Infinitesimal perturbation analysis for risk measures based on the Smith max-stable random field

Erwan Koch* Christian Y. Robert†

December 14, 2018

Abstract

When using risk or dependence measures based on a given underlying model, it is essential to be able to quantify the sensitivity or robustness of these measures with respect to the model parameters. In this paper, we consider an underlying model which is very popular in spatial extremes, the Smith max-stable random field. We study the sensitivity properties of risk or dependence measures based on the values of this field at a finite number of locations. Max-stable fields play a key role, e.g., in the modelling of natural disasters. As their multivariate density is generally not available for more than three locations, the Likelihood Ratio Method cannot be used to estimate the derivatives of the risk measures with respect to the model parameters. Thus, we focus on a pathwise method, the Infinitesimal Perturbation Analysis (IPA). We provide a convenient and tractable sufficient condition for performing IPA, which is intricate to obtain because of the very structure of max-stable fields involving pointwise maxima over an infinite number of random functions. IPA enables the consistent estimation of the considered measures' derivatives with respect to the parameters characterizing the spatial dependence. We carry out a simulation study which shows that the approach performs well in various configurations.

Key words: Infinitesimal perturbation analysis; Max-stable random fields; Monte-Carlo computation; Risk assessment; Robustness; Smith random field.

1 Introduction

In a context of climate change, some extreme events tend to be more frequent, while a constant growth of population and wealth is observed. Hence both the insurance and reinsurance industries are more and more sensitive to natural disasters and need tools to quantify their impacts.

A first step to build such tools is to characterize the behaviour of the maxima of the relevant environmental variables at each point of the region under study. Due to the natural spatial extent of environmental variables, max-stable random fields are ideally suited for modelling purposes and play an important role as spatial fields for extreme

*EPFL (Chair of Statistics STAT): EPFL-SB-MATH-STAT, MA B1 433 (Bâtiment MA), Station 8, 1015 Lausanne, Switzerland.
Email: erwan.koch@epfl.ch

†ISFA Université Lyon 1, 50 Avenue Tony Garnier, 69366 Lyon cedex 07, France.
Email: christian.robert@univ-lyon1.fr

events; see, e.g., Davison et al. (2012). A second step is to use appropriate risk or dependence measures to provide accurate views of the insured risks. The complex structure of max-stable random fields as the pointwise maximum over an infinite number of random functions makes however estimation highly non-trivial. Since the multivariate density function is in general not available for more than two or three locations, standard likelihood-based estimation methods cannot be applied. Composite likelihood estimators are therefore often considered, but they are not asymptotically efficient (in the sense of the Cramér-Rao bound). Hence, compared to efficient estimators, they are more likely to lead to estimates that are far from the true value. It is therefore essential to assess the robustness and sensitivity of the risk or dependence measures with respect to the parameters of the considered spatial model.

Sensitivity Analysis (SA) consists in the quantitative assessment of how changes in a specific model parameter impacts the model output, which in the context of this paper is a risk or dependence measure. Such an analysis is necessary for discovering which parameters are important and which ones are less. If the risk or dependence measure is excessively sensitive to some critical parameters, it should warn the decision maker not to be overly confident in the value of that measure, especially if there is a huge uncertainty on these parameters. It also informs him/her that energy should be invested to find more reliable estimates of these parameters (e.g., by developing new inference methods). For an excellent introduction to SA, see Asmussen and Glynn (2010), Chapter VII.

One aim of SA is to compute the derivatives of the output of a model with respect to its parameters. We assume in this paper that this output can be written as an expectation. We are interested in this study in the case where the risk or dependence measure considered does not have any closed-form, which entails that the derivative cannot be assessed using differentiation or finite differences based on the analytical expression. In such a situation, there are essentially three approaches to estimate the derivatives, all based on simulations. The first one is based on finite-differences. Corresponding estimators involve a bias-variance trade-off and require simulating at multiple parameter values, making this methodology less powerful than the two others. The second one, called Infinitesimal Perturbation Analysis (IPA), relies on computing the derivative with respect to the parameter of each simulated value (and then averaging them). For this reason, IPA is also referred to as a sample path differentiation or pathwise method. The third approach, referred to as Likelihood Ratio Method (LRM), is based on the derivatives of the density associated with the simulation model. Contrary to IPA where the parameter is considered as purely structural, in LRM it is viewed as a parameter of the probability measure. LRM requires the existence of a density which has an explicit expression. Moreover, corresponding estimators often have a larger (or even much larger) variance than IPA estimators (see, e.g., Glasserman, 2013, Section 7.4). For these reasons IPA is generally considered as the best derivative estimator. However, performing IPA is not always possible as it requires that the derivative and the expectation can be interchanged, condition which may be difficult to check.

A huge literature is dedicated to IPA; excellent references are Heidelberger et al. (1988), Glasserman (1991a,b), L'Ecuyer (1991), Asmussen and Glynn (2010), Chapter VII and Glasserman (2013), Chapter 7. IPA (and more generally SA) is widely applied to many fields, for instance to finance where many risk hedging strategies involve computing sensitivities of option prices to the underlying assets' prices and other parameters (the so-called Greeks). Broadie and Glasserman (1996) and Chen and Fu (2001) use IPA for option pricing and mortgage-backed securities, respectively. Motivated by financial applications, Glasserman and Liu (2010) investigate LRM in the case where the relevant densities are only known through their characteristic functions or Laplace transforms. For applications of IPA to other fields such as queue models and stochastic fluid models, see, e.g., Adams (2007).

Our work is motivated by applications to insurance/reinsurance of losses triggered by extreme events having a spatial extent, typically such as weather events. Possible finance-oriented applications concern, e.g., the assessment of sensitivities of the prices of event-linked securities such as CAT bonds. To the best of our knowledge, IPA and more generally SA have not yet been explored in the case of extreme-value models. We consider the Smith max-stable random field (Smith, 1990), also called Gaussian extreme-value random field, and focus on the derivatives with respect to its spatial dependence parameters. LRM is ruled out when the measure considered is based on values of the Smith field at more than three locations as the multivariate density of the latter does not have any analytical expression. Our main contribution is to give a convenient sufficient condition for performing IPA for risk or dependence measures based on the values of the Smith random field at a finite number of locations. This involves showing that the paths of this field are differentiable and that the previously mentioned interchange between derivative and expectation is feasible. This is arduous as max-stable fields arise as the pointwise maxima over an infinite number of fields. Our condition is tractable because it implicates the derivative of the measure with respect to the field values but not to the spatial dependence parameters. Then, we implement the IPA method in a concrete case by adapting existing simulation algorithms and carry out a numerical study which shows that the method performs well.

The remainder of the paper is organized as follows. Section 2 presents the Smith max-stable random field, gives an example of dependence measure and most importantly details our main results. Then, we provide in Section 3 a simulation study. Finally, Section 4 summarizes our main findings and raises some open questions. The proofs of the main results are gathered in the Appendix in order to facilitate the reading.

Throughout the paper, we shall use the following notations. Let $'$ denote transposition. Moreover, for $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}^d$, $\|\mathbf{x}\|^2 = \mathbf{x}'\mathbf{x} = \sum_{i=1}^d x_i^2$. Finally, for a matrix $A \in \mathbb{R}^{d\times d}$, $\|A\| = \sup\{\|A\mathbf{x}\| : \mathbf{x} \in \mathbb{R}^d \text{ such that } \|\mathbf{x}\| = 1\}$ and for a positive definite symmetric matrix $\Sigma \in \mathbb{R}^{d\times d}$, $\|\mathbf{x}\|^2_{\Sigma^{-1}} = \mathbf{x}'\Sigma^{-1}\mathbf{x}$.

2 Main results

2.1 The Smith random field

Let us consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and define a Poisson point process on $(0,\infty) \times \mathbb{R}^d$, $(\xi_i, \mathbf{c}_i)_{i\geq 1}$, with intensity function $\xi^{-2}\nu(\mathrm{d}\xi) \times \nu(\mathrm{d}\mathbf{c})$, where ν denotes the Lebesgue measure. Let f_Σ designate the probability density function of a d-variate Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Σ. The field $\{Y_\Sigma(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^d}$, defined as

$$Y_\Sigma(\mathbf{x}) = \bigvee_{i=1}^{\infty} \{\xi_i f_\Sigma(\mathbf{x} - \mathbf{c}_i)\}, \quad \mathbf{x} \in \mathbb{R}^d, \tag{1}$$

is referred to as the Smith random field with covariance matrix Σ (Smith, 1990). It is a stationary max-stable random field whose univariate marginal distribution is standard Fréchet. The matrix Σ completely characterizes the dependence structure of the spatial field and hence is essential for a proper risk assessment. We denote by Z_Σ the logarithm of Y_Σ, i.e., $Z_\Sigma(\mathbf{x}) = \log Y_\Sigma(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$.

The Smith random field can be simulated exactly at a finite number of locations. For instance, Dombry et al. (2016) provide an algorithm based on simulating only those functions in the representation of the max-stable field that contribute to the pointwise maximum. Other simulation methods are available. For instance, the one by Schlather (2002) provides very accurate approximations.

2.2 Risk and dependence measures

A univariate risk measure is a mapping from a set of random variables to the real numbers. A dependence measure summarises the strength of dependence between several elements of such a set of random variables. In finance, these random variables often represent portfolio returns, and in an insurance context, they might be the claims associated with insurance policies. Our random variables will consist in the values of the Smith random field at a finite number of locations. Let $\mathbf{x}_1, \dots, \mathbf{x}_M$ be locations in $\mathbb{R}^d$. We consider risk or dependence measures of the form

$$R(\Sigma) = \mathbb{E}\left[H_M\left(\mathbf{Z}_\Sigma\right)\right], \tag{2}$$

where H_M is a function from $\mathbb{R}^M$ to $\mathbb{R}$ and $\mathbf{Z}_\Sigma = (Z_\Sigma(\mathbf{x}_1), \dots, Z_\Sigma(\mathbf{x}_M))'$. As they are written as expectations, these measures can easily be computed using a Monte-Carlo approach by simulating the Smith random field. It is worth mentioning that $R(\Sigma)$ does not depend on Σ explicitly but only implicitly through the field Z_Σ.

Many insightful dependence or risk measures can be written as in (2) and we now present some of them with a special focus on actuarial science. First, note that the Smith field Y_Σ defined in (1) has unit Fréchet margins which are not realistic for environmental variables. Accordingly, transformations of the field are generally considered. Moreover, when modelling damage, it is generally necessary to apply a so-called vulnerability function to the field representing the environmental phenomenon. For instance, considering powers of max-stable fields is useful when quantifying the impact of extreme wind speeds; see Koch (2018) and references therein for details.

A toy but relevant example of dependence measure between what happens at two locations $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$ is the covariance (or the correlation) between powers of the Smith random field at $\mathbf{x}_1$ and $\mathbf{x}_2$ (Koch, 2018), i.e.,

$$\operatorname{Cov}\left(Y_\Sigma^{\beta_1}(\mathbf{x}_1), Y_\Sigma^{\beta_2}(\mathbf{x}_2)\right). \tag{3}$$

Covariance and correlation are used a lot by practitioners in the finance/insurance industry. We have to impose the condition $\beta_1, \beta_2 < 1/2$ to ensure the existence of (3). The measure in (3) is of the form (2) with $M = 2$. Denoting Γ the gamma function, we have $\mathbb{E}[Y_\Sigma^\beta(\mathbf{x})] = \Gamma(1-\beta)$ for any $\mathbf{x} \in \mathbb{R}^d$ and $\beta < 1$. Consequently, we obtain

$$\operatorname{Cov}\left(Y_\Sigma^{\beta_1}(\mathbf{x}_1), Y_\Sigma^{\beta_2}(\mathbf{x}_2)\right) = \mathbb{E}\left[H_2\left(\mathbf{Z}_\Sigma\right)\right],$$

where

$$H_2(\mathbf{z}) = \exp\left(\beta_1 z_1 + \beta_2 z_2\right) - \Gamma\left(1-\beta_1\right)\Gamma\left(1-\beta_2\right), \quad \mathbf{z} = (z_1, z_2)' \in \mathbb{R}^2.$$

Many more sophisticated examples arise for instance in insurance/reinsurance pricing or regulation. Premium loadings that are proportional to specific moments of the sum of the values of Y_{Σ}^{β} at two or more locations constitute excellent examples in insurance pricing. In reinsurance, the premium is sometimes based on order statistics of the claims, as in the case of the ECOMOR or LCR treaties. Let us consider, e.g., $M = 20$ locations and assume that each of those is associated with an insurance policy. For $j = 1, \ldots, M$, let $\beta_j < 1$ and $Y_j = Y_{\Sigma}^{\beta_j}(\mathbf{x}_j)$ and consider their ordered values $Y^{(1:M)} \geq Y^{(2:M)} \geq \cdots \geq Y^{(M:M)}$. For instance, the risk measure $\mathbb{E}[(Y^{(1:M)} - Y^{(3:M)}) + (Y^{(2:M)} - Y^{(3:M)})]$ would be involved in the pricing of an ECOMOR reinsurance treaty having the third largest claim as priority. Finally, from a solvency viewpoint, a useful example could be $\mathrm{ES}_{\alpha}(\sum_{j=1}^{M} Y_j)$, $\alpha \in (0,1)$, where ES_{α} stands for Expected Shortfall at level α. All these quantities can be written under the form (2) but behave in a non-linear way and do not have any analytical expression.

2.3 Main contribution

For reasons explained in the introduction, we are interested in the computation of the derivative (i.e., the sensitivity) of measures $R(\Sigma)$ written as in (2) with respect to Σ at some positive definite matrix Σ_0. It is defined by (see, e.g., Dwyer, 1967)

$$\left.\frac{\partial R(\Sigma)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = \left(\left.\frac{\partial R(\Sigma)}{\partial \sigma_{i,j}}\right|_{\Sigma=\Sigma_0} \right)_{i,j}, \tag{4}$$

where $\Sigma = (\sigma_{i,j})_{i,j}$. Assume that the function $\Sigma \mapsto H_M(\mathbf{Z}_{\Sigma})$ appearing in $R(\Sigma)$ is differentiable at Σ_0. Our aim is to be able to evaluate the derivative (4) using IPA. This is possible as soon as we can interchange the derivative and the expectation, i.e., the equality

$$\left.\frac{\partial R(\Sigma)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = \mathbb{E}\left[\left.\frac{\partial H_M(\mathbf{Z}_{\Sigma})}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} \right] \tag{5}$$

is satisfied. A sufficient condition for this in the general case is given, e.g., in Asmussen and Glynn (2010), Chapter VII, Proposition 2.3. This result is formulated immediately below in the case where the parameter is a matrix.

Proposition 1. *Assume that $\Sigma \mapsto H_M(\mathbf{Z}_{\Sigma})$ is an almost surely (a.s.) differentiable function at Σ_0 and that a.s. $\Sigma \mapsto H_M(\mathbf{Z}_{\Sigma})$ satisfies the Lipschitz condition*

$$|H_M(\mathbf{Z}_{\Sigma_1}) - H_M(\mathbf{Z}_{\Sigma_2})| \leq \|\Sigma_1 - \Sigma_2\| B_{\Sigma_0}$$

for Σ_1, Σ_2 in a non-random neighbourhood of Σ_0, where $\mathbb{E}[B_{\Sigma_0}] < \infty$. Then (5) holds.

If $f : I \to \mathbb{R}$ is differentiable on an open set $I \subset \mathbb{R}^d$, and satisfies $\|\partial f(\mathbf{x})/\partial \mathbf{x}\| \leq K$ for all $\mathbf{x}$ in I, then f is Lipschitz continuous with Lipschitz constant at most K over I. Therefore, we immediately deduce that, if there exists a random variable B_{Σ_0} satisfying $\mathbb{E}[B_{\Sigma_0}] < \infty$ and such that a.s.

$$\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial H_M(\mathbf{Z}_{\Sigma})}{\partial \Sigma} \right\| \leq B_{\Sigma_0},$$

where $\mathcal{V}_{\Sigma_0}$ is a non-random neighbourhood of Σ_0, then (5) holds. Results concerning differentiation with respect to a scalar or a vector also hold in the case of differentiation with respect to a matrix.

Assume now that $\mathbf{z} \mapsto H_M(\mathbf{z})$ is differentiable. The differentiability of the function $\Sigma \mapsto \mathbf{Z}_\Sigma$ in a neighbourhood of Σ_0 will be shown in Theorem 2 (Appendix A). By the chain rule, we have that

$$\frac{\partial H_M(\mathbf{Z}_\Sigma)}{\partial \Sigma} = \sum_{j=1}^{M} \frac{\partial H_M(\mathbf{Z}_\Sigma)}{\partial z_j} \frac{\partial Z_\Sigma(\mathbf{x}_j)}{\partial \Sigma}. \tag{6}$$

We shall prove (Theorem 3 in Appendix A) that there exists some non-random neighbourhood of Σ_0, $\mathcal{V}_{\Sigma_0}$, such that, for any $q > 1$, there exists a random variable $C_{\Sigma_0}(\mathbf{x}, q)$ satisfying a.s.

$$\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma} \right\|^q \leq C_{\Sigma_0}(\mathbf{x}, q) \quad \text{and} \quad \mathbb{E}\left[C_{\Sigma_0}(\mathbf{x}, q)\right] < \infty.$$

This technical outcome will allow us to derive our main result:

Theorem 1. *Assume that $\mathbf{z} \mapsto H_M(\mathbf{z})$ is a differentiable function and that there exists $p > 1$ such that*

$$\sup_{j=1,\dots,M} \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left| \frac{\partial H_M(\mathbf{Z}_\Sigma)}{\partial z_j} \right|^p \right] < \infty, \tag{7}$$

where $\mathcal{V}_{\Sigma_0}$ is a non-random neighbourhood of Σ_0. Then (5) *holds.*

This non-trivial theorem provides a sufficient condition to use IPA to compute the derivatives of any risk or dependence measure written as in (2), or even as a differentiable function of the right-hand side of (2). This condition is much more tractable and easier to check than that in Proposition 1. The simplification stems from the fact that we take care in Theorems 2 and 3 of the intricate term of (6), $\partial Z_\Sigma(\mathbf{x}_j)/\partial\Sigma$, which involves the sample path properties with respect to differentiation of the Smith field. Theorems 2 and 3 are delicate to establish precisely due to the inherent structure of max-stable fields.

Remark 1. *The analytical computation of the terms $\partial Z_\Sigma(\mathbf{x}_j)/\partial\Sigma$, $j = 1, \dots, M$, which is necessary to implement IPA (see* (6)*), requires the coordinates of the centers of the "storms" (see Smith, 1990, for the interpretation of the Smith field in terms of storms) realizing the maxima at the locations $\mathbf{x}_j$ (see* (15) *in Appendix A). To the best of our knowledge, these quantities cannot be obtained from the simulation algorithms available on the Web (e.g., in R packages like SpatialExtremes by Ribatet et al. (2018) or in the code by Dombry et al. (2016) available on the Biometrika website). To overcome this impediment and for other technical reasons, we programmed the simulation algorithm of the Smith random field ourselves by adapting the approach of Schlather (2002). Corresponding code will be provided.*

Remark 2. *The Smith random field in* (1) *has standard Fréchet margins, which makes us focus on the spatial dependence parameters in the covariance matrix Σ. In concrete applications, one should also take care of the marginal parameters at each location, i.e., the three parameters of the generalized extreme-value distribution. Showing that it is possible under mild conditions to perform IPA to estimate the derivatives with respect to these parameters is easy and thus not considered in this paper.*

3 Numerical study

In this section, we numerically assess the accuracy of the IPA method in the case of the particular dependence measure (3) for $d = 2$. Let $\{Z_\Sigma(\mathbf{x})\}_{\mathbf{x}\in\mathbb{R}^2}$ be the Smith random field with covariance matrix Σ and consider two locations $\mathbf{x}_1$ and $\mathbf{x}_2 \in \mathbb{R}^2$. Let $\mathbf{h} = (\mathbf{x}_2 - \mathbf{x}_1)'$ and $h_\Sigma = \sqrt{\mathbf{h}'\Sigma^{-1}\mathbf{h}}$. Let $\beta_1, \beta_2 < 1/2$ and Σ_0 be a symmetric positive definite matrix. Since

$$\operatorname{Cov}\left(Y_\Sigma^{\beta_1}(\mathbf{x}_1), Y_\Sigma^{\beta_2}(\mathbf{x}_2)\right) = \mathbb{E}\left[Y_\Sigma^{\beta_1}(\mathbf{x}_1)Y_\Sigma^{\beta_2}(\mathbf{x}_2)\right] - \Gamma(1-\beta_1)\Gamma(1-\beta_2),$$

when considering the derivatives of (3), it is equivalent to focus on

$$R_e(\Sigma) = \mathbb{E}\left[Y_\Sigma^{\beta_1}(\mathbf{x}_1)Y_\Sigma^{\beta_2}(\mathbf{x}_2)\right]. \tag{8}$$

Below we compare the values of $\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0}$ obtained using IPA with their true values. Hereafter, we show that IPA can be used for the dependence measure (8), explain how to obtain the true values, and compare both approaches in different configurations.

3.1 Validity of IPA

Observe that

$$R_e(\Sigma) = \mathbb{E}\left[H_2\left(\mathbf{Z}_\Sigma\right)\right],$$

with

$$H_2\left(\mathbf{z}\right) = \exp\left(\beta_1 z_1 + \beta_2 z_2\right), \quad \mathbf{z} = (z_1, z_2)' \in \mathbb{R}^2.$$

We first show that Condition (7) of Theorem 1 is satisfied, entailing that the derivative of $R(\Sigma)$ with respect to Σ can be computed using IPA. We have

$$\frac{\partial H_2\left(\mathbf{z}\right)}{\partial z_i} = \beta_i H_2\left(\mathbf{z}\right), \quad i = 1, 2, \tag{9}$$

which implies, since $\beta_1, \beta_2 < 1/2$ and Y_Σ is positive, that, for $p > 1$,

$$\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left|\frac{\partial H_2\left(\mathbf{Z}_\Sigma\right)}{\partial z_i}\right|^p \leq \frac{1}{2} \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} Y_\Sigma^{p\beta_1}(\mathbf{x}_1) \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} Y_\Sigma^{p\beta_2}(\mathbf{x}_2). \tag{10}$$

Next result will allow us to prove that Condition (7) holds.

Proposition 2. *There exists a non-random neighbourhood of Σ_0, $\mathcal{V}_{\Sigma_0}$, such that, for any $\beta < 1$ and $\mathbf{x} \in\mathbb{R}^d$,*

$$\mathbb{E}\left[\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} Y_\Sigma^\beta(\mathbf{x})\right] < \infty.$$

Proof. We have

$$Y_\Sigma(\mathbf{x}) = \bigvee_{i=1}^\infty \xi_i f_\Sigma(\mathbf{x} - \mathbf{c}_i) \leq \frac{1}{\pi^{d/2}\det\left(\Sigma\right)^{1/2}} \bigvee_{i=1}^\infty \xi_i$$

and thus

$$\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} Y_\Sigma^\beta(\mathbf{x}) \leq \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left(\frac{1}{\pi^{d/2}\det\left(\Sigma\right)^{1/2}}\right)^\beta \left(\bigvee_{i=1}^\infty \xi_i\right)^\beta.$$

It is well-know that $\bigvee_{i=1}^{\infty} \xi_i$ has a standard Fréchet distribution and therefore $\mathbb{E}[(\bigvee_{i=1}^{\infty} \xi_i)^{\beta}] < \infty$ since $\beta < 1$. We deduce that it suffices to choose $\mathcal{V}_{\Sigma_0}$ such that $\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \det(\Sigma)^{-\beta/2} < \infty$. This is possible since any invertible matrix admits a neighbourhood of invertible matrices. □

By Proposition 2, (10) and the Hölder inequality, there exists a non-random neighbourhood of Σ_0, $\mathcal{V}_{\Sigma_0}$, such that for some $p > 1$ satisfying $2p\beta_1 < 1$ and $2p\beta_2 < 1$, we have

$$\mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left|\frac{\partial H_2(\mathbf{Z}_\Sigma)}{\partial z_i}\right|^p\right] \leq \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} Y_\Sigma^{2p\beta_1}(\mathbf{x}_1)\right]^{1/2} \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} Y_\Sigma^{2p\beta_2}(\mathbf{x}_2)\right]^{1/2} < \infty.$$

We can therefore apply Theorem 1 and obtain

$$\left.\frac{\partial R_e(\Sigma)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = \mathbb{E}\left[\left.\frac{\partial H_2(\mathbf{Z}_\Sigma)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0}\right].$$

It follows from the combination of (6) and (9) that

$$\left.\frac{\partial H_2(\mathbf{Z}_\Sigma)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = \beta_1 H_2(\mathbf{Z}_\Sigma) \left.\frac{\partial Z_\Sigma(\mathbf{x}_1)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} + \beta_2 H_2(\mathbf{Z}_\Sigma) \left.\frac{\partial Z_\Sigma(\mathbf{x}_2)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0}, \tag{11}$$

where $\partial Z_\Sigma(\mathbf{x}_1)/\partial \Sigma|_{\Sigma=\Sigma_0}$ and $\partial Z_\Sigma(\mathbf{x}_2)/\partial \Sigma|_{\Sigma=\Sigma_0}$ can be computed using (15) (see Appendix A). Computing (11) for S realizations of $\mathbf{Z}_\Sigma$ (obtained by simulation) and taking the empirical mean provides an estimate of $\partial R(\Sigma)/\partial \Sigma|_{\Sigma=\Sigma_0}$. In different configurations, we repeat this procedure 100 times and compute the mean estimate, the standard deviation as well as the relative error of the mean estimate.

3.2 True values of the derivative

Let Φ and ϕ denote the standard Gaussian distribution and density functions, respectively. It directly follows from Theorem 2 in Koch (2018) that, provided $h_\Sigma > 0$,

$$\begin{aligned} R_e(\Sigma) = \int_0^\infty & \theta^{\beta_2}[C_2(\theta, h_\Sigma) C_1(\theta, h_\Sigma)^{\beta_1+\beta_2-2}\Gamma(2-\beta_1-\beta_2) \\ & + C_3(\theta, h_\Sigma) C_1(\theta, h_\Sigma)^{\beta_1+\beta_2-1}\Gamma(1-\beta_1-\beta_2)]\nu(\mathrm{d}\theta), \end{aligned} \tag{12}$$

where

$$\begin{aligned} C_1(\theta, h_\Sigma) &= \Phi\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right) + \frac{1}{\theta}\Phi\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right), \\ C_2(\theta, h_\Sigma) &= \left(\Phi\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right) + \frac{1}{h_\Sigma}\phi\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right) - \frac{1}{h_\Sigma\theta}\phi\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right)\right) \\ & \quad \times \left(\frac{1}{\theta^2}\Phi\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right) + \frac{1}{h_\Sigma\theta^2}\phi\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right) - \frac{1}{h_\Sigma\theta}\phi\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right)\right), \\ C_3(\theta, h_\Sigma) &= \frac{1}{h_\Sigma^2\theta}\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right)\phi\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right) + \frac{1}{h_\Sigma^2\theta^2}\left(\frac{h_\Sigma}{2} + \frac{\log(\theta)}{h_\Sigma}\right)\phi\left(\frac{h_\Sigma}{2} - \frac{\log(\theta)}{h_\Sigma}\right). \end{aligned}$$

Therefore, we obtain from (12) that

$$\left.\frac{\partial R_e\left(\Sigma\right)}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = \int_0^\infty \frac{\partial}{\partial \Sigma}\Big(\theta^{\beta_2}[C_2(\theta,h_\Sigma)C_1(\theta,h_\Sigma)^{\beta_1+\beta_2-2}\Gamma(2-\beta_1-\beta_2)$$
$$+ C_3(\theta,h_\Sigma)C_1(\theta,h_\Sigma)^{\beta_1+\beta_2-1}\Gamma(1-\beta_1-\beta_2)]\Big)\Big|_{\Sigma=\Sigma_0}\nu(\mathrm{d}\theta). \qquad (13)$$

Then, we compute the term

$$\frac{\partial}{\partial \Sigma}\Big(\theta^{\beta_2}[C_2(\theta,h_\Sigma)C_1(\theta,h_\Sigma)^{\beta_1+\beta_2-2}\Gamma(2-\beta_1-\beta_2)+C_3(\theta,h_\Sigma)C_1(\theta,h_\Sigma)^{\beta_1+\beta_2-1}\Gamma(1-\beta_1-\beta_2)]\Big)\Big|_{\Sigma=\Sigma_0}$$

and obtain a closed expression. Owing to its complexity, we did not include it in the paper but it is available upon request. Finally, we obtain the true values of $\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0}$ by computing the right-hand side of (13) using adaptive quadrature with a sufficiently high number of sub-intervals N.

3.3 Results

We choose the locations $\mathbf{x}_1 = (0,0)'$ and $\mathbf{x}_2 = (1,1)'$. Additionally, we consider the covariance matrices Σ_0

$$\Sigma_{0,1} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad \Sigma_{0,2} = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.3 \end{pmatrix}, \quad \Sigma_{0,3} = \begin{pmatrix} 0.2 & 0.15 \\ 0.15 & 0.3 \end{pmatrix},$$

in order to assess the impact of the magnitude of the diagonal terms as well as of non-zero off-diagonal terms. We take $\beta_1 = \beta_2 = \beta$ with $\beta = 0.1, 0.25, 0.45$. Furthermore, we consider both the cases $S = 10^5$ and $S = 10^6$ and we take $N = 10^8$ for the adaptive quadrature. Such a value of N ensures a very accurate approximation of the integral appearing in (13).

We first comment the results related to the cases $\beta = 0.1$ and $\beta = 0.25$. Tables 1–3 seem to indicate that the IPA estimator is always unbiased. Note that the mean estimates correspond to the values we would have obtained if we had done only one Monte-Carlo estimation with 10^7 and 10^8 simulations, in the cases $S = 10^5$ and $S = 10^7$, respectively. Moreover, as expected, the standard deviation decreases when increasing S from 10^5 to 10^6. For a given β, the standard deviation is generally lower in the case of $\Sigma_{0,1}$ than $\Sigma_{0,2}$ and $\Sigma_{0,3}$. However, the ratio between the standard deviation and the mean estimate is lower for $\Sigma_{0,2}$ and $\Sigma_{0,3}$ than for $\Sigma_{0,1}$. Thus, the estimation is more accurate in these cases, likely due to larger values of the derivatives. Moreover, the standard errors as well as the ratios standard errors over mean estimates are larger for $\Sigma_{0,3}$ than $\Sigma_{0,2}$. Again, this is probably due to the larger derivatives associated with $\Sigma_{0,2}$. More generally, we expect the accuracy of the estimation to be an increasing function of the value of the derivatives: in some sense, the signal-to-noise ratio in the data increases. Their order of magnitude seems to depend on the relative values of the diagonal coefficients of Σ_0 compared to the distance between the two locations considered. Furthermore, the introduction of a non-zero off-diagonal decreases the derivatives in our situation, but this effect might be different for other choices of $\mathbf{x}_1$ and $\mathbf{x}_2$ (we expect the main factor to be the angle between the line joining $\mathbf{x}_1$ and $\mathbf{x}_2$ and the directions of the eigenvectors of Σ_0). Finally, as awaited, the standard deviations increase when raising β from 0.1 to 0.25. Nevertheless, the ratio standard deviation over mean estimate does not necessarily decrease since the derivatives increase. There is an interesting trade-off: increasing β increases the variability and makes the tails heavier (and at some point invalidates the central limit theorem) but at the same time increases the derivatives and thus makes them easier to assess.

In the case $\beta = 0.45$, the Monte-Carlo estimator has no variance, which explains that the standard deviation often increases from the case $S = 10^5$ to the case $S = 10^6$. The central limit theorem does not hold but the Monte-Carlo estimator is asymptotically unbiased. For instance, in the cases of $\Sigma_{0,2}$ and $\Sigma_{0,3}$, the relative errors on the term $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{1,1}$ are 4.1% and −2.4%, respectively. But in other cases, a very high Monte-Carlo estimate (coming from the fact that $\partial H_2(\mathbf{Z}_\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0}$ is heavy-tailed) dramatically deteriorates the mean estimate. We should increase the number of simulations S to avoid such a bias.

We do not display the table corresponding to the coefficient $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{1,2}$, which is exactly the same as for the coefficient $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{2,1}$ (Table 2), consistently with the theory.

This application of IPA to the dependence measure (8), although very challenging due to the low values of the derivatives, is highly conclusive. We have also tested this approach in the case of the Brown–Resnick field (Kabluchko et al., 2009) with a power variogram (more general max-stable field than the Smith random field) with different powers and the IPA method also performs well. As shown in Koch (2018), Theorem 2, in the case of the Brown–Resnick field, the dependence measure (8) is also written as in (12) when replacing the Mahalanobis distance h_Σ with the square root of the Brown-Resnick field's variogram. Consequently, the analytical derivative is obtained similarly. In situations where the derivatives are very low (typically as in the case of $\Sigma_{0,1}$ in our example), appropriate improvements of the classical Monte-Carlo method (such as importance sampling) might be necessary to reduce the estimator's variability.

Σ_0	β	**Analytical**	**Mean estimate**		**Standard deviation**		**Relative error**	
			$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$
$\Sigma_{0,1}$	0.1	9.1×10^{-4}	9.1×10^{-4}	9.0×10^{-4}	2.4×10^{-4}	7.3×10^{-5}	-3.2×10^{-3}	-1.4×10^{-2}
$\Sigma_{0,1}$	0.25	9.9×10^{-3}	9.8×10^{-3}	9.9×10^{-3}	2.0×10^{-3}	6.7×10^{-4}	-6.6×10^{-3}	-3.2×10^{-3}
$\Sigma_{0,1}$	0.45	2.1×10^{-1}	4.5×10^{-2}	1.8×10^{-1}	5.5×10^{-1}	4.6×10^{-1}	-7.8×10^{-1}	-1.3×10^{-1}
$\Sigma_{0,2}$	0.1	2.1×10^{-2}	2.1×10^{-2}	2.0×10^{-2}	2.0×10^{-3}	5.2×10^{-4}	4.3×10^{-3}	-4.1×10^{-3}
$\Sigma_{0,2}$	0.25	2.7×10^{-1}	2.7×10^{-1}	2.7×10^{-1}	1.6×10^{-2}	6.7×10^{-3}	-3.9×10^{-3}	7.4×10^{-4}
$\Sigma_{0,2}$	0.45	7.5×10^{0}	5.2×10^{0}	7.8×10^{0}	3.5×10^{0}	1.9×10^{1}	-3.1×10^{-1}	4.1×10^{-2}
$\Sigma_{0,3}$	0.1	2.2×10^{-2}	2.2×10^{-2}	2.2×10^{-2}	3.6×10^{-3}	9.2×10^{-4}	3.1×10^{-3}	-6.9×10^{-3}
$\Sigma_{0,3}$	0.25	2.7×10^{-1}	2.7×10^{-1}	2.7×10^{-1}	2.9×10^{-2}	8.8×10^{-3}	4.2×10^{-4}	-9.4×10^{-4}
$\Sigma_{0,3}$	0.45	7.1×10^{0}	5.4×10^{0}	6.9×10^{0}	6.3×10^{0}	1.2×10^{1}	-2.4×10^{-1}	-2.4×10^{-2}

Table 1: This table concerns the coefficient $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{1,1}$. Each line gives the results for a specific combination of the matrix Σ_0 and the value of β. The column entitled "Analytical" gives the true value computed with (13). Columns "Mean estimate" and "Standard deviation" provide the mean and standard deviation of the 100 estimates obtained by IPA, respectively. Column "Relative error" gives the ratio of the gap between the mean estimate and the true value over the true value. Mean estimate, standard deviation and relative error are given in both cases $S = 10^5$ and $S = 10^6$.

Σ_0	β	**Analytical**	**Mean estimate**		**Standard deviation**		**Relative error**	
			$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$
$\Sigma_{0,1}$	0.1	6.1×10^{-4}	6.1×10^{-4}	6.0×10^{-4}	1.3×10^{-4}	4.2×10^{-5}	6.4×10^{-3}	-1.3×10^{-2}
$\Sigma_{0,1}$	0.25	6.6×10^{-3}	6.5×10^{-3}	6.6×10^{-3}	1.1×10^{-3}	4.4×10^{-4}	-7.9×10^{-3}	-2.0×10^{-3}
$\Sigma_{0,1}$	0.45	1.4×10^{-1}	5.4×10^{-2}	2.7×10^{-1}	2.9×10^{-1}	1.5×10^{0}	-6.1×10^{-1}	9.9×10^{-1}
$\Sigma_{0,2}$	0.1	1.4×10^{-2}	1.4×10^{-2}	1.4×10^{-2}	1.2×10^{-3}	3.5×10^{-4}	6.5×10^{-3}	3.3×10^{-4}
$\Sigma_{0,2}$	0.25	1.8×10^{-1}	1.8×10^{-1}	1.8×10^{-1}	9.2×10^{-3}	4.7×10^{-3}	-1.7×10^{-3}	1.6×10^{-3}
$\Sigma_{0,2}$	0.45	5.0×10^{0}	3.5×10^{0}	5.8×10^{0}	2.0×10^{0}	1.9×10^{1}	-2.9×10^{-1}	1.7×10^{-1}
$\Sigma_{0,3}$	0.1	7.3×10^{-3}	7.5×10^{-3}	7.4×10^{-3}	2.3×10^{-3}	7.5×10^{-4}	2.5×10^{-2}	1.2×10^{-2}
$\Sigma_{0,3}$	0.25	9.1×10^{-2}	9.0×10^{-2}	9.2×10^{-2}	1.9×10^{-2}	6.7×10^{-3}	-1.0×10^{-2}	1.2×10^{-2}
$\Sigma_{0,3}$	0.45	2.4×10^{0}	1.1×10^{0}	5.1×10^{0}	5.4×10^{0}	2.9×10^{1}	-5.2×10^{-1}	1.2×10^{0}

Table 2: This table concerns the coefficient $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{2,1}$. For more explanations, see the caption of Table 1.

Σ_0	β	**Analytical**	**Mean estimate**		**Standard deviation**		**Relative error**	
			$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$	$S = 10^5$	$S = 10^6$
$\Sigma_{0,1}$	0.1	4.0×10^{-4}	4.1×10^{-4}	4.0×10^{-4}	1.5×10^{-4}	5.3×10^{-5}	5.8×10^{-3}	-6.3×10^{-3}
$\Sigma_{0,1}$	0.25	4.4×10^{-3}	4.5×10^{-3}	4.3×10^{-3}	1.2×10^{-3}	5.1×10^{-4}	2.9×10^{-2}	-1.2×10^{-2}
$\Sigma_{0,1}$	0.45	9.2×10^{-2}	7.8×10^{-2}	-9.3×10^{-2}	4.7×10^{-1}	1.4×10^{0}	-1.5×10^{-1}	-2.0×10^{0}
$\Sigma_{0,2}$	0.1	9.1×10^{-3}	9.0×10^{-3}	9.1×10^{-3}	1.3×10^{-3}	4.0×10^{-4}	-1.7×10^{-2}	-2.1×10^{-3}
$\Sigma_{0,2}$	0.25	1.2×10^{-1}	1.2×10^{-1}	1.2×10^{-1}	9.9×10^{-3}	3.4×10^{-3}	1.0×10^{-3}	-1.6×10^{-3}
$\Sigma_{0,2}$	0.45	3.3×10^{0}	2.6×10^{0}	2.6×10^{0}	3.5×10^{0}	1.8×10^{0}	-2.2×10^{-1}	-2.3×10^{-1}
$\Sigma_{0,3}$	0.1	2.4×10^{-3}	2.3×10^{-3}	2.4×10^{-3}	2.1×10^{-3}	7.6×10^{-4}	-6.4×10^{-2}	-1.1×10^{-2}
$\Sigma_{0,3}$	0.25	3.0×10^{-2}	3.2×10^{-2}	2.9×10^{-2}	1.7×10^{-2}	6.6×10^{-3}	3.8×10^{-2}	-3.8×10^{-2}
$\Sigma_{0,3}$	0.45	7.9×10^{-1}	9.0×10^{-1}	-2.7×10^{0}	5.4×10^{0}	2.9×10^{1}	1.4×10^{-1}	-4.4×10^{0}

Table 3: This table concerns the coefficient $(\partial R_e(\Sigma)/\partial\Sigma|_{\Sigma=\Sigma_0})_{2,2}$. For more explanations, see the caption of Table 1.

4 Discussion

It is essential in risk assessment to appraise the sensitivity of the considered risk and dependence measures with respect to the parameters of the underlying model. If applicable, IPA is a powerful technique enabling the estimation of the derivatives of a model output with respect to the model parameters. As explained above, max-stable fields are particularly relevant for modelling extreme spatial events. In this paper, we introduce the field of IPA to extreme-value theory by considering the Smith max-stable field, and especially give a tractable and convenient sufficient condition allowing IPA of any measure written as a differentiable function of the right-hand side of (2). This result is non-trivial due to the complex structure of max-stable fields. Then, we implement IPA on a concrete and challenging example of dependence measure and show that this approach performs well. Future appealing work consists in extending our results to other more general Mixed Moving Maxima (M3) random fields (as suggested by our convincing results on some examples of the Brown–Resnick field) and possibly even to max-stable fields which don't possess any M3 representation. Other interesting research directions might involve spotting appropriate refinements of the Monte-Carlo method (such as important sampling) in order to make IPA more efficient in the case of challenging risk or dependence measures such as the one studied above, where the derivatives are very low and therefore difficult to estimate.

Acknowledgements

Both authors gratefully acknowledge Anthony C. Davison for insightful comments. Erwan Koch would like to thank the Swiss National Science Foundation (project number 200021_178824) for financial support.

A Proof of Theorem 1

As in the definition of the Smith random field, let $(\xi_i, \mathbf{c}_i)_{i\geq 1}$ be the points of a Poisson point process on $(0,\infty)\times\mathbb{R}^d$ with intensity function $\xi^{-2}\lambda(\mathrm{d}\xi)\times\lambda(\mathrm{d}\mathbf{c})$. For $i\geq 1$, let $\varphi_{i,\Sigma}$ be the function from $\mathbb{R}^d$ to $\mathbb{R}$ defined by $\varphi_{i,\Sigma}(\mathbf{x}) = \xi_i f_\Sigma(\mathbf{x}-\mathbf{c}_i)$.

Let us begin by noting that for each $j = 1,...,M$, the supremum $Y_\Sigma(\mathbf{x}_j)$ is a.s. attained by a unique function $\varphi_{i,\Sigma}$ at $\mathbf{x}_j$.

Proposition 3. *Let $\mathbf{x}_1,\ldots,\mathbf{x}_M\in\mathbb{R}^d$, and define for $\mathbf{x}\in\mathbb{R}^d$,*

$$\mathcal{I}_{\mathbf{x}} = \left\{k : \bigvee_{i=1}^{\infty}\{\xi_i f_\Sigma(\mathbf{x}-\mathbf{c}_i)\} = \varphi_{k,\Sigma}(\mathbf{x})\right\}.$$

Then a.s., we have $\sharp\mathcal{I}_{\mathbf{x}_j} = 1$ for all $j\geq 1$, where $\sharp$ stands for the cardinal of a set.

Proof. The points $((\xi_i f_\Sigma(\mathbf{x}_j-\mathbf{c}_i))_{j=1,\ldots,M})_{i\geq 1}$ are the points of a Poisson point process on $\mathbb{R}^M$ with mean measure μ_M characterized by

$$\mu_M\left(\prod_{j=1}^M[y_j,\infty)\right) = \int_{\mathbb{R}^M}\max_{j=1,\ldots,M}\left\{\frac{f_\Sigma(\mathbf{x}_j-\mathbf{c})}{y_j}\right\}\nu(\mathrm{d}\mathbf{c}),\quad y_1>0,\cdots,y_M>0.$$

In particular, the joint cumulative distribution function of Y_Σ for the locations $\mathbf{x}_1,\ldots,\mathbf{x}_M\in\mathbb{R}^d$ is

$$\mathbb{P}(Y_\Sigma(\mathbf{x}_1)\leq y_1,...,Y_\Sigma(\mathbf{x}_M)\leq y_M) = \exp\left(-\int_{\mathbb{R}^M}\max_{j=1,\ldots,M}\left\{\frac{f_\Sigma(\mathbf{x}_j-\mathbf{c})}{y_j}\right\}\nu(\mathrm{d}\mathbf{c})\right);$$

see, e.g., Smith (1990).

Hence, the mean measure μ_M is finite on the sets of the form $\prod_{j=1}^M[y_j,\infty)$, $y_1,\ldots,y_M>0$, meaning that there is no accumulation of points around the component-wise maxima of the Poisson point process. This directly leads that the maxima are attained for random indexes which are a.s. unique. ☐

By Proposition 3, we can define, for $j = 1,...,M$, the a.s. unique indexes $i_{\mathbf{x}_j,\Sigma}$ satisfying

$$Y_\Sigma(\mathbf{x}_j) = \varphi_{i_{\mathbf{x}_j,\Sigma}}(\mathbf{x}_j).$$

We now fix $\omega\in\Omega$ and consider one realization of the point process $(\xi_i(\omega),\mathbf{c}_i)_{i\geq 1}$, denoted by $(\xi_i(\omega),\mathbf{c}_i(\omega))_{i\geq 1}$. For ease of exposition, we do not mention ω further below. Let us consider the set

$$\mathcal{B} = \left\{\mathbf{x}\in\mathbb{R}^d : \exists j,k\geq 1, j\neq k, \bigvee_{i=1}^{\infty}\{\xi_i f_\Sigma(\mathbf{x}-\mathbf{c}_i)\} = \varphi_{j,\Sigma}(\mathbf{x}) = \varphi_{k,\Sigma}(\mathbf{x})\right\},$$

i.e., the set of locations $\mathbf{x}\in\mathbb{R}^d$ for which the maximum $\bigvee_{i=1}^{\infty}\{\xi_i f_\Sigma(\mathbf{x}-\mathbf{c}_i)\}$ is attained by at least two distinct functions $\varphi_{j,\Sigma}$ and $\varphi_{k,\Sigma}$.

Remark 3. *If $\Sigma = Id_d$ (identity matrix with dimension $d\times d$), the set $\mathcal{B}$ is the set of boundaries of the cells of a Poisson Laguerre Tessellation (see, e.g., Dombry and Kabluchko, 2018).*

A key result is that $\mathcal{B}$ has a null Lebesgue measure.

Proposition 4. *The set $\mathcal{B}$ has an empty interior.*

Proof. A point $\mathbf{x} \in \mathcal{B}$ is characterized as follows:

$$\begin{aligned}
&\mathbf{x} \in \mathcal{B}\\
&\Leftrightarrow \exists j, k \geq 1, j \neq k : \xi_k f_\Sigma(\mathbf{x} - \mathbf{c}_k) = \xi_j f_\Sigma(\mathbf{x} - \mathbf{c}_j)\\
&\Leftrightarrow \exists j, k \geq 1, j \neq k : (\mathbf{x} - \mathbf{c}_k)' \Sigma^{-1} (\mathbf{x} - \mathbf{c}_k) - 2\log(\xi_k) = (\mathbf{x} - \mathbf{c}_j)' \Sigma^{-1} (\mathbf{x} - \mathbf{c}_j) - 2\log(\xi_j).
\end{aligned} \tag{14}$$

Let $\mathbf{h} \in \mathbb{R}^d$ such that $\mathbf{x} + \mathbf{h}$ is in a neighbourhood of $\mathbf{x}$ but still belongs to $\mathcal{B}$, i.e., such that

$$\|\mathbf{x} + \mathbf{h} - \mathbf{c}_k\|^2_{\mathbf{\Sigma^{-1}}} - 2\log(\xi_k) = \|\mathbf{x} + \mathbf{h} - \mathbf{c}_j\|^2_{\mathbf{\Sigma}^{-1}} - 2\log(\xi_j).$$

Using (14), we obtain that the previous equality is equivalent to

$$\begin{aligned}
&(\mathbf{x} + \mathbf{h} - \mathbf{c}_k)' \Sigma^{-1} (\mathbf{x} + \mathbf{h} - \mathbf{c}_k) - 2\log(\xi_k) = (\mathbf{x} + \mathbf{h} - \mathbf{c}_j)' \Sigma^{-1} (\mathbf{x} + \mathbf{h} - \mathbf{c}_j) - 2\log(\xi_j)\\
&\Leftrightarrow (\mathbf{x} - \mathbf{c}_k)' \Sigma^{-1} (\mathbf{x} - \mathbf{c}_k) + 2(\mathbf{x} - \mathbf{c}_k)' \Sigma^{-1} \mathbf{h} + \mathbf{h}' \Sigma^{-1} \mathbf{h} - 2\log(\xi_k) = (\mathbf{x} - \mathbf{c}_j)' \Sigma^{-1} (\mathbf{x} - \mathbf{c}_j)\\
&\qquad + 2(\mathbf{x} - \mathbf{c}_j)' \Sigma^{-1} \mathbf{h} + \mathbf{h}' \Sigma^{-1} \mathbf{h} - 2\log(\xi_j)\\
&\Leftrightarrow (\mathbf{x} - \mathbf{c}_k)' \Sigma^{-1} \mathbf{h} = (\mathbf{x} - \mathbf{c}_j)' \Sigma^{-1} \mathbf{h}\\
&\Leftrightarrow (\mathbf{c}_j - \mathbf{c}_k)' \Sigma^{-1} \mathbf{h} = 0.
\end{aligned}$$

Therefore, $\mathbf{x} + \mathbf{h} \in \mathcal{B}$ in a neighbourhood of $\mathbf{x}$ implies that $\mathbf{h}$ is orthogonal to the vector $(\mathbf{c}_j - \mathbf{c}_k)$ for the inner product induced by Σ^{-1}. Thus, only one direction is suitable for $\mathbf{h}$, showing that there is no ball around $\mathbf{x}$ belonging to $\mathcal{B}$. This proves that the interior of $\mathcal{B}$ is empty and, hence, that the Lebesgue measure of $\mathcal{B}$ is equal to zero. $\square$

Let $\mathbf{x} \in \mathbb{R}^d / \mathcal{B}$. We are now interested in the existence of a neighbourhood of Σ_0 over which $i_{\mathbf{x},\Sigma}$ is constant and thus $\Sigma \mapsto Z_\Sigma(\mathbf{x})$ becomes differentiable with respect to Σ.

Theorem 2. *Let $\mathbf{x} \in \mathbb{R}^d / \mathcal{B}$. There exists a neighbourhood of Σ_0, $\mathcal{W}_{\Sigma_0}$, such that $i_{\mathbf{x},\Sigma}$ is constant over this neighbourhood. Moreover, the function $\Sigma \mapsto Z_\Sigma(\mathbf{x})$ is differentiable over $\mathcal{W}_{\Sigma_0}$ and*

$$\left.\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right|_{\Sigma = \Sigma_0} = -\frac{1}{2} \left(\Sigma_0^{-1} - \Sigma_0^{-1} (\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma_0}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma_0}})' \Sigma_0^{-1} \right). \tag{15}$$

Proof. It follows from Proposition 3 that

$$Y_\Sigma(\mathbf{x}) = \bigvee_{i=1}^{\infty} \xi_i f_\Sigma(\mathbf{x} - \mathbf{c}_i) = \varphi_{i_{\mathbf{x},\Sigma}}(\mathbf{x})$$

and, for $\mathbf{x} \in \mathbb{R}^d / \mathcal{B}$, we have, for all $j \neq i_{\mathbf{x},\Sigma}$,

$$\xi_{i_{\mathbf{x},\Sigma}} f_\Sigma(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}) > \xi_j f_\Sigma(\mathbf{x} - \mathbf{c}_j),$$

or equivalently

$$2\log\left(\xi_{i_{\mathbf{x},\Sigma}}\right) - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Sigma^{-1}} > 2\log\left(\xi_j\right) - \|\mathbf{x} - \mathbf{c}_j\|^2_{\Sigma^{-1}},$$

i.e.,

$$\|\mathbf{x} - \mathbf{c}_j\|^2_{\Sigma^{-1}} - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Sigma^{-1}} > 2\log\left(\xi_j / \xi_{i_{\mathbf{x},\Sigma}}\right). \tag{16}$$

Let $\zeta > 0$ and define

$$\begin{aligned}
\mathcal{I}_1 &= \left\{ j \neq i_{\mathbf{x},\Sigma} : \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 < \zeta \right\}, \\
\mathcal{I}_2 &= \left\{ j \neq i_{\mathbf{x},\Sigma} : \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 > \zeta, 2\log(\xi_j/\xi_{i_{\mathbf{x},\Sigma}}) > 0 \right\}, \\
\mathcal{I}_3 &= \left\{ j \neq i_{\mathbf{x},\Sigma} : \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 > \zeta > 0 > 2\log(\xi_j/\xi_{i_{\mathbf{x},\Sigma}}) \right\},
\end{aligned}$$

such that $\mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3 = \{j \geq 1, j \neq i_{\mathbf{x},\Sigma}\}$. Moreover it is easily to see, using the definition of $i_{\mathbf{x},\Sigma}$ and the form of the intensity function of the point process $(\xi_i, \mathbf{c}_i)_{i\geq 1}$, that $\sharp\mathcal{I}_1$ is finite, $\sharp\mathcal{I}_2$ is finite, but $\sharp\mathcal{I}_3$ is infinite.

Let Θ be a positive definite matrix of size $d \times d$. We have, for any $j \geq 1$,

$$\begin{aligned}
\|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 &= (\mathbf{x} - \mathbf{c}_j)'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_j) \\
&= (\mathbf{x} - \mathbf{c}_j)' \left(\Sigma^{-1} - \Theta^{-1}\right) (\mathbf{x} - \mathbf{c}_j) + \|\mathbf{x} - \mathbf{c}_j\|_{\Theta^{-1}}^2 .
\end{aligned}$$

In addition,

$$\begin{aligned}
\left|(\mathbf{x} - \mathbf{c}_j)' \left(\Sigma^{-1} - \Theta^{-1}\right) (\mathbf{x} - \mathbf{c}_j)\right| &\leq \|\mathbf{x} - \mathbf{c}_j\| \left\|\left(\Sigma^{-1} - \Theta^{-1}\right)(\mathbf{x} - \mathbf{c}_j)\right\| \\
&\leq \|\mathbf{x} - \mathbf{c}_j\|^2 \left\|\Sigma^{-1} - \Theta^{-1}\right\| .
\end{aligned}$$

Since $\|\cdot\|$ and $\|\cdot\|_{\Sigma^{-1}}$ are equivalent norms, there exists a positive constant D such that

$$\left|(\mathbf{x} - \mathbf{c}_j)' \left(\Sigma^{-1} - \Theta^{-1}\right) (\mathbf{x} - \mathbf{c}_j)\right| \leq D \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 \left\|\Sigma^{-1} - \Theta^{-1}\right\|, \quad \mathbf{x} \in \mathbb{R}^d.$$

It follows that, for any $j \geq 1$, there exists a function $\mathbf{x} \mapsto a_j(\mathbf{x}, \Sigma, \Theta)$ such that

$$\|\mathbf{x} - \mathbf{c}_j\|_{\Theta^{-1}}^2 = \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 \left(1 + a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right) \tag{17}$$

and

$$\sup_{\mathbf{x}, \mathbf{c}_j, j\geq 1} \left|a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right| \leq D \left\|\Sigma^{-1} - \Theta^{-1}\right\| . \tag{18}$$

Using (17), we obtain

$$\begin{aligned}
& \|\mathbf{x} - \mathbf{c}_j\|_{\Theta^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Theta^{-1}}^2 \\
= & \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 \left(1 + a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right) - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 \left(1 + a_{i_{\mathbf{x},\Sigma}}\left(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}, \Sigma, \Theta\right)\right) \\
= & \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 + \|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta) \\
& - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 a_{i_{\mathbf{x},\Sigma}}\left(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}, \Sigma, \Theta\right) \\
= & \left(\|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2\right) \left(1 + a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right) \\
& - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 \left(a_{i_{\mathbf{x},\Sigma}}\left(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}, \Sigma, \Theta\right) - a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right) .
\end{aligned} \tag{19}$$

Using (18), we see that there exists $\kappa > 0$ such that, for $\|\Sigma^{-1} - \Theta^{-1}\| < \kappa$, we have, for all $j \geq 1$ such that $j \neq i_{\mathbf{x},\Sigma}$, that

$$\left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2 \left|a_{i_{\mathbf{x},\Sigma}}\left(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}, \Sigma, \Theta\right) - a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right| < \zeta/2,$$

and, for all $j \in \mathcal{I}_3$,

$$\left(\|\mathbf{x} - \mathbf{c}_j\|_{\Sigma^{-1}}^2 - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|_{\Sigma^{-1}}^2\right) \left(1 + a_j(\mathbf{x} - \mathbf{c}_j, \Sigma, \Theta)\right) > \zeta/2.$$

Hence, using (19), we obtain, for all $j \in \mathcal{I}_3$,

$$\|\mathbf{x} - \mathbf{c}_j\|^2_{\Theta^{-1}} - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Theta^{-1}} > 0 > 2\log(\xi_j/\xi_{i_{\mathbf{x},\Sigma}}). \tag{20}$$

Now, using (16), the continuity of $\Sigma^{-1} \mapsto \|\cdot\|_{\Sigma^{-1}}$ and the fact that $\sharp\mathcal{I}_1$ and $\sharp\mathcal{I}_2$ are finite, there exists $\kappa' > 0$ such that, for $\|\Sigma^{-1} - \Theta^{-1}\| < \kappa'$, we have, for all $j \in \mathcal{I}_1 \cup \mathcal{I}_2$, that

$$\|\mathbf{x} - \mathbf{c}_j\|^2_{\Theta^{-1}} - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Theta^{-1}} > 2\log(\xi_j/\xi_{i_{\mathbf{x},\Sigma}}). \tag{21}$$

Combining (20) and (21), we obtain that, for all positive definite matrix Θ satisfying $\|\Sigma^{-1} - \Theta^{-1}\| < \min\{\kappa, \kappa'\}$, we have, for all $j \geq 1$ such that $j \neq i_{\mathbf{x},\Sigma}$,

$$\|\mathbf{x} - \mathbf{c}_j\|^2_{\Theta^{-1}} - \left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Theta^{-1}} > 2\log(\xi_j/\xi_{n_{\mathbf{x}}}).$$

Accordingly, $i_{\mathbf{x},\Theta} = i_{\mathbf{x},\Sigma}$. Hence we can choose the neighbourhood of Σ_0

$$\mathcal{W}_{\Sigma_0} = \left\{\Theta \text{ positive definite} : \left\|\Sigma_0^{-1} - \Theta^{-1}\right\| < \min\{\kappa, \kappa'\}\right\},$$

to define the derivative of $\Sigma \mapsto Z_\Sigma(\mathbf{x})$ at Σ_0.

We now compute the corresponding derivative. Using (1) and Proposition 3, we have

$$Z_\Sigma(\mathbf{x}) = \log\left(\xi_{i_{\mathbf{x},\Sigma}}\right) - \frac{d}{2}\log(2\pi) - \frac{1}{2}\log(\det(\Sigma)) - \frac{1}{2}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}),$$

and hence

$$\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma} = -\frac{1}{2}\left(\frac{\partial \log(\det(\Sigma))}{\partial \Sigma} + \frac{\partial (\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})}{\partial \Sigma}\right).$$

Formula (11.7) in Dwyer (1967) gives, for any symmetric matrix Σ, that

$$\frac{\partial \log(\det(\Sigma))}{\partial \Sigma} = \Sigma^{-1}. \tag{22}$$

Moreover, since $i_{\mathbf{x},\Sigma}$ is constant over $\mathcal{W}_{\Sigma_0}$, Equation (11.8) in Dwyer (1967) provides, for any symmetric matrix Σ,

$$\frac{\partial (\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})}{\partial \Sigma} = -\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}. \tag{23}$$

Combining (22) and (23), we finally obtain

$$\left.\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right|_{\Sigma=\Sigma_0} = -\frac{1}{2}\left(\Sigma_0^{-1} - \Sigma_0^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma_0}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma_0}})'\Sigma_0^{-1}\right).$$

□

We now prove that the derivative of $\Sigma \mapsto Z_\Sigma(\mathbf{x})$ can be uniformly bounded by an integrable random variable over a neighbourhood of Σ_0.

Theorem 3. *There exists a non-random neighbourhood of Σ_0, $\mathcal{V}_{\Sigma_0}$, such that, for any $q > 1$, there exists a random variable $C_{\Sigma_0}(\mathbf{x}, q)$ satisfying a.s.*

$$\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\|\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right\|^q \leq C_{\Sigma_0}(\mathbf{x}, q)$$

and $\mathbb{E}\left[C_{\Sigma_0}(\mathbf{x}, q)\right] < \infty$.

Proof. First recall that a.s.

$$\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma} = -\frac{1}{2}\left(\Sigma^{-1} - \Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}\right),$$

which gives a.s.

$$\left\|\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right\| \leq \frac{1}{2}\left(\left\|\Sigma^{-1}\right\| + \left\|\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}\right\|\right).$$

Consequently, using the well-known fact that, for all $a, b \in \mathbb{R}$ and $q \geq 1$, $|a-b|^q \leq 2^{q-1}(|a|^q + |b|^q)$, we obtain

$$\left\|\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right\|^q \leq \frac{1}{2}\left(\left\|\Sigma^{-1}\right\|^q + \left\|\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}\right\|^q\right). \tag{24}$$

Let $A = \Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})$, such that $\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1} = AA'$. Note that AA' is a non-negative symmetric matrix of rank 1 and thus it only has one positive eigenvalue given by

$$\lambda = A'A = \|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-2}}$$

(since we have that $(AA')A = A(A'A) = (A'A)A$). It follows that

$$\|AA'\| = \|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-2}}. \tag{25}$$

For a positive definite symmetric matrix Θ, we denote by $\lambda_{\max}(\Theta)$ and $\lambda_{\min}(\Theta)$ respectively the maximum and the minimum of its positive eigenvalues. We have that

$$\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-1}} = (\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}) \geq \lambda_{\min}(\Sigma^{-1})\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2,$$

and

$$\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-2}} = (\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-2}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}) \leq \lambda_{\max}\left(\Sigma^{-2}\right)\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2,$$

which yield

$$\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-2}} \leq \frac{\left[\lambda_{\max}(\Sigma^{-1})\right]^2}{\lambda_{\min}(\Sigma^{-1})}\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^2_{\Sigma^{-1}}. \tag{26}$$

Combining (24), (25) and (26), we have, for any $q > 1$, that a.s.

$$\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\|\frac{\partial Z_\Sigma(\mathbf{x})}{\partial \Sigma}\right\|^q \leq C_{\Sigma_0}(\mathbf{x}, q),$$

where

$$C_{\Sigma_0}(\mathbf{x}, q) = \frac{1}{2}\left(\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}\right\|^q + \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}}\left(\frac{\left[\lambda_{\max}(\Sigma^{-1})\right]^2}{\lambda_{\min}(\Sigma^{-1})}\right)^q \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^{2q}_{\Sigma^{-1}}\right). \tag{27}$$

In order to control $C_{\Sigma_0}(\mathbf{x}, q)$, it is sufficient to control $\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\|^{2q}_{\Sigma^{-1}}$. As the center of the "storm" realizing the maximum at point $\mathbf{x}$, $\mathbf{c}_{i_{\mathbf{x},\Sigma}}$ is characterized by

$$\begin{aligned}
&\xi_{i_{\mathbf{x},\Sigma}}\ f_\Sigma(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}) \geq \xi_i\ f_\Sigma(\mathbf{x} - \mathbf{c}_i) \quad \forall i \geq 1 \\
\Leftrightarrow\ &\log(\xi_{i_{\mathbf{x},\Sigma}}) - \frac{1}{2}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}})'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}) \geq \log(\xi_i) - \frac{1}{2}(\mathbf{x} - \mathbf{c}_i)'\Sigma^{-1}(\mathbf{x} - \mathbf{c}_i),\ \forall i \geq 1 \\
\Leftrightarrow\ &\left\|\mathbf{x} - \mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Sigma^{-1}} \leq 2\log(\xi_{i_{\mathbf{x},\Sigma}}) - 2\log(\xi_i) + \|\mathbf{x} - \mathbf{c}_i\|^2_{\Sigma^{-1}},\ \forall i \geq 1,
\end{aligned} \tag{28}$$

and therefore

$$\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Sigma^{-1}} \leq 2\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \log(\xi_{i_{\mathbf{x},\Sigma}}) + \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_i\|^2_{\Sigma^{-1}} - 2\log(\xi_i),\ \forall i\geq 1. \quad (29)$$

Additionally, we have, for all $i\geq 1$,

$$\|\mathbf{x}-\mathbf{c}_i\|^2_{\Sigma^{-1}} = \|\mathbf{x}-\mathbf{c}_i\|^2_{\Sigma_0^{-1}} + (\mathbf{x}-\mathbf{c}_i)'\left(\Sigma^{-1}-\Sigma_0^{-1}\right)(\mathbf{x}-\mathbf{c}_i),$$

which yields, by equivalence of the norms $\|.\|$ and $\|.\|_{\Sigma^{-1}}$, that

$$\begin{aligned}\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_i\|^2_{\Sigma^{-1}} &\leq \|\mathbf{x}-\mathbf{c}_i\|^2_{\Sigma_0^{-1}} + \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}-\Sigma_0^{-1}\right\| \|\mathbf{x}-\mathbf{c}_i\|^2 \\ &\leq \left(D_0 + \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}_i\|^2 \qquad (30)\end{aligned}$$

for some positive constant D_0. Hence let us now choose $\mathcal{V}_{\Sigma_0}$ such that

$$\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}-\Sigma_0^{-1}\right\| < \infty. \quad (31)$$

Now, for two real-valued random variables U and V, $U+V\geq\lambda$ implies that $U\geq\lambda/2$ or $V\geq\lambda/2$, giving $\mathbb{P}(U+V\geq\lambda)\leq\mathbb{P}\left(U\geq\lambda/2\right)+\mathbb{P}\left(V\geq\lambda/2\right)$. Thus, using (29) and (30), we obtain

$$\begin{aligned}&\mathbb{P}\left(\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\right\|^2_{\Sigma^{-1}}\geq\lambda\right) \\ &\leq \mathbb{P}\left(\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}_i\|^2 - 2\log(\xi_i) + 2\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\log(\xi_{i_{\mathbf{x},\Sigma}})\geq\lambda\ \forall i\geq 1\right) \\ &\leq \mathbb{P}\left(\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}_i\|^2 - 2\log(\xi_i)\geq\frac{\lambda}{2}\ \forall i\geq 1\right) \\ &\quad+\mathbb{P}\left(2\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\log(\xi_{i_{\mathbf{x},\Sigma}})\geq\frac{\lambda}{2}\right). \qquad (32)\end{aligned}$$

We first deal with the first term of the right-hand side of (32). Since the $(\xi_i,\mathbf{c}_i)_{i\geq 1}$ are the points of a Poisson process on $(0,\infty)\times\mathbb{R}^2$ with intensity function $\xi^{-2}\lambda(\mathrm{d}\xi)\times\lambda(\mathrm{d}\nu)$, we have

$$\begin{aligned}&\mathbb{P}\left(\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}_i\|^2 - 2\log(\xi_i)\geq\frac{\lambda}{2}\ \forall i\geq 1\right) \\ &=\exp\left(-\mu\left\{(\xi,\mathbf{c})\in(0,\infty)\times\mathbb{R}^2 : \left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}\|^2-2\log(\xi)<\frac{\lambda}{2}\right\}\right), \\ &\qquad (33)\end{aligned}$$

where

$$\mu\left\{(\xi,\mathbf{c})\in(0,\infty)\times\mathbb{R}^2:\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}\|^2-2\log(\xi)<\frac{\lambda}{2}\right\}$$

$$=\int_{e^{-\frac{\lambda}{4}}}^{\infty}\left(\int_{\|\mathbf{x}-\mathbf{c}\|^2\le\left(\frac{\lambda}{2}+2\log(\xi)\right)\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)^{-1}}\nu(\mathrm{d}\mathbf{c})\right)\xi^{-2}\nu(\mathrm{d}\xi)$$

$$=\frac{\pi^{d/2}}{\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)^d\Gamma\left(d/2+1\right)}\int_{e^{-\frac{\lambda}{4}}}^{\infty}\left(\frac{\lambda}{2}+2\log(\xi)\right)^d\xi^{-2}\nu(\mathrm{d}\xi).$$

Making the change of variable $u=\lambda/2+2\log(\xi)$, yielding $\xi=\exp\left(u/2-\lambda/4\right)$ and $\nu(\mathrm{d}\xi)=\exp\left(u/2-\lambda/4\right)\nu(\mathrm{d}u)/2$, we obtain

$$\mu\left\{(\xi,\mathbf{c})\in(0,\infty)\times\mathbb{R}^2:\|\mathbf{x}-\mathbf{c}\|_\Sigma^2-2\log(\xi)<\frac{\lambda}{2}\right\}$$

$$=\frac{\pi^{d/2}}{\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)^d\Gamma\left(d/2+1\right)}\exp\left(\frac{\lambda}{4}\right)\int_{e^{-\frac{\lambda}{4}}}^{\infty}u^d\exp\left(-\frac{u}{2}\right)\nu(\mathrm{d}u) \quad (34)$$

As λ tends to infinity,

$$\int_{e^{-\frac{\lambda}{4}}}^{\infty}u^d\exp\left(-\frac{u}{2}\right)\nu(\mathrm{d}u)\to\int_0^{\infty}u^d\exp\left(-\frac{u}{2}\right)\nu(\mathrm{d}u),$$

which is a constant. Therefore, it follows from (33) that

$$\mathbb{P}\left(\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)\|\mathbf{x}-\mathbf{c}_i\|^2-2\log(\xi_i)\ge\frac{\lambda}{2}\ \forall i\ge1\right)$$

$$\underset{\lambda\to\infty}{\sim}\exp\left(-\left[\frac{\pi^{d/2}}{\left(D_0+\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\Sigma^{-1}-\Sigma_0^{-1}\right\|\right)^d\Gamma\left(d/2+1\right)}\int_0^{\infty}u^d\exp\left(-\frac{u}{2}\right)\nu(\mathrm{d}u)\right]\exp\left(\frac{\lambda}{4}\right)\right). \quad (35)$$

Let us now deal with the second term of the right-hand side of (32). Observe that

$$\mathbb{P}\left(2\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\log(\xi_{i_{\mathbf{x},\Sigma}})\ge\frac{\lambda}{2}\right)=\mathbb{P}\left(\inf_{\Sigma\in\mathcal{V}_{\Sigma_0}}\xi_{i_{\mathbf{x},\Sigma}}^{-1}\le\exp\left(-\frac{\lambda}{4}\right)\right).$$

The mapping $\xi\to\xi^{-1}$ applied to the points of the Poisson point process yields a new Poisson point process with intensity function $\nu(\mathrm{d}\xi)$. Furthermore, such a Poisson point process on $(0,\infty)$ is homogeneous and can be represented as the sum of independent standard exponential random variables. Thus, we can write $\min_{i\ge1}\{\xi_i^{-1}\}\overset{d}{=}\mathrm{Exp}(1)$, where $\overset{d}{=}$ stands for equality in distribution. Moreover, $\inf_{\Sigma\in\mathcal{V}_{\Sigma_0}}\xi_{i_{\mathbf{x},\Sigma}}^{-1}\ge\min_{i\ge1}\{\xi_i^{-1}\}$, implying

$$\begin{aligned}\mathbb{P}\left(\inf_{\Sigma\in\mathcal{V}_{\Sigma_0}}\xi_{i_{\mathbf{x},\Sigma}}^{-1}\le\exp\left(-\frac{\lambda}{4}\right)\right)&\le\mathbb{P}\left(\min_{i\ge1}\left\{\xi_i^{-1}\right\}\le\exp\left(-\frac{\lambda}{4}\right)\right)\\&=1-\exp\left(-\exp\left(-\frac{\lambda}{4}\right)\right)\\&\underset{\lambda\to\infty}{\sim}\exp\left(-\frac{\lambda}{4}\right),\end{aligned}$$

which gives for large λ

$$\mathbb{P}\left(\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} 2\log(\xi_{i_{\mathbf{x},\Sigma}}) \geq \frac{\lambda}{2}\right) \leq \exp\left(-\frac{\lambda}{4}\right). \tag{36}$$

Now, for any $q > 1$, we have

$$\mathbb{E}\left[\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2q}\right] = \int_0^\infty \mathbb{P}\left(\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2q} \geq u\right) \nu(\mathrm{d}u).$$

We carry out the change of variable $\lambda = u^{1/q}$, which gives $u = \lambda^q$ and hence $\mathrm{d}u = q\lambda^{q-1}\mathrm{d}\lambda$. Accordingly, we obtain

$$\mathbb{E}\left[\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2q}\right] = q\int_0^\infty \lambda^{q-1}\mathbb{P}\left(\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2} \geq \lambda\right) \nu(\mathrm{d}\lambda),$$

and therefore, using (32), (35) and (36),

$$\mathbb{E}\left[\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2q}\right] < \infty.$$

Finally, let us recall that

$$C_{\Sigma_0}(\mathbf{x},q) = \frac{1}{2}\left(\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}\right\|^q + \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left(\frac{\left[\lambda_{\max}(\Sigma^{-1})\right]^2}{\lambda_{\min}(\Sigma^{-1})}\right)^q \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \|\mathbf{x}-\mathbf{c}_{i_{\mathbf{x},\Sigma}}\|_{\Sigma^{-1}}^{2q}\right).$$

Consequently, using (31), we finally deduce that any neighbourhood of Σ_0, $\mathcal{V}_{\Sigma_0}$, satisfying

$$\begin{aligned}
\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}\right\|^q &< \infty,\\
\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left(\frac{\left[\lambda_{\max}(\Sigma^{-1})\right]^2}{\lambda_{\min}(\Sigma^{-1})}\right) &< \infty,\\
\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}} \left\|\Sigma^{-1}-\Sigma_0^{-1}\right\| &< \infty,
\end{aligned}$$

leads to

$$\mathbb{E}\left[C_{\Sigma_0}(\mathbf{x},q)\right] < \infty.$$

$\square$

We finally provide the proof of Theorem 1.

Proof. Let us choose $\mathcal{V}_{\Sigma_0}$ as in the proof of Theorem 3. It follows from (6) that

$$\left\|\frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial\Sigma}\right\| \leq \sum_{i=1}^M \left|\frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i}\right| \left\|\frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial\Sigma}\right\|,$$

which gives

$$\sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial\Sigma}\right\| \leq \sum_{i=1}^M \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left|\frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i}\right| \sup_{\Sigma\in\mathcal{V}_{\Sigma_0}}\left\|\frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial\Sigma}\right\|.$$

Hence we choose

$$B_{\Sigma_0} = \sum_{i=1}^{M} \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left| \frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i} \right| \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial \Sigma} \right\|.$$

We have

$$\mathbb{E}\left[B_{\Sigma_0}\right] = \sum_{i=1}^{M} \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left| \frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i} \right| \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial \Sigma} \right\| \right].$$

Let $q > 1$ such that $p^{-1} + q^{-1} = 1$. By Hölder inequality, we have

$$\begin{aligned} & \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left| \frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i} \right| \sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial \Sigma} \right\| \right] \\ \leq \; & \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left| \frac{\partial H_M\left(\mathbf{Z}_\Sigma\right)}{\partial z_i} \right|^p \right]^{1/p} \mathbb{E}\left[\sup_{\Sigma \in \mathcal{V}_{\Sigma_0}} \left\| \frac{\partial Z_\Sigma(\mathbf{x}_i)}{\partial \Sigma} \right\|^q \right]^{1/q}. \end{aligned}$$

By Theorem 3 and the assumptions, the result follows. □

References

Adams, R. (2007). *Infinitesimal perturbation analysis for active queue management*. PhD thesis, Georgia Institute of Technology.

Asmussen, S. and Glynn, P. W. (2010). *Stochastic Simulation: Algorithms and Analysis*. Springer-Verlag New York.

Broadie, M. and Glasserman, P. (1996). Estimating security price derivatives using simulation. *Management Science*, 42(2):269–285.

Chen, J. and Fu, M. C. (2001). Efficient sensitivity analysis of mortgage backed securities. In *12th Annual Derivatives Securities Conference, New York*.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science*, 27(2):161–186.

Dombry, C., Engelke, S., and Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika*, 103(2):303–317.

Dombry, C. and Kabluchko, Z. (2018). Random tessellations associated with max-stable random fields. *Bernoulli*, 24(1):30–52.

Dwyer, P. S. (1967). Some applications of matrix derivatives in multivariate analysis. *Journal of the American Statistical Association*, 62(318):607–625.

Glasserman, P. (1991a). *Gradient Estimation Via Perturbation Analysis*. Springer US.

Glasserman, P. (1991b). Structural conditions for perturbation analysis derivative estimation: finite-time performance indices. *Operations Research*, 39(5):724–738.

Glasserman, P. (2013). *Monte Carlo Methods in Financial Engineering*. Springer-Verlag New York.

Glasserman, P. and Liu, Z. (2010). Sensitivity estimates from characteristic functions. *Operations Research*, 58(6):1611–1623.

Heidelberger, P., Cao, X.-R., Zazanis, M. A., and Suri, R. (1988). Convergence properties of infinitesimal perturbation analysis estimates. *Management Science*, 34(11):1281–1302.

Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5):2042–2065.

Koch, E. (2018). Spatial risk measures induced by powers of max-stable random fields. *Available at https://arxiv.org/pdf/1804.05694v1.pdf*.

L'Ecuyer, P. (1991). An overview of derivative estimation. In *Proceedings of the 1991 Winter Simulation Conference*, pages 207–217. IEEE.

Ribatet, M., Singleton, R., and R Core team (2018). SpatialExtremes: Modelling Spatial Extremes. *R package version 2.0-7*.

Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1):33–44.

Smith, R. L. (1990). Max-stable processes and spatial extremes. *Unpublished manuscript*, University of North Carolina.