Hybrid simulation scheme for volatility modulated moving average fields

Claudio Heinrich [*]   Mikko S. Pakkanen [†]   Almut E.D. Veraart[‡]

February 15, 2019

Abstract

We develop a simulation scheme for a class of spatial stochastic processes called volatility modulated moving averages. A characteristic feature of this model is that the behaviour of the moving average kernel at zero governs the roughness of realisations, whereas its behaviour away from zero determines the global properties of the process, such as long range dependence. Our simulation scheme takes this into account and approximates the moving average kernel by a power function around zero and by a step function elsewhere. For this type of approach, Bennedsen et al. [8], who considered an analogous model in one dimension, coined the term hybrid simulation scheme. We derive the asymptotic mean square error of the simulation scheme and compare it in a simulation study with several other simulation techniques and exemplify its favourable performance in a simulation study.

Key words: Simulation, random field, moving average, stochastic volatility, Matérn covariance.

[*]Department of Mathematics, Aarhus University & Norwegian Computing Center Oslo
  E-mail: claudio.heinrich@nr.no
[†]Department of Mathematics, Imperial College London,
  E-mail: m.pakkanen@imperial.ac.uk
[‡]Department of Mathematics, Imperial College London,
  E-mail: a.veraart@imperial.ac.uk

1 Introduction

In this article we develop a simulation scheme for real-valued random fields that we call *volatility modulated moving average* (VMMA) fields. A VMMA is defined by the formula

$$X_{\mathbf{t}} = \int_{\mathbb{R}^2} g(\mathbf{t} - \mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}), \tag{1.1}$$

where W is Gaussian white noise, $g \in L^2(\mathbb{R}^2)$ is a deterministic kernel, and σ is a random volatility field. This model has been used for the statistical modelling of spatial phenomena in various disciplines, examples being modelling of vegetation and nitrate deposition [25], of sea surface temperature [32] and of wheat yields [42].

We are interested in the case where the moving average kernel g has a singularity at zero. In this situation, the order of the singularity coincides with a smoothness parameter that specifies the Hausdorff dimension of the random field. Allowing for parametrisation of smoothness is necessary for a model to be useful in spatial prediction, see [37]. Our model exhibits Hausdorff dimension greater 2 and has therefore rough, non-differentiable realisations. Such models are used, for example, in surface modelling, where specific examples include seafloor morphology [19] or the surface modelling of celestial bodies [21].

A particular challenge in simulating volatility modulated moving averages lies in recovering the roughness of the field, while simultaneously capturing the global properties of the field, such as long range dependence. Our hybrid simulation scheme relies on approximating the kernel g by a power function in a small neighbourhood of zero, and by a step function away from zero. This approach allows us to reproduce the explosive behaviour at the origin, while simultaneously approximating the integrand on a large subset of $\mathbb{R}^2$. This idea

is motivated by the recent work of Bennedsen et al. [8], where the authors proposed an analogous scheme for the simulation of the one-dimensional model of Brownian semi-stationary processes. As a consequence, the hybrid simulation scheme preserves the roughness of the random field. The scheme is however limited to simulate the random field on a regular grid.

It is known that any stationary Gaussian random field with a continuous and integrable covariance function has a moving average representation of the form (1.1) with σ constant, cf. Hellmund et al. [24, Proposition 6]. This is for example satisfied for stationary Gaussian fields with Matérn covariance. Our assumptions correspond to a Matérn field with smoothness parameter $\nu \in (0, 1)$. This rough Matérn model has for example been used by Goff and Jordan [19] and in the context of turbulence modelling [39]. Recently, Bolin et al. [12] have developed a scheme for fast and efficient simulation of Gaussian fields with rough Matérn correlation, extending the SPDE approach developed by Lindgren et al. [30]. Our simulation scheme contains an alternative way to simulate such fields as a special case. Introducing random volatility σ gives rise to a method for constructing and simulating non-Gaussian fields with Matérn correlation. Such processes have recently been studied by Bolin [11] and Wallin and Bolin [40].

In a simulation study, we compare the hybrid simulation scheme to other simulation methods for the model (1.1), namely to what we call the Riemann-sum scheme, which corresponds to approximating the integrand by a step function, and to exact simulation using circulant embedding of the covariance matrix, as described by Dietrich and Newsam [16] and Wood and Chan [41]. The hybrid scheme is not exact, as it approximates the integrand only on a compact set. However, using circulant embeddings requires the process to be Gaussian and stationary, which the model (1.1) only satisfies in some special cases. Moreover, in order to apply exact simulation methods, the covariance function of X needs to be known, which is oftentimes costly to compute from the model (1.1). In theory, the asymptotic computational costs

of the hybrid scheme are slightly higher than for the circulant embedding method, see Section 3 for details. However, we found in our simulation study that for a wide range of parameters the hybrid scheme performs in fact faster than exact simulation, even for large values of n.

This article is structured as follows. In Section 2 we introduce our model in detail and discuss some of its properties. In Section 3 we describe the hybrid simulation scheme and derive the asymptotic error of the scheme. Section 4 contains the simulation study comparing the hybrid scheme to other simulation schemes. Proofs for our theoretical results are given in Section 5. The appendix contains technical details and calculations.

2 Volatility modulated moving average fields

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and W white noise on $\mathbb{R}^2$. That is, W is an independently scattered random measure satisfying $W(A) \sim \mathcal{N}(0, \lambda(A))$ for all sets $A \in \mathcal{B}_0 = \{A \in \mathcal{B}(\mathbb{R}^2) : \lambda(A) < \infty\}$, where λ denotes the Lebesgue measure. Recall that a collection of real valued random variables $\Lambda = \{\Lambda(A) : A \in \mathcal{B}_0\}$ is called independently scattered random measure if for every sequence $(A_n)_{n \in \mathbb{N}}$ of disjoint sets with $\lambda(\bigcup_n A_n) < \infty$, the random variables $\Lambda(A_n), n = 1, 2, \dots$ are independent and $\Lambda(\bigcup_n A_n) = \sum_n \Lambda(A_n)$, almost surely.

The kernel function $g : \mathbb{R}^2 \to \mathbb{R}$ is assumed to be of the form

$$g(\mathbf{t}) = \widetilde{g}(\|\mathbf{t}\|) := \|\mathbf{t}\|^\alpha L(\|\mathbf{t}\|) \tag{2.2}$$

for some $\alpha \in (-1, 0)$, and a function $L : (0, \infty) \to (0, \infty)$ that is slowly varying at 0. Here and in the following $\|\cdot\|$ always denotes the Euclidean norm on $\mathbb{R}^2$. Recall that L is said to be slowly varying at 0 if for any $\delta > 0$

$$\lim_{x \to 0} \frac{L(\delta x)}{L(x)} = 1,$$

and that then the function $\widetilde{g}(x) = x^\alpha L(x)$ is called regularly varying at 0 of index α. The explosive behaviour of the kernel at 0 is a crucial feature of this model, as it governs the roughness of the field. Indeed, under weak additional assumptions the Hausdorff dimension of a realisation of X is $2 - \alpha$ with probability 1, see Hansen and Thorarinsdottir [22] and Theorem 2.1, meaning that for $\alpha \to -1$ the realisations of X become extremely rough. In Figure 1 we present samples of realisations of VMMAs for different α.

The roughness of realisations poses a challenge for simulating volatility modulated moving averages. Indeed, possibly the most intuitive way to simulate the model (1.1) is by freezing the integrand over small blocks and simulate the white noise over these blocks as independent centered normal random variables with variance equaling the block size. However, this method does not account for the explosive behaviour of g at 0 and therefore does a poor job in reproducing the roughness of the original process correctly, in particular for values of α close to -1. We will demonstrate this phenomenon in a simulation study in Section 4. The hybrid scheme resolves this issue by approximating g around 0 by a power kernel, and approximating it by a step function away from 0.

The integral in (1.1) is well defined, when σ is measurable with respect to $\mathcal{B}(\mathbb{R}^2) \otimes \mathcal{F}$ and the process $\mathbf{s} \mapsto g(\mathbf{t} - \mathbf{s})\sigma_\mathbf{s}(\omega)$ takes almost surely values in $L^2(\mathbb{R}^2)$. In particular we do not require independence of σ and W or any notion of filtration or predictability for the definition of the integral, as is usually used in the theory of stochastic processes indexed by time. This general theory of stochastic integration dates back to Bichteler [9], see also Kwapień and Woyczyński [29]. A brief discussion can be found in Appendix A. When σ and W are independent, we can realise them on a product space and it is therefore sufficient to define integration with respect to W for deterministic functions, see for example Rajput and Rosiński [33].

The volatility field $(\sigma_\mathbf{s})_{\mathbf{s} \in \mathbb{R}^2}$ is assumed to satisfy $\mathbb{E}[\sigma_\mathbf{s}^2] < \infty$ for all $\mathbf{s}$.
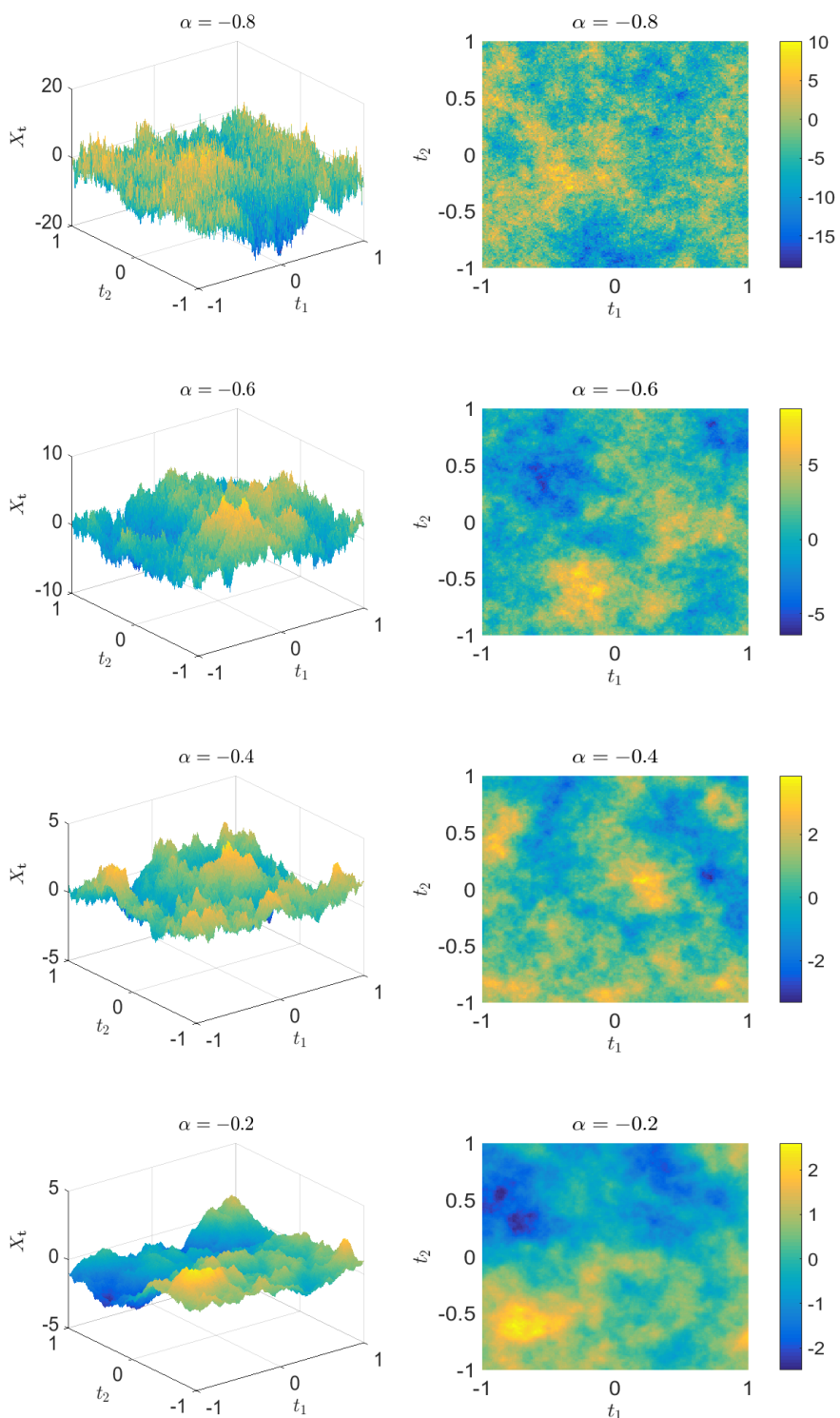
Figure 1: Realisations of volatility modulated moving average fields for different α with Matérn covariance, see Example 2.2. All plots range over $\mathbf{t} \in [-1, 1]^2$ and are generated with constant volatility σ. In Section 4 we present examples of VMMAs with nontrivial volatility.

Moreover, we assume σ to be covariance stationary, meaning that $\mathbb{E}[\sigma_\mathbf{s}]$ does not depend on $\mathbf{s}$ and $\mathrm{cov}(\sigma_{\mathbf{s}+\mathbf{r}}, \sigma_\mathbf{s}) = \mathrm{cov}(\sigma_\mathbf{r}, \sigma_0)$ for all $\mathbf{s}, \mathbf{r} \in \mathbb{R}^2$. In particular $\mathbb{E}[\sigma_\mathbf{s}^2] = \mathbb{E}[\sigma_0^2]$ for all $\mathbf{s} \in \mathbb{R}^2$. For some of our theoretical results we will assume that σ and W are independent, however we show in Appendix A that this is not required for the convergence of the hybrid scheme. We make the assumption that σ is sufficiently smooth such that freezing σ over small blocks will cause an asymptotically negligible error in the simulation. It turns out that this is the case when σ satisfies

$$\mathbb{E}[|\sigma_0 - \sigma_\mathbf{u}|^2] = o(\|\mathbf{u}\|^{2\alpha+2}), \quad \text{for } \mathbf{u} \to 0. \tag{2.3}$$

When σ is independent of the Gaussian noise W, the covariance stationarity of σ implies that the process X is itself covariance stationary and covariance isotropic in the sense that $\mathbb{E}[(X_{\mathbf{t}+\mathbf{s}} - X_\mathbf{t})^2]$ depends only on $\|\mathbf{s}\|$. If σ is in fact stationary, X is stationary and isotropic.

Moreover, we pose the following assumptions on our kernel function g. They ensure in particular that g is square integrable, which together with covariance stationarity of σ ensures the existence of the integral in (1.1).

(A1) The slowly varying function L is continuously differentiable and bounded away from 0 on any interval $(u, \sqrt{2}]$ for $u > 0$.

(A2) $\widetilde{g}'$ is monotonic on $[M, \infty)$ for some constant M, and satisfies $\widetilde{g}'(x) = \mathcal{O}(x^{\beta-1})$, as $x \to \infty$, for some $\beta \in (-\infty, -1)$. This implies, in particular, $\widetilde{g}(x) = \mathcal{O}(x^\beta)$.

(A3) There is a $C > 0$ such that $|L'(x)| < C(1 + x^{-1})$ for all $x \in (0, 1]$.

An appealing feature of the VMMA model is its flexibility in modelling marginal distributions and covariance structure independently. Indeed, assuming that σ is stationary and independent of W, the covariance structure of X is entirely determined by the kernel g, whereas the marginal distribution of X is a centered Gaussian variance mixture with conditional variance

$\int_{\mathbb{R}^2} g(\mathbf{t}-\mathbf{s})^2 \sigma_{\mathbf{s}}^2 d\mathbf{s}$, the distribution of which is governed by the distribution of σ. A popular model choice for the volatility field σ^2 is a moving average process driven by a subordinator (see Example 2.3). In this case, the marginals of $X_{\mathbf{t}}$ follow a type G distribution (i.e. a normal mixture distribution with infinitely divisible conditional variance), and an explicit expression for the characteristic function has been derived in [38].

The behaviour of the kernel at 0 is determined by the exponent α, whereas its behaviour away from 0, e.g. how quickly it decays at ∞, depends on the slowly varying function L. While the behaviour of g at 0 determines local properties of the process X, like the roughness of realisations, the behaviour of g away from 0 governs its global properties, e.g., whether it is long range dependent. Being able to independently choose α and L allows us therefore to model local and global properties of the VMMA independently, which underlines the flexibility of the model. This separation of local and global properties, and the desire to capture both of them correctly, is one of our main motivations to use a hybrid simulation scheme. We now formalise the statement that the roughness of X is determined by the power α.

Theorem 2.1. *(i) Assume independence of σ and W. The variogram of X defined as $V(h) := \mathbb{E}[(X_0 - X_{\mathbf{t}})^2]$, where $h = \|\mathbf{t}\|$, satisfies*

$$h^{-2-2\alpha} L(h)^2 V(h) \to \mathbb{E}[\sigma_0^2] C_\alpha \quad as\ h \to 0,$$

where C_α is a positive constant.

(ii) Assume additionally that the volatility is locally bounded in the sense that it satisfies $\sup_{\|\mathbf{s}\| \leq M+1} \{\sigma_{\mathbf{s}}^2\} < \infty$ almost surely, where M is as in assumption (A2). Then, for all $\varepsilon > 0$, the process X has a version with locally $\alpha + 1 - \varepsilon$-Hölder continuous realisations.

The proof can be found in Section 5. Hansen and Thorarinsdottir [22] analyse the variogram of a closely related model and derive similar results.

We conclude this section by discussing examples of possible choices for kernel functions g and volatility fields σ.

Example 2.2 (Matérn covariance)**.** Originally introduced in the context of tree population modelling in Swedish forests by Matérn [31], the Matérn covariance family has become popular in a variety of different fields such as meteorology, hydrology and machine learning. For an overview we refer to Guttorp and Gneiting [20] and the references therein. It is characterised by the correlation function

$$C(\|\mathbf{r}\|) = \mathbb{E}[(X_{\mathbf{r}} - X_0)^2]/\mathbb{E}[X_0^2] = \frac{(\lambda\|\mathbf{r}\|)^\nu}{2^{\nu-1}\Gamma(\nu)}K_\nu(\lambda\|\mathbf{r}\|), \quad \mathbf{r} \in \mathbb{R}^2,$$

where $\nu > 0$ is usually referred to as the shape parameter, while $\lambda > 0$ is a scale parameter. Here, K_ν denotes the modified Bessel function of the second kind. It has been shown by Jónsdóttir et al. [27], see also Hansen and Thorarinsdottir [22], that the model (1.1) with

$$g(\mathbf{t}) = \|\mathbf{t}\|^{\frac{\nu-1}{2}} K_{\frac{\nu-1}{2}}(\lambda\|\mathbf{t}\|)$$

has Matérn correlation, provided σ is independent of W and covariance stationary. When $\nu \in (0,1)$, the function g satisfies our model assumptions (A1)-(A3) with $\alpha = \nu - 1$, as we argue next. The function

$$L(x) = x^{\frac{1-\nu}{2}} K_{\frac{\nu-1}{2}}(\lambda x)$$

is continuously differentiable on $(0,\infty)$. It holds that $\lim_{x\downarrow 0} L(x) = 2^{-\frac{\nu+1}{2}}\Gamma\left(\frac{\nu-1}{2}\right)$, see [1, Eq. (9.6.9), p.375], which implies that L is slowly varying at 0 and satisfies condition (A3). Condition (A2) is satisfied for any $\beta < -1$, as follows from the identity

$$\frac{d}{dx}(x^{\alpha/2}K_{\alpha/2}(x)) = x^{\frac{\alpha}{2}-1}K_{\frac{\alpha}{2}-1}(x),$$

and the exponential decay of $K_{\frac{\nu-1}{2}}(\lambda x)$, cf. [1, p.378].

Example 2.3 (Ambit fields)**.** In a series of papers Barndorff-Nielsen and Schmiegel [6, 7] proposed to model velocities of particles in turbulent flows by a class of spatio-temporal stochastic processes called *ambit fields*. Over the last years this model found manifold applications throughout various sciences, examples being Barndorff-Nielsen et al. [4] and Jensen et al. [26]. The VMMA model is a purely spatial analogue of an ambit field driven by white noise and can therefore be interpreted as a realisation of an ambit field at a fixed time t. In the framework of turbulence modeling, the squared volatility σ^2 has the physical interpretation of local energy dissipation and it has been argued by Barndorff-Nielsen et al. [5] that it is natural to model σ^2 as (exponential of) an ambit field itself. A specific example that has drawn attention is when σ^2 is a moving average process driven by a subordinator, see [38]. In this case, simulations of σ rely on a shot noise decomposition of the integral, see Rosiński [34] and Cohen et al. [15].

3   The Hybrid Scheme

In this section we present the hybrid simulation scheme. For $r > 0$ and $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$ we introduce the notation $\square_r \mathbf{t}$ for a square with side length $1/r$ centred at $\mathbf{t}$, that is $\square_r \mathbf{t} = \left[t_1 - \frac{1}{2r}, t_1 + \frac{1}{2r}\right] \times \left[t_2 - \frac{1}{2r}, t_2 + \frac{1}{2r}\right]$. We will suppress the index r if it is 1, and will denote $\square_r$ instead of $\square_r 0$. We simulate the process $X_{\mathbf{t}}$ for $\mathbf{t} \in [-1, 1]^2$ on the square grid $\Gamma_n := \left\{\frac{1}{n}(i, j), \ i, j \in \{-n, ..., n\}\right\}$.

A first necessary step for approximating the integral (1.1) is to truncate the range of integration, i.e.

$$X_{\mathbf{t}} \approx \int_{\square_{1/C}\mathbf{t}} g(\mathbf{t} - \mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}),$$

for some large $C > 0$. To ensure convergence of the simulated process as

$n \to \infty$, we increase the range of integration simultaneously with increasing the grid resolution n. We let therefore $C = C_n \approx n^\gamma$ for some $\gamma > 0$. More precisely, it proves to be convenient to choose $C_n = \frac{N_n + 1/2}{n}$ with $N_n = [n^{1+\gamma}]$, where $[x]$ denotes the integer part of x.

An intuitive approach to simulating the model (1.1) is approximating the integrand on $\square_{C_n^{-1}}\mathbf{t}$ by freezing it over squares with side length $1/n$, i.e.

$$X_\mathbf{t}^{R,n} = \sum_{\mathbf{j} \in \mathbf{t} + \{-N_n, \dots, N_n\}^2} g(\mathbf{t} - \mathbf{b_j}/n)\sigma_{\mathbf{j}/n} \int_{\square_n \mathbf{j}} W(d\mathbf{s}), \qquad (3.4)$$

where $\mathbf{b_j} \in \square \mathbf{j}$ are evaluation points chosen such that $\mathbf{t} - \mathbf{b_j}/n \neq 0$ for all $\mathbf{t} \in \Gamma_n$ and $\mathbf{j} \in \mathbb{Z}^2$. Indeed, $X_\mathbf{t}^{R,n}$ can be simulated, assuming that the volatility σ can be simulated on the square grid $\{\frac{1}{n}(i,j), \ i,j \in \mathbb{Z}\}$, since $\{\int_{\square_n \mathbf{j}} W(d\mathbf{s})\}_{\mathbf{j} \in \mathbb{Z}^2} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \frac{1}{n^2})$. We will refer to this simulation method as Riemann-sum scheme. Nguyen and Veraart [32] use this technique to simulate volatility moving averages with bounded moving average kernel and demonstrate that it performs well in this setting. In our framework, however, a crucial weakness of this approach is the inaccurate approximation of the kernel function g around its singularity at 0, which results in a poor recovery of the roughness of X.

This weakness can be overcome by choosing a small $\kappa \in \mathbb{N}_0$ (typically, $\kappa \in \{0,1,2\}$) and approximating g by a power kernel on $\frac{1}{n}[-\kappa-1/2, \kappa+1/2]^2$. More specifically, denoting $K_\kappa = \{-\kappa, \dots, \kappa\}^2$ and $\overline{K}_\kappa = \{-N_n, \dots, N_n\}^2 \setminus K_\kappa$, the hybrid scheme approximates $X_\mathbf{t}$ by

$$X_\mathbf{t}^n := \sum_{\mathbf{j} \in K_\kappa} \sigma_{\mathbf{t}-\mathbf{j}/n} L(\|\mathbf{b_j}\|/n) \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \|\mathbf{t}-\mathbf{s}\|^\alpha W(d\mathbf{s})$$

$$+ \sum_{\mathbf{j} \in \overline{K}_\kappa} \sigma_{\mathbf{t}-\mathbf{j}/n} g(\mathbf{b_j}/n) \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} W(d\mathbf{s}). \qquad (3.5)$$

See Figure 2 for a visualisation. In order to simulate $X_\mathbf{t}$ on the grid $\mathbf{t} \in \Gamma_n$, we simulate the families of centred Gaussian random variables $\mathcal{W}_n^1$ and $\mathcal{W}_n^2$,
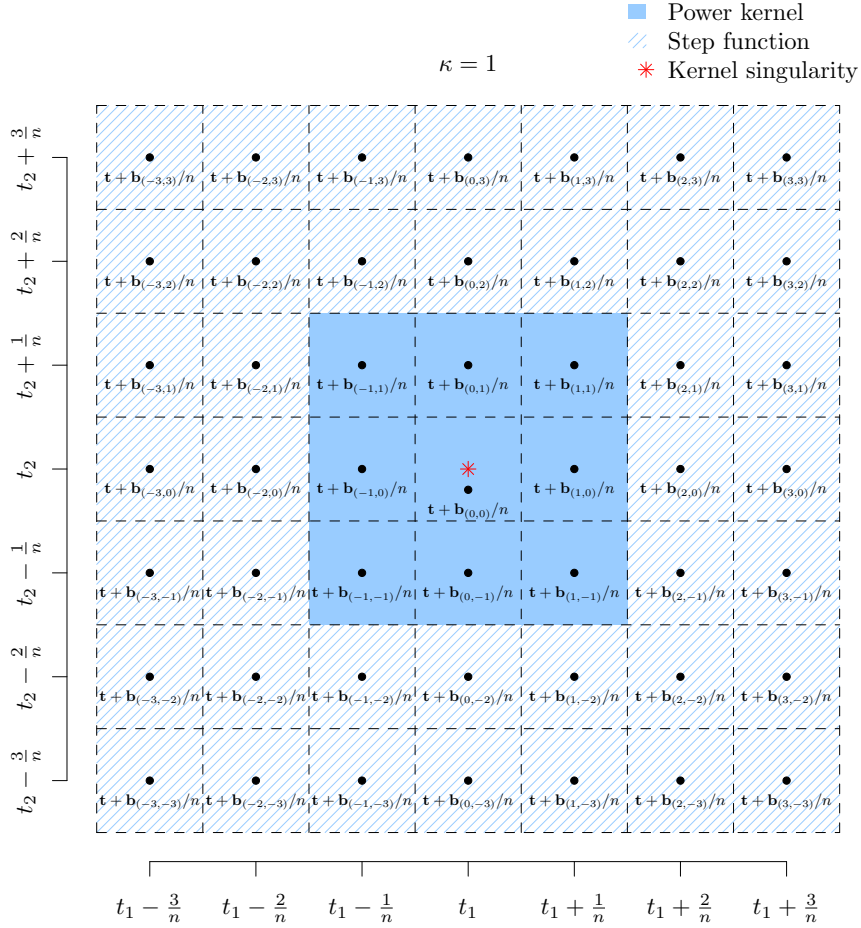
Figure 2: Visualisation of the hybrid scheme. Dividing $\mathbb{R}^2$ into small squares of size $1/n^2$, the kernel function g is approximated by a power kernel in the squares close to the singularity, and by a step function further away. The figure shows the situation for $\kappa = 1$, whereas for $\kappa = 0$ ($\kappa = 2$) the power kernel is used for only the central square (the central 25 squares). Simulating the random variables corresponding to the squares shown in the figure corresponds to simulating the process X at $(t_1, t_2) \in \mathbb{R}^2$ only. For simulating X at a different location (t_1', t_2') we obtain the same pattern shifted and need to account for the covariances of the random variables (not shown).

defined as

$$\mathcal{W}_n^1 := \left\{ W_{\mathbf{i},\mathbf{j}}^n = \int_{\square_n \mathbf{i}/n} \|(\mathbf{i} + \mathbf{j})/n - \mathbf{s}\|^\alpha W(d\mathbf{s}), \ W_{\mathbf{i}}^n = \int_{\square_n \mathbf{i}/n} W(d\mathbf{s}), \right.$$
$$\left. \mathbf{i} \in \{-n - \kappa, \ldots, n + \kappa\}^2 \text{ and } \mathbf{j} \in K_\kappa \right\},$$

$$\mathcal{W}_n^2 := \left\{ W_{\mathbf{i}}^n = \int_{\square_n \mathbf{i}/n} W(d\mathbf{s}), \right.$$
$$\left. \mathbf{i} \in \{-N_n - n, \ldots, N_n + n\}^2 \setminus \{-n - \kappa, \ldots, n + \kappa\}^2 \right\}.$$

Indeed, replacing $\mathbf{t}$ by $\mathbf{i}/n$ in (3.5) yields

$$X_{\mathbf{i}/n}^n = \sum_{\mathbf{j} \in K_\kappa} L(\|\mathbf{b}_{\mathbf{j}}\|) \sigma_{\frac{\mathbf{i}-\mathbf{j}}{n}} W_{\mathbf{i}-\mathbf{j},\mathbf{j}}^n + \sum_{\mathbf{j} \in \overline{K}_\kappa} g(\mathbf{b}_{\mathbf{j}}/n) \sigma_{\frac{\mathbf{i}-\mathbf{j}}{n}} W_{\mathbf{i}-\mathbf{j}}^n$$
$$=: \tilde{X}(\mathbf{i}/n) + \hat{X}(\mathbf{i}/n), \qquad \text{for } \mathbf{i} \in \{-n, \ldots, n\}^2.$$

By definition the random vectors $\left\{ ((W_{\mathbf{i},\mathbf{j}}^n)_{\mathbf{j} \in K_\kappa}, W_{\mathbf{i}}^n), \mathbf{i} \in \mathbb{Z}^2 \right\}$ are independent and identically distributed along $\mathbf{i}$. As a consequence, $\mathcal{W}_n^1$ and $\mathcal{W}_n^2$ are independent and $\mathcal{W}_n^2$ is composed of i.i.d. $\mathcal{N}(0, 1/n^2)$-distributed random variables. In order to simulate $\mathcal{W}_n^1$ we need to compute the covariance matrix of $((W_{0,\mathbf{j}}^n)_{\mathbf{j} \in K_\kappa}, W_0^n)$, which is of size $(|K_\kappa|+1)^2$ with $|K_\kappa| = (2\kappa+1)^2$. In contrast to the purely temporal model considered by Bennedsen et al. [8], computing the covariance structure becomes much more involved in our spatial setting. It relies partially on explicit expressions derived in Appendix B, and partially on numeric integration.

Note that the complexity of computing $\tilde{X}(\frac{\mathbf{i}}{n})$ for all $\mathbf{i} \in \{-n, \ldots, n\}^2$ is $\mathcal{O}(n^2)$, as the number of summands does not increase with n. The sum $\hat{X}(\frac{\mathbf{i}}{n})$ can be written as the two dimensional discrete convolution of the matrices A and B defined by

$$A_{\mathbf{k}} := \begin{cases} 0 & \mathbf{k} \in K_\kappa \\ g(\mathbf{b}_{\mathbf{k}}/n) & \mathbf{k} \in \overline{K}_\kappa \end{cases}, \qquad B_{\mathbf{k}} := \sigma_{\mathbf{k}/n} W_{\mathbf{k}}^n, \quad \text{for } \mathbf{k} \in \{-N-n, \ldots, N+n\}^2.$$

We remark that this expression as convolution is the main motivation that in (3.4) and (3.5) we chose to evaluate σ at the midpoints $\mathbf{t} - \mathbf{j}/n$ of $\square_n(\mathbf{t} - \mathbf{j}/n)$. Using FFT to carry out the convolution leads to a computational complexity of $\mathcal{O}(N^2 \log N) = \mathcal{O}(n^{2+2\gamma} \log n)$ for computing $\{\hat{X}(\frac{\mathbf{i}}{n})\}_{\mathbf{i}\in\{-n,\dots,n\}^2}$. Consequently, the computational complexity of the hybrid scheme is $\mathcal{O}(n^{2+2\gamma} \log n)$, provided the computational complexity of simulating $\{\sigma_{\mathbf{i}/n}\}_{\mathbf{i}\in\{-N-n,\dots,N+n\}^2}$ does not exceed $\mathcal{O}(n^{2+2\gamma} \log n)$. By comparison we recall that the exact simulation of an isotropic Gaussian field using circulant embeddings is of complexity $\mathcal{O}(n^2 \log n)$, see Gneiting et al. [18]. However, exact simulation requires σ to be constant and the covariance structure to be known. If the field is instead specified by its kernel function g, the covariance matrix needs to be computed by numerical integration and the complexity increases to $\mathcal{O}(n^4)$. A general Gaussian field with arbitrary covariance function can be simulated by computing the Cholesky-factorisation of the covariance matrix of $X_{\mathbf{t}}$ realised on the grid Γ_n, leading to a complexity of $\mathcal{O}(n^6)$, see Asmussen and Glynn [3, p. 312].

Next we derive the asymptotics for the mean square error of the hybrid simulation scheme.

Theorem 3.1. *Let* $\alpha \in (-1, 0)$. *Assume that* σ *is independent of* W *and satisfies* (2.3). *If* $\gamma > -(1+\alpha)/(1+\beta)$, *we have for all* $\mathbf{t} \in \mathbb{R}^2$ *that*

$$n^{2(\alpha+1)} L(1/n)^{-2} \mathbb{E}[|X_{\mathbf{t}} - X_{\mathbf{t}}^n|^2] \to \mathbb{E}[\sigma_0^2] J(\alpha, \kappa, \mathbf{b}), \quad as \ n \to \infty.$$

Here the constant $J(\alpha, \kappa, \mathbf{b})$ *is defined as*

$$J(\alpha, \kappa, \mathbf{b}) = \sum_{\mathbf{j}\in\mathbb{Z}^2\setminus\{-\kappa,\dots,\kappa\}^2} \int_{\square_{\mathbf{j}}} (\|\mathbf{x}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2 d\mathbf{x},$$

which is finite for $\alpha < 0$.

The proof is given in Section 5. This theorem and the computational complexity $\mathcal{O}(n^{2+2\gamma} \log n)$ of the hybrid scheme provide guidance how to choose

the cutoff parameter γ. It should be chosen small under the constraint $\gamma > -(1+\alpha)/(1+\beta)$, where β is chosen minimally such that (A2) is satisfied. For example if X is of Matérn type as in Example 2.2, the function $\widetilde{g}$ decays exponentially, and β can be chosen arbitrarily small. In this case the asymptotic of the mean square error given in the theorem applies for any $\gamma > 0$.

The sequence of evaluation points $\mathbf{b} = (\mathbf{b_j})_{\mathbf{j}\in\mathbb{Z}^2}$ can be chosen optimally, such that it minimises the limiting constant $J(\alpha, \kappa, \mathbf{b})$ and thus the asymptotic mean square error of the hybrid scheme. To this end $\mathbf{b_j}$ needs to be chosen in such a way that it minimises

$$\int_{\square_{\mathbf{j}}} (\|\mathbf{x}\|^\alpha - \|\mathbf{b_j}\|^\alpha)^2 d\mathbf{x},$$

for all $\mathbf{j} \in \mathbb{Z}^2$. By standard L^2 theory, $c \in \mathbb{R}$ minimises $\int_{\square_{\mathbf{j}}}(\|\mathbf{x}\|^\alpha - c)^2 d\mathbf{x}$ if and only if the function $\mathbf{x} \mapsto \|\mathbf{x}\|^\alpha - c$ is orthogonal to constant functions, that is, if it satisfies

$$\int_{\square_{\mathbf{j}}} (\|\mathbf{x}\|^\alpha - c)d\mathbf{x} = 0.$$

It follows then that $J(\alpha, \kappa, \mathbf{b})$ becomes minimal if we choose $\mathbf{b}$ such that

$$\|\mathbf{b_j}\| = \left(\int_{\square_{\mathbf{j}}} \|\mathbf{x}\|^\alpha d\mathbf{x} \right)^{1/\alpha}. \tag{3.6}$$

In Appendix B, we derive an explicit expression for this integral involving the Gauß hyperbolic function ${}_2F_1$. However, in our numerical experiments computing these integrals explicitly for all $\mathbf{j} \in \overline{K}_\kappa$ slowed the hybrid scheme down considerably, and we recommend choosing the midpoints $\mathbf{b_j} = \mathbf{j}$ instead. Figure 3 shows the constant $J(\alpha, \kappa, \mathbf{b}^{\mathrm{opt}}) = J_{\mathrm{opt}}$ for optimally chosen evaluation points $\mathbf{b}^{\mathrm{opt}}$ and the error caused by choosing midpoints $\mathbf{b_j} = \mathbf{j}$ instead, giving evidence that choosing midpoints leads to a nearly optimal result.

For $\mathbf{j} \in K_\kappa \setminus \{0\}$, the evaluation points $\mathbf{b_j}$ do not appear in the limiting expression in Theorem 3.1, and we will simply choose the midpoints $\mathbf{b_j} = \mathbf{j}$. However, for $\mathbf{j} = 0$ the expression $L(\|\mathbf{j}\|)$ is not necessarily defined. Indeed,
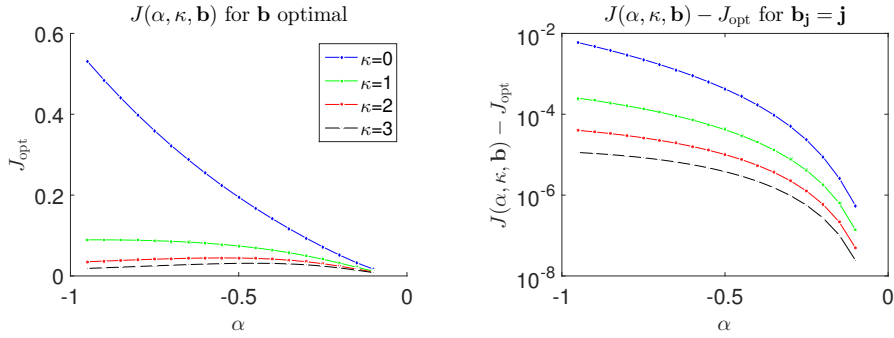
Figure 3: The first figure shows the value of $J(\alpha, \kappa, \mathbf{b}^{\text{opt}}) = J_{\text{opt}}$ for different values of α and κ where the evaluation points $\mathbf{b}^{\text{opt}}$ are chosen optimally, as in (3.6). The second figure shows the absolute error $J(\alpha, \kappa, \mathbf{b}) - J_{\text{opt}}$ for $\mathbf{b}$ chosen as midpoints, i.e. $\mathbf{b_j} = \mathbf{j}$, demonstrating that this choice leads to close to optimal results.

the slowly varying function L might have a singularity at 0. This shows that particular attention should be paid to the choice of $\mathbf{b}_0$, which is optimal if it minimises the L^2 error of the central cell, i.e.,

$$\mathbf{b}_0 = \underset{\mathbf{b} \in \square_n \setminus \{0\}}{\arg\min} \ \mathbb{E}\left(\int_{\square_n} g(\mathbf{s})W(d\mathbf{s}) - L(\|\mathbf{b}\|) \int_{\square_n} \|\mathbf{s}\|^\alpha W(d\mathbf{s}) \right)^2.$$

By straightforward calculation it can be shown that this is equivalent to

$$L(\|\mathbf{b}_0\|) = \left(\int_{\square_n} \|\mathbf{s}\|^{2\alpha} L(\|\mathbf{s}\|) d\mathbf{s} \right) \left(\int_{\square_n} \|\mathbf{s}\|^{2\alpha} d\mathbf{s} \right)^{-1}$$

$$= 8 C_{0,0}^{-1} \int_0^{1/\sqrt{2}} r^{2\alpha+1} L(r/n) \big(\pi/4 - \arccos(\sqrt{2}r) \mathbb{1}_{\{r>1/2\}} \big) dr,$$

where $C_{0,0}$ is defined in Appendix B. The integral on the right hand side is finite for $\alpha > -1$, which follows from the Potter bound (5.7), and can be evaluated numerically.

Before presenting numerical results of the hybrid scheme, let us mention two possible generalisations. A similar scheme can be implemented for general kernels g that are not of the form specified in (2.2). In this case, the covariance matrix of $\mathcal{W}_n^1$ needs to be computed numerically and does not rely on closed form expressions as derived in Section B in the appendix. An obvious draw-

back of this approach, apart from being more computationally involved, is that for general kernels g the roughness of the random field (1.1) cannot be characterised by a single parameter α, limiting the usefulness of such models for application. A second possible generalisation is towards higher-dimensional VMMA processes. However, also in this case the covariance matrix of $\mathcal{W}_n^1$ needs to be computed numerically, even for kernels satisfying (2.2). As the number of covariances to compute is proportional to n^{2d}, where d is the considered dimension, this would substantially increase the computational costs of the scheme.

4 Numerical results

In this section we demonstrate in a simulation study that the hybrid scheme is capable of capturing the roughness of the process correctly, and compare it in that aspect to other simulation schemes. Before doing so, we present in Figure 4 samples of VMMAs highlighting the effect of volatility. The volatility is modelled as $\sigma_{\mathbf{t}}^2 = \exp(X_{\mathbf{t}}')$, where X' is again a volatility modulated moving average, compare Example 2.3. For X' we choose the roughness parameter $\alpha = -0.2$ and the slowly varying function $L(x) = e^{-x}$. For the first realisation we chose $\alpha = -0.3$ and $L(x) = e^{-x}$. For the second we chose $\alpha = -0.7$ and L such that the model has Matérn covariance, see Example 2.2. In both cases it becomes apparent that areas of lower volatility cause the VMMA field to vary less.

For our simulation study we first recall the definition of fractal or Hausdorff dimension. For a set $S \subset \mathbb{R}^d$ and $\varepsilon > 0$, an ε-cover of S is a countable collection of balls $\{B_i\}_{i\in\mathbb{N}}$ with diameter $|B_i| \leq \varepsilon$ such that $S \subset \bigcup_i B_i$. The δ-dimensional Hausdorff measure of S is then defined as

$$H^\delta(S) = \lim_{\varepsilon \to 0} \inf \left\{ \sum_{i=1}^\infty |B_i|^\delta \ : \ \{B_i\}_{i\in\mathbb{N}} \text{ is } \varepsilon\text{-cover of } S \right\},$$
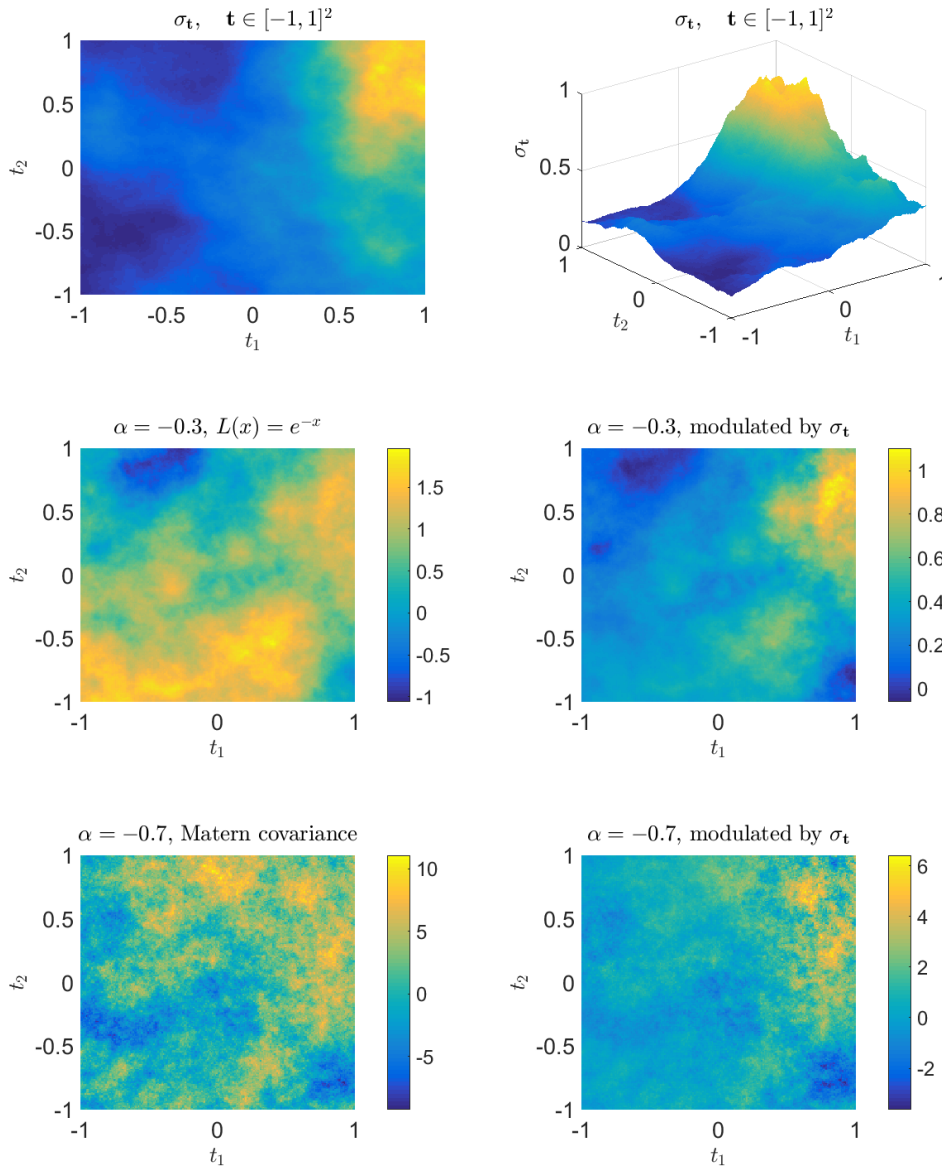
Figure 4: Examples for moving average fields modulated by volatility. The first row shows the volatility $(\sigma_{\mathbf{t}})_{\mathbf{t} \in \mathbb{R}^2}$ modelled as $\sigma_{\mathbf{t}}^2 = \exp(X_{\mathbf{t}}')$, where X' is again a VMMA field. The second and third row show realisations of VMMAs. On the left hand side the field is simulated with constant volatility, the right hand side is generated by the same Gaussian noise and with the same model parameters, but is modulated by $(\sigma_{\mathbf{t}})_{\mathbf{t} \in \mathbb{R}^2}$. For the second row we chose $\alpha = -0.3$ and the slowly varying function $L(x) = e^{-x}$. The third row is generated with $\alpha = -0.7$ and Matérn covariance.

and the fractal or Hausdorff dimension of S is

$$\mathrm{HD}(S) := \inf\{\delta > 0 \; : \; H^\delta(S) = 0\}.$$

The Hausdorff dimension of a spatial stochastic process $(X_{\mathbf{t}})_{\mathbf{t} \in \mathbb{R}^2}$ is the (random) Hausdorff dimension of its graph $\mathrm{HD}(\{(\mathbf{t}, X_{\mathbf{t}}), \; \mathbf{t} \in \mathbb{R}^2\})$, and takes consequently values in $[2, 3]$. For the model (1.1) with constant volatility $\sigma \equiv 1$ it follows easily from a standard result [2, Theorem 8.4.1] and Theorem 2.1 that $\mathrm{HD}(X) = 2 - \alpha$, see also Hansen and Thorarinsdottir [22]. Gneiting et al. [17] give an overview over existing methods for estimating the Hausdorff dimension of both time series data and spatial data, and provide implementations for various estimators in form of the R package fractaldim [36], which we rely on.

We estimate the Hausdorff dimension from simulations of X generated by the hybrid scheme, and compare to estimates from other simulation methods. We consider the model (1.1) with constant volatility σ and Matérn covariance, see Example 2.2. In this case the process X can be simulated exactly using circulant embeddings of the covariance matrix, to which we compare. Note that exact simulation is only available for Gaussian processes with known covariance function and is not applicable for general VMMAs. Moreover we compare to the Riemann-sum scheme introduced in (3.4). For the hybrid scheme we consider $\kappa = 0, 1, 2, 3$. With each technique we simulate 100 i.i.d. Monte-Carlo samples of the process $(X_{\mathbf{t}})_{\mathbf{t} \in [-1,1]^2}$ for every $\alpha \in \{-0.8, -0.7, ..., -0.1\}$. As grid resolution we chose $n = 100$ and, for the hybrid scheme and the Riemann-sum scheme, $N_n = [n^{1+\gamma}]$ with $\gamma = 0.3$, i.e. $N_n = 398$. Thereafter we estimate the roughness of X and average the estimates over the Monte-Carlo samples. There is a variety of different estimators for fractal dimension of spatial data. For a detailed overview and asymptotic properties we refer to Gneiting et al. [17] and the references therein. We apply the square increment estimator ν_{SI} introduced and analysed by Chan and Wood [13, Equation (4.3)] because of

its favourable asymptotic properties. Figure 5 shows the results and compares them to the theoretical value of the Hausdorff dimension $2 - \alpha$, plotted as dashed line. For the second plot in the figure we remark that the sample variance of the roughness estimates was between 0.005 and 0.01, for all values of α and all simulation methods.

Exact simulation using circulant embeddings performs slightly better than the hybrid scheme, in particular when $\alpha \approx 0$. This is not surprising, taking into account that the roughness of the process is governed by the behaviour of the kernel g at 0, which is well approximated by the hybrid scheme but, intuitively speaking, perfectly recovered by exact simulation. Let us stress again that exact simulation using circulant embeddings is only available for the model (1.1) in a few special cases. For $\kappa \geq 1$ the hybrid simulation scheme recovers the roughness very precisely, when $\alpha < -0.3$. When $\alpha \geq -0.3$ or $\kappa = 0$ it still performs reasonably well but tends to overestimate the roughness of the process slightly. This behaviour is likely to be caused by the at 0 slowly varying function, $L(x) = x^{-\alpha/2} K_{\alpha/2}(x)$ in the Matérn covariance case, which, intuitively speaking, varies more at 0 for larger values of α. As expected, the Riemann-sum approximation underestimates the roughness of the field significantly, as it does not account for the explosive behaviour of g at 0.

For the exact simulation via circulant embeddings we used the R package `RandomFields` [35], and refer to Gneiting et al. [18] for more details on this simulation method. For the roughness estimation we relied on the R package `fractaldim` [36]. Our implementation of the hybrid scheme is in MATLAB and is available in the online supplement of this article.

In Table 1 we compare computation times for the hybrid scheme, the circulant embeddings method, and the Riemann-sum scheme. For generating a single realisation, the circulant embedding method and the Riemann-sum scheme perform faster than the hybrid scheme. The main reason for this, however, is the costly computation of the covariance of the family $\mathcal{W}_n^1$, which
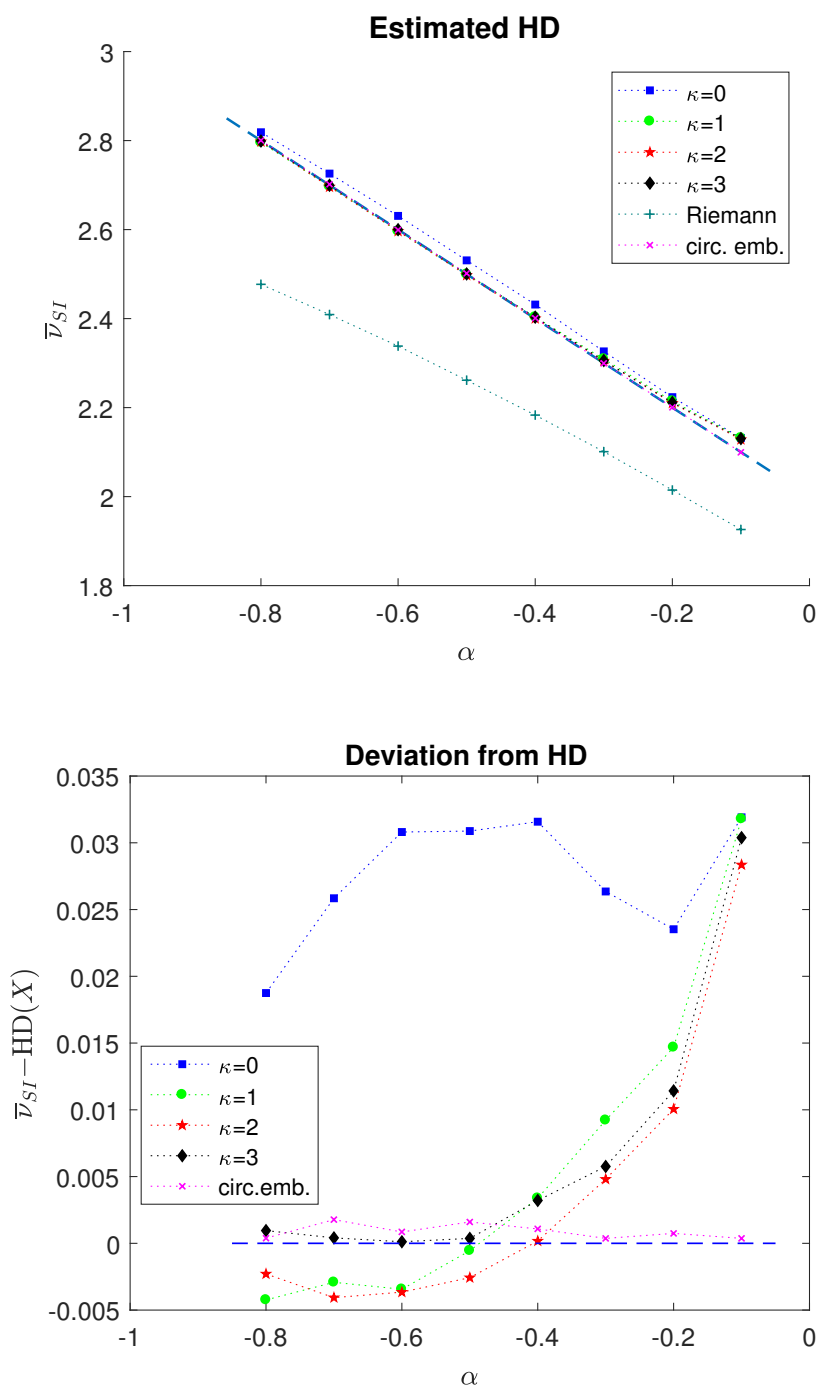
Figure 5: Roughness estimated from samples generated by the hybrid scheme, the Riemann-sum approximation method and by exact simulation using the circulant embedding method for Gaussian fields. The theoretical roughness is marked as a dashed line. The roughness is estimated by the isotropic estimator ν_{SI} introduced by Chan and Wood [13], averaged over 100 i.i.d. samples. The second plot shows in more detail the deviation between the estimation and the theoretical value, not including the Riemann-sum approximation scheme.

MC samples	$\kappa = 0$	$\kappa = 1$	$\kappa = 2$	$\kappa = 3$	circ.emb.	Riemann-sum
1	12.6 s	13.2 s	14.3 s	15.3 s	0.8 s	1.2 s
100	51 s	61.3 s	72.6 s	77.7 s	75.6 s	32.5 s

Table 1: Computation time of the hybrid scheme for different κ, for exact simulation using circulant embeddings, and for the Riemann-sum scheme, for a Matérn covariance Gaussian field. The first row shows the computation time for a single realisation, the second for 100 i.i.d. samples. The parameters of the model were chosen as $n = 100$, $\alpha = -0.6$, and, for the hybrid and the Riemann-sum scheme, $\gamma = 0.3$. The computation time was measured on a computer with with 2.9 GHz CPU and 32 GB RAM.

is only required once when generating i.i.d. Monte-Carlo samples, where the hybrid scheme performs more favourable. In view of the rather long computation times for all algorithms, let us stress that $n = 100$ corresponds to simulating X on a fine grid containing $(2n + 1)^2 = 40,401$ grid points.

5 Proofs

This section is dedicated to the proofs of our theoretical results. We begin by recalling the Potter bound which follows from Bingham et al. [10, Theorem 1.5.6]. For any $\delta > 0$ there exists a constant $C_\delta > 0$ such that

$$L(x)/L(y) \leq C_\delta \max\left\{ \left(\frac{x}{y}\right)^\delta, \left(\frac{x}{y}\right)^{-\delta} \right\}, \quad x, y \in (0, 1]. \qquad (5.7)$$

This bound will play an important role throughout all the proofs in this section.

Proof of Theorem 2.1 (i). The proof is similar to the proof of Bennedsen et al. [8, Proposition 2.1]. We have for $h > 0$ by covariance stationarity of σ that

$$V(h) = \mathbb{E}[\sigma_0^2] \int_{\mathbb{R}^2} \left(g(\mathbf{s} + h\mathbf{e}) - g(\mathbf{s}) \right)^2 d\mathbf{s},$$

where $\mathbf{e}$ is any unit vector and we used transformation into polar coordinates.

We obtain

$$V(h) = \mathbb{E}[\sigma_0^2](A_h + A'_h), \quad \text{where}$$

$$A_h = \int_{\{\|\mathbf{s}\| \leq 1\}} \big(g(\mathbf{s} + h\mathbf{e}/2) - g(\mathbf{s} - h\mathbf{e}/2)\big)^2 d\mathbf{s}, \quad \text{and}$$

$$A'_h = \int_{\{\|\mathbf{s}\| > 1\}} \big(g(\mathbf{s} + h\mathbf{e}/2) - g(\mathbf{s} - h\mathbf{e}/2)\big)^2 d\mathbf{s}.$$

Since the function $\widetilde{g}$ is continuously differentiable on $(0, \infty)$, we obtain by the mean value theorem the following estimate for A'_h.

$$A'_h \leq h^2 \bigg\{ \int_{\{1 < \|\mathbf{s}\| < M+1\}} \sup_{\{\xi \,:\, |\xi - \|\mathbf{s}\|| \leq h/\sqrt{2}\}} (\widetilde{g}'(\xi))^2 \, d\mathbf{s}$$

$$+ 2\pi \int_M^\infty \widetilde{g}'(r)^2 r dr \bigg\},$$

where we used that $|\widetilde{g}'|$ is decreasing on $[M, \infty)$. The term in curly brackets is finite by Assumption (A2), and we obtain that $A'_h = \mathcal{O}(h^2)$, as $h \to 0$. For A_h we make the substitution $\mathbf{x} = \mathbf{s}/h$ and obtain

$$A_h = h^2 \int_{\|\mathbf{x}\| \leq 1/h} \big(g(h(\mathbf{x} + \mathbf{e}/2)) - g(h(\mathbf{x} - \mathbf{e}/2))\big)^2 d\mathbf{x}$$

$$= h^{2+2\alpha} L^2(h) \int_{\|\mathbf{x}\| \leq 1/h} G_h(\mathbf{x}) d\mathbf{x},$$

where

$$G_h(\mathbf{x}) = \left(\|\mathbf{x} + \mathbf{e}/2\|^\alpha \frac{L(h\|\mathbf{x} + \mathbf{e}/2\|)}{L(h)} - \|\mathbf{x} - \mathbf{e}/2\|^\alpha \frac{L(h\|\mathbf{x} - \mathbf{e}/2\|)}{L(h)} \right)^2.$$

Note that $G_h(\mathbf{x}) \to \big(\|\mathbf{x} + \mathbf{e}/2\|^\alpha - \|\mathbf{x} - \mathbf{e}/2\|^\alpha\big)^2$, as $h \to 0$. Moreover, there is an integrable function G satisfying $G(\mathbf{x}) \geq |G_h(\mathbf{x})|$ for all $\mathbf{x}$, for sufficiently small h, allowing us to apply the dominated convergence theorem. Indeed, the existence of G follows since L is bounded away from 0 on $(0, 1]$ and by Assumption (A3). For details we refer to the proof of [8, Proposition 2.2]. Now, an application of the dominated convergence theorem yields the the

statement with

$$C_\alpha := 2\pi \int_{\mathbb{R}^2} \left(\|\mathbf{x} + \mathbf{e}/2\|^\alpha - \|\mathbf{x} - \mathbf{e}/2\|^\alpha \right)^2 d\mathbf{x}. \qquad \square$$

Proof of Theorem 2.1 (ii). The proof relies on the Kolmogorov-Chentsov theorem [28, Theorem 3.23], which requires localisation of the process, as σ does not necessarily have sufficiently high moments. We therefore first show the existence of a Hölder continuous version under two auxiliary assumption that will be relaxed thereafter:

(1) $\sigma_\mathbf{s}^2 \leq m$ for some $m > 0$, for all $\|\mathbf{s}\| \leq M + 1$.

(2) For $\mathbf{t}$ with $\|\mathbf{t}\| \leq 1$ it holds that

$$\int_{\{\|\mathbf{s}\| \geq M+1\}} (g(\mathbf{t} - \mathbf{s}) - g(-\mathbf{s}))^2 \sigma_\mathbf{s}^2 d\mathbf{s} \leq m\|\mathbf{t}\|^2.$$

Here, M denotes the constant introduced in Assumption (A2), and both conditions are assumed to be satisfied pathwise, i.e. with $\sigma_\mathbf{s} = \sigma_\mathbf{s}(\omega)$ for almost all $\omega \in \Omega$.

Under the auxiliary assumptions we have for all $p > 0$, $\|\mathbf{t}\| \leq 1$ that

$$\mathbb{E}[(X_\mathbf{t} - X_0)^p] \leq C_p \mathbb{E}\left[\left(\int_{\mathbb{R}^2} \left(g(\mathbf{t} - \mathbf{s}) - g(-\mathbf{s}) \right)^2 \sigma_\mathbf{s}^2 d\mathbf{s} \right)^{p/2} \right]$$

$$\leq C_p m^{p/2} \left(\int_{\{\|\mathbf{s}\| \leq M+1\}} \left(g(\mathbf{t} - \mathbf{s}) - g(-\mathbf{s}) \right)^2 d\mathbf{s} + \|\mathbf{t}\|^2 \right)^{p/2}$$

$$\leq C_p m^{p/2} \left(V_0(\|\mathbf{t}\|) + \|\mathbf{t}\|^2 \right)^{p/2},$$

where V_0 denotes the variogram of the process $(X_\mathbf{t})_{\mathbf{t} \in \mathbb{R}^2}$ with $\sigma \equiv 1$. In the first inequality we used that σ and W are independent and therefore $X_\mathbf{t} - X_0$ has a Gaussian mixture distribution with the integral on the right hand side being the conditional variance. Applying the first part of the theorem and the Potter bound (5.7) we obtain for any $\delta > 0$ a constant $C_{p,m,\delta}$ such that for all

t with $\|\mathbf{t}\| \leq 1$

$$\mathbb{E}[(X_{\mathbf{t}} - X_0)^p] \leq C_{p,m,\delta}\|\mathbf{t}\|^{p+p\alpha-\delta}.$$

Therefore, the Kolmogorov-Chentsov Theorem [28, Theorem 2.23] implies that X has a Hölder continuous version of any order $\gamma < 1 + \alpha - \frac{\delta}{p} - \frac{2}{p}$. and the result follows for any $\gamma \in (0, 1 + \alpha)$ by letting $p \to \infty$.

Now the proof can be completed by relaxing assumptions (1) and (2) to the weaker assumption $\sup_{\|\mathbf{s}\| \leq M+1} \{\sigma_{\mathbf{s}}^2\} < \infty$. By the mean value theorem we obtain that for all **t** with $\|\mathbf{t}\| \leq 1$

$$\|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\| \geq M+1\}} (g(\mathbf{t} - \mathbf{s}) - g(-\mathbf{s}))^2 \sigma_{\mathbf{s}}^2 d\mathbf{s}$$
$$\leq \|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\| \geq M+1\}} \left|\|\mathbf{t} - \mathbf{s}\| - \|\mathbf{s}\|\right|^2 \sup_{r \in [\|\mathbf{s}\|, \|\mathbf{t}-\mathbf{s}\|]} \left(\widetilde{g}'(r)^2\right) \sigma_{\mathbf{s}}^2 d\mathbf{s}$$
$$\leq \int_{\{\|\mathbf{s}\| \geq M+1\}} \widetilde{g}'(\|\mathbf{s}\| - 1)^2 \sigma_{\mathbf{s}}^2 d\mathbf{s},$$

where we used that $|\widetilde{g}'|$ is decreasing on $[M, \infty)$. By taking expectation it follows from Assumption (A2) that the right hand side is almost surely finite. Using the assumption from the theorem it follows that the random variable

$$Z := \max \left\{ \sup_{\|\mathbf{s}\| \leq M+1} (\sigma_{\mathbf{s}}^2), \sup_{\|\mathbf{t}\| \leq 1} \left(\|\mathbf{t}\|^{-2} \int_{\{\|\mathbf{s}\| \geq M+1\}} (g(\mathbf{t} - \mathbf{s}) - g(-\mathbf{s}))^2 \sigma_{\mathbf{s}}^2 d\mathbf{s} \right) \right\}$$

is finite, almost surely. The process $(X_{\mathbf{t}} \mathbb{1}_{\{Z \leq m\}})_{\mathbf{t} \in \mathbb{R}^2}$ satisfies the auxiliary assumptions (1) and (2) and coincides with X on $\{\omega : Z(\omega) \leq m\}$. Therefore, the existence of a version of X with $\alpha + 1 - \varepsilon$-Hölder continuous sample paths follows by letting $m \to \infty$. $\qquad\square$

For the proof of Theorem 3.1 we need the following auxiliary result. The proof is similar to the proof of Bennedsen et al. [8, Lemma 4.2] and not repeated.

Lemma 5.1. *Let $\alpha \in \mathbb{R}$ and $\mathbf{j} \in \mathbb{Z}^2 \setminus \{(0,0)\}$. If $\mathbf{b_j} \in \square\mathbf{j}$, it holds that*

(i) $\lim\limits_{n\to\infty} \int_{\square\mathbf{j}} \left(\|\mathbf{x}\|^{\alpha} \frac{L(\|\mathbf{x}\|/n)}{L(1/n)} - \|\mathbf{b_j}\| \frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)} \right)^2 d\mathbf{x} = \int_{\square\mathbf{j}} (\|\mathbf{x}\|^{\alpha} - \|\mathbf{b}\|^{\alpha})^2 d\mathbf{x},$

(ii) $\lim\limits_{n\to\infty} \int_{\square\mathbf{j}} \|\mathbf{x}\|^{2\alpha} \left(\frac{L(\|\mathbf{x}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)} \right)^2 d\mathbf{x} = 0.$

The same holds for $\mathbf{j} = (0,0)$ *if* $\mathbf{b}_{(0,0)} \neq (0,0)$ *and* $\alpha > -1$.

Proof of Theorem 3.1. Recall the definition

$$X_{\mathbf{t}}^n := \sum_{\mathbf{j}\in K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \|\mathbf{t}-\mathbf{s}\|^{\alpha} L(\|\mathbf{b_j}\|) \sigma_{\mathbf{t}-\mathbf{j}/n} W(d\mathbf{s})$$

$$+ \sum_{\mathbf{j}\in\overline{K}_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} g(\mathbf{b_j}/n) \sigma_{\mathbf{t}-\mathbf{j}/n} W(d\mathbf{s}).$$

We introduce the auxiliary object X'^n defined as

$$X_{\mathbf{t}}'^n := \sum_{\mathbf{j}\in K_\kappa \cup \overline{K}_\kappa} \sigma_{\mathbf{t}-\mathbf{j}/n} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} g(\mathbf{t}-\mathbf{s}) W(d\mathbf{s})$$

$$+ \int_{\mathbb{R}^2 \setminus \square_{N_n/n}\mathbf{t}} g(\mathbf{t}-\mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}).$$

Denoting $E_n := \mathbb{E}[|X_{\mathbf{t}}^n - X_{\mathbf{t}}'^n|^2]$ and $E_n' := \mathbb{E}[|X_{\mathbf{t}} - X_{\mathbf{t}}'^n|^2]$, Minkowski's inequality yields

$$E_n(1 - \sqrt{E_n'/E_n})^2 \leq \mathbb{E}[|X_{\mathbf{t}}^n - X_{\mathbf{t}}|^2] \leq E_n(1 + \sqrt{E_n'/E_n})^2. \qquad (5.8)$$

We will show later that $E_n'/E_n \to 0$ as $n \to \infty$, and it is thus sufficient to analyse the asymptotic behaviour of E_n.

We have that

$$
\begin{aligned}
E_n &= \sum_{\mathbf{j}\in K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \left(\|\mathbf{t}-\mathbf{s}\|^\alpha L(\|\mathbf{b_j}\|/n) - g(\mathbf{t}-\mathbf{s})\right)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]d\mathbf{s} \\
&\quad + \sum_{\mathbf{j}\in\{-n,\dots,n\}^2\backslash K_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \left(g(\mathbf{t}-\mathbf{s}) - g(\mathbf{b_j}/n)\right)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]d\mathbf{s} \\
&\quad + \sum_{\mathbf{j}\in\overline{K}_\kappa\backslash\{-n,\dots,n\}^2} \int_{\square_n(\mathbf{t}-\mathbf{j})} \left(g(\mathbf{t}-\mathbf{s}) - g(\mathbf{b_j}/n)\right)^2 \mathbb{E}[\sigma^2_{\mathbf{t}-\mathbf{j}/n}]d\mathbf{s} \\
&\quad + \int_{\mathbb{R}^2\backslash\square_{(2N_n+1)/n}\mathbf{t}} g(\mathbf{t}-\mathbf{s})^2 \mathbb{E}[\sigma^2_{\mathbf{s}}]d\mathbf{s} \\
&= \mathbb{E}[\sigma^2_0](D_1 + D_2 + D_3 + D_4).
\end{aligned}
\tag{5.9}
$$

For D_4 we obtain, recalling assumption (A2) and $N_n = n^{\gamma+1}$ that

$$
D_4 \le \int_{\|\mathbf{s}\|>N_n/n} g(\mathbf{s})^2 d\mathbf{s} = \mathcal{O}((N_n/n)^{2\beta+2}) = \mathcal{O}(n^{2\gamma(1+\beta)}).
$$

Therefore, we have

$$
n^{2(1+\alpha)}D_4 \to 0.
\tag{5.10}
$$

For D_3 we obtain

$$
D_3 = \sum_{\mathbf{j}\in\overline{K}_\kappa\backslash\{-n,\dots,n\}^2} \int_{\square_n\mathbf{j}/n} \left(g(\mathbf{s}) - g(\mathbf{b_j}/n)\right)^2 d\mathbf{s}.
$$

Recalling the notation $\tilde{g}(\|\mathbf{s}\|) = g(\mathbf{s})$ we have for $\mathbf{s} \in \square\mathbf{j}$ with $\mathbf{j} \in \overline{K}_\kappa \backslash \{-n,\dots,n\}^2$ by the mean value theorem $\xi \in [\|\mathbf{s}\| \wedge \|\mathbf{b_j}/n\|, \|\mathbf{s}\| \vee \|\mathbf{b_j}/n\|]$. Since $\tilde{g}'$ is decreasing on $[M,\infty)$ by assumption (A2) it follows that

$$
|g(\mathbf{s}) - g(\mathbf{b_j}/n)| = |\tilde{g}'(\xi)(\|\mathbf{s}\| - \|\mathbf{b_j}\|/n)|
$$

$$
\le \begin{cases} \frac{1}{n}\sup_{y\in[1-1/(\sqrt{2}n),M+1/(\sqrt{2}n)]} |\tilde{g}'(y)|, & (\|\mathbf{j}\| - \sqrt{2})/n < M, \\ \frac{1}{n}|\tilde{g}'((\|\mathbf{j}\| - \sqrt{2})/n)|, & (\|\mathbf{j}\| - \sqrt{2})/n \ge M. \end{cases}
$$

Consequently, we obtain with transformation into polar coordinates

$$\limsup_{n \to \infty} n^2 D_3 \tag{5.11}$$

$$\leq \left(\pi(M+1)^2 \sup_{z \in [1/2, M+1/2]} |\tilde{g}'(z)| + C \int_M^\infty r |\tilde{g}'(r)|^2 dr \right) < \infty.$$

For D_1 we have that

$$D_1 = \frac{1}{n^2} \sum_{\mathbf{j} \in K_\kappa} \int_{\square_{\mathbf{j}}} \left(\|\mathbf{s}/n\|^\alpha L(\|\mathbf{b}_{\mathbf{j}}\|/n) - g(\mathbf{s}/n) \right)^2 d\mathbf{s}$$

$$= \frac{L(1/n)}{n^{2+2\alpha}} \sum_{\mathbf{j} \in K_\kappa} \int_{\square_{\mathbf{j}}} \|\mathbf{s}\|^{2\alpha} \left(\frac{L(\|\mathbf{b}_{\mathbf{j}}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{s}\|/n)}{L(1/n)} \right)^2 d\mathbf{s}.$$

Since the number of elements of K_κ does not depend on n, we have by Lemma 5.1

$$\lim_{n \to \infty} \frac{n^{2+2\alpha} D_1}{L(1/n)} = 0. \tag{5.12}$$

The term D_2 can be written as

$$D_2 = \frac{1}{n^2} \sum_{\mathbf{j} \in \{-n,\dots,n\}^2 \setminus K_\kappa} \int_{\square_{\mathbf{j}}} \left(g(\mathbf{s}/n) - g(\mathbf{b}_{\mathbf{j}}/n) \right)^2 d\mathbf{s}$$

$$= \frac{L(1/n)^2}{n^{2+2\alpha}} \sum_{\mathbf{j} \in \{-n,\dots,n\}^2 \setminus K_\kappa} \underbrace{\int_{\square_{\mathbf{j}}} \left(\|\mathbf{s}\|^\alpha \frac{L(\|\mathbf{s}\|/n)}{L(1/n)} - \|\mathbf{b}_{\mathbf{j}}\|^\alpha \frac{L(\|\mathbf{b}_{\mathbf{j}}\|/n)}{L(1/n)} \right)^2 d\mathbf{s}}_{:=A_{\mathbf{j},n}}.$$

From Lemma 5.1 we know that $\lim_{n \to \infty} A_{\mathbf{j},n} = \int_{\square_{\mathbf{j}}} (\|\mathbf{s}\|^\alpha - \|\mathbf{b}_{\mathbf{j}}\|^\alpha)^2 d\mathbf{s}$. Consequently, if we find a dominating sequence $A_{\mathbf{j}}$ such that $A_{\mathbf{j}} \geq A_{\mathbf{j},n}$ for all n and $\sum_{\mathbf{j} \in \mathbb{Z}^2 \setminus K_\kappa} A_{\mathbf{j}} < \infty$, it follows from dominated convergence theorem that

$$\lim_{n \to \infty} \frac{D_2 n^{2\alpha+2}}{L(1/n)^2} = \sum_{\mathbf{j} \in \mathbb{Z}^2 \setminus K_\kappa} \int_{\square_{\mathbf{j}}} (\|\mathbf{s}\|^\alpha - \|\mathbf{b}_{\mathbf{j}}\|^\alpha)^2 d\mathbf{s}, \quad \text{for } \alpha \in (-1, 0). \tag{5.13}$$

It holds that

$$
\begin{aligned}
A_{\mathbf{j},n} = \int_{\square_{\mathbf{j}}} & \left\{ (\|\mathbf{s}\|^{\alpha} - \|\mathbf{b_j}\|^{\alpha}) \frac{L(\|\mathbf{s}\|/n)}{L(1/n)} + \|\mathbf{b_j}\|^{\alpha} \left(\frac{L(\|\mathbf{s}\|/n)}{L(1/n)} - \frac{L(\|\mathbf{b_j}\|/n)}{L(1/n)} \right) \right\}^2 d\mathbf{s} \\
\leq\ & 2 \int_{\square_{\mathbf{j}}} (\|\mathbf{s}\|^{\alpha} - \|\mathbf{b_j}\|^{\alpha})^2 \left(\frac{L(\|\mathbf{s}\|/n)}{L(1/n)} \right)^2 d\mathbf{s} \\
& + 2 \int_{\square_{\mathbf{j}}} \|\mathbf{b_j}\|^{2\alpha} \left(\frac{L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|/n)}{L(1/n)} \right)^2 d\mathbf{s} \\
=:\ & I_{\mathbf{j},n} + I'_{\mathbf{j},n}.
\end{aligned}
$$

For $I'_{\mathbf{j},n}$ we note that $\|\mathbf{b_j}\|^{2\alpha} \leq (\|\mathbf{j}\| - 1/\sqrt{2})^{2\alpha}$ for $\alpha < 0$. By the mean value theorem we have a $\xi \in [\|\mathbf{s}\|/n \wedge \|\mathbf{b_j}\|/n, \|\mathbf{s}\|/n \vee \|\mathbf{b_j}\|/n]$ such that

$$
|L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|/n)| = L'(\xi) |\|\mathbf{s}\|/n - \|\mathbf{b_j}\|/n| \leq \frac{C}{n} + \frac{C}{\|\mathbf{j}\| - 1/\sqrt{2}} \leq \frac{2C}{\|\mathbf{j}\| - 1/\sqrt{2}},
$$

where we used (A3) and that $\|\mathbf{j}\| \leq n$. Consequently, we obtain

$$
\begin{aligned}
I'_{\mathbf{j},n} &\leq \frac{C}{\inf_{x \in (0,1]} L(x)} (\|\mathbf{j}\| - 1/\sqrt{2})^{2\alpha} \int_{\square_{\mathbf{j}}} (L(\|\mathbf{s}\|/n) - L(\|\mathbf{b_j}\|))^2 d\mathbf{s} \\
&\leq C(\|\mathbf{j}\| - 1/\sqrt{2})^{2(\alpha-1)}.
\end{aligned}
$$

For the term $I_{\mathbf{j},n}$ we obtain by the Potter bound and the mean value theorem that

$$
I_{\mathbf{j},n} \leq C_{\delta} \int_{\square_{\mathbf{j}}} \min(\|\mathbf{s}\|, b_{\mathbf{j}})^{2\alpha-2} \|\mathbf{s}\|^{2\delta} d\mathbf{s} \leq C_{\delta} (\|\mathbf{j}\| - 1/\sqrt{2})^{2(\alpha-1+\delta)},
$$

where we choose $\delta \in (0, -\alpha)$. Consequently, we obtain $I_{\mathbf{j},n} + I'_{\mathbf{j},n} \leq C(\|\mathbf{j}\| - 1/\sqrt{2})^{-2}$ for all $n > 0$, and since

$$
\sum_{\mathbf{j} \in \mathbb{Z}^2 \setminus K_{\kappa}} C(\|\mathbf{j}\| - 1/\sqrt{2})^{-2} < \infty,
$$

(5.13) follows from dominated convergence theorem and Lemma 5.1. Now

(5.9) together with (5.10), (5.11), (5.12) and (5.13) show that

$$E_n \sim \mathbb{E}[\sigma_0^2]J(\alpha,\kappa,\mathbf{b})n^{-2(\alpha+1)}L(1/n)^2, \quad n \to \infty.$$

Therefore, recalling (5.8), the proof of statement (i) of the Theorem can be completed by showing that $E'_n/E_n \to 0$ as $n \to \infty$.

Since σ is covariance stationary, we obtain for E'_n

$$E'_n = \sum_{\mathbf{j}\in K_\kappa \cup \overline{K}_\kappa} \int_{\square_n(\mathbf{t}-\mathbf{j}/n)} \mathbb{E}[(\sigma_{\mathbf{t}-\mathbf{j}/n} - \sigma_{\mathbf{s}})^2]g(\mathbf{t}-\mathbf{s})^2 d\mathbf{s}$$
$$\leq \sup_{\mathbf{u}\in\square_n} \mathbb{E}[|\sigma_{\mathbf{u}} - \sigma_0|^2] \int_{\mathbb{R}^2} g(\mathbf{s})^2 d\mathbf{s},$$

and $E'_n/E_n \to 0$ follows by the assumption (2.3) $\qquad\qquad\square$

6 Acknowledgements

A On general stochastic integrals

We recall the definition of general stochastic integrals of the form $\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})$ where H is a real valued stochastic process, not necessarily independent of W. The construction of such integrals dates back to Bichteler [9]. In their recent

paper, Chong and Klüppelberg [14] revisit this theory in a spatio-temporal setting and derive a general integrability criterion for stochastic integrals driven by a random measure that is easy to check. In the context of integrals of the form (1.1), this criterion yields the following statement.

Proposition A.1. *Let $(H_{\mathbf{s}})_{\mathbf{s}\in\mathbb{R}^2}$ be a real valued stochastic process, measurable with respect to $\mathcal{B}(\mathbb{R}^2)\otimes\mathcal{F}$, such that $H \in L^2(\mathbb{R}^2)$, almost surely. Then, the stochastic integral $\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})$ exists in the sense of Bichteler [9].*

Proof. We apply Chong and Klüppelberg [14, Theorem 4.1]. Since this result is formulated in a spatio-temporal framework, we introduce an artificial time component and lift the white noise $W(d\mathbf{s})$ to a space time white noise $\widetilde{W}(dt; d\mathbf{s})$ such that $W(A) = \widetilde{W}([0,1]\times A)$ for all $A \in \mathcal{B}(\mathbb{R}^2)$. Equipping $(\Omega, \mathcal{F}, \mathbb{P})$ with the maximal filtration $\mathcal{F}_t = \mathcal{F}$ for all $t \in [0,1]$, the spatio-temporal process defined as $H_{\mathbf{s}}(t) := H_{\mathbf{s}}$ for all $t \in [0,1]$ is predictable and it holds that

$$\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s}) = \int_{[0,1]\times\mathbb{R}^2} H_{\mathbf{s}}(t)\widetilde{W}(dt; d\mathbf{s})$$

if the latter exists. The random measure $\widetilde{W}$ satisfies the conditions of Chong and Klüppelberg [14, Theorem 4.1] with characteristics $B = \mu = \nu = 0$ and $C(A; B) = \lambda(A\cap B)$ for all $A, B \in \mathcal{B}([0,1]\times\mathbb{R}^2)$, where λ denotes the Lebesgue measure. The theorem then implies that H is integrable with respect to W if and only if it satisfies almost surely $\int_{\mathbb{R}^2} H_{\mathbf{s}}^2 d\mathbf{s} < \infty$. $\qquad\square$

Note that the proofs for some of our theoretical results rely on the isometry

$$\mathbb{E}\left[\left(\int_{\mathbb{R}^2} H_{\mathbf{s}}W(d\mathbf{s})\right)^2\right] = \mathbb{E}\left[\int_{\mathbb{R}^2} H_{\mathbf{s}}^2 d\mathbf{s}\right],$$

which does not necessarily hold when H and W are dependent. In particular, we cannot rely on Theorem 3.1 in this more general framework, and deriving the L^2-error of the hybrid scheme would require to make additional regularity assumptions on the integrand. We are content with arguing that the hybrid

scheme converges also in the dependent case as long as σ is continuous, without further specifying the asymptotics.

Proposition A.2. *Assume that $(\sigma_\mathbf{s})_{\mathbf{s}\in\mathbb{R}^2}$ is covariance stationary and has a continuous version. Then, $X_\mathbf{t}^n \xrightarrow{\mathbb{P}} X_\mathbf{t}$ for all $\mathbf{t} \in \mathbb{R}^2$, i.e. the hybrid scheme converges.*

Proof. For ease of notation we assume $\mathbf{t} = 0$. Denote by g^n and σ^n the discrete approximations to g and σ used in the hybrid scheme, i.e.

$$g^n(\mathbf{s}) := \sum_{\mathbf{j}\in K_\kappa} \|\mathbf{s}\|^\alpha L\left(\frac{b_\mathbf{j}}{n}\right)\mathbb{1}_{\square_n\mathbf{j}/n}(\mathbf{s}) + \sum_{\mathbf{j}\in\overline{K}_\kappa} g\left(\frac{b_\mathbf{j}}{n}\right)\mathbb{1}_{\square_n\mathbf{j}/n}(\mathbf{s}),$$

$$\sigma_\mathbf{s}^n := \sum_{\mathbf{k}\in K_\kappa\cup\overline{K}_\kappa} \sigma_{\mathbf{t}-\mathbf{k}/n}\mathbb{1}_{\square_n(\mathbf{t}-\mathbf{k}/n)}(\mathbf{s}).$$

We consider the two discretized approximations for the integral X_0

$$X_0^n = \int_{\mathbb{R}^2} g^n(-\mathbf{s})\sigma_\mathbf{s}^n W(d\mathbf{s}) \quad\text{and}\quad \widetilde{X}_0^n := \int_{\mathbb{R}^2} g(-\mathbf{s})\sigma_\mathbf{s}^n W(d\mathbf{s}),$$

where X_0^n matches the hybrid scheme discretization introduced in Section 3.

The proof takes several steps. First we argue that continuity of σ implies $\widetilde{X}_0^n \xrightarrow{\mathbb{P}} X_0$. Thereafter we argue that $\widetilde{X}_0^n - X_0^n \xrightarrow{\mathbb{P}} 0$. To this end we restrict the integral to a compact set, arguing that the remainder becomes negligible when the compactum is chosen sufficiently large. For the integral restricted to the compactum the convergence then follows by a version of the dominated convergence theorem for stochastic integrals.

In order to see that $\widetilde{X}_0^n \xrightarrow{\mathbb{P}} X_0$, note that

$$X_0 = \int_{\mathbb{R}^2} g(-\mathbf{s})\sigma_\mathbf{s} W(d\mathbf{s}) = \int_{\mathbb{R}^2} \sigma_\mathbf{s} M_g(d\mathbf{s}),$$

where the random measure M_g is defined as $M_g(A) = \int_A g(-\mathbf{s})W(d\mathbf{s})$. The sequence σ^n consists of simple integrands that pointwise converge to σ by

continuity. By construction of the stochastic integral, this readily implies that

$$X_0 = \int_{\mathbb{R}^2} \sigma_{\mathbf{s}} M_g(d\mathbf{s}) = \lim_{n \to \infty} \int_{\mathbb{R}^2} \sigma_{\mathbf{s}}^n M_g(d\mathbf{s}) = \lim_{n \to \infty} \widetilde{X}_0^n, \qquad (1.14)$$

in probability, see [9]. Note that all integrals exist by covariance stationarity of σ.

Next, we consider for a large constant K the decomposition

$$X_0 = X_0^{\geq K} + X_0^{<K} =: \int_{\|\mathbf{s}\| \geq K} g(-\mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}) + \int_{\|\mathbf{s}\| < K} g(-\mathbf{s})\sigma_{\mathbf{s}} W(d\mathbf{s}),$$

and similarly for X_0^n and $\widetilde{X}_0^n$. We argue that for fixed $\varepsilon, \delta > 0$ there is sufficiently large K such that

$$\limsup_n \mathbb{P}[|\widetilde{X}_0^{n,\geq K} - X_0^{n,\geq K}| > \delta] < \varepsilon. \qquad (1.15)$$

By an application of Cauchy-Schwarz inequality it holds that

$$\mathbb{E}[|X_0^{n,\geq K} - \widetilde{X}_0^{n,\geq K}|] \qquad (1.16)$$
$$\leq \sum_{\|\mathbf{j}\|/n \geq K} \mathbb{E}\left[\left|\sigma_{-\mathbf{j}/n} \int_{\square_n \mathbf{j}/n} \left(g\left(\frac{b_{\mathbf{j}}}{n}\right) - g(\mathbf{s})\right) W(d\mathbf{s})\right|\right]$$
$$\leq \sum_{\|\mathbf{j}\|/n \geq K} (\mathbb{E}[\sigma_{-\mathbf{j}/n}^2])^{1/2} \sqrt{\int_{\square_n \mathbf{j}/n} \left(g\left(\frac{b_{\mathbf{j}}}{n}\right) - g(\mathbf{s})\right)^2 d\mathbf{s}}.$$

An application of the mean value theorem shows that, for $\|\mathbf{j}\|/n$ sufficiently large,

$$\left(g\left(\frac{b_{\mathbf{j}}}{n}\right) - g(\mathbf{s})\right)^2 \leq \frac{C}{n^2} \widetilde{g}'^2(\|\mathbf{j}\|/n), \qquad \text{for } \mathbf{s} \in \square_n \mathbf{j}/n,$$

for some constant C. It follows that

$$\sqrt{\int_{\square_n \mathbf{j}/n} \left(g\left(\frac{b_{\mathbf{j}}}{n}\right) - g(\mathbf{s})\right)^2 d\mathbf{s}} \leq \frac{C}{n^2} \widetilde{g}'(\|\mathbf{j}\|/n).$$

Therefore, (1.16) implies, for suitable constants C_1 and C_2,

$$\mathbb{E}[|X_0^{n,\geq K} - \widetilde{X}_0^{n,\geq K}|] \leq C_1 \sum_{\|\mathbf{j}\|/n \geq K} \frac{\widetilde{g}'(\|\mathbf{j}\|/n)}{n^2}$$

$$\leq C_2 \int_{\|\mathbf{s}\|>K} \widetilde{g}'(\|\mathbf{s}\|)d\mathbf{s} = 2\pi C_2 \int_K^\infty \widetilde{g}'(r)r\,dr.$$

By assumption (A2), $g'(r)r$ is integrable and (1.15) follows.

It remains to argue that $X_0^{n,<K} \xrightarrow{\mathbb{P}} X_0^{<K}$. Consider for some $C > 0$ the decomposition $\Omega = \Omega_C \cup \Omega_C^c$ of the probability space, where

$$\Omega_C := \{\omega \,:\, |\sigma_{\mathbf{s}}(\omega)| < C, \text{for all } \|\mathbf{s}\| < K\}.$$

Since $\{\|\mathbf{s}\| < K\}$ is compact and σ has continuous paths, it holds that $\mathbb{P}[\Omega_C^c] \to 0$, as $C \to \infty$. Moreover, $g^n(-\mathbf{s})\sigma_{\mathbf{s}}^n \mathbb{1}_{\Omega_C}$ converges pointwise to $g(-\mathbf{s})\sigma_{\mathbf{s}}\mathbb{1}_{\Omega_C}$, where we used continuity of σ, and is uniformly bounded for all $\mathbf{s} \in K$ and all n. Consequently it follows from a dominated convergence theorem for stochastic integrals [14, Theorem 2.3], that $X_0^{n,<K}\mathbb{1}_{\Omega_C} \xrightarrow{\mathbb{P}} X_0^{<K}\mathbb{1}_{\Omega_C}$. By letting $C \to \infty$ it follows that $X_0^{n,<K} \xrightarrow{\mathbb{P}} X_0^{<K}$ for fixed K.

Now, combining this convergence with (1.14) and (1.15) shows that for any $\delta, \varepsilon > 0$ we can choose K as in (1.15), and obtain

$$\mathbb{P}[|X_0^n - X_0| > 3\delta] \leq \mathbb{P}[|X_0^{n,\geq K} - \widetilde{X}_0^{n,\geq K}| > \delta] + \mathbb{P}[|\widetilde{X}_0^{n,\geq K} - X_0^{\geq K}| > \delta]$$

$$+ \mathbb{P}[|X_0^{n,<K} - X_0^{<K}| > \delta] \leq 3\varepsilon,$$

for sufficiently large n. This shows the claim of the proposition. $\qquad\square$

B The covariance of $\mathcal{W}_n^1$

In this section we analyse the covariance structure of the Gaussian family $\mathcal{W}_n^1$ introduced in Section 3. For a wide range of covariances we are able to derive closed expressions, whereas the remaining covariances are computed by

numerical integration. Let us remark that, in addition to the symmetry of the covariance matrix, the isotropy of the process adds 8 more spatial symmetries (corresponding to the linear transformations in the orthogonal group $O(2)$ that map the grid Γ_n onto itself), which reduces the number of necessary computations drastically. Since the random variables in $\mathcal{W}_n^1$ are i.i.d. along $\mathbf{i}$, it is sufficient to derive the covariance matrix for

$$\left\{W_{0,\mathbf{j}}^n,\ W_0^n\right\}_{\mathbf{j}\in K_\kappa}.$$

For $\mathbf{j}_1, \mathbf{j}_2 \in \{-\kappa, \dots, \kappa\}^2$ it holds that

$$\begin{aligned}
C_{1,1} &:= \operatorname{var}(W_0^n) = \frac{1}{n^2}, \\
C_{1,\mathbf{j}_1} &:= \operatorname{cov}(W_0^n, W_{0,\mathbf{j}_1}^n) = \frac{1}{n^{2+\alpha}} \int_\square \|\mathbf{j}_1 - \mathbf{s}\|^\alpha d\mathbf{s}, \\
C_{\mathbf{j}_1,\mathbf{j}_2} &:= \operatorname{cov}(W_{0,\mathbf{j}_1}^n, W_{0,\mathbf{j}_2}^n) = \frac{1}{n^{2+2\alpha}} \int_\square \|\mathbf{j}_1 - \mathbf{s}\|^\alpha \|\mathbf{j}_2 - \mathbf{s}\|^\alpha d\mathbf{s}.
\end{aligned}$$

We now derive explicit expressions for $C_{\mathbf{j},\mathbf{j}}$ using the Gauss hypergeometric function $_2F_1$. Clearly, these expressions can be applied to compute $C_{1,\mathbf{j}}$ by replacing α with $\alpha/2$. Using symmetries we may assume without loss of generality that $\mathbf{j} = (j_1, j_2)$ with $j_1 \geq j_2 \geq 0$. We introduce the notation $\triangleleft \mathbf{j}$ for the area $\{(x_1, x_2)\ :\ j_2 \leq x_1 \leq j_1,\ j_2 \leq x_2 \leq x_1\}$, that is a right triangle with lower right vertex (j_1, j_2) and hypotenuse lying on the diagonal $\{(x_1, x_2)\ :\ x_1 = x_2\}$. In order to obtain explicit expressions for $C_{\mathbf{j},\mathbf{j}}$, we first derive explicit expressions for

$$\int_{\triangleleft \mathbf{j}} \|\mathbf{x}\|^{2\alpha} d\mathbf{x}, \quad \text{for all } \mathbf{j} = (j_1, j_2) \in \mathbb{R}^2, 0 \leq j_2 < j_1. \tag{2.17}$$

Thereafter we give for all $\mathbf{j} = (j_1, j_2) \in \mathbb{Z}^2$ with $0 \leq j_2 \leq j_2$ an explicit formula to write $C_{\mathbf{j},\mathbf{j}}$ as linear combination of such integrals.

Transforming into polar coordinates we obtain that

$$
\int_{\triangleleft \mathbf{j}} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} = \int_{\arctan(j_2/j_1)}^{\pi/4} \int_{j_2/\sin(\theta)}^{j_1/\cos(\theta)} r^{2\alpha+1} dr d\theta
$$

$$
= \frac{1}{2\alpha + 2} \int_{\arctan(j_2/j_1)}^{\pi/4} \left(\frac{j_1}{\cos(\theta)} \right)^{2\alpha+2} - \left(\frac{j_2}{\sin(\theta)} \right)^{2\alpha+2} d\theta. \quad (2.18)
$$

It holds that $\arctan(j_2/j_1) = \arccos(\frac{j_1}{\|\mathbf{j}\|})$, and consequently we obtain by substituting $\cos(\theta) = z$ the following expression for the first summand:

$$
\frac{j_1^{2\alpha+2}}{2\alpha + 2} \int_{\arctan(j_2/j_1)}^{\pi/4} \cos(\theta)^{-2\alpha-2} d\theta
$$

$$
= -\frac{j_1^{2\alpha+2}}{2\alpha + 2} \int_{j_1/\|\mathbf{j}\|}^{\cos(\pi/4)} z^{-2\alpha-2}(1 - z^2)^{-1/2} dz
$$

$$
= \frac{j_1^{2\alpha+2}}{4(\alpha + 1)} \int_{1/2}^{j_1^2/\|\mathbf{j}\|^2} z^{-\alpha-\frac{3}{2}}(1 - z)^{-1/2} dz
$$

$$
= \frac{j_1^{2\alpha+2}}{4(\alpha + 1)} \int_{j_2^2/\|\mathbf{j}\|^2}^{1/2} (1 - z)^{-\alpha-\frac{3}{2}} z^{-1/2} dz
$$

$$
= \frac{j_1^{2\alpha+2}}{4(\alpha + 1)} (B(1/2; 1/2, -\alpha - 1/2) - B(j_2^2/\|\mathbf{j}\|^2; 1/2, -\alpha - 1/2))
$$

$$
= \frac{j_1^{2\alpha+2}}{2^{3/2}(\alpha + 1)} {}_2F_1(1/2, 3/2 + \alpha; 3/2; 1/2)
$$

$$
- \frac{j_1^{2\alpha+2} j_2}{2\|\mathbf{j}\|(\alpha + 1)} {}_2F_1(1/2, 3/2 + \alpha; 3/2; j_2^2/\|\mathbf{j}\|^2).
$$

Here, $B(x; p, q)$ denotes the incomplete beta function, satisfying $B(x; p, q) = \frac{x^p}{p} {}_2F_1(p, 1 - q; p + 1; x)$. For the first equality we used that $d/dz(\arccos(z)) = -(1 - z^2)^{-1/2}$. For the second summand in (2.18) we argue similarly, using

that $\arctan(j_2/j_1) = \arcsin(\frac{j_2}{\|\mathbf{j}\|})$,

$$-\frac{j_2^{2\alpha+2}}{2\alpha+2}\int_{\arctan(j_2/j_1)}^{\pi/4}\sin(\theta)^{-2\alpha-2}d\theta$$

$$=-\frac{j_2^{2\alpha+2}}{2\alpha+2}\int_{j_2/\|\mathbf{j}\|}^{\sin(\pi/4)}z^{-2\alpha-2}(1-z^2)^{-1/2}dz$$

$$=-\frac{j_2^{2\alpha+2}}{4(\alpha+1)}\int_{j_2^2/\|\mathbf{j}\|^2}^{1/2}z^{-\alpha-\frac{3}{2}}(1-z)^{-1/2}dz$$

$$=-\frac{j_2^{2\alpha+2}}{4(\alpha+1)}\int_{1/2}^{j_1^2/\|\mathbf{j}\|^2}(1-z)^{-\alpha-\frac{3}{2}}z^{-1/2}dz$$

$$=-\frac{j_2^{2\alpha+2}}{4(\alpha+1)}(B(j_1^2/\|\mathbf{j}\|^2;1/2,-\alpha-1/2)-B(1/2;1/2,-\alpha-1/2))$$

$$=\frac{j_2^{2\alpha+2}}{2^{3/2}(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;1/2)$$

$$-\frac{j_2^{2\alpha+2}j_1}{2\|\mathbf{j}\|(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;j_1^2/\|\mathbf{j}\|^2).$$

This leads to

$$\int_{\triangleleft\mathbf{j}}\|\mathbf{x}\|^{2\alpha}d\mathbf{x}=\frac{j_2^{2\alpha+2}+j_1^{2\alpha+2}}{2^{3/2}(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;1/2)$$

$$-\frac{j_1 j_2^{2\alpha+2}}{2\|\mathbf{j}\|(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;j_1^2/\|\mathbf{j}\|^2)$$

$$-\frac{j_1^{2\alpha+2}j_2}{2\|\mathbf{j}\|(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;j_2^2/\|\mathbf{j}\|^2),$$

for all $0 \le j_2 < j_1$. For implementation we remark that in the case $j_2 = 0$ the hypergeometric function in the second line is not defined since in this case $j_1^2/\|\mathbf{j}\|^2 = 1$, and we use

$$\int_{\triangleleft(j_1,0)}\|\mathbf{x}\|^{2\alpha}d\mathbf{x}=\frac{\sqrt{2}j_1^{2\alpha+2}}{4(\alpha+1)}{}_2F_1(1/2,3/2+\alpha;3/2;1/2).$$

Thus, we have explicit expressions for integrals of the form (2.17) and all that remains to do is to argue that for $0 \le j_2 < j_1$ we can write $C_{\mathbf{j},\mathbf{j}}$ as linear

combinations of such integrals. By symmetry we obtain that

$$C_{(0,0),(0,0)} = \frac{1}{n^{2+2\alpha}} \int_{\square} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} = \frac{8}{n^{2+2\alpha}} \int_{\triangleleft (1/2,0)'} \|\mathbf{x}\|^{2\alpha} d\mathbf{x}.$$

For $j > 0$ we obtain

$$C_{(j,j),(j,j)} = \frac{2}{n^{2+2\alpha}} \int_{\triangleleft (j+1/2,j-1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x}, \quad \text{and}$$

$$C_{(j,0),(j,0)} = \frac{2}{n^{2+2\alpha}} \left(\int_{\triangleleft (j+1/2,0)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} - \int_{\triangleleft (j-1/2,0)'} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} \right.$$
$$\left. - \int_{\triangleleft (j+1/2,1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} + \int_{\triangleleft (j-1/2,1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} \right).$$

For $0 < j_2 < j_1$ we obtain

$$C_{(j_1,j_2),(j_1,j_2)} = \frac{1}{n^{2+2\alpha}} \left(\int_{\triangleleft (j_1+1/2,j_2-1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} - \int_{\triangleleft (j_1-1/2,j_2-1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} \right.$$
$$\left. - \int_{\triangleleft (j_1+1/2,j_2+1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} + \int_{\triangleleft (j_1-1/2,j_2+1/2)} \|\mathbf{x}\|^{2\alpha} d\mathbf{x} \right).$$

This covers all possible choices for $0 \leq j_2 < j_1$, and consequently we obtain explicit expressions for $C_{\mathbf{j},\mathbf{j}}$ and $C_{\mathbf{j},1}$ for all $\mathbf{j}$.

References

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* U.S. Government Printing Office, Washington, D.C.

[2] Adler, R. J. (1981). *The geometry of random fields.* John Wiley & Sons, Ltd., Chichester. Wiley Series in Probability and Mathematical Statistics.

[3] Asmussen, S. and Glynn, P. W. (2007). *Stochastic simulation: algorithms and analysis.* Springer, New York.

[4] Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. (2014).

Modelling electricity forward markets by ambit fields. *Adv. in Appl. Probab.*, 46(3):719–745.

[5] Barndorff-Nielsen, O. E., Hedevang, E., Schmiegel, J., and Szozda, B. (2016). Some recent developments in ambit stochastics. In *Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015*, volume 138, pages 3–25. Springer.

[6] Barndorff-Nielsen, O. E. and Schmiegel, J. (2007). Ambit processes; with applications to turbulence and tumour growth. In *Stochastic analysis and applications*, pages 93–124. Springer.

[7] Barndorff-Nielsen, O. E. and Schmiegel, J. (2008). Time change, volatility, and turbulence. In *Mathematical control theory and finance*, pages 29–53. Springer, Berlin.

[8] Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2017). Hybrid scheme for Brownian semistationary processes. *Finance and Stochastics*, 21(4):931–965.

[9] Bichteler, K. (2002). *Stochastic integration with jumps*. Cambridge University Press, Cambridge.

[10] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular variation*. Cambridge University Press, Cambridge.

[11] Bolin, D. (2014). Spatial Matérn fields driven by non-Gaussian noise. *Scand. J. Stat.*, 41(3):557–579.

[12] Bolin, D., Kirchner, K., and Kovács, M. (2019). Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *Available at arXiv:1705.06565*.

[13] Chan, G. and Wood, A. T. A. (2000). Increment-based estimators of fractal dimension for two-dimensional surface data. *Stat. Sinica*, 10:343–376.

[14] Chong, C. and Klüppelberg, C. (2015). Integrability conditions for space-time stochastic integrals: theory and applications. *Bernoulli*, 21(4):2190–2216.

[15] Cohen, S., Lacaux, C., and Ledoux, M. (2008). A general framework for simulation of fractional fields. *Stochastic Process. Appl.*, 118(9):1489–1517.

[16] Dietrich, C. R. and Newsam, G. N. (1993). A fast and exact method for multidimensional Gaussian stochastic simulations. *Water Resour. Res.*, 29(8):2861–2869.

[17] Gneiting, T., Ševčíková, H., and Percival, D. B. (2012). Estimators of fractal dimension: assessing the roughness of time series and spatial data. *Statist. Sci.*, 27(2):247–277.

[18] Gneiting, T., Ševčíková, H., Percival, D. B., Schlather, M., and Jiang, Y. (2006). Fast and exact simulation of large Gaussian lattice systems in $\mathbf{R^2}$: exploring the limits. *J. Comput. Graph. Statist.*, 15(3):483–501.

[19] Goff, J. A. and Jordan, T. H. (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *J. Geophys. Res.-Sol. Ea.*, 93(B11):13589–13608.

[20] Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika*, 93(4):989–995.

[21] Hansen, L., Thorarinsdottir, T., Ovcharov, E., Gneiting, T., and Richards, D. (2015). Gaussian random particles with flexible Hausdorff dimension. *Adv. in Appl. Probab.*, 47(2):307–327.

[22] Hansen, L. V. and Thorarinsdottir, T. L. (2013). A note on moving average models for Gaussian random fields. *Statist. Probab. Lett.*, 83(3):850–855.

[23] Heinrich, C. (2017). *Fine scale properties of ambit fields – limit theory and simulation.* PhD thesis, Aarhus University. Available at *data.math.au.dk/publications/phd/2017/math-phd-2017-ch.pdf.*

[24] Hellmund, G., Prokešová, M., and Jensen, E. B. V. (2008). Lévy-based Cox point processes. *Adv. in Appl. Probab.*, 40(3):603–629.

[25] Huang, W., Wang, K., Breidt, F. J., and Davis, R. A. (2011). A class of stochastic volatility models for environmental applications. *J. Time Series Anal.*, 32(4):364–377.

[26] Jensen, E. B. V., Jónsdóttir, K. Y., Schmiegel, J., and Barndorff-Nielsen, O. E. (2006). Spatio-temporal modelling-with a view to biological growth. *Mg. Stat. Pro.*, 107:47.

[27] Jónsdóttir, K. Y., Rønn-Nielsen, A., Mouridsen, K., and Jensen, E. B. V. (2013). Lévy-based modelling in brain imaging. *Scand. J. Stat.*, 40(3):511–529.

[28] Kallenberg, O. (2002). *Foundations of modern probability.* Springer-Verlag, New York Berlin Heidelberg, second edition.

[29] Kwapień, S. and Woyczyński, W. A. (1992). *Random series and stochastic integrals: single and multiple.* Birkhäuser Boston, Inc., Boston, MA.

[30] Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. B*, 73(4):423–498.

[31] Matérn, B. (1960). *Spatial Variation: Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations.* Statens skogsforskningsinstitut.

[32] Nguyen, M. and Veraart, A. E. D. (2017). Modelling spatial heteroskedasticity by volatility modulated moving averages. *Spatial Statistics*, 20:148–190.

[33] Rajput, B. S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Rel.*, 82(3):451–487.

[34] Rosiński, J. (1990). On series representations of infinitely divisible random vectors. *Ann. Probab.*, 18(1):405–430.

[35] Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Strokorb, K., Engelke, S., Martini, J., Ballani, F., Moreva, O., Auel, J., Menck, P. J., Gross, S., Ober, U., Christoph Berreth, Burmeister, K., Manitz, J., Ribeiro, P., Singleton, R., Pfaff, B., and R Core Team (2017). *RandomFields: Simulation and Analysis of Random Fields*. R package version 3.1.50.

[36] Ševčíková, H., Percival, D., and Gneiting, T. (2014). *fractaldim: Estimation of fractal dimensions*. R package version 0.8-4.

[37] Stein, M. L. (2012). *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media.

[38] Veraart, A. (2015). Stationary and multi-self-similar random fields with stochastic volatility. *Stochastics*, 87(5):848–870.

[39] Von Karman, T. (1948). Progress in the statistical theory of turbulence. *P. Natl. Acad. Sci. USA*, 34(11):530–539.

[40] Wallin, J. and Bolin, D. (2015). Geostatistical modelling using non-Gaussian Matérn fields. *Scand. J. Stat.*, 42(3):872–890.

[41] Wood, A. T. A. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *J. Comput. Graph. Stat.*, 3(4):409–432.

[42] Yan, J. (2007). Spatial stochastic volatility for lattice data. *J. Agric. Biol. Environ. Stat.*, 12(1):25–40.