC}\right] &= \sum_{k=0}^{h^{-1}-1} h \sum_{j=1}^{nh} \|\Phi_{jk}\|_n^{-2} w_{jk}\, \mathbb{E}\left[\langle \Delta X, n\Phi_{jk}\rangle_n \langle \Delta Y, n\Phi_{jk}\rangle_n + \left[\varepsilon^X, \varphi_{jk}\right]_n \left[\varepsilon^Y, \varphi_{jk}\right]_n - \frac{\eta_{XY}}{n}\right] \\ &= \sum_{k=0}^{h^{-1}-1} h \sum_{j=1}^{nh} \|\Phi_{jk}\|_n^{-2} w_{jk} \left(\mathbb{E}\left[n\langle \Delta X, \Delta Y\rangle_{nh;k}\right] \|\Phi_{jk}\|_n^2 + \frac{\eta_{XY}}{n}\left([\varphi_{jk},\varphi_{jk}]_n - 1\right)\right) \\ &= \sum_{k=0}^{h^{-1}-1} h\rho_{kh}\sigma^X_{kh}\sigma^Y_{kh} = \int_0^1 \rho_t \sigma_t^X \sigma_t^Y \, dt + o(1)\,,\end{aligned}$$

in view of Parseval identity, Itô isometry, the orthogonality relations (8a) and (8b) and $\sum_{j=1}^{nh} w_{jk} = 1$. The variance calculation is simplified by the independent block structure:

$$\mathbb{V}\mathrm{ar}\left(\widehat{IC}\right) = \sum_{k=0}^{h^{-1}-1} h^2 \mathbb{V}\mathrm{ar}\left(\sum_{l=1}^{nh} \|\Phi_{jk}\|_n^{-2} w_{jk}\tilde{x}_{jk}\tilde{y}_{jk}\right).$$

Consider the variance on the kth block. By the orthogonality relations (8a) and (8b) of the φ_{jk}'s and Φ_{jk}'s and application of (25) to Σ_{kh} and $\mathbf{H}$, the evaluation of the variance on the block yields

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\sum_{l=1}^{nh} \|\Phi_{jk}\|_n^{-2} w_{jk}\tilde{x}_{jk}\tilde{y}_{jk}\right) &= \sum_{l=1}^{nh} \|\Phi_{jk}\|_n^{-4} w_{jk}^2 \mathbb{V}\mathrm{ar}\left(\left[\varepsilon^X, \varphi_{jk}\right]_n \left[\varepsilon^Y, \varphi_{jk}\right]_n\right) \\ &\quad + \sum_{j=1}^{nh} \|\Phi_{jk}\|_n^{-4} w_{jk}^2 \left(\mathbb{E}\left[\langle \Delta X, n\Phi_{jk}\rangle_n^2 \langle \Delta Y, n\Phi_{jk}\rangle_n^2\right] - \left(\mathbb{E}\left[\langle \Delta X, n\Phi_{jk}\rangle_n \langle \Delta Y, n\Phi_{jk}\rangle_n\right]\right)^2\right) \\ &\quad + \sum_{j=1}^{nh} \|\Phi_{jk}\|_n^{-4} w_{jk}^2 \left(\mathbb{E}\left[\left[\varepsilon^X, \varphi_{jk}\right]_n^2 \langle \Delta Y, n\Phi_{jk}\rangle_n^2\right] + \mathbb{E}\left[\left[\varepsilon^Y, \varphi_{jk}\right]_n^2 \langle \Delta X, n\Phi_{jk}\rangle_n^2\right]\right) \\ &= \sum_{j=1}^{nh} \|\Phi_{jk}\|_n^{-4} w_{jk}^2 \left(\frac{\eta_X^2\eta_Y^2 + \eta_{XY}^2}{n^2}\left([\varphi_{jk},\varphi_{jk}]_n\right)^2 + \mathbb{E}\left[n^2(\langle \Delta X, \Delta Y\rangle_{nh;k})^2\right] \|\Phi_{jk}\|_n^4 \right. \\ &\quad \left. + \frac{\eta_X^2}{n}[\varphi_{jk},\varphi_{jk}]_n \,\mathbb{E}\left[\langle n\Delta Y, \Delta Y\rangle_{nh;k}\right] \|\Phi_{jk}\|_n^2 + \frac{\eta_Y^2}{n}[\varphi_{jk},\varphi_{jk}]_n \,\mathbb{E}\left[\langle n\Delta X, \Delta X\rangle_{nh;k}\right] \|\Phi_{jk}\|_n^2\right) \\ &= \sum_{j=1}^{nh} w_{jk}^2 \left(\|\Phi_{jk}\|_n^{-4} \frac{\eta_X^2\eta_Y^2 + \eta_{XY}^2}{n^2}[\varphi_{jk},\varphi_{jk}]_n^2 + (1+\rho_{kh}^2)(\sigma^X_{kh}\sigma^Y_{kh})^2 \right. \\ &\quad \left. + n^{-1}\left(\eta_X^2(\sigma^Y_{kh})^2 + \eta_Y^2(\sigma^X_{kh})^2\right)\|\Phi_{jk}\|_n^{-2}[\varphi_{jk},\varphi_{jk}]_n\right)\end{aligned}$$

$$= \sum_{j=1}^{nh} w_{jk}^2 \underbrace{\left(\|\Phi_{jk}\|_n^{-4} \, \frac{\eta_X^2 \eta_Y^2 + \eta_{XY}^2}{n^2} + (1+\rho_{kh}^2)(\sigma_{kh}^X \sigma_{kh}^Y)^2 + \|\Phi_{jk}\|_n^{-2} \left(\frac{(\sigma_{kh}^Y \eta_X)^2}{n} + \frac{(\sigma_{kh}^X \eta_Y)^2}{n} \right) \right)}_{=I_{jk}^{-1}}$$

We have used Itô isometry, the features of model $(\mathcal{E}_3)$ and Proposition 1. To increase the readability we introduce the shortcut I_{jk}. Selecting appropriate weights with $\sum_{j=1}^{nh} w_{jk} = 1$ on the blocks gives rise to an optimization problem with side condition. Minimizing the asymptotic variance yields oracle weights

$$w_{jk} = \frac{I_{jk}}{\sum_{l=1}^{nh} I_{lk}} \,. \tag{26}$$

Plugging in these weights the asymptotic variance on the kth block becomes $\sum_j I_{jk}^{-1} (I_{jk}^2 / (\sum_l I_{lk})^2) = (\sum_l I_{lk})^{-1}$. Next, consider

$$\frac{1}{\sqrt{n}h} \sum_{j=1}^{nh} I_{jk} = \frac{1}{\sqrt{n}h} \sum_{j=1}^{nh} \left(a + bn^2 \sin^4 \left(\frac{j\pi}{2nh} \right) + cn \sin^2 \left(\frac{j\pi}{2nh} \right) \right)^{-1} ,$$

with the shortcuts $a = (1+\rho_{kh}^2)(\sigma_{kh}^X \sigma_{kh}^Y)^2$, $b = 16(\eta_X^2 \eta_Y^2 + \eta_{XY}^2)$ and $c = 4\left((\eta_X \sigma_{kh}^Y)^2 + (\eta_Y \sigma_{kh}^X)^2\right)$. For $0 < \alpha < 3/8$, we obtain the bound

$$\begin{aligned} \frac{1}{\sqrt{n}h} \sum_{j=n^{5/8+\alpha}h}^{nh} I_{jk} &\leq \frac{1}{\sqrt{n}h} nh \left(a + \frac{b}{2} n^2 \frac{\pi^4 n^{20/8+4\alpha} h^4}{16 n^4 h^4} + \frac{c}{2} n \frac{\pi^2 n^{10/8+2\alpha} h^2}{4 n^2 h^2} \right)^{-1} \\ &= \sqrt{n} \left(a + \frac{b}{32} n^{1/2+4\alpha} + \frac{c}{8} n^{1/4+2\alpha} \right)^{-1} = o(1) \,, \end{aligned}$$

where we use that $\sin x \geq x/2$ on $(0,1)$, and further that by Taylor

$$\begin{aligned} \frac{1}{\sqrt{n}h} \sum_{j=1}^{n^{5/8+\alpha}h} I_{jk} &= \frac{1}{\sqrt{n}h} \sum_{j=1}^{n^{5/8+\alpha}h} \left(a + bn^2 \left(\frac{\pi^4 j^4}{16 n^4 h^4} + \mathcal{O}\left(j^6 n^{-6} h^{-6} \right) \right) + cn \left(\frac{j^2 \pi^2}{4 n^2 h^2} + \mathcal{O}\left(j^4 n^{-4} h^{-4} \right) \right) \right)^{-1} \\ &= \frac{1}{\sqrt{n}h} \sum_{j=1}^{n^{5/8+\alpha}h} \left(a + (b/16)\pi^4 (j/\sqrt{n}h)^4 + (c/4)\pi^2 (j/\sqrt{n}h)^2 \right)^{-1} + o(1) \,. \end{aligned}$$

This means that uniformly for all k, the high frequencies $j \gtrsim n^{5/8}$ do not contribute to the variance due to their decreasing weights and thus the sine functions may be approximated by the first order Taylor expansion. The overall variance is with $h_0 := h\sqrt{n} \left(\eta_X^2 \eta_Y^2 + \eta_{XY}^2 \right)^{-1/4}$

$$\mathrm{VAR}_n = \sum_{k=0}^{h^{-1}-1} h \frac{h_0}{\sqrt{n}} \left(\eta_X^2 \eta_Y^2 + \eta_{XY}^2 \right)^{1/4} \left(\sum_l I_{lk} \right)^{-1}$$

and hence, $\sqrt{n} \left(\eta_X^2 \eta_Y^2 + \eta_{XY}^2 \right)^{-1/4} \mathrm{VAR}_n$ and $n^{-1/2} h^{-1} \sum_l I_{lk}$ have the structure of Riemann sums. Because of $h_0 \to \infty$ we can replace j/h_0 by an integration variable z and we expect

$$n^{-1/2} h^{-1} \sum_{j=1}^{nh} I_{jk} \approx \int_0^{nh/h_0} \frac{1}{f_1(z)} dz \tag{27}$$

with

$$f_1(z) = f_1(\Sigma, \mathbf{H}; z) = \pi^4 z^4 + \pi^2 z^2 \left(\frac{\eta_Y^2 (\sigma^X)^2}{\sqrt{\eta_X^2 \eta_Y^2 + \eta_{XY}^2}} + \frac{\eta_X^2 (\sigma^Y)^2}{\sqrt{\eta_X^2 \eta_Y^2 + \eta_{XY}^2}} \right) + (1+\rho^2)(\sigma^X \sigma^Y)^2. \tag{28}$$

From $\int_0^\infty |(f_1^{-1})'(z)|\,dz < \infty$ and

$$\int_{j/h_0}^{(j+1)/h_0} |f_1(z)^{-1} - f_1(j/h_0)^{-1}|\,dz \le \int_{j/h_0}^{(j+1)/h_0}\int_{j/h_0}^{z} |(f_1^{-1})'(u)|\,du dz \le h_0^{-1}\int_{j/h_0}^{(j+1)/h_0} |(f_1^{-1})'(z)|\,dz$$

we infer that the approximation error in (27) is indeed of order $\mathcal{O}(h_0^{-1})$ and thus tends to zero. Notice that the upper integration limit tends to infinity by our choice of h.

The computation of the integral over the positive axis is provided in Proposition 3 below. The convergence of the total variance as given in Theorem 3, using $h \to 0$, follows in the same way.

Concerning the adaptive weights obtained from a preestimation step, we can work conditionally on them and by independence just assume to dispose of a deterministic sequence $\Sigma_{t,n}$ converging uniformly to Σ_t. By definition of the weights summing up to one the bias is (asymptotically) zero as in the oracle case. The variance on a block is correspondingly given by

$$V_k^{(n)} := \frac{\sum_{j=1}^{nh} (I_{jk}^{(n)})^2 I_{jk}^{-1}}{\left(\sum_{r=1}^{nh} I_{rk}^{(n)}\right)^2}$$

with $I_{jk}^{(n)}$ defined like I_{jk}, but in terms of the approximate values $\Sigma_{t,n}$ instead of Σ_t. In view of the order $n^{-1}\|\Phi_{jk}\|_n^{-2} \sim j/h$, compare (9), the asymptotically dominating term in I_{jk}^{-1} is $n^{-2}\|\Phi_{jk}\|_n^{-4}(\eta_X^2\eta_Y^2 + \eta_{XY}^2)$, which is independent of Σ_t. Together with the uniform convergence of $\Sigma_{t,n}$ this observation shows $|(I_{jk}^{(n)})^{-1} - I_{jk}^{-1}| = \mathcal{o}(I_{jk}^{-1})$ uniformly over all j and k. Consequently, we also have $|I_{jk}^{(n)} - I_{jk}| = \mathcal{o}(I_{jk})$ uniformly and we conclude uniformly in k

$$V_k^{(n)} = V_k(1 + \mathcal{o}(1)) \text{ with } V_k := \Big(\sum_{r=1}^{nh} I_{rk}\Big)^{-1},$$

the oracle variance over a block. Pursuing the same calculus as in the oracle case, the rescaled total variance converges to the same integral.

In the literature on nonparametric estimation methods for related and more general models, much mathematical effort is put in the proof of (stable) central limit theorems. For our Gaussian models the conclusion of asymptotic normality is direct. We can apply a standard i. i. d. triangular central limit theorem like Corollary 3. 1 from Hall & Heyde (1980), verifying a Lyapunov condition with fourth moments.

The following Proposition completes the proof of Theorem 3. For the computation of the Riemann integrals we use some concepts from complex analysis.

Proposition 3. *Consider the functions $f_1 : \mathbb{C} \to \mathbb{C}$ generalizing the function* (28) *and $f_2 : \mathbb{C} \to \mathbb{C}$ with*

$$f_2(z) := f_2(\rho, \sigma; z) = \pi^4 z^4 + 2\sigma^2\pi^2 z^2 + (1+\rho^2)\sigma^4 = f_1\left(\begin{pmatrix}\sigma & \rho\\ \rho & \sigma\end{pmatrix}, \begin{pmatrix}\eta & 0\\ 0 & \eta\end{pmatrix}; z\right),$$

which depend on parameters ρ and positive $\sigma^{\cdot}, \eta_{\cdot}$. For the improper integrals along the positive real line in the case $\sigma^{\cdot} > 0, \rho \neq 0$, the following identities hold true:

$$\int_0^\infty \frac{1}{f_2(x)}\,dx = \frac{1}{2\sigma^3\rho(1+\rho^2)^{1/4}}\,\sin\left(\frac{1}{2}\left(\operatorname{Arg}(\mathtt{i}-\rho) - \frac{\pi}{2}\right)\right), \tag{29a}$$

$$\int_0^\infty \frac{1}{f_1(x)}\,dx = \frac{1}{\sqrt{2}\sqrt{A^2-B}\sqrt{B}}\left(\sqrt{A+\sqrt{A^2-B}} - \operatorname{sgn}(A^2-B)\sqrt{A-\sqrt{A^2-B}}\right), \tag{29b}$$

where A and B are short expressions for the terms

$$A = \left(\frac{\eta_Y^2}{\sqrt{\eta_X^2\eta_Y^2+\eta_{XY}^2}}(\sigma^X)^2 + \frac{\eta_X^2}{\sqrt{\eta_X^2\eta_Y^2+\eta_{XY}^2}}(\sigma^Y)^2\right), \quad B = 4(\sigma^X\sigma^Y)^2(1+\rho^2),$$

$\mathrm{Arg}(z)$, *with* $\mathrm{Arg}(z) = \arctan(\mathrm{Im}(z)/\,\mathrm{Re}(z))$ *for* $\mathrm{Re}(z) > 0$, *denotes the argument of a complex number and we determine to take the root located in the upper half plane* $\mathbb{H} = \{z \in \mathbb{C} | \,\mathrm{Im}(z) > 0\}$ *in* (29b) *in the case that* $A^2 - B < 0$.
For $\rho = 0$ *and strictly positive* $\sigma^{\cdot}$ *(and* $\eta_{\cdot}$*), the integrals yield*

$$\int_0^\infty \frac{1}{f_2(x)}\,dx = \frac{1}{4\sigma^3}\,, \tag{30a}$$

$$\int_0^\infty \frac{1}{f_1(x)}\,dx = \frac{1}{2\eta_Y(\eta_X^2\eta_Y^2+\eta_{XY}^2)^{-1/4}(\sigma^X)^2\sigma^Y + 2\eta_X(\eta_X^2\eta_Y^2+\eta_{XY}^2)^{-1/4}(\sigma^Y)^2\sigma^X}\,. \tag{30b}$$

Proof. The meromorphic functions $f_1^{-1} : \mathbb{C} \to \mathbb{C}$, $f_2^{-1} : \mathbb{C} \to \mathbb{C}$ have four simple poles in the complex plane, since f_1, f_2 each has four simple non-real zeros. We can apply a specific version of the residue theorem (cf. Theorem 7.10 in Chapter III of Freitag & Busam (2005)) to evaluate the above improper real integrals. We restrict ourselves to the case $\rho \neq 0$ for which the solutions of (29a) and (29b) are not feasible using algebra programs or standard integral tables.
We first give the proof of (29a) for the simplified function f_2^{-1}. The zeros of f_2 are

$$z_{2;1,2,3,4} = \pm\frac{\exp(\mathrm{i}\,\pi/4)}{\pi}\sqrt{\mathrm{i}\pm\rho}\,\sigma$$

and are located symmetrically on a disk around the null in the complex plane. The residue theorem allows to calculate the integral $\int f_2^{-1}$ along the real line by the limit of a curve integral over a half-disk in the upper half plane. Since f_2 is even on the real line and $z_{2;1}$ and $z_{2;4}$ are the poles in the upper half plane, we obtain:

$$\begin{aligned}
\int_0^\infty \frac{1}{f_2(x)}\,dx &= \pi\mathrm{i}\left(\mathrm{Res}\left(f_2^{-1}; z_{2;1}\right) + \mathrm{Res}\left(f_2^{-1}; z_{2;4}\right)\right) \\
&= \pi\mathrm{i}\left(\left(4\exp(\mathrm{i}\pi/4)\pi\sqrt{\rho+\mathrm{i}}\,\sigma^3 + 4\exp(\mathrm{i}\,3\pi/4)\pi(\rho+\mathrm{i})^{3/2}\sigma^3\right)^{-1}\right. \\
&\quad \left. -\left(4\exp(\mathrm{i}\pi/4)\pi\sqrt{\mathrm{i}-\rho}\,\sigma^3 + 4\exp(\mathrm{i}\,3\pi/4)\pi(\mathrm{i}-\rho)^{3/2}\sigma^3\right)^{-1}\right) \\
&= \frac{1}{4\sigma^3\rho}(-1)^{3/4}\left((\mathrm{i}-\rho)^{-1/2} - (\mathrm{i}+\rho)^{-1/2}\right) \\
&= \left(2\sigma^3\rho(\rho^2+1)^{1/4}\right)^{-1}\sin\left(\frac{1}{2}\left(\mathrm{Arg}(\mathrm{i}-\rho) - \frac{\pi}{2}\right)\right).
\end{aligned}$$

In this proof we always use the unique square root in the upper half plane of complex numbers (and the usual definition for real numbers).
The analysis for the general case f_1 is a bit more involved, since depending on the parameters $\rho, \sigma^{\cdot}$ and the ratios $\eta_X^2/\sqrt{\eta_X^2\eta_Y^2+\eta_{XY}^2}$, $\eta_Y^2/\sqrt{\eta_X^2\eta_Y^2+\eta_{XY}^2}$ for the zeros of f_1:

$$z_{1;1,2,3,4} = \pm\frac{1}{\sqrt{2}\,\pi}\sqrt{-A\pm\sqrt{A^2-B}}\,,$$

it holds true that either $z_{1;1}$("++") and $z_{1;4}$("- -") or $z_{1;1}$("++") and $z_{1;2}$("+ -") are located in the upper half plane. This role change dependent on whether $A^2 - B$ is positive or negative is illustrated in Figure 5, in which the interesting factor appearing in the solution of the integral is depicted for a possible range of values for A and B in a certain codomain of $\rho, \sigma^X, \sigma^Y, \eta_X/\eta_Y$ for $\eta_{XY} = 0$. Using the above convention for square roots, the left-hand side of (29b) yields

$$\begin{aligned}
\int_0^\infty f_1^{-1}(x)\,dx &= \frac{\mathrm{i}}{\sqrt{2}}\frac{1}{\sqrt{A^2-B}}\left((-A+\sqrt{A^2-B})^{-1/2} \pm (-A-\sqrt{A^2-B})^{-1/2}\right) \\
&= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{B-A^2}}\frac{1}{\sqrt{B}}\left(\sqrt{-A-\mathrm{i}\sqrt{B-A^2}} + \mathrm{sgn}(B-A^2)\sqrt{-A+\mathrm{i}\sqrt{B-A^2}}\right) \\
&= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{A^2-B}}\frac{1}{\sqrt{B}}\left(\sqrt{A+\sqrt{A^2-B}} - \mathrm{sgn}(A^2-B)\sqrt{A-\sqrt{A^2-B}}\right).
\end{aligned}$$

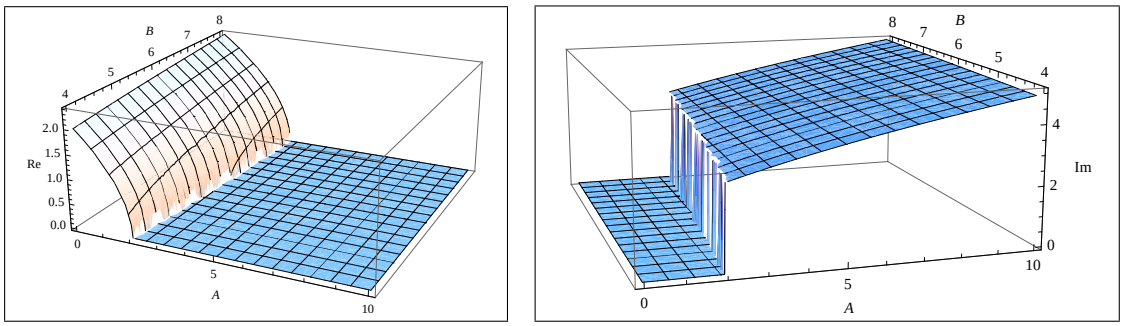

Figure 5: Real and imaginary part of $\sqrt{-A-\mathrm{i}\sqrt{B-A^2}}+\sqrt{-A+\mathrm{i}\sqrt{B-A^2}}$.

In the first line "$\pm$" indicates that depending on the parameters there are two different solutions. As visualized in Figure 5, the right factor in the second line is purely real if $B-A^2>0$ and purely imaginary if $B-A^2<0$. The expressions in the second and third line hence give the positive real solution in each case. In the case $B-A^2>0$, we can write the solution $\sqrt{2}(B-A^2)^{-1/2}B^{-1/4}\cos\left(\frac{1}{2}\operatorname{Arg}(A+\mathrm{i}\sqrt{B-A^2})\right)$ similiarly to (29a) above. □

References

Aït-Sahalia, Y., J. Fan, & D. Xiu (2010). High-frequency estimates with noisy and asynchronous financial data. *Journal of the American Statistical Association 105*(492), 1504–1516.

Aït-Sahalia, Y., L. Zhang, & P. A. Mykland (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies 18*, 351–416.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, & N. Shephard (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, & N. Shephard (2011). Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *Journal of Econometrics 162*(2), 149–169.

Bibinger, M. (2011a). Efficient covariance estimation for asynchronous noisy high-frequency data. *Scandinavian Journal of Statistics 38*, 23–45.

Bibinger, M. (2011b). An estimator for the quadratic covariation of asynchronously observed itô processes with noise: Asymptotic distribution theory. *preprint version, Humboldt-Universität zu Berlin*, URL=`sfb649.wiwi.hu--berlin.de/papers/pdf/SFB649DP2011--034.pdf`.

Christensen, K., S. Kinnebrock, & M. Podolskij (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics, 159*(1), 116–133.

Curci, G. & F. Corsi (2011). Discrete sine transform for multi-scales realized volatility measures. *Quantitative Finance forthcoming*.

Freitag, E. & R. Busam (2005). *Complex Analysis 1* (4 ed.). Springer, Heidelberg.

Gloter, A. & J. Jacod (2001). Diffusions with measurement errors 1 and 2. *ESAIM Probability and Statistics 5*, 225–242.

Grama, I. & M. Nussbaum (2002). Asymptotic equivalence for nonparametric regression. *Mathematical Methods of Statistics 11*, 1–36.

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

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, & M. Vetter (2009). Microstructure noise in the continous case: the pre-averaging approach. *Stochastic Processes and their Applications 119*, 2803–2831.

Le Cam, L. & L. G. Yang (2000). *Asymptotics in Statistics. Some basic concepts* (2 ed.). Springer, New York.

Reiß, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Annals of Statistics* (2), 772–802.

Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics 159*, 235–250.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli 12*(6), 1019–1043.

Zhang, L., P. A. Mykland, & Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

Markus Bibinger, Institute of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, E-mail: bibinger@math.hu-berlin.de
Markus Reiß, Institute of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, E-mail: mreiss@math.hu-berlin.de

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Localising temperature risk" by Wolfgang Karl Härdle, Brenda López Cabrera, Ostap Okhrin and Weining Wang, January 2011.
002 "A Confidence Corridor for Sparse Longitudinal Data Curves" by Shuzhuan Zheng, Lijian Yang and Wolfgang Karl Härdle, January 2011.
003 "Mean Volatility Regressions" by Lu Lin, Feng Li, Lixing Zhu and Wolfgang Karl Härdle, January 2011.
004 "A Confidence Corridor for Expectile Functions" by Esra Akdeniz Duran, Mengmeng Guo and Wolfgang Karl Härdle, January 2011.
005 "Local Quantile Regression" by Wolfgang Karl Härdle, Vladimir Spokoiny and Weining Wang, January 2011.
006 "Sticky Information and Determinacy" by Alexander Meyer-Gohde, January 2011.
007 "Mean-Variance Cointegration and the Expectations Hypothesis" by Till Strohsal and Enzo Weber, February 2011.
008 "Monetary Policy, Trend Inflation and Inflation Persistence" by Fang Yao, February 2011.
009 "Exclusion in the All-Pay Auction: An Experimental Investigation" by Dietmar Fehr and Julia Schmid, February 2011.
010 "Unwillingness to Pay for Privacy: A Field Experiment" by Alastair R. Beresford, Dorothea Kübler and Sören Preibusch, February 2011.
011 "Human Capital Formation on Skill-Specific Labor Markets" by Runli Xie, February 2011.
012 "A strategic mediator who is biased into the same direction as the expert can improve information transmission" by Lydia Mechtenberg and Johannes Münster, March 2011.
013 "Spatial Risk Premium on Weather Derivatives and Hedging Weather Exposure in Electricity" by Wolfgang Karl Härdle and Maria Osipenko, March 2011.
014 "Difference based Ridge and Liu type Estimators in Semiparametric Regression Models" by Esra Akdeniz Duran, Wolfgang Karl Härdle and Maria Osipenko, March 2011.
015 "Short-Term Herding of Institutional Traders: New Evidence from the German Stock Market" by Stephanie Kremer and Dieter Nautz, March 2011.
016 "Oracally Efficient Two-Step Estimation of Generalized Additive Model" by Rong Liu, Lijian Yang and Wolfgang Karl Härdle, March 2011.
017 "The Law of Attraction: Bilateral Search and Horizontal Heterogeneity" by Dirk Hofmann and Salmai Qari, March 2011.
018 "Can crop yield risk be globally diversified?" by Xiaoliang Liu, Wei Xu and Martin Odening, March 2011.
019 "What Drives the Relationship Between Inflation and Price Dispersion? Market Power vs. Price Rigidity" by Sascha Becker, March 2011.
020 "How Computational Statistics Became the Backbone of Modern Data Science" by James E. Gentle, Wolfgang Härdle and Yuichi Mori, May 2011.
021 "Customer Reactions in Out-of-Stock Situations – Do promotion-induced phantom positions alleviate the similarity substitution hypothesis?" by Jana Luisa Diels and Nicole Wiebach, May 2011.

SFB 649, Spandauer Str. 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

022 "Extreme value models in a conditional duration intensity framework" by Rodrigo Herrera and Bernhard Schipp, May 2011.
023 "Forecasting Corporate Distress in the Asian and Pacific Region" by Russ Moro, Wolfgang Härdle, Saeideh Aliakbari and Linda Hoffmann, May 2011.
024 "Identifying the Effect of Temporal Work Flexibility on Parental Time with Children" by Juliane Scheffel, May 2011.
025 "How do Unusual Working Schedules Affect Social Life?" by Juliane Scheffel, May 2011.
026 "Compensation of Unusual Working Schedules" by Juliane Scheffel, May 2011.
027 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappus and Markus Reiß, May 2011.
028 "Asymptotic equivalence and sufficiency for volatility estimation under microstructure noise" by Markus Reiß, May 2011.
029 "Pointwise adaptive estimation for quantile regression" by Markus Reiß, Yves Rozenholc and Charles A. Cuenod, May 2011.
030 "Developing web-based tools for the teaching of statistics: Our Wikis and the German Wikipedia" by Sigbert Klinke, May 2011.
031 "What Explains the German Labor Market Miracle in the Great Recession?" by Michael C. Burda and Jennifer Hunt, June 2011.
032 "The information content of central bank interest rate projections: Evidence from New Zealand" by Gunda-Alexandra Detmers and Dieter Nautz, June 2011.
033 "Asymptotics of Asynchronicity" by Markus Bibinger, June 2011.
034 "An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory" by Markus Bibinger, June 2011.
035 "The economics of TARGET2 balances" by Ulrich Bindseil and Philipp Johann König, June 2011.
036 "An Indicator for National Systems of Innovation - Methodology and Application to 17 Industrialized Countries" by Heike Belitz, Marius Clemens, Christian von Hirschhausen, Jens Schmidt-Ehmcke, Axel Werwatz and Petra Zloczysti, June 2011.
037 "Neurobiology of value integration: When value impacts valuation" by Soyoung Q. Park, Thorsten Kahnt, Jörg Rieskamp and Hauke R. Heekeren, June 2011.
038 "The Neural Basis of Following Advice" by Guido Biele, Jörg Rieskamp, Lea K. Krugel and Hauke R. Heekeren, June 2011.
039 "The Persistence of "Bad" Precedents and the Need for Communication: A Coordination Experiment" by Dietmar Fehr, June 2011.
040 "News-driven Business Cycles in SVARs" by Patrick Bunk, July 2011.
041 "The Basel III framework for liquidity standards and monetary policy implementation" by Ulrich Bindseil and Jeroen Lamoot, July 2011.
042 "Pollution permits, Strategic Trading and Dynamic Technology Adoption" by Santiago Moreno-Bromberg and Luca Taschini, July 2011.
043 "CRRA Utility Maximization under Risk Constraints" by Santiago Moreno-Bromberg, Traian A. Pirvu and Anthony Réveillac, July 2011.

SFB 649, Spandauer Str. 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

044 "Predicting Bid-Ask Spreads Using Long Memory Autoregressive Conditional Poisson Models" by Axel Groß-Klußmann and Nikolaus Hautsch, July 2011.
045 "Bayesian Networks and Sex-related Homicides" by Stephan Stahlschmidt, Helmut Tausendteufel and Wolfgang K. Härdle, July 2011.
046 "The Regulation of Interdependent Markets", by Raffaele Fiocco and Carlo Scarpa, July 2011.
047 "Bargaining and Collusion in a Regulatory Model", by Raffaele Fiocco and Mario Gilli, July 2011.
048 "Large Vector Auto Regressions", by Song Song and Peter J. Bickel, August 2011.
049 "Monetary Policy, Determinacy, and the Natural Rate Hypothesis", by Alexander Meyer-Gohde, August 2011.
050 "The impact of context and promotion on consumer responses and preferences in out-of-stock situations", by Nicole Wiebach and Jana L. Diels, August 2011.
051 "A Network Model of Financial System Resilience", by Kartik Anand, Prasanna Gai, Sujit Kapadia, Simon Brennan and Matthew Willison, August 2011.
052 "Rollover risk, network structure and systemic financial crises", by Kartik Anand, Prasanna Gai and Matteo Marsili, August 2011.
053 "When to Cross the Spread: Curve Following with Singular Control" by Felix Naujokat and Ulrich Horst, August 2011.
054 "TVICA - Time Varying Independent Component Analysis and Its Application to Financial Data" by Ray-Bing Chen, Ying Chen and Wolfgang K. Härdle, August 2011.
055 "Pricing Chinese rain: a multi-site multi-period equilibrium pricing model for rainfall derivatives" by Wolfgang K. Härdle and Maria Osipenko, August 2011.
056 "Limit Order Flow, Market Impact and Optimal Order Sizes: Evidence from NASDAQ TotalView-ITCH Data" by Nikolaus Hautsch and Ruihong Huang, August 2011.
057 "Optimal Display of Iceberg Orders" by Gökhan Cebiroğlu and Ulrich Horst, August 2011.
058 "Optimal liquidation in dark pools" by Peter Kratz and Torsten Schöneborn, September 2011.
059 "The Merit of High-Frequency Data in Portfolio Allocation" by Nikolaus Hautsch, Lada M. Kyj and Peter Malec, September 2011.
060 "On the Continuation of the Great Moderation: New evidence from G7 Countries" by Wenjuan Chen, September 2011.
061 "Forward-backward systems for expected utility maximization" by Ulrich Horst, Ying Hu, Peter Imkeller, Anthony Réveillac and Jianing Zhang.
062 "On heterogeneous latent class models with applications to the analysis of rating scores" by Aurélie Bertrand and Christian M. Hafner, October 2011.
063 "Multivariate Volatility Modeling of Electricity Futures" by Luc Bauwens, Christian Hafner and Diane Pierret, October 2011.

SFB 649, Spandauer Str. 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

064 "Semiparametric Estimation with Generated Covariates" by Enno Mammen, Christoph Rothe and Melanie Schienle, October 2011.
065 "Linking corporate reputation and shareholder value using the publication of reputation rankings" by Sven Tischer and Lutz Hildebrandt, October 2011.
066 "Monitoring, Information Technology and the Labor Share" by Dorothee Schneider, October 2011.
067 "Minimal Supersolutions of BSDEs with Lower Semicontinuous Generators" by Gregor Heyne, Michael Kupper and Christoph Mainberger, October 2011.
068 "Bargaining, Openness, and the Labor Share" by Dorothee Schneider, October 2011.
069 "The Labor Share: A Review of Theory and Evidence" by Dorothee Schneider, October 2011.
070 "The Power of Sunspots: An Experimental Analysis" by Dietmar Fehr, Frank Heinemann and Aniol Llorente-Saguer, October 2011.
071 "Econometric analysis of volatile art markets" by Fabian Y. R. P. Bocart and Christian M. Hafner, October 2011.
072 "Financial Network Systemic Risk Contributions" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, October 2011.
073 "Calibration of self-decomposable Lévy models" by Mathias Trabs, November 2011.
074 "Time-Varying Occupational Contents: An Additional Link between Occupational Task Profiles and Individual Wages" by Alexandra Fedorets, November 2011.
075 "Changes in Occupational Demand Structure and their Impact on Individual Wages" by Alexandra Fedorets, November 2011.
076 "Nonparametric Nonstationary Regression with Many Covariates" by Melanie Schienle, November 2011.
077 "Increasing Weather Risk: Fact or Fiction?" by Weining Wang, Ihtiyor Bobojonov, Wolfgang Karl Härdle and Martin Odening, November 2011.
078 "Spatially Adaptive Density Estimation by Localised Haar Projections" by Florian Gach, Richard Nickl and Vladimir Spokoiny, November 2011.
079 "Martingale approach in pricing and hedging European options under regime-switching" by Grigori N. Milstein and Vladimir Spokoiny, November 2011.
080 "Sparse Non Gaussian Component Analysis by Semidefinite Programming" by Elmar Diederichs, Anatoli Juditsky, Arkadi Nemirovski and Vladimir Spokoiny, November 2011.
081 "Parametric estimation. Finite sample theory" by Vladimir Spokoiny, November 2011.
082 "Continuous Equilibrium under Base Preferences and Attainable Initial Endowments" by Ulrich Horst, Michael Kupper, Andrea Macrinay and Christoph Mainberger, November 2011.
083 "Equilibrium Pricing in Incomplete Markets under Translation Invariant Preferences" by Patrick Cheridito, Ulrich Horst, Michael Kupper and Traian A. Pirvu, November 2011.

SFB 649, Spandauer Str. 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

084 "Competition and regulation in a differentiated good market" by Raffaele Fiocco, November 2011.
085 "Risk Patterns and Correlated Brain Activities. Multidimensional statistical analysis of fMRI data with application to risk patterns" by Alena Myšičková, Song Song, Piotr Majer, Peter N.C. Mohr, Hauke R. Heekeren and Wolfgang K. Härdle, December 2011.
086 "Spectral estimation of covolatility from noisy observations using local weights" by Markus Bibinger and Markus Reiß, December 2011.

SFB 649, Spandauer Str. 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".