MODERATE DEVIATION FOR RANDOM ELLIPTIC PDES WITH SMALL NOISE

XIAOOU LI, JINGCHEN LIU, JIANFENG LU, AND XIANG ZHOU

Abstract. Partial differential equations with random inputs have become popular models to characterize physical systems with uncertainty coming from, e.g., imprecise measurement and intrinsic randomness. In this paper, we perform asymptotic rare event analysis for such elliptic PDEs with random inputs. In particular, we consider the asymptotic regime that the noise level converges to zero suggesting that the system uncertainty is low, but does exists. We develop sharp approximations of the probability of a large class of rare events.

1. Introduction

The study of rare events due to system uncertainty, for example the failure of materials due to intrinsic randomness, is crucial and yet challenging. While those events do not often occur, they lead to catastrophic consequences. Therefore it is important to estimate the probabilities of such events and to characterize those events which help finding interventions to prevent them from happening. In this paper, we consider the following classical continuum mechanical model in the form of a linear elliptic partial differential equation (PDE) defined on a domain $U \subset \mathbb{R}^d$,

$$-\nabla \cdot (a(x)\nabla u(x)) = f(x), \tag{1}$$

subject to certain boundary conditions that will be specified in the sequel. The solution to the above equation u is the displacement field of the elastic material, ∇u is the strain, a is the elasticity tensor, $a(x)\nabla u(x)$ is the stress tensor, and f is the external body force. The elasticity tensor $a(x)$ (which is uniformly positive definite) is determined by the property of the specific material. Instead of assuming that a is deterministic, we are interested in the situations when the tensor a contains randomness. The randomness is introduced to incorporate the uncertainties of simple elastic materials at the macroscopic level or heterogeneity in the microstructures of complex materials. Under this setting, the solution $u(x)$ (as a function of $a(x)$) is also a stochastic process whose law is determined by that of $a(x)$.

Besides material mechanics, the elliptic PDE (1) arises also in many other fields of applications, such as hydrogeology and porous medium. The tensor $a(x)$ carries different names such as conductivity and permeability. It is recognized that the modeling of the random field $a(x)$ is of primal

Date: March 17, 2017.

importance for the analysis. In this paper, we consider that the random function $a(x)$ follows a log-normal distribution, that is,

$$a(x) = a_0(x)e^{-\sigma\xi(x)} \qquad x \in U, \tag{2}$$

where $\xi(x)$ is a Gaussian random field defined on U and $a_0(x)$ is a deterministic function. In elasticity, in general $a(\cdot)$ is a function of 4-tensor. For simplicity of notation, we consider a scalar field here (i.e., an isotropic material). The technique and result for a general $a(\cdot)$ is similar. The scalar $\sigma > 0$ is a parameter indexing the noise level. Many studies by practioners, e.g., [Freeze, 1975, Bear and Verruijt, 1987, Charbeneau, 2000], have shown that the best fit of the empirical data is the log-normal distribution. Hence, the log-normal assumption is well justified in applications and is used in mathematical analysis and numerical computation of the random PDE (1). In our paper, we follow this convention of log-normal assumption for the rare-event analysis.

In this work, we consider the small noise asymptotic regime, that σ tends to zero. Yet, even small noise can lead to drastic difference of the PDE solution from that of the deterministic case when the noise level is zero. Our results characterize such rare events, more precisely, the deviation of the solution of the random elliptic PDE in the presence of small noise. In particular, we focus on the deviation from the deterministic solution as the uncertainty level goes to 0. Let $\mathcal{H}$ be a mapping from $C(\bar{U})$ to $\mathbb{R}$. Of primary interest

$$\omega(\sigma) = \mathbb{P}\{\mathcal{H}(u) > \mathcal{H}(u_0) + b_\sigma\} \quad \text{as } \sigma \to 0.$$

where u is the solution to equation (1) and u_0 is the solution when the noise level is zero, i.e., $a(x) = a_0(x)$. The level b_σ will be sent to zero as the noise level σ goes to zero, which will be specified in the sequel. The main contribution of this paper is to derive sharp asymptotic approximations of $\omega(\sigma)$ as $\sigma \to 0$.

Given that $\mathcal{H}(u)$ is a (complicated) functional of the input Gaussian process $\xi(x)$, the analysis of the tail probability $\omega(\sigma)$ links naturally to the rare-event analysis of Gaussian random field. The study of the extremes of Gaussian random fields focuses mostly on the tail probabilities of the supremum of the field. The results contain general bounds on $P(\max \xi(x) > b)$ as well as sharp asymptotic approximations as $b \to \infty$. A partial literature contains [Landau and Shepp, 1970, Marcus and Shepp, 1970, Sudakov and Tsirelson, 1974, Borell, 1975, Borell, 2003, Ledoux and Talagrand, 1991, Talagrand, 1996, Berman, 1985]. Several methods have been introduced to obtain bounds and asymptotic approximations. A general upper bound for the tail of $\max \xi(x)$ is developed in [Borell, 1975, Tsirelson et al., 1976], which is known as the Borel–TIS inequality. For asymptotic results, there are several methods, such as the double sum method ([Piterbarg, 1996]) , the Euler–Poincaré characteristics of the excursion set approximation ([Adler, 1981, Taylor et al., 2005, Adler and Taylor, 2007, Taylor and Adler, 2003]), the tube method ([Sun, 1993]), and the Rice method ([Azais and Wschebor, 2008, Azais and Wschebor, 2009]). Recently, the exact tail approximation of integrals of exponential functions of Gaussian random fields is developed by [Liu, 2012, Liu and Xu, 2012]. Efficient computations via importance sampling has been developed by [Adler et al., 2008, Adler et al., 2012]. For the analysis of the tail probabilities of lognormal random fields with small noise, refer to the recent work in [Li et al., 2016]. There are also existing work in the context of PDE with random coefficients. [Liu and Zhou, 2013, Liu and Zhou, 2014] derive asymptotic analysis of one-dimensional elliptic PDE. [Liu et al., 2015] presents the corresponding rare-event simulation algorithms. These works focused on the asymptotic regime that the noise level σ is fixed. Furthermore, [Xu et al., 2014] presents asymptotic analysis for stochastic KdV equation.

The rest of the paper is organized as follows. Section 2 presents the problem setup and the main asymptotic results. The technical proofs are given in Section 3.

2. Main results

2.1. **The problem setup.** We consider the following elliptic PDE. Let $U \subset \mathbb{R}^d$ be an open domain with a smooth boundary. The differential equation concerning $u : U \to \mathbb{R}$ with Dirichlet boundary condition is given by

$$\begin{cases} -\nabla \cdot (a(x)\nabla u(x)) = f(x) & \text{for } x \in U; \\ u(x) = 0 & \text{for } x \in \partial U. \end{cases} \tag{3}$$

In the context of elastic mechanics, u characterizes the material deformation due to external force f and $a : U \to \mathbb{R}$ gives the stiffness of the material. Throughout this paper, we assume $u(\cdot)$ to be a scalar function for simplicity. We assume that the material is clamped to a frame on the boundary ∂U and hence the Dirichlet boundary condition $u(\partial U) = 0$ in (3) is assumed. The external force f is sufficiently smooth and bounded, that is, there exists a constant $c \in \mathbb{R}$ such that

$$|f(x)| \leq c, \qquad \forall x \in U. \tag{4}$$

We study the behavior of the material under the influence of internal randomness, which may be the result of manufacturing processing or the uncertainty of the material properties at the microscopic level. We adopt a probabilistic viewpoint of the complexity and heterogeneity inherent in the material and view the coefficient $a(x)$ as a random field. The process $a(x)$ is physically restricted to be positive and is modeled as a lognormal random field given as in (2). Furthermore, the Gaussian random function ξ has mean zero and its covariance function is denoted by

$$C(x, y) = \mathbf{E}\{\xi(x)\xi(y)\}, \tag{5}$$

which is certainly independent of σ. In addition, C admits the normalization condition $C(x,x) \equiv 1$.

The solution $u(x)$ depends implicitly on $a(x)$ through equation (3) and further $\xi(x)$ via a logarithmic change of variable. It is useful to define a mapping from the coefficient ξ to the solution u

$$\mathbf{J}[\xi] \triangleq u_\xi$$

where u_ξ is the solution to equation (3) with $a(x) = a_0(x)e^{-\xi(x)}$. This mapping depends only on the deterministic function a_0, the external force f, the domain U, and the boundary condition. In this paper, we are interested in the asymptotic regime that the amplitude of the uncertainty level σ tends to zero. Then the failure problem concerns the random solution $u_{\sigma\xi} = \mathbf{J}(\sigma\xi)$ by noting the definition of $\mathbf{J}$ above. As $\sigma \to 0$, the process $a(x)$ tends to its limiting field $a_0(x)$. Let $u_0(x)$ be the corresponding limiting solution satisfying equation

$$\begin{cases} -\nabla \cdot (a_0(x)\nabla u_0(x)) = f(x) & \text{for } x \in U; \\ u_0(x) = 0 & \text{for } x \in \partial U. \end{cases} \tag{6}$$

Then, under mild conditions, we have $u(x) \to u_0(x)$ as $\sigma \to 0$.

We provide asymptotic analysis of the event that u deviates from its limiting solution u_0. Let $\mathcal{H}$ be a functional from $C(\bar{U})$ to $\mathbb{R}$ characterizing the deviation. For instance $\mathcal{H}(u) = \int_U (u(x) - u_0(x))\, dx$. Let $\mathcal{G}$ be the composition of $\mathbf{J}$ and $\mathcal{H}$, that is,

$$\mathcal{G}(\xi) = \mathcal{H}(\mathbf{J}[\xi]).$$

To simplify notation, we always choose $\mathcal{H}$ such that $\mathcal{G}(\mathbf{0}) = \mathcal{H}(u_0) = 0$. We are interested the tail probability of $\mathcal{G}(\sigma\xi)$ as $\sigma \to 0$. In particular, we derive asymptotic approximations for

$$\omega(\sigma) = \mathbb{P}\{\mathcal{G}(\sigma\xi) > b\} \quad \text{as } \sigma \to 0, \tag{7}$$

where the deviation level is chosen to be $b = \kappa\sigma^\alpha$ for some fixed $\alpha \in (0,1)$ and $\kappa > 0$. In particular, the deviation level b also goes to 0 as the uncertainty vanishes.

2.2. **Asymptotic results.** We first introduce some notation that will be used in the sequel. Throughout this analysis, we consider $\mathcal{G}$ to be a differentiable function and let $\mathcal{G}'$ be its Fréchet derivative, that is,

$$\mathcal{G}(\xi + \varepsilon\eta) = \mathcal{G}(\xi) + \varepsilon \int_U \mathcal{G}'[\xi](x)\eta(x)dx + o(\varepsilon), \quad \text{as } \varepsilon \to 0,\ \forall\, \xi, \eta \in C(\bar{U}).$$

For $0 < \beta < 1$, we say that a function w is Hölder continuous with order β if the Hölder coefficient

$$[w]_\beta = \sup_{x,y\in\bar{U},x\neq y} \frac{|w(x)-w(y)|}{|x-y|^\beta} < \infty. \tag{8}$$

We use $C^k(\bar{U})$ to denote the space containing all k-time continuously differentiable functions. For nonnegative integer k and $0 \le \beta < 1$, we use $C^{k,\beta}(\bar{U})$ to denote the set of functions in $C^k(\bar{U})$ whose k-th order partial derivatives are Hölder continuous with coefficient β. For simplicity, we write $C^{0,\beta}(\bar{U}) = C^\beta(\bar{U})$. We proceed to the definition of norms over $C^{k,\beta}(\bar{U})$. We first define the seminorms

$$[w]_{k,0} = \max_{|\gamma|=k} \sup_{\bar{U}} |D^\gamma w| \quad \text{and} \quad [w]_{k,\beta} = \max_{|\gamma|=k} [D^\gamma w]_\beta,$$

where γ is a multi-index $\gamma = (\gamma_1, ..., \gamma_d)$, $|\gamma| = \sum_{i=1}^d \gamma_i$, and $D^\gamma w = \frac{\partial^{|\gamma|} w}{\partial^{\gamma_1} x_1 ... \partial^{\gamma_d} x_d}$. We further define the norms

$$\|w\|_{C^k(\bar{U})} = \sum_{j=0}^k [w]_{j,0} \quad \text{and} \quad \|w\|_{C^{k,\beta}(\bar{U})} = \|w\|_{C^k(\bar{U})} + [w]_{k,\beta}.$$

Equipped with $\|\cdot\|_{C^{k,\beta}(\bar{U})}$, the space $C^{k,\beta}(\bar{U})$ is a Banach space for all non-negative integer k and $0 \le \beta < 1$. To simplify notation, we write

$$|w|_k = \|w\|_{C^k(\bar{U})}, \quad |w|_{k,\beta} = \|w\|_{C^{k,\beta}(\bar{U})}, \quad |w|_\beta = |w|_{0,\beta}.$$

We now present sharp asymptotic approximations of the tail probabilities $w(\sigma)$ under the following assumptions on the functional $\mathcal{G}$ and the covariance function $C(x,y)$.

Assumption.

A1. There exist constants $k, \beta, \delta_G, \kappa_G$ such that k is a non-negative integer, $0 \le \beta < 1$, $\delta_G > 0$ and for all $|w|_{k,\beta} \le \delta_G$, $\mathcal{G}'[w] \in C^{k,\beta}(\bar{U})$. In addition $\mathcal{G}'$ is a (local) Lipschitz operator in the sense that for all $|w_1|_{k,\beta}, |w_2|_{k,\beta} \le \delta_G$, we have

$$\left|\mathcal{G}'[w_1] - \mathcal{G}'[w_2]\right|_{k,\beta} \le \kappa_G |w_1 - w_2|_{k,\beta}.$$

A2. There exists $x \in \bar{U}$ such that $\mathcal{G}'[\mathbf{0}](x) \neq 0$.

A3. The Gaussian random field $\{\xi(x) : x \in U\}$ has a Hölder continuous sample path and belongs to the space $C^{k,\beta}(\bar{U})$ almost surely, that is, $\mathbb{P}(|\xi|_{k,\beta} < \infty) = 1$. The covariance function $C(\cdot,\cdot)$ is positive definite and satisfies $\sup_{y\in\bar{U}} |C(\cdot,y)|_{k,2\beta} < \infty$. Moreover, we assume that $\sup_{y\in\bar{U}} |C_{D^\gamma\xi}(\cdot,y)|_{2\beta} < \infty$ for all γ such that $|\gamma| \le k$, where we define

$$C_{D^\gamma\xi}(x,y) \triangleq \mathbb{E}\{D^\gamma\xi(x) D^\gamma\xi(y)\}. \tag{9}$$

Define a mapping $\mathbf{C} : C(\bar{U}) \to C(\bar{U})$

$$\mathbf{C}w \triangleq \int C(\cdot, y)w(y)\, \mathrm{d}y.$$

We consider the optimization problem

$$\min_{\xi \in \mathcal{B}, \mathcal{G}(\sigma \mathbf{C}\xi) = b} \mathcal{K}(\xi) \tag{10}$$

where the functional $\mathcal{K} : C^0(\bar{U}) \to \mathbb{R}$ is

$$\mathcal{K}(w) \triangleq \int_U w(x)C(x, y)w(y)dxdy,$$

and the set $\mathcal{B}$ is defined as

$$\mathcal{B} \triangleq \left\{ w \in C^{k,\beta}(\bar{U}) : |w|_{k,\beta} \leq \sigma^{\alpha-1-\varepsilon} \right\} \tag{11}$$

for some $\varepsilon > 0$ and α is given as below (7). Because $\mathcal{B}$ is a compact subset of $C^{k,\beta}(\bar{U})$ and the functionals $\mathcal{K}$ and $\mathcal{G}$ are continuous over $\mathcal{B}$, the above optimization problem has at least one solution. Later in the current section, we will show that this solution is also unique. With the above optimization, we have the following sharp asymptotic approximation for the tail probability of $\omega(\sigma)$.

Theorem 1. *Under Assumptions A1-A3, for $0 < \alpha < 1$ and $b = \kappa\sigma^\alpha$, we have*

$$\mathbb{P}\{\mathcal{G}(\sigma\xi) > b\} = (c_1 + o(1))\sigma^{1-\alpha} \exp\Big(-\frac{1}{2}K^*_\sigma \Big) \text{ as } \sigma \to 0,$$

where $c_1 = \kappa^{-1}\{(2\pi)^{-1}\mathcal{K}(\mathcal{G}'[\mathbf{0}])\}^{\frac{1}{2}}$ and

$$K^*_\sigma = \min_{w \in \mathcal{B}, \mathcal{G}(\sigma \mathbf{C} w) = b} \mathcal{K}(w).$$

The constants k and β in Assumptions A1-A3 are problem-dependent. For example, [Li et al., 2015] consider the functional

$$\mathcal{G}(\xi) = \int_U e^{\sigma\xi(t)+\mu(t)}dt - \int_U e^{\mu(t)}dt,$$

where $\mu(\cdot) \in C^0(\bar{U})$ is a deterministic function. This particular $\mathcal{G}$ satisfies Assumptions A1 and A2 with $k = 0$ and $\beta = 0$. In the context of elliptic PDE, the following theorem presents sufficient conditions for Assumptions A1-A3 with $k = 1$ and $0 < \beta < 1$.

Theorem 2. *Let the functional $\mathcal{G}(\xi) = \mathcal{H}(u_\xi)$, where u_ξ is the solution to (3). Suppose that the following assumptions hold.*

H1. There exist constants $\beta, \delta_H, \kappa_H$ *such that* $\delta_H > 0$, $0 < \beta < 1$ *and* $\mathcal{H}'(u) \in C^{\beta}(\bar{U})$ *for all* $|u - u_0|_{2,\beta} \le \delta_G$*. In addition,* $\mathcal{H}'$ *is Lipschitz in the sense that*

$$|\mathcal{H}'[u_1] - \mathcal{H}'[u_2]|_{\beta} \le \kappa_H |u_1 - u_2|_{2,\beta}$$

for all $|u_1 - u_0|_{2,\beta}, |u_2 - u_0|_{2,\beta} \le \delta_H$*. Here,* $u_0 \in C^{2,\beta}(\bar{U})$ *is the solution to* (6) *when* ξ *is set to be* $\mathbf{0}$*.*

H2. There exists $x \in \bar{U}$ *such that* $\nabla g_0(x) \cdot \nabla u_0(x) \neq 0$*, where* $g_0 \in C^{2,\beta}(\bar{U})$ *is the solution to the PDE*

$$\begin{cases} -\nabla \cdot (a_0(x) \nabla g_0(x)) = \mathcal{H}'[u_0](x) & \text{for } x \in U; \\ g_0(x) = 0 & \text{for } x \in \partial U, \end{cases} \tag{12}$$

H3. U *is a bounded domain with a* $C^{2,\beta}$ *boundary* ∂U*,* $a_0 \in C^{1,\beta}(\bar{U})$*,* $\min_{x \in \bar{U}} a_0(x) > 0$ *and* $f \in C^{\beta}(\bar{U})$*.*

H4. The Gaussian random field $\{\xi(x), x \in U\}$ *is Hölder continuous and belongs to the space* $C^{k,\beta}(\bar{U})$ *almost surely. Its covariance function* $C(\cdot,\cdot)$ *is positive definite and satisfies* $\sup_{y \in \bar{U}} |C(\cdot, y)|_{1,2\beta} < \infty$*. Moreover, we assume that* $\sup_{y \in \bar{U}} |C_{D^{\gamma}\xi}(\cdot, y)|_{2\beta} < \infty$ *for all* γ *such that* $|\gamma| \le 1$*, where* $C_{D^{\gamma}\xi}$ *is defined in* (9)*.*

Then Assumptions A1-A3 are satisfied with $k = 1$ *and the Hölder coefficient being* β*.*

Under Assumption H3, the PDE (3) has a unique solution $u_0 \in C^{2,\beta}(\bar{U})$ when ξ is set to be $\mathbf{0}$. Furthermore, under Assumptions H1 and H3, (12) also has a unique solution in $C^{2,\beta}(\bar{U})$. Therefore, g_0 and u_0 in the above theorem are well defined. See Lemma 5 on page 13 for the existence and the uniqueness of the Hölder continuous solution to elliptic PDEs. Combining Theorems 1 and 2, we arrive at the next corollary.

Corollary 1. *Under the assumptions of Theorem 2, for* $0 < \alpha < 1$ *and* $b = \kappa \sigma^{\alpha}$*, we have*

$$\mathbb{P}\{\mathcal{G}(\sigma\xi) > b\} = (c_2 + o(1))\sigma^{1-\alpha} \exp\Big(-\frac{1}{2}K_{\sigma}^*\Big) \quad \text{as } \sigma \to 0,$$

where $c_2 = \kappa^{-1}\{(2\pi)^{-1}\mathcal{K}(a\nabla g_0 \cdot \nabla u_0)\}^{\frac{1}{2}}$ *and* K_{σ}^* *is the minimum obtained in* (10)*.*

2.3. **Numerical approximation.** Now we proceed to characterizing the solution to the optimization (10).

Theorem 3. *Under Assumptions A1-A3,*

(i) the optimization problem (10) *has a unique solution for* σ *sufficiently small, denoted by* ξ^**;*

(ii) we have the following approximation as $\sigma \to 0$

$$\xi^* = (1 + o_{k,\beta}(1))\kappa\sigma^{\alpha-1}\frac{\mathcal{G}'[\mathbf{0}]}{\mathcal{K}(\mathcal{G}'[\mathbf{0}])},$$

where we write $h_\sigma(\cdot) = o_{k,\beta}(1)$ *if* $|h_\sigma|_{k,\beta} = o(1)$ *as* $\sigma \to 0$.

The solution of the optimization in (10) is generally not in a closed form. Theorem 3 presents its first order approximation. It is not accurate enough for a sharp asymptotic approximation. We present further a numerical approximation for ξ^* in the following section.

In this section, we present a numerical method for computing the solution ξ^* to (10). To solve theoptimization, we introduce the Lagrangian multiplier $\lambda \in \mathbb{R}$ and define the Lagrangian function L

$$L(\xi) = \iint \xi(x)C(x,y)\xi(y)dxdy - 2\frac{\lambda}{\sigma}(\mathcal{G}(\sigma\mathbf{C}\xi) - b).$$

The first order condition $\frac{\partial L}{\partial \xi} \equiv 0$ implies the KKT condition for λ and ξ

$$\mathbf{C}\xi = \lambda\mathbf{C}\mathcal{G}'[\sigma\mathbf{C}\xi].$$

Since the covariance function $C(x,y)$ is positive definite and thus the linear map $\mathbf{C}$ is a bijection. The above condition becomes

$$\xi = \lambda\mathcal{G}'[\sigma\mathbf{C}\xi]. \tag{13}$$

The solution $(\xi^*,\ \lambda^*)$ to the constrained optimization problem is determined by

$$\begin{cases} \xi^* = \lambda^*\mathcal{G}'[\sigma\mathbf{C}\xi^*], & \text{(14a)} \\ \mathcal{G}(\sigma\mathbf{C}\xi^*) = b. & \text{(14b)} \end{cases}$$

Our strategy is to first find λ given ξ to satisfy the constraint (14b); and then we look for ξ and the corresponding $\lambda = \Lambda(\xi)$ determined by the previous step to satisfy the fix point equation (14a). Motivated by this, we define a functional

$$\Lambda : \mathcal{B} \to [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$$

such that for each $w \in \mathcal{B}$, $\lambda = \Lambda(w)$ solves the following equation

$$\mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w]) = b. \tag{15}$$

To see that $\Lambda(\cdot)$ is well defined, for each $w \in \mathcal{B}$ we define the function $T_w : [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}] \to \mathbb{R}$,

$$T_w(\lambda) = \lambda - \mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{-1}\Big(\mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w]) - b\Big).$$

Clearly, solutions to (15) are fixed points of the function $T_w(\cdot)$. The well-posedness of the function $\Lambda(\cdot)$ is then established by the next proposition.

Proposition 1. *For σ sufficiently small, $w \in \mathcal{B}$, and $|\lambda_1|, |\lambda_2| \leq \sigma^{\alpha-1-\varepsilon}$, we have that $|T_w(\lambda_1)|, |T_w(\lambda_2)| \leq \sigma^{\alpha-1-\varepsilon}$ and there exists a constant κ_T independent of σ and w, such that*

$$|T_w(\lambda_1) - T_w(\lambda_2)| \leq \kappa_T \sigma^{\alpha-\varepsilon} |\lambda_1 - \lambda_2|.$$

The above proposition and the contraction mapping theorem guarantee that for each $w \in \mathcal{B}$, $T_w(\cdot)$ has a unique fixed point in $[-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$. Therefore, there is a unique solution $\Lambda[w] \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$ satisfying (15). Furthermore, it ensures the convergence of the iterative algorithm based on the contraction mapping $T_w(\lambda)$. We further define an operator Ξ.

$$\Xi[w] = \Lambda[w]\mathcal{G}'[\sigma \mathbf{C} w]. \tag{16}$$

Proposition 2. *For σ sufficiently small, Ξ is a contraction mapping over $\mathcal{B}$. More specifically, there exists a constant κ_Ξ such that for all $w_1, w_2 \in \mathcal{B}$, we have*

$$|\Xi[w_1] - \Xi[w_2]|_{k,\beta} \leq \kappa_\Xi \sigma^\alpha |w_1 - w_2|_{k,\beta}.$$

The above proposition and the contraction mapping theorem guarantee that (14) has a unique solution (λ^*, ξ^*) in $[-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}] \times \mathcal{B}$. Furthermore, this solution can be computed numerically via the following iterative algorithm.

1. Initialize $\hat{\xi}_0^* = \kappa \sigma^{\alpha-1} \frac{\mathcal{G}'[\mathbf{0}]}{\mathcal{K}(\mathcal{G}'[\mathbf{0}])}$.
2. At l-th iteration, update $\hat{\xi}_l^*$ by
$$\hat{\xi}_l^* = \Xi[\hat{\xi}_{l-1}^*].$$

According to the contraction mapping theorem, the rate of convergence is

$$|\hat{\xi}_l^* - \xi^*|_{k,\beta} \leq (\kappa_\Xi \sigma^\alpha)^l |\hat{\xi}_0^* - \xi^*|_{k,\beta} = O(\sigma^{\alpha l + \alpha - 1}).$$

Therefore, if we run $l > \frac{2(1-\alpha)}{\alpha}$ iterations, then $|\hat{\xi}_l^* - \xi^*|_{k,\beta} = o(\sigma^{1-\alpha})$, and we could use $\mathcal{K}(\hat{\xi}_l^*)$ to approximate K_σ^* in Theorem 1.

3. Technical proofs

Throughout the proof we will use κ_0 as generic notation for large and not-so-important constants whose value may vary from place to place. Similarly, we use ε_0 as generic notation for small positive constants. Furthermore, for two sequences a_σ and b_σ, we write $a_\sigma = o(b_\sigma)$ if $b_\sigma / a_\sigma \to 0$ as σ tend to zero and $a_\sigma = O(b_\sigma)$ if b_σ / a_σ is bounded when σ varies. Moreover, for two sequences of functions $a_\sigma(\cdot)$ and $b_\sigma(\cdot)$, we write $a_\sigma = o_{k,\beta}(b_\sigma)$ if $|a_\sigma|_{k,\beta} = o(|b_\sigma|_{k,\beta})$ and $a_\sigma = O_{k,\beta}(b_\sigma)$ if $|a_\sigma|_{k,\beta} = O(|b_\sigma|_{k,\beta})$.

The proofs in this sections are organized as follows. The proof of Theorem 1 is presented in Section 3.1. Section 3.2 shows the proof of Theorem 2. Section 3.3 presents proofs of Proposition 1, 2, and 3. The proofs of supporting lemmas are postponed to Appendix A.

3.1. **Proof of Theorem 1.** We start with a useful lemma that restrict our analysis on the event $\mathcal{L} = \{\xi - \mathbf{C}\xi^* \in \mathcal{B}\}$, whose proof will be presented in Section A.

Lemma 1. *There exists positive constant ε_0 such that*

$$\mathbb{P}(\xi - \mathbf{C}\xi^* \in \mathcal{B}^c) \leq e^{-\varepsilon_0 \sigma^{2\alpha-2-2\varepsilon}}.$$

Proof for Theorem 1. Let ξ^* be the solution to (10). We define an exponential change of measure

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\Big(\int_U \xi^*(x)\xi(x)dx - \frac{1}{2}\int_U\int_U \xi^*(x)C(x,y)\xi^*(y)dsdt\Big). \tag{17}$$

Under measure $\mathbb{Q}$, $\xi(x)$ is a Gaussian random field with mean function $\mathbf{C}\xi^*(x)$ and covariance function $C(x,y)$. Let

$$\mathcal{L} = \{\xi - \mathbf{C}\xi^* \in \mathcal{B}\}.$$

According to Lemma 1, we only need to consider the event restricted to $\mathcal{L}$. By means of the change of measure $\mathbb{Q}$, we have

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{G}(\sigma\xi) > b, \mathcal{L}\right)\\
=\ & \mathbf{E}^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}};\ \mathcal{G}(\sigma\xi) > b,\ \mathcal{L}\right]\\
=\ & \exp\left(\frac{1}{2}\int_{U\times U}\xi^*(x)C(x,y)\xi^*(y)dsdt\right)\mathbf{E}^{\mathbb{Q}}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; \mathcal{G}(\sigma\xi) > b, \mathcal{L}\right],
\end{aligned} \tag{18}$$

where $\mathbf{E}^{\mathbb{Q}}$ denotes the expectation with respect to the measure $\mathbb{Q}$. It is easy to check that the random field $\mathbf{C}\xi^*(x) + \xi(x)$ under $\mathbb{P}$ has the same distribution as $\xi(x)$ under $\mathbb{Q}$. Thus, we replace the probability measure $\mathbb{Q}$ and ξ with $\mathbb{P}$ and $\mathbf{C}\xi^* + \xi$ in (18) and obtain

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{G}(\sigma\xi) > b, \mathcal{L}\right)\\
=\ & \exp\left(\frac{1}{2}\int_{U\times U}\xi^*(x)C(x,y)\xi^*(y)dxdy\right)\mathbf{E}\left[e^{-\int_U \xi^*(x)(\mathbf{C}\xi^*(x)+\xi(x))dx}; \mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) > b, \xi \in \mathcal{B}\right]\\
=\ & \exp\left(-\frac{1}{2}\int_{U\times U}\xi^*(x)C(x,y)\xi^*(y)dsdt\right)\mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; \mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*) > 0, \xi \in \mathcal{B}\right]\\
=\ & e^{-\frac{1}{2}K_\sigma^*} \times \mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; \mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*) > 0, \xi \in \mathcal{B}\right].
\end{aligned}$$

We define two events

$$F = \{\mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*) > 0\}, \text{ and } F_1 = \Big\{ \int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\sigma\xi(x)dx > 0 \Big\}.$$

Let the event $\mathcal{L}_1 = \{\xi \in \mathcal{B}\}$. We will present an approximation for

$$I_1 = \mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; F_1\right]$$

and show that

$$I_2 = \mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; (F_1 \triangle F) \cap \mathcal{L}_1\right]$$

is ignorable, where "$\triangle$" denotes the symmetric difference between two sets. First, we compute

$$I_1 = \mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; \int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx > 0\right]. \tag{19}$$

According to Proposition 2, ξ^* is the fixed point of the contraction map Ξ and thus

$$\xi^* = \Xi[\xi^*] = \Lambda[\xi^*]\mathcal{G}'[\sigma\mathbf{C}\xi^*].$$

Therefore, ξ^* and $\mathcal{G}'[\sigma\mathbf{C}\xi^*]$ are different only by a factor of $\Lambda[\xi^*]$. Thus, $\int_U \xi^*(x)\xi(x)dx$ and $\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx > 0$ are different by a factor $\Lambda[\xi^*]$. The following lemma establishes an approximation for $\Lambda[\xi^*]$.

Lemma 2. *For all $w \in \mathcal{B}$, $\Lambda[w] = \kappa\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{\alpha-1}(1 + o(1))$. This approximation is uniform in w.*

Thanks to Lemma 2, we have

$$\Lambda[\xi^*] = (1 + o(1))\frac{\kappa\sigma^{\alpha-1}}{\mathcal{K}(\mathcal{G}'[\mathbf{0}])}.$$

Let $Z_1 = \int_U \xi^*(x)\xi(x)dx$, then Z_1 is a normally distributed random variable with a zero mean. The expectation (19) can be computed as follows

$$\begin{aligned} &\mathbf{E}\left[e^{-Z_1}; Z_1 > 0,\right] \\ =\ & \int_0^\infty \frac{1}{\sqrt{2\pi \operatorname{Var}(Z_1)}} e^{-\frac{z_1^2}{2\operatorname{Var}(Z_1)} - z_1} dz_1 \\ =\ & \frac{1}{\sqrt{2\pi \operatorname{Var}(Z_1)}} E[e^{-\frac{V^2}{2\operatorname{Var}(Z_1)}}], \end{aligned} \tag{20}$$

where V is a random variable following the exponential distribution with rate 1. Notice that

$$\operatorname{Var}(Z_1) = \int_{U\times U} \xi^*(x)C(x,y)\xi^*(y)dxdy = (1 + o(1))\kappa^2\sigma^{2\alpha-2}\mathcal{K}^{-1}[\mathcal{G}'[\mathbf{0}]]. \tag{21}$$

The second equality is obtained with the aid of Proposition 3(ii). The above display, (20) and dominated convergence theorem give

$$I_1 = \kappa^{-1}\{(2\pi)^{-1}\mathcal{K}(\mathcal{G}'[\mathbf{0}])\}^{1/2}\sigma^{1-\alpha}(1+o(1)).$$

Now, we proceed to the term I_2.

Lemma 3. *Under Assumption A1, we have that for $|w_1|_{k,\beta}, |w_2|_{k,\beta} \le \delta_G$,*

$$|w_1 - w_2|_{k,\beta}^{-2}\Big|\mathcal{G}(w_1) - \mathcal{G}(w_2) - \int_U \mathcal{G}'[w_2](x)(w_1(x) - w_2(x))dx\Big| \le meas(U)\kappa_G,$$

where $meas(U)$ is the Lebesgue measure of U and $k, \beta, \delta_G, \kappa_G$ are constants appeared in Assumption A1.

According to Lemma 3, we have that for σ sufficiently small and $\xi \in \mathcal{B}$,

$$\Big|\mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*) - \sigma\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx\Big| \le meas(U)\kappa_G\sigma^2|\xi|_{k,\beta}^2. \tag{22}$$

Note that on the event $F_1 \triangle F$, $\mathcal{G}(\sigma(\xi+\mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*)$ and $\sigma\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx$ have opposite signs and thus

$$\Big|\mathcal{G}(\sigma(\xi + \mathbf{C}\xi^*)) - \mathcal{G}(\sigma\mathbf{C}\xi^*) - \sigma\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx\Big| \ge |\sigma\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx|. \tag{23}$$

We combine (22) and (23) and arrive at

$$(F\triangle F_1)\cap\mathcal{L}_1 \subset \Big\{meas(U)\kappa_G\|\xi\|_{k,\beta}^2 \ge \sigma^{-1}|\int_U \mathcal{G}'[\sigma\mathbf{C}\xi^*](x)\xi(x)dx|\Big\}\cap\mathcal{L}_1.$$

We write $Z_2 = \|\xi\|_{k,\beta}^2$, then the above display implies that

$$(F\triangle F_1)\cap\mathcal{L}_1 \subset \{meas(U)\kappa_G Z_2 \ge \sigma^{-1}\Lambda[\xi^*]|Z_1|\}\cap\mathcal{L}_1.$$

This gives an upper bound of the expectation

$$\mathbf{E}\left[e^{-\int_U \xi^*(x)\xi(x)dx}; (F\triangle F_1)\cap\mathcal{L}_1\right] \le \mathbf{E}\left[e^{-Z_1}; \kappa_G Z_2 \ge \sigma^{-1}\Lambda[\xi^*]|Z_1|, \mathcal{L}_1\right].$$

On the event $\{0 < |Z_1| \le \sigma^\varepsilon\}$, this expectation is negligible compared to I_1, that is,

$$E[e^{Z_1}; 0 < |Z_1| < \sigma^\varepsilon] = O(\mathbb{P}(0 < |Z_1| < \sigma^\varepsilon)) = O(\sigma^{1-\alpha+\varepsilon}). \tag{24}$$

The second equality in the above display is due to (21). Furthermore, on the set $\mathcal{L}_1$, we have $|Z_1| \le |\xi^*|_0|\xi|_0\kappa_0 \le \kappa_0\sigma^{2\alpha-2-\varepsilon}$, where κ_0 is a sufficiently large constant. Therefore, we only need to focus on the expectation

(25)

$$\mathbf{E}\left[e^{Z_1};\sigma^\varepsilon<|Z_1|<\kappa_0\sigma^{2\alpha-2-\varepsilon},Z_2>\Lambda(\xi^*)|Z_1/\sigma|\right]=\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^z\mathbb{P}(Z_2>\Lambda(\xi^*)z/\sigma|Z_1=z)p_{Z_1}(z)dz$$
$$+\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^z\mathbb{P}(Z_2>\Lambda(\xi^*)z/\sigma|Z_1=-z)p_{Z_1}(z)dz,$$

where $p_{Z_1}(z)$ is the density function of Z_1.

Lemma 4. *For $z\in[\sigma^\varepsilon,\kappa_0\sigma^{2\alpha-2-\varepsilon}]$, there exists a constant $\varepsilon_0>0$ such that*

$$\mathbb{P}(Z_2>\Lambda(\xi^*)z/\sigma|Z_1=z)+\mathbb{P}(Z_2>\Lambda(\xi^*)z/\sigma|Z_1=-z)\le e^{-\varepsilon_0\sigma^{\alpha-2}z}. \tag{26}$$

With the above lemma, the expectation (25) is bounded by

$$\begin{aligned}(25)&\le\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^{-(\varepsilon_0\sigma^{\alpha-2}-1)z}p_{Z_1}(z)dz\\&=\frac{1}{\sqrt{2\pi\operatorname{Var}(Z_1)}}\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^{-(\varepsilon_0\sigma^{\alpha-2}-1)z-\frac{z^2}{2\operatorname{Var}(Z_1)}}dz\\&\le\frac{1}{\sqrt{2\pi\operatorname{Var}(Z_1)}}\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^{-\frac{\varepsilon_0}{2}\sigma^{\alpha-2}z}dz,\end{aligned}$$

for σ sufficiently small so that $\varepsilon_0\sigma^{\alpha-2}-1>\frac{\varepsilon_0}{2}\sigma^{\alpha-2}$. The above inequality is further bounded by

$$(25)\le\frac{1}{\sqrt{2\pi\operatorname{Var}(Z_1)}}\int_{\sigma^\varepsilon}^{\kappa_0\sigma^{2\alpha-2-\varepsilon}}e^{-\frac{\varepsilon_0}{2}\sigma^{\alpha-2}z}dz\le\frac{1}{\sqrt{2\pi\operatorname{Var}(Z_1)}}\kappa_0\sigma^{2\alpha-2-\varepsilon}e^{-\frac{\varepsilon_0}{2}\sigma^{\alpha-2+\varepsilon}}=O(e^{-\frac{\varepsilon_0}{2}\sigma^{\alpha-2+\varepsilon}}).$$

Therefore,

$$(25)=o(\sigma^{1-\alpha}).$$

We combine our analysis for I_1 and I_2 and conclude our proof for Theorem 1. □

3.2. **Proof of Theorem 2.**

Proof of Theorem 2. We first present two useful lemmas. The following lemma guarantees the existence and uniqueness of the Hölder continuous solution to the elliptic PDE.

Lemma 5. *Suppose that U is a bounded domain with a $C^{2,\beta}$ boundary ∂U for $0<\beta<1$. Assume that there exist positive constants δ and M such that $\min_{x\in\bar U}a(x)>\delta>0$, and $|a|_{1,\beta}\le M$, and*

$f \in C^{\beta}(\bar U)$. Then the elliptic PDE

$$\begin{cases} -\nabla \cdot (a(x)\nabla u(x)) = f(x) & \textit{for } x \in U; \\ u(x) = 0 & \textit{for } x \in \partial U, \end{cases} \tag{27}$$

has a unique solution in $C^{2,\beta}(\bar U)$. Denote this solution by $u_{a,f}$, then

$$|u_{a,f}|_{2,\beta} \le \kappa(\delta, M, d, U)|f|_{\beta}, \tag{28}$$

where $\kappa(\delta, M, d, U)$ is a positive constant, depending only on δ, M, d and the domain U.

We will also need the following lemma on the stability of the solution.

Lemma 6. *Suppose that U is a bounded domain with a $C^{2,\beta}$ boundary ∂U for $0 < \beta < 1$. Let a_1, a_2, f_1 and f_2 be functions over the domain U such that*

$$\min_{x\in\bar U} a_1(x) \ge \delta, \qquad \min_{x\in\bar U} a_2(x) \ge \delta, \qquad |a_1|_{1,\beta}, |a_2|_{1,\beta} \le M, \ \textit{and } f_1, f_2 \in C^{\beta}(\bar U).$$

Then,

$$|u_{a_1,f_1} - u_{a_2,f_2}|_{2,\beta} \le \tilde\kappa(\delta, M, d, U)\{|f_1 - f_2|_{\beta} + |a_1 - a_2|_{1,\beta}|f_1|_{\beta}\},$$

where the constant $\tilde\kappa(\delta, M, d, U)$ depends only on δ, M, d and the domain U.

The Fréchet derivative $\mathcal{G}'[w]$ has the following expression.

$$\mathcal{G}'[w](x) = a_w(x)\nabla g_w(x) \cdot \nabla u_w(x),$$

where $a_w(x) = a_0 e^{-w(x)}$, $u_w \in C^{2,\beta}(\bar U)$ is the unique solution to

$$\begin{cases} -\nabla \cdot (a_w(x)\nabla u_w(x)) = f(x) & \text{for } x \in U; \\ u_w(x) = 0 & \text{for } x \in \partial U, \end{cases}$$

and $g_w(x) \in C^{2,\beta}(\bar U)$ is the unique solution to

$$\begin{cases} -\nabla \cdot (a_w(x)\nabla g_w(x)) = \mathcal{H}'[u_w](x) & \text{for } x \in U; \\ g_w(x) = 0 & \text{for } x \in \partial U. \end{cases}$$

For $w_1, w_2 \in C^{1,\beta}(\bar{U})$, we are going to establish an upper bound for $|\mathcal{G}'[w_1] - \mathcal{G}'[w_2]|_{1,\beta}$. Note that

$$\begin{aligned}
& \mathcal{G}'[w_1](x) - \mathcal{G}'[w_2](x) \\
= \; & (a_{w_1}(x) - a_{w_2}(x))\nabla g_{w_1}(x) \cdot \nabla u_{w_1}(x) \\
& + a_{w_2}\nabla g_{w_2}(x)\nabla(u_{w_1} - u_{w_2}(x)) + a_{w_2}(x)\nabla(g_{w_1}(x) - g_{w_2}(x)) \cdot \nabla u_{w_1}(x).
\end{aligned}$$

Thus,

$$\begin{aligned}
& |\mathcal{G}'[w_1] - \mathcal{G}'[w_2]|_{1,\beta} \\
\leq \; & |(a_{w_1} - a_{w_2})\nabla g_{w_1} \cdot \nabla u_{w_1}|_{1,\beta} \\
& + |a_{w_2}\nabla g_{w_2}\nabla(u_{w_1} - u_{w_2})|_{1,\beta} + |a_{w_2}\nabla(g_{w_1} - g_{w_2}) \cdot \nabla u_{w_1}|_{1,\beta}.
\end{aligned} \tag{29}$$

We will establish upper bounds for the three terms on the right-hand side in the above expression separately. First, note that $a_{w_k} = a_0 e^{-w_k}$, $k = 1, 2$. Thus, there exists a constant $\varepsilon_0 > 0$ such that for all $|w_1|_{1,\beta}, |w_2|_{1,\beta} \leq \varepsilon_0$,

$$|a_{w_1} - a_{w_2}|_{1,\beta} \leq \kappa_0 |w_1 - w_2|_{1,\beta}. \tag{30}$$

Therefore,

$$|(a_{w_1} - a_{w_2})\nabla g_{w_1} \cdot \nabla u_{w_1}|_{1,\beta} \leq |a_{w_1} - a_{w_2}|_{1,\beta} |\nabla g_{w_1}|_{1,\beta} |\nabla u_{w_1}|_{1,\beta} \leq \kappa_0 |w_1 - w_2|_{1,\beta} |g_{w_1}|_{2,\beta} |u_{w_1}|_{2,\beta}. \tag{31}$$

Now we present upper bounds for $|g_{w_1}|_{2,\beta}$ and $|u_{w_1}|_{2,\beta}$. Let ε_0 be sufficiently small such that for all $|w|_{1,\beta} \leq \varepsilon_0$, $\min_{x \in \bar{U}} a_w(x) \geq \frac{1}{2} \min_{x \in \bar{U}} a_0(x)$ and $|a_w|_{1,\beta} \leq 2|a_0|_{1,\beta}$. According to Lemma 5, we have that for all $|w|_{1,\beta} \leq \delta_0$

$$|u_w|_{2,\beta} \leq \kappa(\delta, M, d, U)|f|_\beta, \tag{32}$$

where $\delta = \frac{\min_{x \in \bar{U}} a_0(x)}{2}$ and $M = 2|a_0|_{1,\beta}$. Furthermore, according to Assumption H1, we have that for $|u_w - u_0|_{2,\beta} \leq \delta_H$

$$|\mathcal{H}'[u_w]|_\beta \leq |\mathcal{H}'[u_0]|_\beta + \kappa_H |u_w - u_0|_{2,\beta} \leq |\mathcal{H}'[u_0]|_\beta + \kappa_H \delta_H. \tag{33}$$

Set $f = \mathcal{H}'[u_w]$ in Lemma 5 we have

$$|g_w|_{2,\beta} \leq \kappa(\delta, M, d, U)|\mathcal{H}'[u_w]|_\beta \leq \kappa(\delta, M, d, U)(|\mathcal{H}'[u_0]|_\beta + \kappa_H \delta_H). \tag{34}$$

Combine this with (31) and (32), we have that for $|w_1|_{1,\beta}, |w_2|_{1,\beta} \leq \varepsilon_0$

$$|(a_{w_1} - a_{w_2})\nabla g_{w_1} \cdot \nabla u_{w_1}|_{1,\beta} \leq \kappa_0 |w_1 - w_2|_{1,\beta}, \tag{35}$$

with a possibly different κ_0. We proceed to the second term on the right-hand side of (29).

(36)
$$|a_{w_2}\nabla g_{w_2}\cdot\nabla(u_{w_1}-u_{w_2})|_{1,\beta}\le|a_{w_2}|_{1,\beta}|\nabla g_{w_2}|_{1,\beta}|\nabla(u_{w_1}-u_{w_2})|_{1,\beta}\le|a_{w_2}|_{1,\beta}|g_{w_2}|_{2,\beta}|u_{w_1}-u_{w_2}|_{2,\beta}.$$

For $|w_2|_{1,\beta}\le\varepsilon_0$, we have $|a_{w_2}|_{1,\beta}\le 2|a_0|_{1,\beta}$. Moreover, $|g_{w_2}|_{2,\beta}$ is bounded above by a constant according to (34). Therefore,

$$|a_{w_2}\nabla g_{w_2}\cdot\nabla(u_{w_1}-u_{w_2})|_{1,\beta}\le\kappa_0|u_{w_1}-u_{w_2}|_{2,\beta}, \tag{37}$$

for a possibly different κ_0. Taking $a_1=a_{w_1}$, $a_2=a_{w_2}$, and $f_1=f_2=f$ in Lemma 6, we have

$$|u_{w_1}-u_{w_2}|_{2,\beta}\le\tilde\kappa(\delta,M,d,U)|a_1-a_2|_{1,\beta}|f_{body}|_\beta\le\kappa_0|w_1-w_2|_{1,\beta}. \tag{38}$$

(37) and (38) give

$$|a_{w_2}\nabla g_{w_2}\cdot\nabla(u_{w_1}-u_{w_2})|_{1,\beta}\le\kappa_0^2|w_1-w_2|_{1,\beta}. \tag{39}$$

We proceed to the third term on the right-hand side of (29).

(40)
$$|a_{w_2}\nabla(g_{w_1}-g_{w_2})\cdot\nabla u_{w_1}|_{1,\beta}\le|a_{w_2}|_{1,\beta}|\nabla(g_{w_1}-g_{w_2})|_{1,\beta}|\nabla u_{w_1}|_{1,\beta}\le|a_{w_2}|_{1,\beta}|g_{w_1}-g_{w_2}|_{2,\beta}|u_{w_1}|_{2,\beta}.$$

According to the definition of a_{w_2} and (38), we have that for $|w_1|_{1,\beta},|w_2|_{1,\beta}\le\varepsilon_0$,

$$|a_{w_2}\nabla(g_{w_1}-g_{w_2})\cdot\nabla u_{w_1}|_{1,\beta}\le\kappa_0|g_{w_1}-g_{w_2}|_{2,\beta}. \tag{41}$$

Motivated by the definition of g_{w_1} and g_{w_2}, we take $f_1=\mathcal{H}'[w_1]$, $f_2=\mathcal{H}'[w_2]$, $a_1=a_{w_1}$ and $a_2=a_{w_2}$ in Lemma 6, then

$$|g_{w_1}-g_{w_2}|_{2,\beta}\le\tilde\kappa(\delta,M,d,U)\{|\mathcal{H}'[w_1]-\mathcal{H}'[w_2]|_\beta+|a_{w_1}-a_{w_2}|_{1,\beta}|\mathcal{H}'[w_1]|_\beta\}. \tag{42}$$

According to Assumption H1, for $|w_1|_{1,\beta},|w_2|_{1,\beta}\le\delta_H$, we have

$$|\mathcal{H}'[w_1]-\mathcal{H}'[w_2]|_\beta\le\kappa_H|w_1-w_2|_{1,\beta} \tag{43}$$

(43), (30), (33) and (42) give

$$|g_{w_1}-g_{w_2}|_{2,\beta}\le\kappa_0|w_1-w_2|_{1,\beta}.$$

The above inequality and (41) give

$$|a_{w_2}\nabla(g_{w_1}-g_{w_2})\cdot\nabla u_{w_1}|_{1,\beta}\le\kappa_0^2|w_1-w_2|_{1,\beta} \tag{44}$$

We combine (29), (35), (39), and (44), and arrive at

$$|\mathcal{G}'[w_1] - \mathcal{G}'[w_2]|_{1,\beta} \le \kappa_0 |w_1 - w_2|_{1,\beta}, \tag{45}$$

for ε_0 sufficiently small, $|w_1|_{1,\beta}, |w_2|_{1,\beta} \le \varepsilon_0$ and a possibly different κ_0. Thus, Assumption A1 is satisfied with $k = 1$. According to the definition of $\mathcal{G}'$, Assumption A2 is a dirrect application of Assumption H2. Assumption A3 is the same Assumption H4 for $k = 1$. Now we have already checked all the Assumptions A1-A3. □

3.3. **Proof of propositions.**

Proof of Proposition 1 . Note that as σ tends to zero, we have $\sigma \mathbf{C} w = o_{k,\beta}(1)$, $\mathcal{G}'[\sigma \mathbf{C} w] = \mathcal{G}'[\mathbf{0}] + o_{k,\beta}(1)$ and $\sigma \mathbf{C} \lambda \mathcal{G}'[\sigma \mathbf{C} w] = o_{k,\beta}(1)$ for all $|\lambda| \le \sigma^{\alpha - 1 - \varepsilon}$ and $w \in \mathcal{B}$. This allow us to expand $\mathcal{G}(\sigma \mathbf{C} \lambda \mathcal{G}'[\sigma \mathbf{C} w])$ near the origin. We elaborate this expansion as follows. First, according to Assumption A1, we have that there exists a constant ε_0 such that for all $w \in \mathcal{B}$ and $\sigma \le \varepsilon_0$,

$$\mathcal{G}'[\sigma \mathbf{C} w] = \mathcal{G}'[\mathbf{0}] + O_{k,\beta}(\sigma \mathbf{C} w). \tag{46}$$

Second, with the aid of (46) we have that for all $|\lambda_1|, |\lambda_2| \le \sigma^{\alpha - 1 - \varepsilon}$ and $w \in \mathcal{B}$,

$$\sigma \mathbf{C} \lambda_1 \mathcal{G}'[\sigma \mathbf{C} w] - \sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w] = \sigma(\lambda_1 - \lambda_2) \mathbf{C} \{\mathcal{G}'[\mathbf{0}] + O_{k,\beta}(\sigma \mathbf{C} w)\}. \tag{47}$$

Thanks to Lemma 3 on page 12 and (47), we have that for all $|\lambda_1|, |\lambda_2| \le \sigma^{\alpha - 1 - \varepsilon}$ and $w \in \mathcal{B}$,

$$\mathcal{G}(\sigma \mathbf{C} \lambda_1 \mathcal{G}'[\sigma \mathbf{C} w]) - \mathcal{G}(\sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]) = \int_U \mathcal{G}'[\sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]](x) v(x) dx + O(|v|_{k,\beta}^2), \tag{48}$$

where we define

$$v(x) = \sigma \mathbf{C} \lambda_1 \mathcal{G}'[\sigma \mathbf{C} w](x) - \sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w](x).$$

Setting w as $\lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]$ in (46), we have

$$\mathcal{G}'[\sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]] = \mathcal{G}'[\mathbf{0}] + O_{k,\beta}(\sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]) = \mathcal{G}'[\mathbf{0}] + O_{k,\beta}(\sigma \lambda_2 \mathcal{G}'[\mathbf{0}]). \tag{49}$$

The last equality in the above display is due to (46) and the fact $O_{k,\beta}(\sigma \mathbf{C} w) = o_{k,\beta}(1)$. According to (47) and (49), we have

$$\begin{aligned}
& \int_U \mathcal{G}'[\sigma \mathbf{C} \lambda_2 \mathcal{G}'[\sigma \mathbf{C} w]](x) v(x) dx \\
= \; & \sigma(\lambda_1 - \lambda_2) \Big\{ \int_U \mathbf{C} \mathcal{G}'[\mathbf{0}](x) \mathcal{G}'[\mathbf{0}](x) dx + O(\int_U \sigma^2 \lambda_2 \mathcal{G}'[\mathbf{0}](x) \mathcal{G}'(x) dx) + O(\int_U \sigma \mathcal{G}'[\mathbf{0}](x) \mathbf{C} w(x) dx) \\
& + O(\sigma \lambda_2 \sigma \int_U \mathcal{G}'[\mathbf{0}](x) \mathbf{C} w(x) dx) \Big\}.
\end{aligned}$$

Note that for $\lambda_2 \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$ the above expression is simplified as

$$\begin{aligned}\int_U \mathcal{G}'[\sigma\mathbf{C}\lambda_2\mathcal{G}'[\sigma\mathbf{C}w]](x)v(x)dx &= \sigma(\lambda_1-\lambda_2)\Big\{\int_U \mathbf{C}\mathcal{G}'[\mathbf{0}](x)\mathcal{G}'[\mathbf{0}](x)dx + O(\sigma^{\alpha-\varepsilon})\Big\} \\ &= \sigma(\lambda_1-\lambda_2)\{\mathcal{K}(\mathcal{G}'[\mathbf{0}]) + O(\sigma^{\alpha-\varepsilon})\}. \end{aligned} \tag{50}$$

Combining the above expression with (48), we have that for $|\lambda_1|, |\lambda_2| \le \sigma^{\alpha-1-\varepsilon}$ and $w \in \mathcal{B}$.

$$\mathcal{G}(\sigma\mathbf{C}\lambda_1\mathcal{G}'[\sigma\mathbf{C}w]) - \mathcal{G}(\sigma\mathbf{C}\lambda_2\mathcal{G}'[\sigma\mathbf{C}w]) = \sigma(\lambda_1-\lambda_2)\{\mathcal{K}(\mathcal{G}'[\mathbf{0}]) + O(\sigma^{\alpha-\varepsilon})\} + O(\sigma^2(\lambda_1-\lambda_2)^2),$$

which can be simplified as

$$\mathcal{G}(\sigma\mathbf{C}\lambda_1\mathcal{G}'[\sigma\mathbf{C}w]) - \mathcal{G}(\sigma\mathbf{C}\lambda_2\mathcal{G}'[\sigma\mathbf{C}w]) = \sigma(\lambda_1-\lambda_2)\{\mathcal{K}(\mathcal{G}'[\mathbf{0}]) + O(\sigma^{\alpha-\varepsilon})\}. \tag{51}$$

Recall the definition of $T_w(\lambda)$, we plug the above expression into the difference $T_w(\lambda_1) - T_w(\lambda_2)$, and arrive at

$$T_w(\lambda_1) - T_w(\lambda_2) = \lambda_1 - \lambda_2 - \mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{-1} \times \sigma(\lambda_1-\lambda_2)\{\mathcal{K}(\mathcal{G}'[\mathbf{0}]) + O(\sigma^{\alpha-\varepsilon})\},$$

which is simplified as

$$T_w(\lambda_1) - T_w(\lambda_2) = -\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}(\lambda_1-\lambda_2) \times O(\sigma^{\alpha-\varepsilon}).$$

The above expression implies that for $|\lambda_1|, |\lambda_2| \le \sigma^{\alpha-1-\varepsilon}$,

$$T_w(\lambda_1) - T_w(\lambda_2) = (\lambda_1-\lambda_2) \times O(\sigma^{\alpha-\varepsilon}). \tag{52}$$

This shows that $T_w(\lambda)$ is a contraction mapping for $\lambda \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$. To see $T_w(\lambda) \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$ for $\lambda \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}]$ and $w \in \mathcal{B}$, we let $\lambda_2 = 0$ and $\lambda_1 = \lambda$ in (52) and obtain that

$$T_w(\lambda) - T_w(0) = \lambda O(\sigma^{\alpha-\varepsilon}) = O(\sigma^{2\alpha-1-2\varepsilon}).$$

Recall that $b = \kappa\sigma^\alpha$, and $T_w(0) = -\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{-1}b = -\kappa\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{\alpha-1}$. This implies

$$T_w(\lambda) = \kappa\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{\alpha-1}(1+o(1)) \in [-\sigma^{\alpha-1-\varepsilon}, \sigma^{\alpha-1-\varepsilon}] \tag{53}$$

and concludes our proof. □

Proof of Proposition 2. According to the definition of Ξ,

$$\Xi[w_1] - \Xi[w_2] = \Lambda[w_1](\mathcal{G}'[\sigma\mathbf{C}w_1] - \mathcal{G}'[\sigma\mathbf{C}w_2]) + (\Lambda[w_1] - \Lambda[w_2])\mathcal{G}'[\sigma\mathbf{C}w_2].$$

Therefore, we have

$$|\Xi[w_1] - \Xi[w_2]|_{k,\beta} \leq |\Lambda[w_1]| \times |(\mathcal{G}'[\sigma \mathbf{C} w_1] - \mathcal{G}'[\sigma \mathbf{C} w_2])|_{k,\beta} + |\Lambda[w_1] - \Lambda[w_2]| \times |\mathcal{G}'[\sigma \mathbf{C} w_2]|_{k,\beta}. \tag{54}$$

We establish upper bound for the first and second terms on the right-hand-side of the above inequality separately. To start with, according to Assumptions A1 and A3 that $\sup_{y \in \bar{U}} |C(\cdot, y)|_{k,2\beta} < \infty$, for $w_1, w_2 \in \mathcal{B}$, we have

$$|\Lambda[w_1]| \times |(\mathcal{G}'[\sigma \mathbf{C} w_1] - \mathcal{G}'[\sigma \mathbf{C} w_2])|_{k,\beta} = O(\sigma |\Lambda[w_1]| |w_1 - w_2|_{k,\beta}) = O(\sigma^\alpha) |w_1 - w_2|_{k,\beta}. \tag{55}$$

The second equality in the above expression is due to Lemma 2 on page 11. We proceed to the second term on the right-hand-side of (54). Because $\Lambda[w]$ is the fixed point of $T_w(\cdot)$, we have

$$T_{w_1}(\Lambda[w_1]) = \Lambda[w_1] \text{ and } T_{w_2}(\Lambda[w_2]) = \Lambda[w_2].$$

Taking differencing between the above two equalities, we have

$$T_{w_1}(\Lambda[w_1]) - T_{w_2}(\Lambda[w_2]) = \Lambda[w_1] - \Lambda[w_2].$$

Adding and subtracting the term $T_{w_1}(\Lambda[w_2])$ in the above equality, we have

$$\Lambda[w_1] - \Lambda[w_2] = T_{w_1}(\Lambda[w_1]) - T_{w_1}(\Lambda[w_2]) + T_{w_1}(\Lambda[w_2]) - T_{w_2}(\Lambda[w_2]).$$

Consequently,

$$|\Lambda[w_1] - \Lambda[w_2]| \leq |T_{w_1}(\Lambda[w_1]) - T_{w_1}(\Lambda[w_2])| + |T_{w_1}(\Lambda[w_2]) - T_{w_2}(\Lambda[w_2])|. \tag{56}$$

According to Proposition 1, the first term on the right-hand-side of the above expression is bounded above by $O(\sigma^{\alpha-\varepsilon})|\Lambda[w_1] - \Lambda[w_2]|$.

Lemma 7. *For all $|\lambda| = O(\sigma^{\alpha-1})$ and $w_1, w_2 \in \mathcal{B}$, we have*

$$|T_{w_1}(\lambda) - T_{w_2}(\lambda)| = O(\sigma^\alpha)|w_1 - w_2|_{k,\beta}.$$

According to Lemma 7, the second term on the right-hand-side of (56) is bounded above by $O(\sigma^\alpha)|w_1 - w_2|_{k,\beta}$. Therefore, we have

$$|\Lambda[w_1] - \Lambda[w_2]| \leq O(\sigma^{\alpha-\varepsilon})|\Lambda[w_1] - \Lambda[w_2]| + O(\sigma^\alpha)|w_1 - w_2|_{k,\beta}.$$

Consequently, we have that for $w_1, w_2 \in \mathcal{B}$,

$$|\Lambda[w_1] - \Lambda[w_2]| = O(\sigma^\alpha)|w_1 - w_2|_{k,\beta}. \tag{57}$$

According to (46),

$$|\mathcal{G}'[\sigma \mathbf{C} w_2]|_{k,\beta} = O(1).$$

The above approximation and (57) give

$$|\Lambda[w_1] - \Lambda[w_2]| \times |\mathcal{G}'[\sigma \mathbf{C} w_2]|_{k,\beta} = O(\sigma^{\alpha})|w_1 - w_2|_{k,\beta}.$$

Combining the above display with (54) and (55), we complete our proof. □

Proof of Proposition 3. (i) is a direct application of Proposition 2, contraction mapping theorem and the KKT condition (14). We proceed to the proof of (ii). Because ξ^* is the fixed point of Ξ in $\mathcal{B}$, we have

$$\Xi[\xi^*] = \Lambda(\Xi^*)\mathcal{G}'[\sigma \mathbf{C}\xi^*] = \kappa\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{\alpha-1}(1+o(1))(\mathcal{G}'[\mathbf{0}] + O_{k,\beta}(\sigma\xi^*)) = (1+o_{k,\beta}(1))\frac{\kappa\mathcal{G}'[\mathbf{0}]}{\mathcal{K}(\mathcal{G}'[\mathbf{0}])}.$$

To obtain the second equality in the above display, we use approximation in Lemma 2 on page 11 and (46). □

Acknowledgement

Jingchen Liu is partially supported by the National Science Foundation (SES-1323977, IIS-1633360) and Army Grant (W911NF-15-1-0159). Jianfeng Lu is partially supported by National Science Foundation (DMS-1454939). Xiang Zhou acknowledges the support from Hong Kong General Research Fund (109113, 11304314, 11304715).

References

[Adler, 1981] Adler, R. (1981). *The Geometry of Random Fields*. Wiley, Chichester, U.K.; New York, U.S.A.

[Adler et al., 2008] Adler, R., Blanchet, J., and Liu, J. (2008). Efficient simulation for tail probabilities of Gaussian random fields. In *Proceeding of Winter Simulation Conference*.

[Adler et al., 2012] Adler, R., Blanchet, J., and Liu, J. (2012). Efficient Monte Carlo for large excursions of Gaussian random fields. *Ann. Appl. Probab.*, 22(3):1167–1214.

[Adler and Taylor, 2007] Adler, R. and Taylor, J. (2007). *Random fields and geometry*. Springer.

[Azais and Wschebor, 2008] Azais, J. M. and Wschebor, M. (2008). A general expression for the distribution of the maximum of a Gaussian field and the approximation of the tail. *Stochastic Processes and their Applications*, 118(7):1190–1218.

[Azais and Wschebor, 2009] Azais, J. M. and Wschebor, M. (2009). *Level sets and extrema of random processes and fields*. Wiley, Hoboken, N.J.

[Bear and Verruijt, 1987] Bear, J. and Verruijt, A. (1987). *Modeling Groundwater Flow and Pollution*. D. Reidel Publishing Company, Holland.

[Berman, 1985] Berman, S. M. (1985). An asymptotic formula for the distribution of the maximum of a Gaussian process with stationary increments. *Journal of Applied Probability*, 22(2):454–460.

[Borell, 1975] Borell, C. (1975). The brunn-minkowski inequality in gauss space. *Inventiones mathematicae*, 30:207–216.

[Borell, 2003] Borell, C. (2003). The Ehrhard inequality. *Comptes Rendus Mathematique*, 337(10):663–666.

[Charbeneau, 2000] Charbeneau, R. J. (2000). *Groundwater Hydraulics and Pollutant Transport*. Prentice Hall.

[Cirel'son et al., 1976] Cirel'son, B. S., Ibragimov, I. A., and Sudakov, V. N. (1976). Norms of gaussian sample functions. In Maruyama, G. and Prokhorov, J. V., editors, *Proceedings of the Third Japan — USSR Symposium on Probability Theory*, pages 20–41. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Freeze, 1975] Freeze, R. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resour. Res.*, 11.

[Gilbarg and Trudinger, 2015] Gilbarg, D. and Trudinger, N. S. (2015). *Elliptic partial differential equations of second order*. springer.

[Landau and Shepp, 1970] Landau, H. J. and Shepp, L. A. (1970). Supremum of a Gaussian process. *Sankhya-the Indian Journal of Statistics Series A*, 32(Dec):369–378.

[Ledoux and Talagrand, 1991] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach spaces: isoperimetry and processes*. Ergebnisse der Mathematik und ihrer Grenzgebiete 3. Folge, Bd. 23. Springer-Verlag, Berlin ; New York.

[Li et al., 2015] Li, X., Liu, J., and Xu, G. (2015). On the tail probabilities of aggregated lognormal random fields with small noise. *Mathematics of Operations Research*.

[Li et al., 2016] Li, X., Liu, J., and Xu, G. (2016). On the tail probabilities of aggregated lognormal random fields with small noise. *Mathematics of Operations Research*, 41(1):236–246.

[Liu, 2012] Liu, J. (2012). Tail approximations of integrals of Gaussian random fields. *Ann. Probab.*, 40:1069–1104.

[Liu et al., 2015] Liu, J., Lu, J., and Zhou, X. (2015). Efficient rare event simulation for failure problems in random media. *SIAM Journal on Scientific Computing*, 37(2):A609–A624.

[Liu and Xu, 2012] Liu, J. and Xu, G. (2012). Some asymptotic results of Gaussian random fields with varying mean functions and the associated processes. *Ann. Stat.*, 40:262–293.

[Liu and Zhou, 2013] Liu, J. and Zhou, X. (2013). On the failure probability for one dimensional random material under delta external force. *Commun. Math. Sci.*, 11(2):499 – 521.

[Liu and Zhou, 2014] Liu, J. and Zhou, X. (2014). Extreme analysis of a random ordinary differential equation. *J. Appl. Probab.*

[Marcus and Shepp, 1970] Marcus, M. B. and Shepp, L. A. (1970). Continuity of Gaussian processes. *Transactions of the American Mathematical Society*, 151(2):377–391.

[Piterbarg, 1996] Piterbarg, V. I. (1996). *Asymptotic methods in the theory of Gaussian processes and fields*. American Mathematical Society, Providence, R.I.

[Sudakov and Tsirelson, 1974] Sudakov, V. and Tsirelson, B. (1974). Extremal properties of half spaces for spherically invariant measures. *Zap. Nauchn. Sem. LOMI*, 45:75–82.

[Sun, 1993] Sun, J. Y. (1993). Tail probabilities of the maxima of Gaussian random-fields. *Annals of Probability*, 21(1):34–71.

[Talagrand, 1996] Talagrand, M. (1996). Majorizing measures: The generic chaining. *Annals of Probability*, 24(3):1049–1103.

[Taylor and Adler, 2003] Taylor, J. and Adler, R. (2003). Euler characteristics for Gaussian fields on manifolds. *Annals of Probability*, 31(2):533–563.

[Taylor et al., 2005] Taylor, J., Takemura, A., and Adler, R. J. (2005). Validity of the expected Euler characteristic heuristic. *Annals of Probability*, 33(4):1362–1396.

[Tsirelson et al., 1976] Tsirelson, B., Ibragimov, I., and Sudakov, V. (1976). Norms of Gaussian sample functions. *Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975)*, 550:20–41.

[Xu et al., 2014] Xu, G., Lin, G., and Liu, J. (2014). Rare-event simulation for the stochastic korteweg–de vries equation. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):698–716.

Appendix A. Proof of supporting lemmas

Proof of Lemma 1. Note that the event $\{\xi - \mathbf{C}\xi^* \notin \mathcal{B}\} = \{|\xi - \mathbf{C}\xi^*|_{k,\beta} > \sigma^{\alpha-1-\varepsilon}\}$ implies the event $\{|\xi| > \sigma^{\alpha-1-\varepsilon} - |\xi^*|_{k,\beta}\}$. According to Proposition 3, $|\xi^*|_{k,\beta} = O(\sigma^{\alpha-1})$. Thus,

$$\{\xi - \mathbf{C}\xi^* \notin \mathcal{B}\} \subset \{|\xi|_{k,\beta} > \varepsilon_0 \sigma^{\alpha-1-\varepsilon}\}, \tag{58}$$

for a positive constant ε_0 and σ sufficiently small. Recall the definition

$$|\xi|_{k,\beta} = \sum_{l=1}^{k} \sup_{|\gamma|=l} \sup_{x\in\bar{U}} |D^\gamma \xi(x)| + \sup_{|\gamma|=k} [D^\gamma \xi]_\beta.$$

Consequently,

$$\begin{aligned}
\{\xi - \mathbf{C}\xi^* \notin \mathcal{B}\} \quad &\subset \quad \bigcup_{l=1}^{k} \{\sup_{|\gamma|=l} \sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}\} \bigcup \{\sup_{|\gamma|=l} [D^\gamma \xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}\} \\
&= \quad \bigcup_{l=1}^{k} \bigcup_{|\gamma|=l} \{\sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}\} \bigcup_{|\gamma|=l} \{[D^\gamma \xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}\}.
\end{aligned}$$

The equality in the above display is due to the fact that $\{\sup_{l=1}^m X_l \geq \eta\} = \cup_{l=1}^m \{X_l \geq \eta\}$ for any random variable X_l, $l = 1, ..., m$ and constant η. According to the above display, we arrive at a upper bound of probability.

$$\mathbb{P}(\xi - \mathbf{C}\xi^* \notin \mathcal{B}) \leq \sum_{l=1}^{k} \sum_{|\gamma|=l} \mathbb{P}(\sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) + \sum_{|\gamma|=l} \mathbb{P}([D^\gamma \xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}). \tag{59}$$

We establish upper bounds for $\mathbb{P}(\sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1})$ and $\mathbb{P}([D^\gamma \xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1})$ separately. We first analyze the term $\mathbb{P}(\sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1})$. We will need the following lemma, known as the Borell-TIS inequality, which was proved independently by [Borell, 1975] and [Cirel'son et al., 1976].

Lemma 8 (Borell-TIS inequality)**.** *Let $g(x)$ be a centered and almost surely bounded Gaussian random field. Then, $\mathbf{E}\sup_{x\in U} |g(x)| < \infty$. Furthermore, for any $t > \mathbf{E}\sup_{x\in U} |g(x)|$, we have*

$$\mathbb{P}\Big(\sup_{x\in U} |g(x)| - \mathbf{E}\sup_{x\in U} |g(x)| > t\Big) \leq 2\exp\left\{-\frac{t^2}{2\sup_{x\in U} \mathrm{Var}(g(x))}\right\}.$$

According to Lemma 8, we have that for all $|\gamma| \leq k$, $\mathbf{E}\sup_{x\in\bar{U}} |D^\gamma \xi(x)| < \infty$ and

$$\mathbb{P}(\sup_{x\in\bar{U}} |D^\gamma \xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) \leq 2\exp\left\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2 \sup_{x\in\bar{U}} C_{D^\gamma\xi}(x,x)}\right\}, \tag{60}$$

for σ sufficiently small such that $\sigma^{2\alpha-2-2\varepsilon} > 2\mathbf{E}\sup_{x\in\bar U}|D^\gamma\xi(x)|$, and $C_{D^\gamma\xi}$ is defined (9). According to Assumption A3, there exists a constant κ_0 such that for all $|\gamma|\leq k$,

$$\sup_{x\in\bar U} C_{D^\gamma\xi}(x,x) \leq \sup_{y\in\bar U}|C_{D^\gamma\xi}(\cdot,y)|_\beta < \kappa_0.$$

The above display together with (60) give

$$\mathbb{P}(\sup_{x\in\bar U}|D^\gamma\xi(x)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) \leq 2\exp\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2\kappa_0}\}$$

Combine this with (59), we have

$$\mathbb{P}(\xi - \mathbf{C}\xi^* \notin \mathcal{B}) \leq \kappa_0\exp\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2\kappa_0}\} + \sum_{|\gamma|=l}\mathbb{P}([D^\gamma\xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}), \tag{61}$$

for a possibly different κ_0 such that $\kappa_0 \geq 2\mathrm{Card}\{\gamma : |\gamma|\leq k\}$. We proceed to establishing upper bounds for $\mathbb{P}([D^\gamma\xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1})$, $|\gamma| = k$. Recall that

$$[D^\gamma\xi]_\beta = \sup_{x,y\in\bar U, x\neq y}\frac{|D^\gamma\xi(x) - D^\gamma\xi(y)|}{|x-y|^\beta}.$$

Motivated by this definition, we define another centered Gaussian random field double indexed by $x,y\in\bar U$

$$g(x,y) = \begin{cases} \frac{D^\gamma\xi(x)-D^\gamma\xi(y)}{|x-y|^\beta} & \text{for} \quad x\neq y \\ 0 & \text{for} \quad x = y \end{cases}. \tag{62}$$

According to Assumption A3 $\xi\in C^{k,\beta}(\bar U)$ almost surely. Thus, $g(\cdot,\cdot)$ is bounded almost surely. According to Lemma 8, we have that $\mathbf{E}\sup_{x,y\in\bar U,x\neq y}|g(x,y)| < \infty$, and

$$\mathbb{P}(\sup_{x,y\in\bar U}|g(x,y)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) \leq 2\exp\left\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2\sup_{x,y\in\bar U}\mathrm{Var}\, g(x,y)}\right\},$$

for σ sufficiently small such that $\sigma^{2\alpha-2-2\varepsilon} > 2\mathbf{E}\sup_{x,y\in\bar U}|g(x,y)|$. The variance of $g(x,y)$ in the above expression is bounded above as follows.

$$\begin{aligned}\mathrm{Var}\, g(x,y) &= |x-y|^{-2\beta}\{C_{D^\gamma\xi}(x,x) - C_{D^\gamma\xi}(x,y) + C_{D^\gamma\xi}(y,y) - C_{D^\gamma\xi}(x,y)\} \\ &\leq [C_{D^\gamma\xi}(x,\cdot)]_{2\beta} + [C_{D^\gamma\xi}(y,\cdot)]_{2\beta},\end{aligned}$$

which is bounded above by a constant κ_0 according to Assumption A3. Thus, we have

$$\mathbb{P}(\sup_{x,y\in\bar U}|g(x,y)| > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) \leq 2\exp\left\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2\kappa_0}\right\}.$$

Note that $[D^\gamma \xi]_\beta = \sup_{x,y\in\bar U} |g(x,y)|$. Therefore, the above display is equivalent to

$$\mathbb{P}([D^\gamma\xi]_\beta > \frac{\sigma^{\alpha-1-\varepsilon}}{k+1}) \leq 2\exp\left\{-\frac{\sigma^{2\alpha-2-2\varepsilon}}{8(k+1)^2\kappa_0}\right\}. \tag{63}$$

We conclude our proof by combining the above inequality with (61). □

Proof of Lemma 2. Because $\Lambda[w]$ is a fixed point of $T_w(\cdot)$, this lemma is a direct application of (53). □

Proof of Lemma 3. We define a function $h : [0,1] \to \mathbb{R}$,

$$h(s) = \mathcal{G}(w_2 + s(w_1 - w_2)) - \mathcal{G}(w_2) - s\int_U \mathcal{G}'[w_2](x)\{w_1(x) - w_2(x)\}dx.$$

Notice that $h(0) = 0$ and $h(1) = \mathcal{G}(w_1) - \mathcal{G}(w_2) - \int_U \mathcal{G}'[w_2](x)(w_1(x) - w_2(x))dx$. Apply mean value theorem to h, we have

$$\mathcal{G}(w_1) - \mathcal{G}(w_2) - \int_U \mathcal{G}'[w_2](x)(w_1(x) - w_2(x))dx = h(1) - h(0) = h'(\tilde s), \tag{64}$$

for some $\tilde s \in [0,1]$. According to the definition of Fréchet derivative, it is easy to check that

$$h'(s) = s\int_U \{\mathcal{G}'[w_1 + s(w_1 - w_2)](x) - \mathcal{G}'[w_2](x)\}(w_1(x) - w_2(x))dx.$$

Furthermore, we have

$$\begin{aligned}
&\left|s\int_U \{\mathcal{G}'[w_1 + s(w_1 - w_2)](x) - \mathcal{G}'[w_2](x)\}(w_1(x) - w_2(x))dx\right| \\
\leq\ & meas(U)|w_1 - w_2|_0 \times |\mathcal{G}'[w_1 + s(w_1 - w_2)](x) - \mathcal{G}'[w_2](x)|_0 \\
\leq\ & meas(U)|w_1 - w_2|_0 \times \kappa_G |w_1 - w_2|_{k,\beta} \\
\leq\ & meas(U)|w_1 - w_2|^2_{k,\beta}.
\end{aligned}$$

Here, $meas(U)$ is the Lebesgue measure of the set U, the second inequality is due to Assumption A1, and the third inequality is due to the fact that w, $|w|_0 \leq |w|_{k,\beta}$. Combine the above inequality and (64) we obtain the desired result. □

Proof of Lemma 4. We prove the lemma by induction. We first prove this lemma for the case where $k = 0$ and $\beta > 0$. We consider the conditional random field $\{\xi(x), x \in \bar U \mid Z_1 = z\}$. It can be shown that there exists a continuous Gaussian random field, denoted by $\{\chi(x), x \in \bar U\}$, who has the same distribution as $\{\xi(x), x \in \bar U \mid Z_1 = z\}$ and belongs to $C^\beta(\bar U)$ almost surely. The mean and covariance function of $\chi(x)$ satisfy

$$\begin{aligned}
\mu_\chi(x) &= \mathrm{Var}(Z_1)^{-1}\,\mathrm{Cov}(Z_1,\xi(x))z = \mathrm{Var}(Z_1)^{-1}\int_U \xi^*(y)C(x,y)dy,\\
C_\chi(x,y) &= C(x,y) - \mathrm{Var}(Z_1)^{-1}\,\mathrm{Cov}(\xi(x),Z_1)\,\mathrm{Cov}(\xi(y),Z_1)\\
&= C(x,y) - \mathrm{Var}(Z_1)^{-1}\int_U \xi^*(z)C(x,z)dz\int_U \xi^*(z)C(y,z)dz.
\end{aligned}$$

According to the expression (21) and $\sup_{y\in\bar U}|C(\cdot,y)|_{2\beta}\in<\infty$, we have that,

$$|\mu_\chi|_\beta = O(\sigma^{1-\alpha}z) \text{ and } \sup_{y\in\bar U}|C_\chi(,y)|_{2\beta}<\infty. \tag{65}$$

Let $\zeta(x)=\chi(x)-\mu_\chi(x)$ be a centered Gaussian random field. Then event $\{|\chi|_\beta^2>\frac{\Lambda(\xi^*)z}{\sigma}\}$ implies that $\{|\zeta|_\beta > (\frac{\Lambda(\xi^*)z}{\sigma})^{\frac12} - |\mu_\chi|_\beta\}$. Furthermore, according to (65) and Lemma 2 on page 11, we have

$$\{|\chi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}\}\subset\{|\zeta|_\beta > \varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z} - O(\sigma^{1-\alpha}z)\}.$$

Because $z\le\kappa_0\sigma^{2\alpha-2-\varepsilon}$, we have $\sigma^{\frac{\alpha}{2}-1}\sqrt{z}-O(\sigma^{1-\alpha}z)\ge\varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z}$ for a possibly different ε_0. Therefore,

$$\{|\chi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}\}\subset\{|\zeta|_\beta > \varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z}\}.$$

Consequently, we have

$$\mathbb{P}(|\chi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}) \le \mathbb{P}(|\zeta|_\beta > \varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z}). \tag{66}$$

According to the definition of the norm $|\zeta|_\beta=\sup_{x\in\bar U}|\zeta(x)|+[\zeta]_\beta$. Therefore, an upper bound for (66) is

$$\mathbb{P}(|\chi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}) \le \mathbb{P}(\sup_{x\in\bar U}|\zeta(x)|\ge\frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}) + \mathbb{P}([\zeta]_\beta\ge\frac{\varepsilon_2}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}). \tag{67}$$

We will present upper bounds for the first and second terms in the above display separately. We start with the first term. Because ζ is a centered and continuous Gaussian random field, with the aid of Lemma 8, we have that $\mathbf{E}\sup_{x\in\bar U}|\zeta(x)|<\infty$ and

$$\mathbb{P}(\sup_{x\in\bar U}|\zeta(x)|>\frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}) \le 2\exp\left\{-\frac{\varepsilon_0^2\sigma^{\alpha-2}z}{32\sup_{x\in\bar U}\mathrm{Cov}_\chi(x,x)}\right\},$$

for σ and z such that $\varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z}>2\mathbf{E}\sup_{x\in\bar U}|\zeta(x)|$. Because $z\ge\sigma^\varepsilon$, $\varepsilon_0\sigma^{\frac{\alpha}{2}-1}\sqrt{z}>2\mathbf{E}\sup_{x\in\bar U}|\zeta(x)|$ is satisfied for σ sufficiently small. Consequently, for σ sufficiently small, we have

$$\mathbb{P}(\sup_{x\in\bar U}|\zeta(x)|>\frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}) < e^{-\varepsilon_0\sigma^{\alpha-2}z} \tag{68}$$

for a sufficiently small and possibly different ε_0. We proceed to the second term on the right-hand-side of (67). Because $\zeta \in C^\beta(\bar{U})$ almost surely, we obtain an upper bound for $\mathbb{P}([\zeta]_\beta > \frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z})$ using similar arguments as those for (63) on page 25

$$\mathbb{P}([\zeta]_\beta > \frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}) < 2e^{-\varepsilon_0\sigma^{\alpha-2}z}, \tag{69}$$

for σ sufficiently small and a positive constant ε_0. Combine (67), (68) and (69), we have

$$\mathbb{P}(|\chi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}) < 2e^{-\varepsilon_0\sigma^{\alpha-2}z}. \tag{70}$$

Recall that χ has the same distribution as $\{\zeta(x) : x \in \bar{U}|Z_1 = z\}$, thus (70) implies

$$\mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) < 2e^{-\varepsilon_0\sigma^{\alpha-2}z}. \tag{71}$$

Using similar arguments, we have that for σ sufficiently small

$$\mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = -z) < 2e^{-\varepsilon_0\sigma^{\alpha-2}z}. \tag{72}$$

Combing the above inequality with (71), we have

$$\mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) + \mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) < 4e^{-\varepsilon_0\sigma^{\alpha-2}z} < e^{-\varepsilon_0'\sigma^{\alpha-2}z}$$

for $\varepsilon_0' < \varepsilon_0$ and σ sufficiently small. This completes our proof for the case where $k = 0$ and $\beta > 0$. For the case $k = 0$ and $\beta = 0$, $|\xi|_\beta = |\xi|_0$. With similar proof as those for (68), we have

$$\mathbb{P}(|\xi|_0^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) \leq \mathbb{P}(\sup_{x\in\bar{U}}|\zeta(x)| \geq \frac{\varepsilon_0}{2}\sigma^{\frac{\alpha}{2}-1}\sqrt{z}) < 2e^{-\varepsilon_0\sigma^{\alpha-2}z}. \tag{73}$$

We also have similar results conditional on $Z_1 = -z$. Therefore, for $\beta = 0$ we also have

$$\mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) + \mathbb{P}(|\xi|_\beta^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) < 4e^{-\varepsilon_0\sigma^{\alpha-2}z}.$$

This completes our proof for the case that $k = 0$. We now proceed to prove the lemma for $k \geq 1$. Assuming that for $k = m$,

$$\mathbb{P}(|\xi|_{k,\beta}^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) + \mathbb{P}(|\xi|_{k,\beta}^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) < e^{-\varepsilon_0\sigma^{\alpha-2}z} \tag{74}$$

for some positive constant ε_0 that is independent with σ and z but possibly depend on k. We will prove that the following inequality holds for σ sufficiently small and a positive constant ε_0,

$$\mathbb{P}(|\xi|_{m+1,\beta}^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) + \mathbb{P}(|\xi|_{m+1,\beta}^2 > \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z) < e^{-\varepsilon_0'\sigma^{\alpha-2}z}. \tag{75}$$

According to the definition of the norm $|\cdot|_{m+1,\beta}$, we know that for $\beta > 0$

$$|\xi|_{m+1,\beta} = |\xi|_m + \sup_{|\gamma|=m+1}\sup_{x\in\bar U}|D^\gamma\xi(x)| + \sup_{|\gamma|=m+1}[D^\gamma\xi]_\beta.$$

Therefore,

$$\begin{aligned}&\Big\{|\xi|^2_{m+1,\beta} \ge \frac{\Lambda(\xi^*)z}{\sigma}\Big\}\\ \subset\ & \{|\xi|^2_m \ge \frac{\Lambda(\xi^*)z}{2\sigma}\}\bigcup\Big(\bigcup_{|\gamma|=m+1}\bigcup_{|\gamma'|=m+1}\{(\sup_{x\in\bar U}|D^\gamma\xi(x)| + [D^{\gamma'}\xi]_\beta)^2 \ge \frac{\Lambda(\xi^*)z}{2\sigma}\}\Big).\end{aligned}$$

Consequently, we arrive at an upper bound

$$\begin{aligned}&\mathbb{P}\Big(|\xi|^2_{m+1,\beta} \ge \frac{\Lambda(\xi^*)z}{\sigma}|Z_1 = z\Big)\\ \le\ & \mathbb{P}(|\xi|^2_m \ge \frac{\Lambda(\xi^*)z}{2\sigma}|Z_1 = z)\\ &+\sum_{|\gamma|=m+1}\sum_{|\gamma'|=m+1}\mathbb{P}\Big((\sup_{x\in\bar U}|D^\gamma\xi(x)| + [D^{\gamma'}\xi]_\beta)^2 \ge \frac{\Lambda(\xi^*)z}{2\sigma}|Z_1 = z\Big),\end{aligned}\tag{76}$$

We present upper bounds for the first and second terms on the right-hand-side of the above display separately. For the first term, according to (74), we have

$$\mathbb{P}(|\xi|^2_m \ge \frac{\Lambda(\xi^*)z}{2\sigma}|Z_1 = z) \le e^{-\varepsilon_0\sigma^{\alpha-2}z}.\tag{77}$$

For the second term, notice that

$$\begin{aligned}&\{(\sup_{x\in\bar U}|D^\gamma\xi(x)| + [D^{\gamma'}\xi]_\beta)^2 \ge \frac{\Lambda(\xi^*)z}{2\sigma}\}\\ =\ & \{\sup_{x\in\bar U}|D^\gamma\xi(x)| + [D^{\gamma'}\xi]_\beta \ge \sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}\}\\ \subset\ & \{\sup_{x\in\bar U}|D^\gamma\xi(x)| \ge \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}\}\cup\{[D^{\gamma'}\xi]_\beta \ge \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}\}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{P}\Big((\sup_{x\in\bar U}|D^\gamma\xi(x)| + [D^{\gamma'}\xi]_\beta)^2 \ge \frac{\Lambda(\xi^*)z}{2\sigma}|Z_1 = z\Big)\\ \le\ & \mathbb{P}\Big(\sup_{x\in\bar U}|D^\gamma\xi(x)| \ge \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}|Z_1 = z\Big) + \mathbb{P}\Big([D^{\gamma'}\xi]_\beta \ge \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}|Z_1 = z\Big).\end{aligned}\tag{78}$$

Now we present upper bounds for the two terms on the right-hand-side of the above inequality for γ and γ' such that $|\gamma| = m+1$ and $|\gamma'| = m+1$. To do so, we consider a continuous

Gaussian random field χ_1 that belongs to $C^{\beta}(\bar{U})$ almost surely, and it has the same distribution as $\{D^{\gamma}\xi(x), x \in \bar{U} | Z_1 = z\}$.

Lemma 9. *Let $C_{\chi_1}(s,t) = \mathbf{E}\chi_1(s)\chi_1(t)$ and $\mu_{\chi_1}(t) = \mathbf{E}\chi_1(t)$, then we have*

$$|\mu_{\chi_1}|_{\beta} = O(\sigma^{1-\alpha} z) \text{ and } \sup_{y \in \bar{U}} |C_{\chi_1}(\cdot, y)| < \infty.$$

The above expressions are uniform in γ for $|\gamma| = m+1$.

Notice that the above lemma has the same form as (65), so with similar arguments as those for (68), we have

$$\mathbb{P}\Big(\sup_{x \in \bar{U}} |D^{\gamma}\chi_1(x)| \geq \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}\Big) \leq e^{-\varepsilon_0 \sigma^{\alpha-2} z}. \tag{79}$$

Also, similar as arguments before (69), we have

$$\mathbb{P}\Big([D^{\gamma'}\chi_1]_{\beta} \geq \frac{1}{2}\sqrt{\frac{\Lambda(\xi^*)z}{2\sigma}}\Big) \leq e^{-\varepsilon_0 \sigma^{\alpha-2} z}. \tag{80}$$

Combining (79) and (80) and (78), we have

$$\mathbb{P}\Big((\sup_{x \in \bar{U}} |D^{\gamma}\xi(x)| + [D^{\gamma'}\xi])^2 \geq \frac{\Lambda(\xi^*)z}{2\sigma} | Z_1 = z\Big) \leq 2e^{-\varepsilon_0 \sigma^{\alpha-2} z}.$$

Combining the above display with (76) and (77), we have

$$\mathbb{P}\Big(|\xi|^2_{m+1,\beta} \geq \frac{\Lambda(\xi^*)z}{\sigma} | Z_1 = z\Big) \leq e^{-\varepsilon_0 \sigma^{\alpha-2} z},$$

for σ sufficiently small and a possibly different constant ε_0. Similarly, conditional on $Z_1 = -z$, we have

$$\mathbb{P}\Big(|\xi|^2_{m+1,\beta} \geq \frac{\Lambda(\xi^*)z}{\sigma} | Z_1 = -z\Big) \leq e^{-\varepsilon_0 \sigma^{\alpha-2} z}.$$

Thus,

$$\mathbb{P}\Big(|\xi|^2_{m+1,\beta} \geq \frac{\Lambda(\xi^*)z}{\sigma} | Z_1 = z\Big) + \mathbb{P}\Big(|\xi|^2_{m+1,\beta} \geq \frac{\Lambda(\xi^*)z}{\sigma} | Z_1 = -z\Big) \leq 2e^{-\varepsilon_0 \sigma^{\alpha-2} z},$$

and we complete the proof for (75) for the case where $\beta > 0$. For $\beta = 0$, $|\xi|_{m+1} = |\xi|_m + \sup_{|\gamma|=m+1} \sup_{x \in \bar{U}} |D^{\gamma}\xi(x)|$. We obtain the proof for the case where $\beta = 0$ by ignoring all the $[D^{\gamma'}\xi]_{\beta}$ terms in the proof for the case where $\beta > 0$. This completes the induction. □

Proof of Lemma 5. According to Theorem 6.14 in [Gilbarg and Trudinger, 2015], we have that the PDE (27) has a unique solution in $C^{2,\beta}(\bar{U})$. Denote this solution by $u_{a,f}$, then according to Theorem

6.6 in [Gilbarg and Trudinger, 2015], we have the upper bound

$$|u_{a,f}|_{2,\beta} \le \kappa(\delta, M, d, U)(|u_{a,f}|_0 + |f|_0).$$

We conclude the proof with the following upper bound provided by Theorem 3.7 in [Gilbarg and Trudinger, 2015],

$$|u_{a,f}|_0 \le \kappa_0 |f|_0$$

for a constant κ_0 depending only on the domain U and $|a|_1$. □

Proof of Lemma 6. According to the definition of u_{a_1,f_1} and u_{a_2,f_2}, we have that

$$-\nabla \cdot (a_1(x)\nabla u_{a_1,f_1}(x)) = f_1 \text{ and } -\nabla \cdot (a_2(x)\nabla u_{a_2,f_2}(x)) = f_2.$$

Taking difference between the above two equalities, we have

$$-\nabla \cdot (a_1 \nabla u_{a_1,f_1}) + \nabla \cdot (a_2(x)\nabla u_{a_2,f_2}) = f_1(x) - f_2(x) \text{ for } x \in U.$$

Rearranging terms in the above expression, we have

$$-\nabla \cdot \Big(a_2(x)\nabla(u_{a_2,f_2}(x) - u_{a_1,f_1}(x))\Big) = f_2(x) - f_1(x) - \nabla \cdot \{(a_1(x) - a_2(x))\nabla u_{a_1,f_1}(x)\}.$$

Therefore, $\bar{u} = u_{a_2,f_2} - u_{a_1,f_1} \in C^{2,\beta}(\bar{U})$ is a solution to the elliptic PDE

$$\begin{cases} -\nabla \cdot (a_2(x)\nabla \bar{u}(x)) = \bar{f}(x) & \text{for } x \in U; \\ \bar{u}(x) = 0 & \text{for } x \in \partial U, \end{cases}$$

where $\bar{f}(x) = f_2(x) - f_1(x) - \nabla \cdot \{(a_1(x) - a_2(x))\nabla u_{a_1,f_1}(x)\}$. According to Lemma 5, we have

$$|u_{a_2,f_2} - u_{a_1,f_1}|_{2,\beta} \le \kappa(\delta, M, d, U)|\bar{f}|_\beta. \tag{81}$$

We further establish an upper bound for $|\bar{f}|_\beta$,

$$|\bar{f}|_\beta \le |f_2 - f_1|_\beta + |a_2 - a_1|_{1,\beta}|u_{a_1,f_1}|_{2,\beta}. \tag{82}$$

According to Lemma 5,

$$|u_{a_1,f_1}|_{2,\beta} \le \kappa(\delta, M, d, U)|f_1|_\beta.$$

Combining this with (81) and (82), we have

$$|u_{a_2,f_2} - u_{a_1,f_1}|_{2,\beta} \le \kappa(\delta, M, d, U)\{|f_2 - f_1|_\beta + \kappa(\delta, M, d, U)|a_2 - a_1|_{1,\beta}|f_1|_\beta\}.$$

We complete the proof by setting $\tilde{\kappa}(\delta, M, d, U) = \max(\kappa(\delta, M, d, U), \kappa(\delta, M, d, U)^2)$. □

Proof of Lemma 7. We take difference between $T_{w_1}(\lambda)$ and $T_{w_2}(\lambda)$,

$$T_{w_1}(\lambda) - T_{w_2}(\lambda) = -\mathcal{K}(\mathcal{G}'[\mathbf{0}])^{-1}\sigma^{-1}\{\mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w_1]) - \mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w_2])\}$$

Therefore,

$$|T_{w_1}(\lambda) - T_{w_2}(\lambda)| = O(\sigma^{-1}\{|\mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w_1]) - \mathcal{G}(\sigma\mathbf{C}\lambda\mathcal{G}'[\sigma\mathbf{C}w_2])|\}). \tag{83}$$

According to Lemma 3, we have

$$\begin{aligned} & \mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma\mathbf{C}w_1]) - \mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma\mathbf{C}w_2]) \\ = \ & \sigma\lambda\int_U \mathcal{G}'[\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma\mathbf{C}w_2]](x)\mathbf{C}\{\mathcal{G}'[\sigma\mathbf{C}w_1](x) - \mathcal{G}'[\sigma\mathbf{C}w_2](x)\}dx \\ & +O(\sigma^2\lambda^2|\mathcal{G}'[\sigma\mathbf{C}w_1] - \mathcal{G}'[\sigma\mathbf{C}w_2]|_{k,\beta}^2). \end{aligned}$$

According to (46) and Assumption A1, the above display can be further simplified as

$$\mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma w_1]) - \mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma w_2]) = O(\sigma\lambda|\mathcal{G}'[\sigma w_1] - \mathcal{G}'[\sigma w_2]|_{k,\beta}),$$

which is further simplified as

$$\mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma w_1]) - \mathcal{G}(\sigma\lambda\mathbf{C}\mathcal{G}'[\sigma w_2]) = O(\sigma\lambda\sigma|w_1 - w_2|_{k,\beta}).$$

The above expression and (83) give

$$|T_{w_1}(\lambda) - T_{w_2}(\lambda)| = O(\sigma\lambda|w_1 - w_2|_{k,\beta}) = O(\sigma^{\alpha})|w_1 - w_2|_{k,\beta}.$$

The last inequality in the above expression is due to $\lambda = O(\sigma^{\alpha-1})$. □

Proof of Lemma 9. We need the next lemma for the current proof.

Lemma 10. *We define the covariance function*

$$C_{D^\gamma\xi,\xi}(x,y) = \mathrm{Cov}(D^\gamma\xi(x), \xi(y)).$$

Then $\sup_{y\in\bar{U}} |C_{D^\gamma\xi,\xi}(\cdot,y)|_{2\beta} < \infty$ *for all* $|\gamma| \le k$ *under Assumption A3.*

Now we compute the mean and covariance of χ_1.

$$\mu_{\chi_1}(x) = \mathbb{E}[D^\gamma\xi(x)|Z_1 = z] = \mathrm{Var}(Z_1)^{-1}\,\mathrm{Cov}(D^\gamma\xi(x), Z_1)z = \mathrm{Var}(Z_1)^{-1}\int_U C_{D^\gamma\xi,\xi}(x,y)\xi^*(y)dtz,$$

and

$$\begin{aligned} C_{\chi_1}(x,y) &= C_{D^\gamma \xi}(x,y) - \mathrm{Var}(Z_1)^{-1}\,\mathrm{Cov}(D^\gamma \xi(x), Z_1)\,\mathrm{Cov}(D^\gamma \xi(y), Z_1) \\ &= C_{D^\gamma \xi}(x,y) - \frac{\int_U C_{D^\gamma \xi,\xi}(x,r)\xi^*(y)dr \int_U C_{D^\gamma \xi,\xi}(y,r)\xi^*(y)dr}{\mathrm{Var}(Z_1)}. \end{aligned}$$

Recall that $\mathrm{Var}(Z_1) \geq \varepsilon_0 \sigma^{2\alpha-2}$ for some positive constant ε, and $|\xi^*|_{k,\beta} = O(\sigma^{\alpha-1})$. With the aid of Lemma 10, we simplify the mean and covariance of χ_1.

$$|\mu_{\chi_1}|_\beta = O(\sigma^{2-2\alpha}|\xi^*|_0 \sup_y |C_{D^\gamma \xi,\xi}(\cdot,y)|_{2\beta} z) = O(\sigma^{1-\alpha} z),$$

and

$$\sup_{y\in\bar U} |C_{\chi_1}(\cdot,y)|_{2\beta} = O(\sup_{y\in\bar U} |C_{D^\gamma\xi}(\cdot,y)|_{2\beta} + \sigma^{2-2\alpha}|\xi^*|_0^2 \sup_{y\in\bar U} |C_{D^\gamma \xi,\xi}(\cdot,y)|_{2\beta}^2) = O(1).$$

□

Proof of Lemma 10. We will use induction to prove that for all $l = 0, 1, ..., k$, $|\gamma| = l$,

$$\sup_{y\in\bar U} |C_{D^\gamma \xi,\xi}(\cdot,y)|_{2\beta} < \infty. \tag{84}$$

To start with, for $l = 0$ and $|\gamma| = l$, (84) holds because of Assumption A3 and

$$C_{D^\gamma \xi,\xi}(s,t) = C(s,t).$$

Suppose that for all $|\gamma'| = l$,

$$\sup_{y\in\bar U} |C_{D^{\gamma'} \xi,\xi}(\cdot,y)|_{k-l,2\beta} < \infty. \tag{85}$$

For $|\gamma| = l+1$, we want to show that

$$\sup_{y\in\bar U} |C_{D^\gamma \xi,\xi}(\cdot,y)|_{k-l-1,2\beta} < \infty. \tag{86}$$

Without loss of generality, we assume that $\gamma = (\gamma_1, ..., \gamma_d)$ and $\gamma_1 \geq 1$. Let $e_1 = (1, ..., 0)$ be a d-dimensional basis vector, and $\gamma' = \gamma - e_1$, then $|\gamma'| = l$. We compute $C_{D^\gamma \xi,\xi}$.

$$\begin{aligned} C_{D^\gamma \xi,\xi}(x,y) &= \lim_{\varepsilon_1\to 0} \mathrm{Cov}(\frac{D^{\gamma'}\xi(x+\varepsilon_1 e_1) - D^{\gamma'}\xi(x)}{\varepsilon_1}, \xi(y)) \\ &= \lim_{\varepsilon_1\to 0} \varepsilon_1^{-1}\{C_{D^{\gamma'}\xi,\xi}(x+\varepsilon_1 e_1, y) - C_{D^{\gamma'}\xi,\xi}(x,y)\} \\ &= \frac{\partial}{\partial x_1} C_{D^{\gamma'}\xi}(x,y). \end{aligned}$$

Consequently,

$$|C_{D^\gamma\xi,\xi}(\cdot,y)|_{k-l-1,2\beta} = |\frac{\partial}{\partial x_1} C_{D^{\gamma'}\xi}(\cdot,y)|_{k-l-1,2\beta} \leq |C_{D^{\gamma'}\xi}(\cdot,y)|_{k-l,2\beta}.$$

Thus,

$$\sup_{y\in\bar{U}} |C_{D^\gamma\xi,\xi}(\cdot,y)|_{k-l-1,2\beta} \leq \sup_{y\in\bar{U}} |C_{D^{\gamma'}\xi}(\cdot,y)|_{k-l,2\beta} < \infty.$$

The second inequality of the above display is due to (85). The lemma is proved by induction. □

School of Statistics, University of Minnesota, Minneapolis, MN 55455

E-mail address: `lixx1766@umn.edu`

Department of Statistics, Columbia University, New York, NY 10027

E-mail address: `jcliu@stat.columbia.edu`

Department of Mathematics, Department of Physics, and Department of Chemistry, Duke University, Durham, NC 27708 USA

E-mail address: `jianfeng@math.duke.edu`

Department of Mathematics, City University of Hong Kong, Tat Chee Ave, Kowloon, Hong Kong SAR

E-mail address: `xizhou@cityu.edu.hk`