The Annals of Probability
2014, Vol. 42, No. 3, 1020–1053
DOI: 10.1214/12-AOP794
© Institute of Mathematical Statistics, 2014

ON THE EXISTENCE OF PATHS BETWEEN POINTS IN HIGH LEVEL EXCURSION SETS OF GAUSSIAN RANDOM FIELDS

BY ROBERT J. ADLER[1,4], ELINA MOLDAVSKAYA[3,5] AND GENNADY SAMORODNITSKY[1,2]

Technion, Technion and Cornell University

The structure of Gaussian random fields over high levels is a well researched and well understood area, particularly if the field is smooth. However, the question as to whether or not two or more points which lie in an excursion set belong to the same connected component has constantly eluded analysis. We study this problem from the point of view of large deviations, finding the asymptotic probabilities that two such points are connected by a path laying within the excursion set, and so belong to the same component. In addition, we obtain a characterization and descriptions of the most likely paths, given that one exists.

1. Introduction.

Let $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ be a real-valued sample continuous Gaussian random field. Given a level u, the excursion set of $\mathbf{X}$ above the level u is the random set

$$A_u = \{\mathbf{t} \in \mathbb{R}^d : X(\mathbf{t}) > u\}. \tag{1.1}$$

Understanding the structure of the excursion sets of random fields is a mathematical problem with many applications, and it has generated significant interest, with several recent books on the subject (e.g., [1] and [2]) and with considerable emphasis on the topology of these sets. One very natural question in this setting which has until now eluded solution but which we study in this paper is the following: given that two points in $\mathbb{R}^d$ belong to the excursion set, what is the probability that they belong to the same path-connected component of the excursion set? Specifically, let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, $\mathbf{a} \neq \mathbf{b}$. Recall that a path in $\mathbb{R}^d$ connecting $\mathbf{a}$ and $\mathbf{b}$ is a continuous map $\xi : [0, 1] \to \mathbb{R}^d$ with $\xi(0) = \mathbf{a}$, $\xi(1) = \mathbf{b}$. We denote the collection of all such paths by $\mathcal{P}(\mathbf{a}, \mathbf{b})$ and are interested in the conditional probability

$$P\bigl(\exists \xi \in \mathcal{P}(\mathbf{a}, \mathbf{b}) : X\bigl(\xi(v)\bigr) > u, \text{ for all } 0 \le v \le 1 \,|\, X(\mathbf{a}) > u, X(\mathbf{b}) > u\bigr).$$

Received March 2012; revised August 2012.

[1] Supported in part by US-Israel Binational Science Foundation, 2008262.
[2] Supported in part by ARO Grant W911NF-07-1-0078, NSF Grant DMS-1005903 and NSA Grant H98230-11-1-0154 at Cornell University.
[3] Supported in part by Israel Science Foundation, 853/10.
[4] Supported in part by AFOSR FA8655-11-1-3039.
[5] Supported in part by Office for Absorption of New Scientists and EC-FP7-230804 Grant.
MSC2010 subject classifications. Primary 60G15, 60F10; secondary 60G60, 60G70, 60G17.
Key words and phrases. Gaussian process, excursion set, large deviations, exceedence probabilities, connected component, optimal path, energy of measures.

It is straightforward to check that we are considering measurable collections of outcomes, so this probability is well defined.

Of course, the conditional probability above is a ratio of two probabilities, the denominator being no more than a bivariate Gaussian probability, which is well understood. Therefore, we will concentrate on the unconditional probability

$$\Psi_{\mathbf{a},\mathbf{b}}(u) \triangleq P\big(\exists \xi \in \mathcal{P}(\mathbf{a},\mathbf{b}) : X\big(\xi(v)\big) > u, \text{ for all } 0 \le v \le 1\big). \tag{1.2}$$

If the random field is stationary, we may, without loss of generality, assume that $\mathbf{b} = \mathbf{0}$, in which case we will use the simpler notation $\Psi_{\mathbf{a}}$ in (1.2).

When the domain of a random field is restricted to a (compact) subset $T \subset \mathbb{R}^d$, the points $\mathbf{a}$ and $\mathbf{b}$ will be assumed to be in T, and the entire path in (1.2) will be required to lie in T as well (the implicit assumption being that T contains *some* path between $\mathbf{a}$ and $\mathbf{b}$). Nevertheless, we will use the same notation and also write

$$\Psi_{\mathbf{a},\mathbf{b}}(u) = P\big(\exists \xi \in \mathcal{P}(\mathbf{a},\mathbf{b}) : \xi(v) \in T \text{ and } X\big(\xi(v)\big) > u, \text{ for all } 0 \le v \le 1\big).$$

Which of the two interpretations of $\Psi_{\mathbf{a},\mathbf{b}}$ is intended at any point will be clear from the context.

We will study the logarithmic behavior of the probability $\Psi_{\mathbf{a},\mathbf{b}}(u)$ for high levels u, that is, as $u \to \infty$. We start with a large deviations approach, which, as usual, will not only describe the probability but also give us insight into the highest probability configurations. This makes up Sections 3 and 4, which follow a brief technical Section 2 collecting some results on the reproducing kernel Hilbert space of a Gaussian process. Throughout we will treat the general and the stationary cases in parallel, but separately, since the stationary case is somewhat more transparent and more readily provides illustrative and illuminating special cases. In particular, we will look at a number of one-dimensional examples in Sections 5–7, where we can compute quite a lot. Even in this case the results are new and rather unexpected. We look at the multidimensional case in Section 8. While this section also contains some interesting and surprising examples, it turns out that typical examples involve nonconvex optimization problems that we do not, at this stage, know how to solve in general.

2. Some technical preliminaries. In this section we introduce much of the notation we will use in the rest of the paper and recall certain important notions, concentrating in particular on the reproducing kernel Hilbert (RKHS) space of a Gaussian process.

Our main reference for the RKHS is van der Vaart and van Zanten [9], and we use it selectively so as to prepare the background for using the large deviations theory of Deuschel and Stroock [3]. An alternative route would be to have followed the new notes by Lifshits [6].

We consider a real-valued centered continuous Gaussian random field $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$. When needed (particularly, in the nonstationary case) we may restrict the domain of the random field to a compact subset T of $\mathbb{R}^d$. We denote the covariance function of $\mathbf{X}$ by $R_{\mathbf{X}}(\mathbf{s},\mathbf{t}) = \operatorname{cov}(X(\mathbf{s}), X(\mathbf{t}))$.

As is customary, when the random field is stationary, we will use the single variable notation $R_{\mathbf{X}}(\mathbf{t}) = R_{\mathbf{X}}(\mathbf{0}, \mathbf{t})$ for the covariance function. In this case we denote the spectral measure of $\mathbf{X}$ by $F_{\mathbf{X}}$, this being the symmetric, finite, Borel probability measure on $\mathbb{R}^d$ satisfying

$$R_{\mathbf{X}}(\mathbf{t}) = \int_{\mathbb{R}^d} e^{i(\mathbf{t},\mathbf{x})} F_{\mathbf{X}}(d\mathbf{x}), \qquad \mathbf{t} \in \mathbb{R}^d. \tag{2.1}$$

If $\mathbf{X}$ is stationary, then this and local boundedness imply that

$$\lim_{\|\mathbf{t}\|\to\infty} \frac{X(\mathbf{t})}{\|\mathbf{t}\|} = 0$$

with probability 1, so that almost all the sample paths of $\mathbf{X}$ belong to the space

$$C_0 = \Bigl\{\boldsymbol{\omega} = \bigl(\omega(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d\bigr), \text{ continuous, such that } \lim_{\|\mathbf{t}\|\to\infty} \omega(\mathbf{t})/\|\mathbf{t}\| = 0\Bigr\}.$$

Equipped with the norm

$$\|\boldsymbol{\omega}\|_{C_0} = \sup_{\mathbf{t}\in\mathbb{R}^d} \frac{|\omega(\mathbf{t})|}{1 + \|\mathbf{t}\|}, \tag{2.2}$$

C_0 becomes a separable Banach space, with dual space

$$C_0^* = \Bigl\{\text{finite signed Borel measures } \mu \text{ on } \mathbb{R}^d \text{ with } \int_{\mathbb{R}^d} \|\mathbf{t}\| \|\mu\|(d\mathbf{t}) < \infty\Bigr\}.$$

We view the stationary random field $\mathbf{X}$ as a Gaussian random element of C_0, generating a Gaussian probability measure $\mu_{\mathbf{X}}$ on that space.

In the absence of stationarity, we will usually consider a continuous Gaussian random field $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in T)$, for a compact set $T \subset \mathbb{R}^d$. In that case we view the random field $\mathbf{X}$ as a Gaussian random element in the space $C(T)$ of continuous functions on T, equipped with the supremum norm, thus generating a Gaussian probability measure $\mu_{\mathbf{X}}$ on $C(T)$.

The reproducing kernel Hilbert space (henceforth RKHS) $\mathcal{H}$ of the Gaussian measure $\mu_{\mathbf{X}}$ (or of the random field $\mathbf{X}$) is a subspace of C_0 or $C(T)$, depending on the parameter space of $\mathbf{X}$, obtained as follows. In the general case we identify $\mathcal{H}$ with the closure $\mathcal{L}$ in the mean square norm of the space of finite linear combinations $\sum_{j=1}^k a_j X(\mathbf{t}_j)$ of the values of the process, $a_j \in \mathbb{R}, \mathbf{t}_j \in \mathbb{R}^d$ (or T) for $j = 1, \dots, k$, $k = 1, 2, \dots$ via the injection $\mathcal{L} \to C(T)$ given by

$$H \to w_H = \bigl(E\bigl(X(\mathbf{t})H\bigr), \mathbf{t} \in T\bigr). \tag{2.3}$$

When $\mathbf{X}$ is stationary, the RKHS $\mathcal{H}$ can also be identified with the subspace of functions, with even real parts and odd imaginary parts, of the L^2 space of the spectral measure $F_{\mathbf{X}}$ in (2.1), via the injection $L^2(F_{\mathbf{X}}) \to C_0$ given by

$$h \to S(h) = \left(\int_{\mathbb{R}^d} e^{i(\mathbf{t},\mathbf{x})} \bar{h}(\mathbf{x}) F(d\mathbf{x}), \mathbf{t} \in \mathbb{R}^d\right). \tag{2.4}$$

We denote by $(\cdot,\cdot)_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{H}}$ the inner product and the norm in the RKHS $\mathcal{H}$. Since both injections described above are isometric, we have the important equalities:

$$E(H^2) = \|w_H\|_{\mathcal{H}}^2. \tag{2.5}$$

In the stationary case, these can be written somewhat more informatively as

$$\|h\|_{L^2(F_{\mathbf{X}})}^2 = \int_{\mathbb{R}^d} \|h(x)\|^2 F_{\mathbf{X}}(d\mathbf{x}) = \|S(h)\|_{\mathcal{H}}^2. \tag{2.6}$$

We shall use these equalities heavily in what follows.

Note that for every $\mathbf{s} \in \mathbb{R}^d$, the fixed $\mathbf{s}$ covariance function $R_{\mathbf{s}} = R(\cdot, \mathbf{s})$ is in $\mathcal{H}$, and for every $w_H \in \mathcal{H}$, and $\mathbf{t} \in \mathbb{R}^d$, $w_H(\mathbf{t}) = (w_H, R_{\mathbf{t}})_{\mathcal{H}}$, meaning that the coordinate projections are continuous operations on the RKHS. This is also the *reproducing property* of the RKHS. Note also that the quadruple $(C(T), \mathcal{H}, w, \mu_{\mathbf{X}})$ in general, or $(C_0, \mathcal{H}, S, \mu_{\mathbf{X}})$ in the stationary case, is a Wiener quadruple in the sense of Section 3.4 in [3].

In the sequel we will use the notation $M^+(E)$ [resp., $M_1^+(E)$] for the collection of all Borel finite (resp., probability) measures on a topological space E.

3. The basic large deviations result. We start with a large deviation result for the probability $\Psi_{\mathbf{a},\mathbf{b}}$ there exists a path between $\mathbf{a}$ and $\mathbf{b}$ wholly within a connected component of an excursion set.

THEOREM 3.1. (i) *Let $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in T)$ be a continuous Gaussian random field on a compact set $T \subset \mathbb{R}^d$. Then*

$$\lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{\mathbf{a},\mathbf{b}}(u) = -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}), \tag{3.1}$$

where

$$\begin{aligned}\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}) \triangleq \inf\{&EH^2 : H \in \mathcal{L},\ \text{and, for some } \xi \in \mathcal{P}(\mathbf{a},\mathbf{b}),\\ &\xi(v) \in T \text{ and } w_H(\xi(v)) > 1, 0 \le v \le 1\}.\end{aligned} \tag{3.2}$$

(ii) *Let $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ be a continuous stationary Gaussian random field, with covariance function satisfying*

$$\limsup_{\|\mathbf{t}\|\to\infty} R_{\mathbf{X}}(\mathbf{t}) \le 0. \tag{3.3}$$

Then

$$\lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{\mathbf{a}}(u) = -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a}), \tag{3.4}$$

where

$$\begin{aligned}\mathcal{C}_{\mathbf{X}}(\mathbf{a}) \triangleq \inf\biggl\{&\int_{\mathbb{R}^d} \|h(\mathbf{x})\|^2 F_{\mathbf{X}}(d\mathbf{x}) : \text{for some } \xi \in \mathcal{P}(\mathbf{0},\mathbf{a})\\ &\int_{\mathbb{R}^d} e^{i(\xi(v),\mathbf{x})}\bar{h}(\mathbf{x}) F_{\mathbf{X}}(d\mathbf{x}) > 1, 0 \le v \le 1\biggr\}.\end{aligned}$$

PROOF. We start with putting our problem into the large deviation setup for Gaussian measures of [3]. We will use the language of part (i) of the theorem, but the setup for part (ii) is completely parallel. Observe that for $u > 0$

$$\Psi_{\mathbf{a},\mathbf{b}}(u) = P\big(u^{-1}\mathbf{X} \in A\big),$$

where A is the open subset of $C(T)$ given by

$$A \equiv A_{\mathbf{a},\mathbf{b}} \stackrel{\Delta}{=} \big\{\boldsymbol{\omega} \in C(T) : \exists \xi \in \mathcal{P}(\mathbf{a},\mathbf{b}) \text{ such that } \omega\big(\xi(v)\big) > 1, 0 \le v \le 1\big\}.$$

Therefore, by Theorem 3.4.5 in [3], we conclude that

$$\begin{aligned} -\inf_{\boldsymbol{\omega}\in A} I(\boldsymbol{\omega}) &\le \liminf_{u\to\infty} \frac{1}{u^2}\log \Psi_{\mathbf{a},\mathbf{b}}(u) \le \limsup_{u\to\infty}\frac{1}{u^2}\log\Psi_{\mathbf{a},\mathbf{b}}(u) \\ &\le -\inf_{\boldsymbol{\omega}\in\bar{A}} I(\boldsymbol{\omega}) \end{aligned} \tag{3.5}$$

for the rate function I which, by Theorem 3.4.12 of [3], can be written as

$$I(\boldsymbol{\omega}) = \begin{cases} \frac{1}{2}\|\boldsymbol{\omega}\|^2_{\mathcal{H}}, & \text{if } \boldsymbol{\omega} \in \mathcal{H}, \\ \infty, & \text{if } \boldsymbol{\omega} \notin \mathcal{H}, \end{cases} \tag{3.6}$$

for $\boldsymbol{\omega} \in C(T)$. Then (3.5) already proves the lower limit statement

$$\liminf_{u\to\infty}\frac{1}{u^2}\log\Psi_{\mathbf{a},\mathbf{b}}(u) \ge -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}),$$

valid for both parts of the theorem. Therefore, it remains to prove the matching upper limit. Here the argument is more involved in part (ii) of the theorem, since noncompactness of the domain of the field requires us to rule out the possibility of increasingly long ranging paths. We present the argument in this case. The proof for part (i) is similar, and easier (since we do not have to worry about paths which "escape to infinity" as in the following).

As is common with large deviation arguments, although we know that $A = A^\circ \ne \bar{A}$, this is not per se important. All that we need show is that the ω in the set difference $\bar{A} \setminus A$ do not contribute to the infimum on the far right of (3.5).

We start by checking that

$$\bar{A} \subseteq \Big(\bigcap_{0<\delta<1}(1-\delta)A\Big) \cup \Big(\bigcap_{0<\delta<1}(1-\delta)A_1\Big) \tag{3.7}$$

(in the sense of the usual multiplication of a set of functions by a real number), where $A_1 \subset C_0$ is given by

$$\begin{aligned} A_1 = \big\{\boldsymbol{\omega} \in C_0 : &\text{ for every } r > 0 \text{ there is } \mathbf{t} \in \mathbb{R}^d \text{ with } \|\mathbf{t}\| \ge r \\ &\text{and a path } \xi \in \mathcal{P}(\mathbf{0},\mathbf{t}) \text{ such that } \boldsymbol{\omega}\big(\xi(v)\big) > 1, 0 \le v \le 1\big\}. \end{aligned}$$

To see this, let $\boldsymbol{\omega} \in \bar{A}$, so that there is a sequence $\boldsymbol{\omega}_n \in A, n = 1, 2, \ldots,$ with $\boldsymbol{\omega}_n \to \boldsymbol{\omega}$ in C_0. Suppose first that there is $r > 0$ such that for a subsequence

$n_k \uparrow \infty$, for each $k = 1, 2, \dots$ there is a path $\xi_k \in \mathcal{P}(\mathbf{0}, \mathbf{a})$ satisfying $\|\xi_k(v)\| \le r$ and $\boldsymbol{\omega}_{n_k}(\xi_k(v)) > 1, 0 \le v \le 1$. Given $0 < \delta < 1$, choose k so large that

$$\|\boldsymbol{\omega}_{n_k} - \boldsymbol{\omega}\|_{C_0} \le \delta/(1+r).$$

Then for every $\mathbf{t} \in \mathbb{R}^d$ with $\|\mathbf{t}\| \le r$ we have $|\boldsymbol{\omega}_{n_k}(\mathbf{t}) - \boldsymbol{\omega}(\mathbf{t})| \le \delta$, so that $\boldsymbol{\omega}(\xi_k(v)) > 1 - \delta$ for $0 \le v \le 1$, and $\boldsymbol{\omega} \in (1-\delta)A$.

Alternatively, suppose that such an $r > 0$ does not exist. Then for every $r > 0$, for all but finitely many n, there is a path $\xi_n \in \mathcal{P}(\mathbf{0}, \mathbf{a})$, going through a point $\mathbf{t}_n$ with $\|\mathbf{t}_n\| = r$, lying within the ball of radius r centered at the origin prior to hitting the point $\mathbf{t}_n$, and such that $\boldsymbol{\omega}_n(\xi_n(v)) > 1, 0 \le v \le 1$. Given $r > 0$ and $0 < \delta < 1$, choose n outside of the above exceptional finite set, and so large that

$$\|\boldsymbol{\omega}_n - \boldsymbol{\omega}\|_{C_0} \le \delta/(1+r).$$

As before, we conclude that there is a path connecting $\mathbf{0}$ and $\mathbf{t}_n$ such that the function $\boldsymbol{\omega}$ takes values above $1-\delta$ along this path. Therefore, $\boldsymbol{\omega} \in (1-\delta)A_1$, and so we have shown (3.7).

Now note that since

$$\inf_{\boldsymbol{\omega} \in (1-\delta)A} I(\boldsymbol{\omega}) = (1-\delta)^2 \inf_{\boldsymbol{\omega} \in A} I(\boldsymbol{\omega})$$

for any $0 < \delta < 1$, the upper limit part in (3.4), and so the result, will follow from (3.7) once we check that $I(\boldsymbol{\omega}) = \infty$ for any $\omega \in A_1$, which we establish by showing that $A_1 \cap \mathcal{H} = \varnothing$.

Suppose that, to the contrary, there is a $\boldsymbol{\omega} = S(h) \in A_1$ for some $h \in \mathcal{H}$. Fix an arbitrary $\varepsilon > 0$. Assumption (3.3) guarantees the existence of a $r_\varepsilon > 0$ such that $R_{\mathbf{X}}(\mathbf{t}) \le \varepsilon$ if $\|\mathbf{t}\| \ge r_\varepsilon$. By the definition of A_1, for every $n = 1, 2, \dots$ there is $\mathbf{t}_n$ with $\|\mathbf{t}_n\| = nr_\varepsilon$ and a path ξ connecting $\mathbf{0}$ and $\mathbf{t}_n$ such that $\boldsymbol{\omega}(\xi(v)) > 1, 0 \le v \le 1$. We can choose $0 < v_1 < \cdots < v_n \le 1$ such that $\|\xi(v_j)\| = jr_\varepsilon$ for $j = 1, \dots, n$. Then

$$\begin{aligned} 1 < \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{\omega}\big(\xi(v_j)\big) &= \int_{\mathbb{R}^d} \left(\frac{1}{n}\sum_{j=1}^{n} e^{i(\xi(v_j),\mathbf{x})}\right) \bar{h}(\mathbf{x}) F_{\mathbf{X}}(d\mathbf{x}) \\ &\le \left\| \frac{1}{n}\sum_{j=1}^{n} e^{i(\xi(v_j),\cdot)} \right\|_{L^2(F_{\mathbf{X}})} \|h\|_{L^2(F_{\mathbf{X}})}. \end{aligned}$$

However,

$$\begin{aligned} \left\| \frac{1}{n}\sum_{j=1}^{n} e^{i(\xi(v_j),\cdot)} \right\|^2_{L^2(F_{\mathbf{X}})} &= \frac{1}{n^2}\left(nR_{\mathbf{X}}(0) + 2\sum_{j_1=1}^{n-1}\sum_{j_2=j_1+1}^{n} R_{\mathbf{X}}\big(\xi(v_{j_1}) - \xi(v_{j_2})\big)\right) \\ &\le \frac{1}{n}R_{\mathbf{X}}(0) + \varepsilon, \end{aligned}$$

so that

$$\|h\|^2_{L^2(F_{\mathbf{X}})} > \frac{1}{\frac{1}{n}R_{\mathbf{X}}(0) + \varepsilon}.$$

Sending first $n \to \infty$ and then $\varepsilon \to 0$, we obtain $\|h\|_{L^2(F_{\mathbf{X}})} = \infty$, which is impossible.

This contradiction proves the rightmost inequality in (3.4) and so we are done. □

Theorem 3.1 describes the logarithmic asymptotic of the path existence probability $\Psi_{\mathbf{a},\mathbf{b}}$ in terms of a solution to an optimization problem in the Hilbert space. The next result contains the dual version of this optimization problem and relates $\Psi_{\mathbf{a},\mathbf{b}}$ to the problem of finding a path of minimal capacity between **a** and **b**.

THEOREM 3.2. (i) *Let* $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in T)$ *be a continuous Gaussian random field on a compact set* $T \subset \mathbb{R}^d$. *Then*

$$\begin{aligned}\lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{\mathbf{a},\mathbf{b}}(u) &= -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}) \\ &= -\frac{1}{2}\biggl[\sup_{\xi\in\mathcal{P}(\mathbf{a},\mathbf{b})} \min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(\xi(u),\xi(v)\bigr)\mu(du)\mu(dv)\biggr]^{-1}.\end{aligned} \tag{3.8}$$

(ii) *Let* $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d)$ *be a continuous stationary Gaussian random field, with covariance function satisfying* (3.3). *Then*

$$\begin{aligned}\lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{\mathbf{a}}(u) &= -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a}) \\ &= -\frac{1}{2}\biggl[\sup_{\xi\in\mathcal{P}(\mathbf{0},\mathbf{a})} \min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(\xi(u)-\xi(v)\bigr)\mu(du)\mu(dv)\biggr]^{-1}.\end{aligned} \tag{3.9}$$

Note that the space $M_1^+([0,1])$ is weakly compact, and the covariance function $R_{\mathbf{X}}$ is continuous. Therefore, for a fixed path ξ, the function

$$\mu \to \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(\xi(u),\xi(v)\bigr)\mu(du)\mu(dv)$$

is weakly continuous on compacts. Hence, it achieves its infimum, and it is legitimate to write “min” in (3.8) and in (3.9).

PROOF OF THEOREM 3.2. The proofs of the two parts are only notationally different, so we will suffice with a proof for part (i) only. We use the Lagrange duality approach of Section 8.6 in [7]. Writing

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}) = \inf_{\xi\in\mathcal{P}(\mathbf{a},\mathbf{b})} \mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi),$$

where, for $\xi \in \mathcal{P}(\mathbf{a},\mathbf{b})$,

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi) \stackrel{\Delta}{=} \inf\{EH^2 : H \in \mathcal{L} \text{ and } w_H(\xi(v)) > 1, 0 \le v \le 1\}, \tag{3.10}$$

we see that it is enough to prove that for every $\xi \in \mathcal{P}(\mathbf{a},\mathbf{b})$,

$$\begin{aligned} &\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi) \\ &\quad = \left[\min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}(\xi(u),\xi(v))\mu(du)\mu(dv)\right]^{-1}. \end{aligned} \tag{3.11}$$

To this end, let $\mathtt{Z} = C([0,1]$. Then $\mathtt{P} \stackrel{\Delta}{=} \{z \in \mathtt{Z} : z(v) \ge 0, 0 \le v \le 1\}$ is a closed convex cone in $\mathtt{Z}$. Its dual cone $\mathtt{P}^{\oplus} \subset \mathtt{Z}^*$ [defined as the collection of $z^* \in \mathtt{Z}^*$ such that $z^*(z) \ge 0$ for all $z \in \mathtt{P}$] can be naturally identified with $M^+([0,1])$. Fix $\xi \in \mathcal{P}(\mathbf{a},\mathbf{b})$, and define $G : \mathcal{L} \to \mathtt{Z}$ by

$$G(H) = G_\xi(H) \stackrel{\Delta}{=} (1 - w_H(\xi(v)), 0 \le v \le 1).$$

Then G is, clearly, a convex mapping. We can also write

$$(\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi))^{1/2} = \inf\{(EH^2)^{1/2} : H \in \mathcal{L}, G(H) \in -\mathtt{P}\}, \tag{3.12}$$

and so our task now is to show that (3.12) implies (3.11).

Suppose first that the feasible set in the optimization problem (3.10) is not empty. Then there is $H \in \mathcal{L}$ such that $G(H)$ belongs to the interior of the cone $-\mathtt{P}$, so by Theorem 1, page 224 of [7], we conclude that

$$\begin{aligned} &(\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi))^{1/2} \\ &\quad = \max_{\mu\in M^+([0,1])} \inf_{H\in\mathcal{L}} \left[(EH^2)^{1/2} + \int_0^1 G(H)(v)\mu(dv)\right], \end{aligned} \tag{3.13}$$

and we may use "max" instead of "sup" because an optimal $\mu \in M^+([0,1])$ exists. For a fixed $\mu \in M^+([0,1])$ with total mass $\|\mu\|$, we let $\hat{\mu} = \mu/\|\mu\| \in M_1^+([0,1])$. Then

$$\begin{aligned} &\inf_{H\in\mathcal{L}} \left[(EH^2)^{1/2} + \int_0^1 G(H)(v)\mu(dv)\right] \\ &\quad = \|\mu\| + \inf_{H\in\mathcal{L}} \left[(EH^2)^{1/2} - \|\mu\| \int_0^1 w_H(\xi(v))\hat{\mu}(dv)\right] \end{aligned} \tag{3.14}$$

$$
\begin{aligned}
&= \|\mu\| + \inf_{a\ge 0} a\left[1 - \|\mu\| \sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 w_H\big(\xi(v)\big)\hat{\mu}(dv)\right] \\
&= \begin{cases} -\infty, & \text{if } \|\mu\| > \left[\displaystyle\sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 w_H\big(\xi(v)\big)\hat{\mu}(dv)\right]^{-1}, \\ \|\mu\|, & \text{if } \|\mu\| \le \left[\displaystyle\sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 w_H\big(\xi(v)\big)\hat{\mu}(dv)\right]^{-1}. \end{cases}
\end{aligned}
$$

Therefore,

$$
\big(\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b};\xi)\big)^{1/2} = \left[\inf_{\mu\in M_1^+([0,1])} \sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 w_H\big(\xi(v)\big)\mu(dv)\right]^{-1},
$$

and (3.11) follows, since by the reproducing property of the RKHS, for every $\mu \in M_1^+([0,1])$,

$$
\begin{aligned}
\sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 w_H\big(\xi(v)\big)\mu(dv) &= \sup_{H\in\mathcal{L}:\, EH^2=1} \int_0^1 \big(w_H, R_X\big(\xi(v),\cdot\big)\big)_{\mathcal{H}}\mu(dv) \\
&= \sup_{w\in\mathcal{H}:\, \|w\|_{\mathcal{H}}=1} \left(w, \int_0^1 R_X\big(\xi(v),\cdot\big)\mu(dv)\right)_{\mathcal{H}} \\
&= \left(\int_0^1\int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu(du)\mu(dv)\right)^{1/2}.
\end{aligned}
$$

In the last step we have used the fact that

$$
w_\mu \triangleq \int_0^1 R_X\big(\xi(v),\cdot\big)\mu(dv) \in \mathcal{H},
$$

so the supremum of the inner product is achieved at $w = w_\mu/\|w_\mu\|_{\mathcal{H}}$, and

$$
\|w_\mu\| = \left(\int_0^1\int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu(du)\mu(dv)\right)^{1/2}.
$$

This establishes (3.11) for the case that the feasible set in (3.10) is not empty. We now turn to the case in which this set is, indeed, empty. This will complete the proof of the theorem. In this case (3.11) reduces to the statement

$$
I_* \triangleq \min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu(du)\mu(dv) = 0. \tag{3.15}
$$

Suppose that, to the contrary, $I_* > 0$. Let $\mu_0 \in M_1^+([0,1])$ achieve the minimum value in the integral defining I_*. Consider the continuous real-valued function

$$
W(u) = \int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu_0(dv), \qquad 0 \le u \le 1.
$$

If this function never vanishes, then, by continuity and compactness, it is bounded away from zero, so a sufficiently large in absolute value multiple of the random variable in $\mathcal{L}$ given by

$$H = \int_0^1 X(\xi(v))\mu_0(dv)$$

is feasible for the optimization problem (3.10), contradicting the assumption that the set of feasible solutions is empty.

Hence, there is $u_0 \in [0, 1]$ such that $W(u_0) = 0$. For $0 < \varepsilon < 1$ define a probability measure in $M_1^+([0, 1])$ by

$$\mu_\varepsilon = (1 - \varepsilon)\mu_0 + \varepsilon\delta_{u_0},$$

where δ_a denotes the point mass at a. Note that

$$\begin{aligned} I(\varepsilon) &\triangleq \int_0^1 \int_0^1 R_{\mathbf{X}}(\xi(u), \xi(v))\mu_\varepsilon(du)\mu_\varepsilon(dv) \\ &= (1 - \varepsilon)^2 I_* + 2\varepsilon(1 - \varepsilon)W(u_0) + \varepsilon^2 R_{\mathbf{X}}(0) \\ &= (1 - \varepsilon)^2 I_* + \varepsilon^2 R_{\mathbf{X}}(0). \end{aligned}$$

Since I_* was assumed to be positive, we see that

$$\frac{d}{d\varepsilon} I(\varepsilon)\Big|_{\varepsilon=0} < 0,$$

which contradicts the minimality of I_*. This proves (3.15) and so the theorem. $\square$

Observe that an alternative way of stating the result of Theorem 3.2 is

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) = \inf_{\xi \in \mathcal{P}(\mathbf{a},\mathbf{b})} \left[\min_{\mu \in M_1^+(\xi)} \int_\xi \int_\xi R_{\mathbf{X}}(\mathbf{t}, \mathbf{s})\mu(d\mathbf{t})\mu(d\mathbf{s}) \right]^{-1}, \tag{3.16}$$

where $M_1^+(\xi)$ is the set of all probability measures in $\mathbb{R}^d$ supported by the path ξ (strictly speaking, by the compact image of the interval [0, 1] under ξ). For a fixed path $\xi \in \mathcal{P}(\mathbf{a}, \mathbf{b})$, the quantity

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi) = \left[\min_{\mu \in M_1^+(\xi)} \int_\xi \int_\xi R_{\mathbf{X}}(\mathbf{t}, \mathbf{s})\mu(d\mathbf{t})\mu(d\mathbf{s}) \right]^{-1} \tag{3.17}$$

is known as *the capacity of the path* ξ *with respect to the kernel* $R_{\mathbf{X}}$; see [4]. Therefore, we can treat the problem of solving (3.16) as one of finding a path between the points $\mathbf{a}$ and $\mathbf{b}$ of minimal capacity.

4. Fixed paths and measures of minimal energy. The dual formulation (3.4) of the optimization problem required to find the asymptotics of the path existence probability $\Psi_{\mathbf{a},\mathbf{b}}(u)$ involves solving fixed path ξ optimization problems (3.10) or (3.11). For a fixed path we have the following version of Theorems 3.1 and 3.2.

THEOREM 4.1. (i) *For a* $\xi \in \mathcal{P}(\mathbf{a}, \mathbf{b})$ *let*

$$\Psi_{\mathbf{a},\mathbf{b}}(u; \xi) = P\big(X\big(\xi(v)\big) > u, 0 \le v \le 1\big).$$

Then

$$\lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{\mathbf{a},\mathbf{b}}(u; \xi) = -\frac{1}{2}\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi). \tag{4.1}$$

(ii) *The primal problem* (3.10) *can be rewritten in the form*

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi) = \inf\big\{EH^2 : H \in \mathcal{L}, E\big[X\big(\xi(v)\big)H\big] \ge 1, 0 \le v \le 1\big\}. \tag{4.2}$$

Further, if the feasible set in (4.2) *is nonempty, then the infimum in* (4.2) *is achieved at a unique* $H_\xi \in \mathcal{L}$.

(iii) *The set* $\mathcal{W}_\xi$ *of* $\mu \in M_1^+([0, 1])$ *over which the minimum in the dual problem* (3.11) *is achieved is a weakly compact convex subset of* $M_1^+([0, 1])$. *Furthermore, if the feasible set in* (4.2) *is nonempty, then, for every* $\mu \in \mathcal{W}_\xi$,

$$\mu\big(\big\{0 \le v \le 1 : E\big[X\big(\xi(v)\big)H_\xi\big] > 1\big\}\big) = 0. \tag{4.3}$$

(iv) *Suppose that the feasible set in* (4.2) *is nonempty. Then for every* $\varepsilon > 0$,

$$P\bigg(\sup_{0\le v\le 1}\bigg|\frac{1}{u}X\big(\xi(v)\big) - x_\xi(v)\bigg| \ge \varepsilon \bigg| X\big(\xi(v)\big) > u, 0 \le v \le 1\bigg) \to 0 \tag{4.4}$$

as $u \to \infty$. *Here*

$$x_\xi(v) = E\big[X\big(\xi(v)\big)H_\xi\big], \qquad 0 \le v \le 1. \tag{4.5}$$

The probability measures $\mu \in \mathcal{W}_\xi$ are called *capacitary measures*, or *measures of minimal energy*; see [4].

PROOF OF THEOREM 4.1. Part (i) of the theorem can be proved in the same way as Theorem 3.1. The fact that the primal formulations (3.10) and (4.2) are equivalent is an immediate consequence of the definition of w_H. Suppose now that the feasible set in (4.2) is nonempty, and let $H_n \in \mathcal{L}$, $n = 1, 2, \dots$ be a sequence of feasible solutions such that $EH_n^2 \to \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$. The weak compactness of the unit ball in $\mathcal{L}$ shows that this sequence has a subsequential weak limit H_ξ with $EH_\xi^2 = \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$. Since the set of feasible solutions is weakly closed, H_ξ is feasible. The uniqueness of the optimal solution to (4.2) follows from convexity of the norm.

Convexity and weak compactness of the set $\mathcal{W}_\xi$ follow from the nonnegative definiteness and continuity of $R_{\mathbf{X}}$; see, for example, Remark 2, page 160, in [4]. The statement (4.3) is a part of the relation between the dual and primal optimal solutions; see Theorem 1, page 224, in [7].

For part (iv) of the theorem, note that by the Gaussian large deviation principle of Theorem 3.4.5 in [3],

$$
\begin{aligned}
&\limsup_{u\to\infty} \frac{1}{u^2} \log P\left(X(\xi(v)) > u, 0 \le v \le 1, \sup_{0\le v\le 1} \left| \frac{1}{u} X(\xi(v)) - x_\xi(v) \right| \ge \varepsilon \right) \\
&\quad \le -\frac{1}{2} \inf\Big\{ EH^2, H \in \mathcal{L} : E[X(\xi(v))H] \ge 1, 0 \le v \le 1, \\
&\qquad\qquad \sup_{0\le v\le 1} \big| E[X(\xi(v))H] - E[X(\xi(v))H_\xi] \big| \ge \varepsilon \Big\}.
\end{aligned}
\tag{4.6}
$$

Therefore, the statement (4.4) will follow from Parts (i) and (ii) of the theorem once we prove that the infimum in (4.6) is strictly larger than $\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$. Suppose that, to the contrary, the two infima are equal. By the weak compactness of the unit ball in $\mathcal{L}$ and the fact that the feasible set in (4.6) is weakly closed, this would imply existence of H_* feasible for (4.6) such that $EH_*^2 = EH_\xi^2$. Since H_ξ is not feasible for (4.6), we know that $H_* \ne H_\xi$. Since H_* is feasible for (4.2), we have obtained a contradiction to the uniqueness of H_ξ proved above. This completes the proof of the theorem. □

REMARK 4.2. Theorem 4.1 has the following important interpretation. Assuming that the feasible set in (4.2) is nonempty, part (iv) of the theorem implies that the nonrandom function x_ξ in (4.5) is the most likely choice for the normalized sample path $u^{-1}X(\xi(v)), 0 \le v \le 1$ along ξ, given that $\{X(\xi(v)) > u, 0 \le v \le 1\}$. Part (iii) of the theorem implies that the values of the random field along the path ξ have to (nearly) touch the level u at the points of the support of any measure of minimal energy. In other words, the sample path needs to be "supported," or "held," at the level u at the points of the support in order to achieve the highest probability of exceeding the high level u along the entire path ξ. We will see explicit examples of how this works in the following section, when we more closely investigate the one-dimensional case.

The duality relation of the optimization problems (4.2) and (3.11) immediately provides upper and lower bounds on $\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$ of the form

$$
\left[\int_0^1 \int_0^1 R_{\mathbf{X}}(\xi(u), \xi(v)) \mu(du) \mu(dv) \right]^{-1} \le \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi) \le EH^2 \tag{4.7}
$$

for any $\mu \in M_1^+([0, 1])$ and any $H \in \mathcal{L}$ feasible for (4.2). In particular, if

$$
\left[\int_0^1 \int_0^1 R_{\mathbf{X}}(\xi(u)\xi(v)) \mu(du) \mu(dv) \right]^{-1} = EH^2 \tag{4.8}
$$

for some μ and H as above, then $\mu \in \mathcal{W}_\xi$, $H = H_\xi$, and the common value in (4.8) is equal to $\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$.

Finding a measure of minimal energy, $\mu \in \mathcal{W}_\xi$, is, in general, a difficult problem. The following theorem includes a characterization of these measures.

THEOREM 4.3. *Assume that the feasible set in* (4.2) *is nonempty.*

(i) *For every* $\mu \in \mathcal{W}_\xi$ *we have*

$$H_\xi = \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi) \int_0^1 X(\xi(v))\mu(dv)$$

with probability 1.

(ii) *A probability measure* $\mu \in M_1^+([0,1])$ *is a measure of minimal energy* (*i.e.*, $\mu \in \mathcal{W}_\xi$) *if and only if*

$$\begin{aligned} &\min_{0\le v\le 1} \int_0^1 R_{\mathbf{X}}(\xi(u), \xi(v))\mu(du) \\ &\quad = \int_0^1 \int_0^1 R_{\mathbf{X}}(\xi(u_1), \xi(u_2))\mu(du_1)\mu(du_2) > 0. \end{aligned} \tag{4.9}$$

Note that part (ii) of the theorem also says that the integral in the left-hand side of (4.9) is equal to the double integral in its right-hand side for μ-almost every $0 \le v \le 1$.

PROOF OF THEOREM 4.3. For part (i), let $\mu \in \mathcal{W}_\xi$. The calculations following the maximization problem (3.13) show that $\mu_\xi = \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)^{1/2}\mu$ is an optimal measure for that problem. It follows from Theorem 1, page 224, in [7] that H_ξ solves the minimization problem in (3.14), when any measure in $M^+([0,1])$ optimal for (3.13) is used. Using the measure μ_ξ, we see that

$$H_\xi = a\int_0^1 X(\xi(v))\mu_\xi(dv) = a\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)^{1/2} \int_0^1 X(\xi(v))\mu(dv)$$

for some $a > 0$. Testing all random variables of the type

$$H_\xi = b\int_0^1 X(\xi(v))\mu(dv), \qquad b > 0,$$

in (4.2) leads to the conclusion that

$$b = \left[\min_{0\le v\le 1} \int_0^1 R_{\mathbf{X}}(\xi(u), \xi(v))\mu(du)\right]^{-1}.$$

The fact that $b = \mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi)$ follows now from the optimality of μ and the general properties of measures of minimal energy for bounded symmetric kernels; see, for example, Theorem 2.4 in [4].

We now prove part (ii). Suppose first that μ satisfies (4.9), and define $H \in \mathcal{L}$ by

$$H = \big(K(\mu)\big)^{-1} \int_0^1 X\big(\xi(v)\big)\mu(dv),$$

where $K(\mu)$ is the double integral in the right-hand side of (4.9). Note that for any $0 \le v \le 1$,

$$E\big[X\big(\xi(v)\big)H\big] = \big(K(\mu)\big)^{-1} \int_0^1 R_{\mathbf{X}}\big(\xi(u), \xi(v)\big)\mu(du) \ge 1$$

by (4.9). Therefore, H is feasible for (4.2). However,

$$EH^2 = \frac{1}{K(\mu)^2} E\bigg(\int_0^1 X\big(\xi(v)\big)\mu(dv)\bigg)^2 = \frac{1}{K(\mu)},$$

so that μ and H satisfy the relation (4.8). Hence, $\mu \in \mathcal{W}_\xi$ (and $H = H_\xi$).

In the opposite direction, if $\mu \in \mathcal{W}_\xi$, then the equality in (4.9) is a general property of measures of minimal energy for bounded symmetric kernels, as in the proof of part (i). The fact that the equal terms in (4.9) are positive follows from the fact that the feasible set in (4.2) is nonempty, so $\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}; \xi) < \infty$. □

REMARK 4.4. If the random field $\mathbf{X}$ is stationary, then the results of this section can be restated in the language used in Section 3 in the stationary case. In particular, the primal problem (4.2) becomes

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}; \xi) = \inf\bigg\{\int_{\mathbb{R}^d} \|h(\mathbf{x})\|^2 F_{\mathbf{X}}(d\mathbf{x}) : \int_{\mathbb{R}^d} e^{i(\xi(v), \mathbf{x})}\bar{h}(\mathbf{x}) F_{\mathbf{X}}(d\mathbf{x}) > 1, 0 \le v \le 1\bigg\},$$

while the optimal solution of the primal problem in part (i) of Theorem 4.3 becomes

$$h_\xi(\mathbf{x}) = \mathcal{C}_{\mathbf{X}}(\mathbf{a}; \xi) \int_0^1 e^{i(\xi(v), \mathbf{x})}\mu(dv) \qquad \text{for } F_{\mathbf{X}}\text{-almost all } \mathbf{x} \in \mathbb{R}^d$$

and for any $\mu \in \mathcal{W}_\xi$. This relation can also be restated in terms on the measures supported by the (image of) path ξ instead of the unit interval, as in (3.17). If μ is an optimal measure in (3.17), then we have

$$h_\xi(\mathbf{x}) = \mathcal{C}_{\mathbf{X}}(\mathbf{a}; \xi) \int_\xi e^{i(\mathbf{t}, \mathbf{x})}\mu(d\mathbf{t}) \qquad \text{for } F_{\mathbf{X}}\text{-almost all } \mathbf{x} \in \mathbb{R}^d. \tag{4.10}$$

Note that the function in the right-hand side of (4.10) is, up to a constant, the characteristic function of the measure μ. If the support of the spectral measure $F_{\mathbf{X}}$ happens to be the entire space $\mathbb{R}^d$, then the characteristic functions of all optimal measures in (3.17) are equal and, hence, the uniqueness of a characteristic function shows that, in this case [and as long as the feasible set in (4.2) is nonempty], there is exactly one probability measure $\mu \in M_1^+(\xi)$ of minimal energy.

REMARK 4.5. An immediate conclusion of part (ii) of Theorem 4.3 and the assumed continuity of the covariance function is that the function

$$v \mapsto \int_0^1 R_{\mathbf{X}}\big(\xi(u), \xi(v)\big)\mu(du), \qquad 0 \le v \le 1,$$

is constant on the support of any measure $\mu \in \mathcal{W}_\xi$. This seems to indicate that the support of any measure of minimal energy may not be "large." In the examples below, however, this intuition holds only in some cases.

5. The one-dimensional case. In this and the following two sections we specialize to the one-dimensional case $d = 1$. Let $a < b$. As before, we are interested in the probability

$$\Psi_{a,b}(u) = P\big(X(t) > u, a \le t \le b\big).$$

There is, essentially, a single path between a and b, and the results of the previous two sections immediately specialize to yield the following special case. [Note that condition (3.3) is superfluous in the one-dimensional nonstationary case.]

THEOREM 5.1. *Let* $\mathbf{X}$ *be a continuous Gaussian process on an interval including* $[a, b]$. *Then the limit*

$$-\frac{1}{2}C_{\mathbf{X}}(a,b) \triangleq \lim_{u\to\infty} \frac{1}{u^2} \log \Psi_{a,b}(u)$$

exists, *and*

$$\begin{aligned} &C_{\mathbf{X}}(a,b) \\ &= \inf\big\{EH^2 : H \in \mathcal{L}, E\big[X\big(a + (b-a)v\big)H\big] \ge 1, 0 \le v \le 1\big\} \qquad (5.1) \\ &= \left[\min_{\mu \in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a + (b-a)u, a + (b-a)v\big)\mu(du)\mu(dv)\right]^{-1}. \qquad (5.2) \end{aligned}$$

If the process $\mathbf{X}$ *is stationary, an alternative expression for* $C_{\mathbf{X}}(a) \triangleq C_{\mathbf{X}}(0, a)$, $a > 0$, *is given by*

$$\begin{aligned} C_{\mathbf{X}}(a) = \inf\bigg\{ &\int_{-\infty}^{\infty} \|h(x)\|^2 F_{\mathbf{X}}(dx) : \int_{-\infty}^{\infty} e^{ivax} \bar{h}(x) F_{\mathbf{X}}(dx) > 1, \\ &0 \le v \le 1\bigg\}. \end{aligned} \qquad (5.3)$$

The set $\mathcal{W}_{a,b}$ *of* $\mu \in M_1^+([0, 1])$ *over which the minimum in* (5.2) *is achieved is a weakly compact convex subset of* $M_1^+([0, 1])$. *The measures in* $\mathcal{W}_{a,b}$ *are characterized by the relation*

$$\begin{aligned} &\min_{0\le v\le 1} \int_0^1 R_{\mathbf{X}}\big(a + (b-a)u, a + (b-a)v\big)\mu(du) \\ &\quad = \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a + (b-a)u_1, a + (b-a)u_2\big)\mu(du_1)\mu(du_2). \end{aligned} \qquad (5.4)$$

Suppose, further, that the problem (5.1) *has a feasible solution. In this case the double integral in* (5.4) *is positive for any* $\mu \in \mathcal{W}_{a,b}$, *and the problem* (5.1) *has a unique optimal solution*, $H_{a,b}$. *For each* $\mu \in \mathcal{W}_{a,b}$,

$$H_{a,b} = C_{\mathbf{X}}(a, b) \int_0^1 X\bigl(a + (b - a)v\bigr)\mu(dv)$$

with probability 1. *In the stationary case, the problem* (5.3) *has a unique optimal solution*, h_a. *For each* $\mu \in \mathcal{W}_a \stackrel{\Delta}{=} \mathcal{W}_{0,a}$

$$h_a(x) = C_{\mathbf{X}}(a) \int_0^1 e^{iavx} \mu(dv) \qquad \text{for } F_{\mathbf{X}}\text{-almost all } -\infty < x < \infty.$$

The conditional law on $C[a, b]$ *of the scaled process* $u^{-1}\mathbf{X}$ *restricted to the interval* $[a, b]$, *given that* $X(t) > u, a \le t \le b$, *converges as* $u \to \infty$ *to the Dirac measure at*

$$x_{a,b}(t) = C_{\mathbf{X}}(a, b) \int_0^1 R_{\mathbf{X}}\bigl(t, a + (b - a)v\bigr)\mu(dv), \qquad a \le t \le b \tag{5.5}$$

and

$$\mu\bigl(\{0 \le v \le 1 : x_{a,b}\bigl(a + (b - a)v\bigr) > 1\}\bigr) = 0.$$

Finally, if the process $\mathbf{X}$ *is stationary, and the support of* $F_{\mathbf{X}}$ *is the entire real line, then the set* $\mathcal{W}_a$ *consists of a single probability measure*, μ_a.

REMARK 5.2. Suppose that the process $\mathbf{X}$ is stationary. For $\mu \in M_1^+([0, 1])$ define $\hat{\mu} = \mu \circ T^{-1}$ with $T : [0, 1] \to [0, 1]$ being the reflection map $Tx = 1 - x$, $0 \le x \le 1$. If $\mu \in \mathcal{W}_a$, then $\hat{\mu}$ satisfies conditions (5.4) because μ does, hence $\hat{\mu} \in \mathcal{W}_a$ as well. By convexity of $\mathcal{W}_a$, so does the symmetric (around $x = 1/2$) probability measure $1/2(\mu + \hat{\mu})$. Therefore, $\mathcal{W}_a$ always contains a symmetric measure. In particular, if $\mathcal{W}_a$ is a singleton, then the unique measure of minimal energy is symmetric.

In the remainder of this section we concentrate on the stationary case. We will investigate how the probability measure μ_a, the function h_a and the limiting shape $x_a \stackrel{\Delta}{=} x_{0,a}$ change as functions of a. This will help us understand the order of magnitude of the probability $\Psi_a(u)$ for varying lengths a of the interval and, according to part (iv) of Theorem 4.1, it will tell us the most likely shape the process $\mathbf{X}$ takes when it exceeds a high level u along the entire interval $[0, a]$.

Our first result describes the situation occurring for some, but not all, stationary Gaussian processes on short intervals.

PROPOSITION 5.3. *Let* $\mathbf{X}$ *be a stationary continuous Gaussian process. Suppose that for some* $a > 0$ *the following condition holds*:

$$R_{\mathbf{X}}(t) + R_{\mathbf{X}}(a - t) \ge R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a) > 0 \qquad \text{for all } 0 \le t \le a. \tag{5.6}$$

Then a measure in $\mathcal{W}_a$ is given by

$$\mu^{(1)} \triangleq \tfrac{1}{2}\delta_0 + \tfrac{1}{2}\delta_1. \tag{5.7}$$

Furthermore,

$$C_{\mathbf{X}}(a) = \frac{2}{R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a)}, \tag{5.8}$$

$$h_a(x) = \frac{1 + e^{iax}}{R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a)} \qquad \textit{for } F_{\mathbf{X}}\textit{-almost all } -\infty < x < \infty \tag{5.9}$$

and

$$x_a(t) = \frac{R_{\mathbf{X}}(t) + R_{\mathbf{X}}(a-t)}{R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a)}, \qquad 0 \le t \le a. \tag{5.10}$$

PROOF. Once we show that $\mu^{(1)} \in \mathcal{W}_a$, the rest of the statements will follow from Theorem 5.1. In order to prove (5.7), we need to check conditions (5.4). These follow immediately from (5.6) and the fact that

$$\int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u_1 - u_2)\big)\mu^{(1)}(du_1)\mu^{(1)}(du_2) = \frac{1}{2}R_{\mathbf{X}}(0) + \frac{1}{2}R_{\mathbf{X}}(a),$$

while for $0 \le v \le 1$,

$$\int_0^1 R_{\mathbf{X}}\big(a(u - v)\big)\mu^{(1)}(du) = \frac{1}{2}R_{\mathbf{X}}(av) + \frac{1}{2}R_{\mathbf{X}}(a - av). \qquad \square$$

REMARK 5.4. Note that a sufficient (but not necessary) condition for (5.6) is concavity of the covariance function $R_{\mathbf{X}}$ on the interval $[0, a]$. Indeed, for a concave covariance function the derivative exists apart from a countable set of points and is monotone. Therefore,

$$R_{\mathbf{X}}(t) - R_{\mathbf{X}}(0) = \int_0^t R'_{\mathbf{X}}(s)\,ds \ge \int_0^t R'_{\mathbf{X}}(a - t + s)\,ds = R_{\mathbf{X}}(a) - R_{\mathbf{X}}(a - t).$$

In particular, if the process $\mathbf{X}$ has a finite second spectral moment, then the second derivative of the covariance function exists, is continuous and negative at zero (unless the covariance function is constant). Therefore, the derivative stays negative on an interval around the origin, hence, the covariance function is concave on $[0, a]$, and (5.6) holds, for $a > 0$ small enough.

On the other hand, apart from degenerate cases, the situation described in Proposition 5.3 cannot continue to hold for arbitrarily large a. For example, if the covariance function vanishes at infinity, then (5.6) fails for a large enough and $t = a/2$, say.

In addition, a simple calculation shows that it is always true that

$$\lim_{u\to\infty} \frac{1}{u^2} \log P\big(\mathbf{X}(0) > u, \mathbf{X}(a) > u\big) = -\big(R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a)\big)^{-1}. \tag{5.11}$$

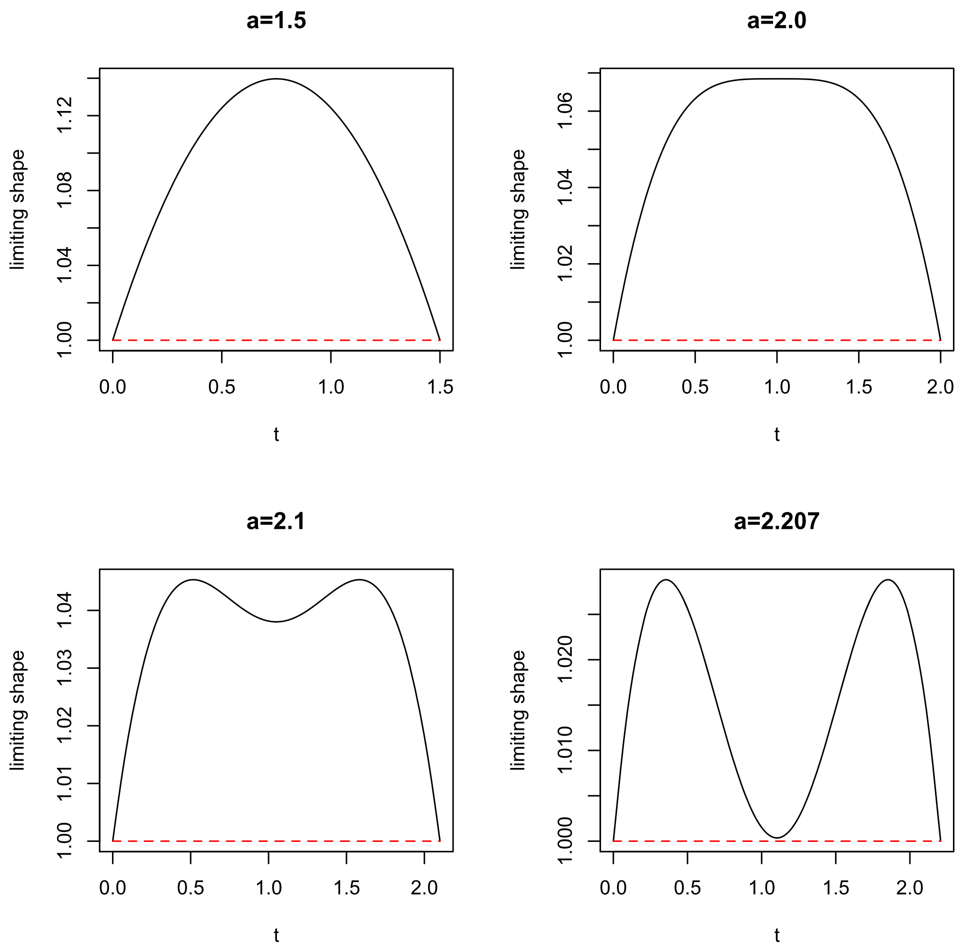

FIG. 1. *Limiting shapes x_a for the stationary Gaussian process with covariance function $R_{\mathbf{X}}(t) = \exp(-t^2/2)$ when Proposition 5.3 applies.*

Combining this with (5.8) shows that, in the scenario of Proposition 5.3, the probability that $\mathbf{X}$ exceeds a high level over an entire interval and the probability that it does so only at the endpoints of the interval are, at a logarithmic scale, the same.

The plots of Figure 1 show the limiting shape x_a for the stationary Gaussian process with covariance function $R_{\mathbf{X}}(t) = \exp(-t^2/2)$, for a range of a for which Proposition 5.3 applies. In this case the largest such a is approximately equal to 2.2079. See Example 6.1 for more details.

The plots of Figure 1 indicate that as a approaches a critical value (approximately 2.2079 in this case), the limiting curve x_a "attempts" to cross the level 1 at the midpoint of $[0, a]$. Equivalently, the normalized process $u^{-1}\mathbf{X}$ attempts to drop below level 1 at that point and so, speaking heuristically, it has to be "supported" at the midpoint $t = a/2$. The interpretation of Theorem 4.1 in Remark 4.2 calls for adding a mass to the measure $\mu^{(1)}$ for the critical value of a at the midpoint of the interval. The next result shows that, in certain cases, this is indeed the optimal thing to do.

PROPOSITION 5.5. *Let* $\mathbf{X}$ *be a stationary continuous Gaussian process. Suppose that, for some* $a > 0$,

$$R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a) > 2R_{\mathbf{X}}(a/2), \tag{5.12}$$

and let

$$\varepsilon_a = \frac{R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a) - 2R_{\mathbf{X}}(a/2)}{3R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a) - 4R_{\mathbf{X}}(a/2)} \in (0, 1]. \tag{5.13}$$

Suppose that for all $0 \le t \le a/2$,

$$\begin{aligned} R_{\mathbf{X}}(t) &+ R_{\mathbf{X}}(a-t) - R_{\mathbf{X}}(0) - R_{\mathbf{X}}(a) \\ &\ge \varepsilon_a\big[R_{\mathbf{X}}(t) + R_{\mathbf{X}}(a-t) - R_{\mathbf{X}}(0) - R_{\mathbf{X}}(a) \\ &\qquad - 2\big(R_{\mathbf{X}}(a/2 - t) - R_{\mathbf{X}}(a/2)\big)\big]. \end{aligned} \tag{5.14}$$

Then a measure in $\mathcal{W}_a$ *is given by*

$$\mu^{(2)} \triangleq \frac{1-\varepsilon_a}{2}\delta_0 + \frac{1-\varepsilon_a}{2}\delta_1 + \varepsilon_a \delta_{1/2}. \tag{5.15}$$

Furthermore,

$$C_{\mathbf{X}}(a) = \frac{3R_{\mathbf{X}}(0) + R_{\mathbf{X}}(a) - 4R_{\mathbf{X}}(a/2)}{R_{\mathbf{X}}(0)^2 + R_{\mathbf{X}}(0)R_{\mathbf{X}}(a) - 2R_{\mathbf{X}}(a/2)^2}, \tag{5.16}$$

$$h_a(x) = C_{\mathbf{X}}(a)\left[\frac{1-\varepsilon_a}{2}\big(1 + e^{iax}\big) + \varepsilon_a e^{iax/2}\right] \tag{5.17}$$

for $F_{\mathbf{X}}$*-almost all* $-\infty < x < \infty$, *and*

$$x_a(t) = C_{\mathbf{X}}(a)\left[\frac{1-\varepsilon_a}{2}\big(R_{\mathbf{X}}(t) + R_{\mathbf{X}}(a-t)\big) + \varepsilon_a R_{\mathbf{X}}\big(|t - a/2|\big)\right], \tag{5.18}$$

$0 \le t \le a$.

PROOF. The proof is identical to that of Proposition 5.3 once we observe that, under (5.12), $\mu^{(2)}$ is a legitimate probability measure. $\square$

The plots of Figure 2 show the limiting shape x_a for the stationary Gaussian process with covariance function $R_{\mathbf{X}}(t) = \exp(-t^2/2)$, for a range of a for which Proposition 5.5 applies. In this case the range of a is, approximately, between 2.2079 and 3.9283. See Example 6.1 for more details.

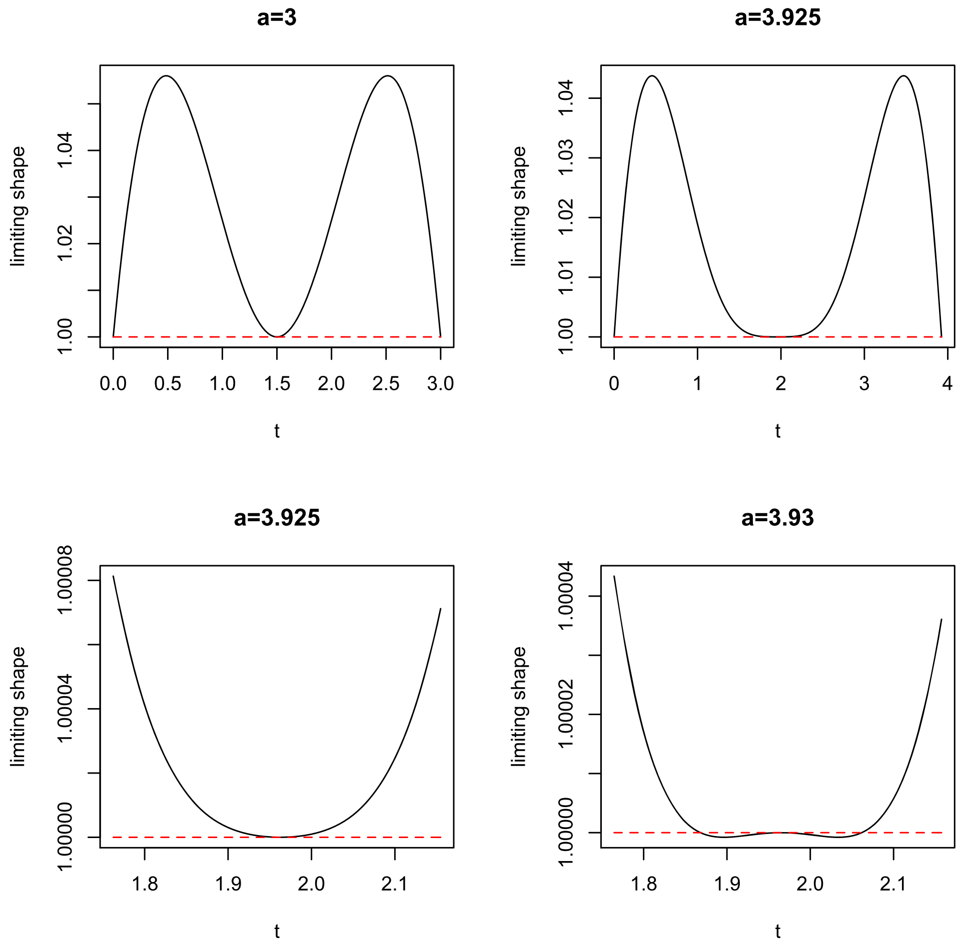

FIG. 2. *Limiting shapes for* $R_{\mathbf{X}}(t) = \exp(-t^2/2)$ *when Proposition* 5.5 *applies* (*the top row*). *The left plot in the bottom row is a blowup of the right plot in the top row. The right plot in the bottom row shows how the constraints are violated soon after the upper critical value of* a.

6. Specific covariance functions. In the previous section we saw some general results for one-dimensional processes, with some illustrative figures for what happens in the case of a Gaussian covariance function. In this section we look more carefully at this case, and also look at what can be said for an exponential covariance.

EXAMPLE 6.1. Consider the centered stationary Gaussian process with the Gaussian covariance function

$$R(t) = e^{-t^2/2}, \qquad t \in \mathbb{R}. \tag{6.1}$$

For this process the spectral measure has a Gaussian spectral density which is of full support in $\mathbb{R}$. In particular, for every $a > 0$ there is a unique (symmetric) measure of minimal energy. Furthermore, the second spectral moment is finite, so that, according to Remark 5.4, for $a > 0$ sufficiently small this process satisfies the conditions of Proposition 5.3. To find the range of a for which this happens, note that conditions (5.6) become, in this case,

$$e^{-t^2/2} + e^{-(a-t)^2/2} \geq 1 + e^{-a^2/2}, \qquad 0 \leq t \leq a. \tag{6.2}$$

Since the function

$$g(t) = e^{-t^2/2} + e^{-(a-t)^2/2}, \qquad 0 \leq t \leq a,$$

is concave if $0 \leq a \leq 2$, and has a unique local minimum, at $t = a/2$, when $a > 2$, it is only necessary to check (6.2) at the midpoint $t = a/2$. At that point the condition becomes

$$\psi(a) = 2e^{-a^2/8} - 1 - e^{-a^2/2} \geq 0.$$

The function ψ crosses 0 at $a_1 \approx 2.2079$, which is the limit of the validity of the situation of Proposition 5.3 in this case. The plots of Figure 1 show the limiting shape x_a for this process in the situation of Proposition 5.3.

Somewhat longer (and numerical) calculations show that the conditions of Proposition 5.5 hold for the process with the covariance function (6.1) for an interval of values of a after the conditions of Proposition 5.3 break down. The conditions of Proposition 5.5 continue to hold until the second derivative at the midpoint $t = a/2$ of the limiting function in (5.18) becomes negative (so that the function takes values smaller than 1 in a neighborhood of the midpoint). To find when this happens, we solve the equation

$$\frac{1-\varepsilon_a}{2}\big(R''_{\mathbf{X}}(t) + R''_{\mathbf{X}}(a-t)\big) + \varepsilon_a R''_{\mathbf{X}}\big(|t - a/2|\big) = 0$$

at $t = a/2$. The resulting equation

$$(1-\varepsilon_a)\big(a^2/4 - 1\big)e^{-a^2/8} - \varepsilon_a = 0$$

has the solution $a_2 \approx 3.9283$, which is the limit of the validity of the situation of Proposition 5.5 in this case. The plots of Figure 2 shed some light on the above discussion. This discussion indicates, and calculations confirm, that, in the next regime, the mass in the middle for the optimal measure splits into two parts that start to move away from the center. Heuristically, this is needed "to support" the trajectory that, otherwise, would "dip" below 1 outside of the midpoint.

These calculations rapidly become complicated. They seem to indicate that the next regime continues to hold until around $a_3 \approx 5.4508$. In this regime the optimal measure takes the form

$$\mu^{(3)} \triangleq \frac{1-\varepsilon_a}{2}\delta_0 + \frac{1-\varepsilon_a}{2}\delta_1 + \frac{\varepsilon_a}{2}\delta_{1/2-d_a} + \frac{\varepsilon_a}{2}\delta_{1/2+d_a}, \tag{6.3}$$

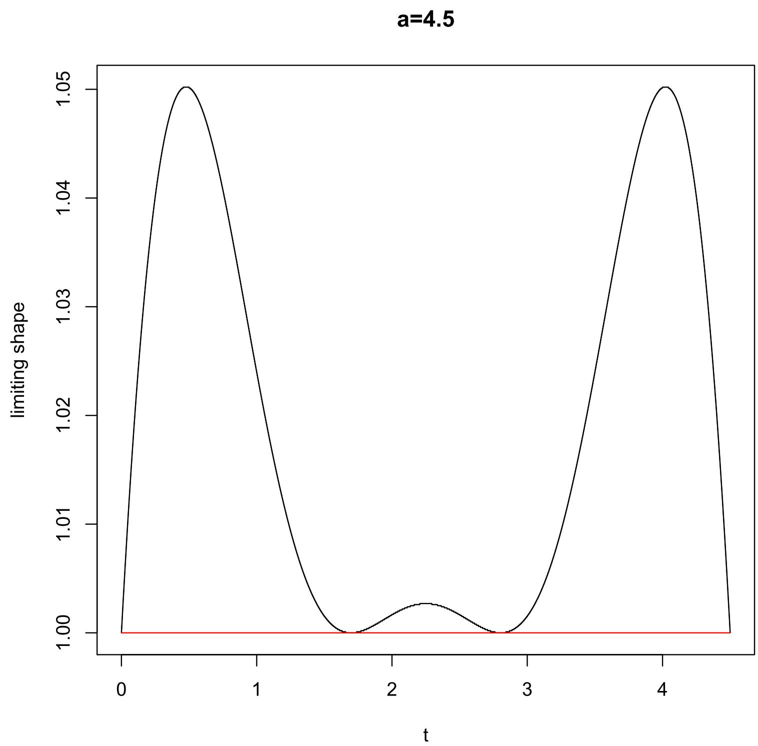

FIG. 3. *The limiting shape in the case* $a = 4.5$ *for* $R_{\mathbf{X}}(t) = \exp(-t^2/2)$.

where d_a is the distance of two internal masses from the midpoint. When $a = 4.5$, $\varepsilon_{4.5} = 0.36632$ and $d_{4.5} = 0.12285$, so that the internal atoms are at 0.37715 and 0.62285, and the rest of the support is concentrated at the endpoints of the interval with probabilities 0.31684. Figure 3 shows the limiting shape $x_{4.5}$.

It would be nice to understand all regimes, but we do not yet know how to find a general structure. On the other hand, Section 7 gives asymptotic results for $a \to \infty$.

Finally, Figure 4 shows the growth of the exponent $C_{\mathbf{X}}(a)$ with a for as long as either Proposition 5.3 or Proposition 5.5 applies.

The next example shows a situation very different from that of Example 6.1.

EXAMPLE 6.2. Consider an Ornstein–Uhlenbeck process, that is, a centered stationary Gaussian process with the covariance function

$$R(t) = e^{-|t|}, \qquad t \in \mathbb{R}. \tag{6.4}$$

For this process the spectral measure has a Cauchy spectral density, so it is also of full support in $\mathbb{R}$. Therefore, for every $a > 0$ there is a unique (symmetric) measure

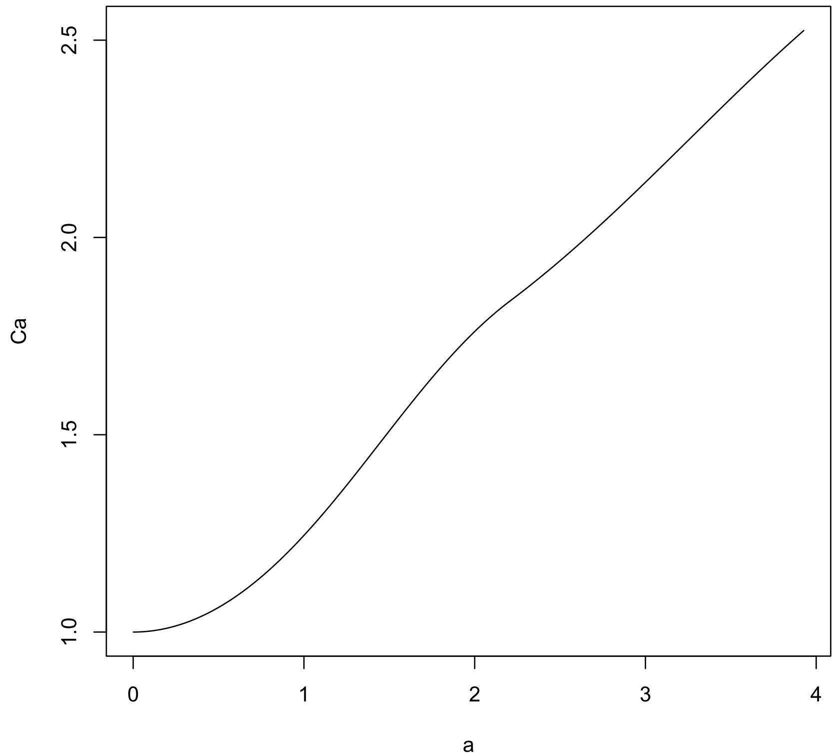

FIG. 4. *The exponent $C_{\mathbf{X}}(a)$ as a function of a for $R_{\mathbf{X}}(t) = \exp(-t^2/2)$.*

of minimal energy. In this case, however, even the first spectral moment is infinite. The covariance function is actually *convex* on the positive half-line so, in particular, the conditions of Proposition 5.3 fail for all $a > 0$. In fact, it is elementary to check that for the probability measure

$$\mu = \frac{1}{a+2}\delta_0 + \frac{1}{a+2}\delta_1 + \frac{a}{a+2}\lambda \tag{6.5}$$

[where λ is the Lebesgue measure on $(0, 1)$], the integrals

$$\int_0^1 R_{\mathbf{X}}\bigl(a(u-v)\bigr)\mu(du), \qquad 0 \le v \le 1,$$

have a constant value, equal to $2/(a+2)$. Therefore, the measure μ in (6.5) is the measure of minimal energy, and $C_{\mathbf{X}}(a) = (a+2)/2$ for all $a > 0$.

By Theorem 5.1 we conclude that the limiting function x_a is equal to 1 almost everywhere in $[0, a]$ with respect to the Lebesgue measure. Since x_a is continuous, it is identically equal to 1 on $[0, a]$.

Examples 6.1 and 6.2 demonstrate a number of the ways a stationary Gaussian process "prefers," in the large deviations sense, to stay above a high level over an interval. The process of Example 6.1 with covariance function (6.1) is smooth; the most likely way for it to stay above a level is to force it to be "slightly" above that level at a properly chosen finite set of time points; after that it is "held" above the level at the rest of the interval $[0, a]$ by the correlations of the process. The optimal configuration of the finite set of points depends on the length of the interval $[0, a]$, and it appears to undergo phase transitions at certain critical interval lengths. The complete picture of this "dynamical system" of finite sets remains unclear. On the other hand, the Ornstein–Uhlenbeck process of Example 6.2 is continuous, but not smooth. In fact, it behaves locally like a Brownian motion. Therefore, "holding" it "slightly" above a level at a discrete point does not help, since it "wants" immediately to go below that level. This explains the nature of the optimal measure μ in (6.5), and this nature stays the same no matter how short or long the interval $[0, a]$ is. In particular, phase transitions do not happen for this process.

It remains to be investigated whether other types of behavior are possible, and under what exact conditions on the Gaussian process each type of behavior occurs. It is also likely that minimal energy measures in $\mathcal{W}_a$ carry additional information, describing how "slightly" above the level u a Gaussian process is most likely to be, given that it is above that level along the interval. The exact nature of this information also remains to be investigated.

7. Asymptotics for long intervals. In this section we investigate the asymptotics of the exponent $C_{\mathbf{X}}(a)$ for large a. We start with a result showing that, for certain short memory stationary Gaussian processes, the exponent $C_{\mathbf{X}}(a)$ grows linearly with a over long intervals. Furthermore, the energy of the uniform distribution λ on $[0, 1]$ becomes, asymptotically, minimal.

THEOREM 7.1. *Let* $\mathbf{X}$ *be a stationary continuous Gaussian process. Assume that* $R_{\mathbf{X}}$ *is positive, and satisfies the following condition*:

$$\int_0^\infty R(t)\,dt < \infty. \tag{7.1}$$

Then, with λ *denoting the uniform probability measure on* $[0, 1]$,

$$\begin{aligned}\lim_{a\to\infty} \frac{1}{a} C_{\mathbf{X}}(a) &= \Big(\lim_{a\to\infty} a \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u-v)\big)\lambda(du)\lambda(dv)\Big)^{-1} \\ &= \frac{1}{2\int_0^\infty R(t)\,dt}.\end{aligned} \tag{7.2}$$

PROOF. By Theorem 5.1, the statement of the present theorem is equivalent to the following pair of claims:

$$\lim_{a\to\infty} a \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u-v)\big)\lambda(du)\lambda(dv) = 2\int_0^\infty R(t)\,dt \tag{7.3}$$

and

$$\liminf_{a\to\infty} a \min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a(u-v)\bigr)\mu(du)\mu(dv) \ge 2\int_0^\infty R(t)\,dt. \tag{7.4}$$

Since

$$\begin{aligned}\int_0^1\int_0^1 & R_{\mathbf{X}}\bigl(a(u-v)\bigr)\lambda(du)\lambda(dv)\\ &= \frac{1}{a}\int_0^1\biggl[\int_0^{av} R(t)\,dt + \int_0^{a(1-v)} R(t)\,dt\biggr]dv,\end{aligned}$$

(7.3) immediately follows from (7.1) and the bounded convergence theorem. Therefore, it only remains to prove (7.4). Suppose that, to the contrary, (7.4) fails, and choose a sequence $a_n\to\infty$ such that

$$\lim_{n\to\infty} a_n \min_{\mu\in M_1^+([0,1])} \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\mu(du)\mu(dv) < 2\int_0^\infty R(t)\,dt.$$

For each n choose a symmetric $\mu_{a_n}\in\mathcal{W}_{a_n}$, so that

$$\lim_{n\to\infty} a_n \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\mu_{a_n}(du)\mu_{a_n}(dv) < 2\int_0^\infty R(t)\,dt. \tag{7.5}$$

We claim that, for every $\gamma>0$,

$$\lim_{n\to\infty}\mu_{a_n}([0,\gamma a_n^{-1}]) = 0. \tag{7.6}$$

Indeed, by the positivity of $R_{\mathbf{X}}$, for any $\gamma>0$,

$$\begin{aligned}\int_0^1\int_0^1 & R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\mu_{a_n}(du)\mu_{a_n}(dv)\\ &\ge \int_0^{\gamma a_n^{-1}}\int_0^{\gamma a_n^{-1}} R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\mu_{a_n}(du)\mu_{a_n}(dv)\\ &\ge \{\mu_{a_n}([0,\gamma a_n^{-1}])\}^2 \inf_{0\le t\le\gamma} R(t),\end{aligned}$$

so that (7.5) necessitates (7.6). Next, define a sequence of signed measures on $[0,1]$ by $\hat\mu_n = \mu_{a_n}-\lambda$. Note that

$$\hat\mu_n\bigl([0,1]\bigr) = 0 \qquad \text{for each } n. \tag{7.7}$$

By the nonnegative definiteness of $R_{\mathbf{X}}$,

$$\begin{aligned}&\int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\mu_{a_n}(du)\mu_{a_n}(dv)\\ &\quad = \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\lambda(du)\lambda(dv) + \int_0^1\int_0^1 R_{\mathbf{X}}\bigl(a_n(u-v)\bigr)\hat\mu_n(du)\hat\mu_n(dv)\end{aligned}$$

$$
\begin{aligned}
&+2\int_0^1\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\lambda(du)\hat{\mu}_n(dv)\\
\geq{}&\int_0^1\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\lambda(du)\lambda(dv)\\
&+2\int_0^1\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\lambda(du)\hat{\mu}_n(dv).
\end{aligned}
$$

We will show that

$$\lim_{n\to\infty} a_n \int_0^1\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\lambda(du)\hat{\mu}_n(dv)=0. \tag{7.8}$$

Together with (7.3) this will provide the necessary contradiction to (7.5). Let $\gamma>0$. Write the integral in (7.8) as

$$
\begin{aligned}
&\int_{\gamma a_n^{-1}}^{1-\gamma a_n^{-1}}\left[\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\,du\right]\hat{\mu}_n(dv)\\
&\quad+2\int_0^{\gamma a_n^{-1}}\left[\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\,du\right]\hat{\mu}_n(dv)\\
&\quad\triangleq J_n^{(1)}+2J_n^{(2)}.
\end{aligned}
$$

Observe that

$$
\begin{aligned}
\big|J_n^{(2)}\big| &= \frac{1}{a_n}\left|\int_0^{\gamma a_n^{-1}}\left[\int_0^{a_n v} R_{\mathbf{X}}(t)\,dt+\int_0^{a_n(1-v)} R_{\mathbf{X}}(t)\,dt\right]\hat{\mu}_n(dv)\right|\\
&\leq \frac{2\int_0^\infty R_{\mathbf{X}}(t)\,dt}{a_n}\|\hat{\mu}_n\|\big([0,\gamma a_n^{-1}]\big),
\end{aligned}
$$

so that by (7.6) we obtain

$$\lim_{n\to\infty} a_n J_n^{(2)}=0 \tag{7.9}$$

for every $\gamma>0$. Next, we write

$$
\begin{aligned}
J_n^{(1)} &= \frac{1}{a_n}\int_{\gamma a_n^{-1}}^{1-\gamma a_n^{-1}}\left[\int_0^{a_n v} R_{\mathbf{X}}(t)\,dt+\int_0^{a_n(1-v)} R_{\mathbf{X}}(t)\,dt\right]\hat{\mu}_n(dv)\\
&= \frac{2\int_0^\infty R_{\mathbf{X}}(t)\,dt}{a_n}\hat{\mu}_n\big([\gamma a_n^{-1},1-\gamma a_n^{-1}]\big)\\
&\quad-\frac{1}{a_n}\int_{\gamma a_n^{-1}}^{1-\gamma a_n^{-1}}\left[\int_{a_n v}^{\infty} R_{\mathbf{X}}(t)\,dt+\int_{a_n(1-v)}^{\infty} R_{\mathbf{X}}(t)\,dt\right]\hat{\mu}_n(dv)\\
&\triangleq J_n^{(11)}-J_n^{(12)}.
\end{aligned}
$$

It follows from (7.7) that

$$|a_n J_n^{(11)}| = 2\int_0^\infty R_{\mathbf{X}}(t)\,dt\,\big|\hat{\mu}_n([0,\gamma a_n^{-1}]) + \hat{\mu}_n([1-\gamma a_n^{-1},1])\big| \to 0 \tag{7.10}$$

as $n \to \infty$, by (7.6). Finally,

$$|a_n J_n^{(12)}| \le 4\int_\gamma^\infty R_{\mathbf{X}}(t)\,dt, \tag{7.11}$$

and we obtain by (7.9), (7.10) and (7.11) that

$$\limsup_{n\to\infty} a_n \left|\int_0^1\int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\lambda(du)\hat{\mu}_n(dv)\right| \le 4\int_\gamma^\infty R_{\mathbf{X}}(t)\,dt.$$

Letting $\gamma \to \infty$ proves (7.8) and, hence, completes the proof of the theorem. □

The next theorem is the counterpart of Theorem 7.1 for certain long memory stationary Gaussian processes. In this case, the uniform distribution on [0, 1] is no longer, asymptotically, optimal. We will assume that the covariance function of the process is regularly varying at infinity:

$$R_{\mathbf{X}}(t) = \frac{L(t)}{|t|^\beta}, \qquad 0 < \beta < 1, \tag{7.12}$$

where L is slowly varying at infinity. Before stating the theorem, we introduce new notation.

Consider the minimization problem

$$\min_{\mu \in M_1^+([0,1])} \int_0^1\int_0^1 \frac{\mu(du)\mu(dv)}{|u-v|^\beta}, \qquad 0 < \beta < 1. \tag{7.13}$$

This is a minimization problem of the same nature as in (5.2) with $a = 0, b = 1$, and the covariance function $R_{\mathbf{X}}$ replaced by the Riesz kernel $K_\beta(u,v) = |u-v|^{-\beta}$, $u, v \in [0,1]$. The general theory of energy of measures in [4] applies to the Riesz kernel. In particular, the minimum in (7.13) is well defined, is finite and positive. Let $\mathcal{W}_\beta$ be the set of measures in $M_1^+([0,1])$ of minimal energy with respect to the Riesz kernel. Note that the uniform measure $\lambda \notin \mathcal{W}_\beta$ since it does not satisfy the optimality conditions in Theorem 2.4 in [4].

THEOREM 7.2. *Let* $\mathbf{X}$ *be a continuous stationary Gaussian process. Assume that* $R_{\mathbf{X}}$ *is positive and satisfies assumption* (7.12) *of regular variation. Then for any* $\mu_\beta \in \mathcal{W}_\beta$,

$$\lim_{a\to\infty} R_{\mathbf{X}}(a)\mathcal{C}_{\mathbf{X}}(\mathbf{a}) = \left(\int_0^1\int_0^1 \frac{\mu_\beta(du)\mu_\beta(dv)}{|u-v|^\beta}\right)^{-1}. \tag{7.14}$$

PROOF. Suppose first that there is a sequence $a_n \uparrow \infty$ such that

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{R_{\mathbf{X}}(a_n)} \min_{\mu\in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a_n(u-v)\big)\mu(du)\mu(dv) \\ < \int_0^1 \int_0^1 \frac{\mu_\beta(du)\mu_\beta(dv)}{|u-v|^\beta}.\end{aligned} \tag{7.15}$$

For each n choose $\mu_n \in \mathcal{W}_{a_n}$, let $n_k \uparrow \infty$ be a subsequence such that $\mu_{n_k} \Rightarrow \hat{\mu}$ weakly as $k \to \infty$ for some $\hat{\mu} \in M_1^+([0,1])$. By Fatou's lemma and the regular variation of $R_{\mathbf{X}}$,

$$\begin{aligned}&\liminf_{k\to\infty} \frac{1}{R_{\mathbf{X}}(a_{n_k})} \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a_{n_k}(u-v)\big)\mu_{n_k}(du)\mu_{n_k}(dv) \\ &\qquad \ge \int_0^1 \int_0^1 \frac{\hat{\mu}(du)\hat{\mu}(dv)}{|u-v|^\beta} \ge \int_0^1 \int_0^1 \frac{\mu_\beta(du)\mu_\beta(dv)}{|u-v|^\beta},\end{aligned}$$

since μ_β has the smallest energy with respect to the Riesz kernel. This contradicts (7.15), thus proving that

$$\begin{aligned}\liminf_{a\to\infty} \frac{1}{R_{\mathbf{X}}(a)} \min_{\mu\in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u-v)\big)\mu(du)\mu(dv) \\ \ge \int_0^1 \int_0^1 \frac{\mu_\beta(du)\mu_\beta(dv)}{|u-v|^\beta}.\end{aligned} \tag{7.16}$$

In order to finish the proof, we need to establish a matching upper limit bound.

To this end, let $\theta > 0$ be a small number. We define a probability measure $\nu_\theta \in M_1^+([0,1])$ by convolving μ_β with the uniform distribution on $[0,\theta]$ and rescaling the resulting convolution back to the unit interval. More explicitly, if X and U are independent random variables, whose laws are μ_β and λ, respectively, then ν_θ is the law of $(X+\theta U)/(1+\theta)$. Note that

$$\nu_\theta \ll \lambda \qquad \text{and} \qquad \frac{d\nu_\theta}{d\lambda} \le \frac{1+\theta}{\theta} \qquad \text{a.e. on } [0,1]. \tag{7.17}$$

Given $0 < \varepsilon < 1-\beta$, by Potter's bounds (see, e.g., Proposition 0.8 in [8]), there is $t_0 > 0$ sufficiently large to ensure

$$\frac{R_{\mathbf{X}}(tx)}{R_{\mathbf{X}}(t)} > (1-\varepsilon)x^{-\beta-\varepsilon}$$

for all $t \ge t_0$ and $x \ge 1$. We have

$$\begin{aligned}&\min_{\mu\in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u-v)\big)\mu(du)\mu(dv) \\ &\qquad \le \int_0^1 \int_0^1 R_{\mathbf{X}}\big(a(u-v)\big)\nu_\theta(du)\nu_\theta(dv)\end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \int_0^1 \mathbf{1}\bigl(|u-v| \le t_0/a\bigr) R_{\mathbf{X}}\bigl(a(u-v)\bigr)\nu_\theta(du)\nu_\theta(dv) \\
&\quad + \int_0^1 \int_0^1 \mathbf{1}\bigl(|u-v| > t_0/a\bigr) R_{\mathbf{X}}\bigl(a(u-v)\bigr)\nu_\theta(du)\nu_\theta(dv) \\
&\stackrel{\Delta}{=} I_1(a) + I_2(a).
\end{aligned}
$$

By the definition of t_0,

$$
\mathbf{1}\bigl(|u-v| > t_0/a\bigr)\frac{R_{\mathbf{X}}(a(u-v))}{R_{\mathbf{X}}(a)} \le \frac{1}{1+\varepsilon}|u-v|^{-(\beta+\varepsilon)},
$$

so that by the dominated convergence theorem we have

$$
\lim_{a\to\infty} \frac{1}{R_{\mathbf{X}}(a)} I_2(a) = \int_0^1 \int_0^1 \frac{\nu_\theta(du)\nu_\theta(dv)}{|u-v|^\beta}.
$$

On the other hand, by (7.17),

$$
I_1(a) \le R_{\mathbf{X}}(0)\frac{2t_0}{a}\frac{1+\theta}{\theta} = o\bigl(R_{\mathbf{X}}(a)\bigr)
$$

as $a \to \infty$. We conclude that

$$
\begin{aligned}
&\limsup_{a\to\infty} \frac{1}{R_{\mathbf{X}}(a)} \min_{\mu\in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}\bigl(a(u-v)\bigr)\mu(du)\mu(dv) \\
&\quad \le \int_0^1 \int_0^1 \frac{\nu_\theta(du)\nu_\theta(dv)}{|u-v|^\beta}.
\end{aligned}
$$

Once we show that

$$
\lim_{\theta\to 0} \int_0^1 \int_0^1 \frac{\nu_\theta(du)\nu_\theta(dv)}{|u-v|^\beta} = \int_0^1 \int_0^1 \frac{\mu_\beta(du)\mu_\beta(dv)}{|u-v|^\beta}, \tag{7.18}
$$

we will have established an upper bound matching (7.16). This will complete the proof of the theorem. Recall that (7.18) is equivalent to

$$
\lim_{\theta\to 0} E\bigl|X_1 - X_2 + \theta(U_1 - U_2)\bigr|^{-\beta} = E|X_1 - X_2|^{-\beta},
$$

where X_1, X_2, U_1, U_2 are independent random variables, X_1 and X_2 with the law μ_β, while U_1 and U_2 are uniformly distributed on $[0, 1]$. This, however, follows by the dominated convergence theorem and the following fact, that can be checked by elementary calculations: there is $r_\beta \in (0, \infty)$ such that for any $0 < b < 1$ and $0 < \theta < 1$,

$$
E\bigl|b + \theta(U_1 - U_2)\bigr|^{-\beta} \le r_\beta b^{-\beta}. \qquad \square
$$

REMARK 7.3. It follows from Proposition A.3 in [5] that the energy of the measure μ_β with respect to the Riesz kernel cannot be smaller than one half of the energy of the uniform measure. Hence,

$$
\lim_{a\to\infty} R_{\mathbf{X}}(a)\mathcal{C}_{\mathbf{X}}(\mathbf{a}) \in \bigl((1-\beta)(2-\beta)/2, (1-\beta)(2-\beta)\bigr).
$$

8. The multidimensional case. Our understanding of the one-dimensional case described in the previous three sections, while incomplete, is nevertheless quite significant. In contrast, there is much less we can say about the multivariate problem of Section 3. The problem lies, in part, in the nonconvexity of the feasible set in (3.1) which leads, in turn, to the "max-min" problem in Theorem 3.2.

The following proposition is a multivariate version of Proposition 5.6. Note that stationarity of the random field is not required.

PROPOSITION 8.1. *Let $\mathbf{X} = (X(\mathbf{t}), \mathbf{t} \in T)$ be a continuous Gaussian random field on a compact set $T \subset \mathbb{R}^d$, and suppose that $\mathbf{a}, \mathbf{b}$ are in T. Suppose that there is a path ξ_0 in T connecting $\mathbf{a}$ and $\mathbf{b}$ such that*

$$R_{\mathbf{X}}(\mathbf{a}, \xi_0(u)) + R_{\mathbf{X}}(\xi_0(u), \mathbf{b}) \geq \frac{R_{\mathbf{X}}(\mathbf{a}, \mathbf{a}) + 2R_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) + R_{\mathbf{X}}(\mathbf{b}, \mathbf{b})}{2} > 0 \tag{8.1}$$

for all $0 \leq u \leq 1$. Then the supremum in (3.8) is achieved on the path ξ_0 and

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) = \frac{4}{R_{\mathbf{X}}(\mathbf{a}, \mathbf{a}) + 2R_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) + R_{\mathbf{X}}(\mathbf{b}, \mathbf{b})}. \tag{8.2}$$

REMARK 8.2. Using $u = 0$ and $u = 1$ in (8.1) shows that conditions of Proposition 8.1 cannot be satisfied unless $R_{\mathbf{X}}(\mathbf{a}, \mathbf{a}) = R_{\mathbf{X}}(\mathbf{b}, \mathbf{b})$. Correspondingly, we can restate (8.2) as

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) = \frac{2}{R_{\mathbf{X}}(\mathbf{a}, \mathbf{a}) + R_{\mathbf{X}}(\mathbf{a}, \mathbf{b})}.$$

Recall (5.11), which shows that this implies the logarithmic equivalence of the probabilities of $\mathbf{X}$ being above the level u along a curve or at its endpoints.

PROOF OF PROPOSITION 8.1. Consider the fixed path ξ_0. The assumption (8.1) shows that the measure

$$\mu_0 = \tfrac{1}{2}\delta_a + \tfrac{1}{2}\delta_b$$

satisfies conditions (4.9) and, hence, is in $\mathcal{W}_{\xi_0}$ by Theorem 4.3. Therefore,

$$\begin{aligned}
\min_{\mu \in M_1^+([0,1])} & \int_0^1 \int_0^1 R_{\mathbf{X}}(\xi_0(u), \xi_0(v)) \mu(du)\mu(dv) \\
&= \int_0^1 \int_0^1 R_{\mathbf{X}}(\xi_0(u), \xi_0(v)) \mu_0(du)\mu_0(dv) \\
&= \frac{R_{\mathbf{X}}(\mathbf{a}, \mathbf{a}) + 2R_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) + R_{\mathbf{X}}(\mathbf{b}, \mathbf{b})}{4}.
\end{aligned}$$

On the other hand, for any other path in T connecting $\mathbf{a}$ and $\mathbf{b}$,

$$\begin{aligned}\min_{\mu\in M_1^+([0,1])} &\int_0^1\int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu(du)\mu(dv)\\ &\le \int_0^1\int_0^1 R_{\mathbf{X}}\big(\xi(u),\xi(v)\big)\mu_0(du)\mu_0(dv)\\ &= \frac{R_{\mathbf{X}}(\mathbf{a},\mathbf{a})+2R_{\mathbf{X}}(\mathbf{a},\mathbf{b})+R_{\mathbf{X}}(\mathbf{b},\mathbf{b})}{4}.\end{aligned}$$

Therefore, the supremum in (3.8) is achieved on the path ξ_0, and (8.2) follows by Theorem 3.2. □

Even for the most common Gaussian random fields, the assumptions of Proposition 8.1 may be satisfied on some path but not on the straight line connecting the two points. In that case, the straight line, clearly, fails to be optimal.

EXAMPLE 8.3. Consider a Brownian sheet in $d \ge 2$ dimensions. This is the continuous centered Gaussian random field $\mathbf{X}$ on $[0,\infty)^d$ with covariance function

$$R_{\mathbf{X}}(\mathbf{s},\mathbf{t}) = \prod_{j=1}^d \min(s_j,t_j), \qquad \mathbf{s},\mathbf{t}\in[0,\infty)^d.$$

We restrict the random field to the hypercube $T=[0,d]^d$, and let

$$\mathbf{a}=(1,2,\dots,d-1,d), \qquad \mathbf{b}=(d,1,2,\dots,d-1).$$

It is elementary to check that the path

$$\xi_0(u) = \begin{cases} \big(1+d(d-1)u,2,\dots,d-1,d\big), & \text{for } 0\le u\le \dfrac{1}{d},\\ (d,1,\dots,j-2,2j-1-du,j+1,\dots,d), & \text{for } \dfrac{j-1}{d}\le u\le\dfrac{j}{d},\\ & j=2,\dots,d,\end{cases}$$

satisfies (8.1) and, hence, the supremum in (3.8) is achieved on that path. Therefore, by Proposition 8.1,

$$\mathcal{C}_{\mathbf{X}}(\mathbf{a},\mathbf{b}) = \frac{2}{d!+(d-1)!}.$$

On the other hand, if we consider the straight line connecting the points $\mathbf{a}$ and $\mathbf{b}$,

$$\xi(u) = \big(1+(d-1)u, 2-u, 3-u,\dots,d-u\big), \qquad 0\le u\le 1,$$

then the sum in the right-hand side of (8.1) becomes

$$L(u) = \prod_{j=2}^d (j-u) + \big(1+(d-1)u\big)(d-1)!, \qquad 0\le u\le 1.$$

The function L achieves the value $d! + (d-1)!$ at the endpoints $u = 0$ and $u = 1$, and is strictly convex if $d \geq 3$. Therefore, it takes values strictly smaller than $d! + (d-1)!$ over $0 < u < 1$. That is, (8.1) fails, and the straight line is not optimal. If $d = 2$, however, then L is a constant function, condition (8.1) holds over the straight line path, and the straight line is optimal.

We also note that, if $d = 1$, then the Brownian sheet becomes the Brownian motion in one dimension. In that case it is, clearly, impossible to find two positive points $a < b$ in which the process has the same variance, so Proposition 8.1 does not apply. In this case, however, we are in the situation of Theorem 5.1, so if $0 < a < b < \infty$, then the measure $\mu = \delta_a$ satisfies (5.4) and, hence, is optimal.

The above example notwithstanding, under certain assumptions on the random field, the straight line path between two points turns out to be optimal for the optimization problem (3.8). The next result describes one such situation.

Recall that a random field on $\mathbb{R}^d$ is *isotropic* if its law is invariant under rigid motions of the parameter space. A centered Gaussian random field $\mathbf{X}$ is isotropic if and only if its covariance function is a function of the Euclidian distance between two points. With the usual abuse of notation we will write $R_{\mathbf{X}}(\mathbf{a}, \mathbf{b}) = R_{\mathbf{X}}(\|\mathbf{b} - \mathbf{a}\|)$, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$.

PROPOSITION 8.4. *Let $\mathbf{X}$ be a continuous centered isotropic Gaussian random field, such that the covariance function $R_{\mathbf{X}}$ is nonincreasing. Then for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$, the straight path connecting the points $\mathbf{a}$ and $\mathbf{b}$ is optimal for the optimization problem* (3.8).

PROOF. We may and will assume, without loss of generality, that $\mathbf{a} = (a, 0, \ldots, 0)$ and $\mathbf{b} = \mathbf{0}$ for some $a > 0$. We start with showing that the supremum over $\xi \in \mathcal{P}(\mathbf{0}, \mathbf{a})$ is achieved over paths in

$$\mathcal{P}_l = \{\xi : [0, 1] \to \{(x, 0, \ldots, 0) : x \geq 0\}, \text{continuous}, \xi(0) = \mathbf{0}, \xi(1) = \mathbf{a}\}.$$

To this end, it is enough to show that for each $\xi \in \mathcal{P}(\mathbf{0}, \mathbf{a})$ there is $\hat{\xi} \in \mathcal{P}_l$ such that

$$\begin{aligned} &\min_{\mu \in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}(\|\xi(u) - \xi(v)\|)\mu(du)\mu(dv) \\ &\quad \leq \min_{\mu \in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}(\|\hat{\xi}(u) - \hat{\xi}(v)\|)\mu(du)\mu(dv). \end{aligned} \tag{8.3}$$

To see this, define for $\xi \in \mathcal{P}(\mathbf{0}, \mathbf{a})$

$$\hat{\xi}(u) = (\|\xi(u)\|, 0, \ldots, 0), \qquad 0 \leq u \leq 1.$$

Clearly, $\hat{\xi} \in \mathcal{P}_l$, and (8.3) follows by the monotonicity of $R_{\mathbf{X}}$ and the triangle inequality

$$\big|\|\xi(u)\| - \|\xi(v)\|\big| \leq \|\xi(u) - \xi(v)\|.$$

Next, any $\xi \in \mathcal{P}_l$ is of the form

$$\xi(u) = (\varphi(u), 0, \dots, 0), \qquad 0 \le u \le 1, \tag{8.4}$$

with $\varphi : [0, 1] \to [0, \infty)$ a continuous function, satisfying $\varphi(0) = 0$, $\varphi(1) = a$. Defining $\hat{\varphi}(u) = \min(\varphi(u), a)$, $0 \le u \le 1$, and

$$\hat{\xi}(u) = (\hat{\varphi}(u), 0, \dots, 0), \qquad 0 \le u \le 1,$$

we see that the supremum over paths in $\mathcal{P}_l$ is, actually, achieved over paths whose image is exactly the interval $[0, a]$. Finally, for any path $\xi \in \mathcal{P}_l$ of the latter type, given in the form (8.4), define

$$r(v) = \inf\{u \in [0, 1] : \varphi(u) = av\}, \qquad 0 \le v \le 1.$$

Then r is a measurable map from $[0, 1]$ to itself, so for any $\mu \in M_1^+([0, 1])$, we can define $\mu_1 \in M_1^+([0, 1])$ by $\mu_1 = \mu \circ r^{-1}$. Then

$$\begin{aligned}
&\int_0^1 \int_0^1 R_{\mathbf{X}}(|\varphi(u) - \varphi(v)|)\mu_1(du)\mu_1(dv) \\
&\quad = \int_0^1 \int_0^1 R_{\mathbf{X}}(|\varphi(r(u)) - \varphi(r(v))|)\mu(du)\mu(dv) \\
&\quad = \int_0^1 \int_0^1 R_{\mathbf{X}}(a|u - v|)\mu(du)\mu(dv).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\min_{\mu \in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}(\|\xi(u) - \xi(v)\|)\mu(du)\mu(dv) \\
&\quad \le \min_{\mu \in M_1^+([0,1])} \int_0^1 \int_0^1 R_{\mathbf{X}}(a|u - v|)\mu(du)\mu(dv),
\end{aligned}$$

and the statement of the proposition follows. □

REFERENCES

[1] ADLER, R. J. and TAYLOR, J. E. (2007). *Random Fields and Geometry*. Springer, New York. MR2319516

[2] AZAÏS, J.-M. and WSCHEBOR, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Wiley, Hoboken, NJ. MR2478201

[3] DEUSCHEL, J.-D. and STROOCK, D. W. (1989). *Large Deviations*. *Pure and Applied Mathematics* **137**. Academic Press, Boston, MA. MR0997938

[4] FUGLEDE, B. (1960). On the theory of potentials in locally compact spaces. *Acta Math.* **103** 139–215. MR0117453

[5] KHOSHNEVISAN, D., XIAO, Y. and ZHONG, Y. (2003). Measuring the range of an additive Lévy process. *Ann. Probab.* **31** 1097–1141. MR1964960

[6] LIFSHITS, M. (2012). *Lectures on Gaussian Processes*. Springer, Heidelberg. MR3024389

[7] LUENBERGER, D. G. (1969). *Optimization by Vector Space Methods*. Wiley, New York. MR0238472

[8] RESNICK, S. I. (1987). *Extreme Values*, *Regular Variation*, *and Point Processes*. *Applied Probability. A Series of the Applied Probability Trust* **4**. Springer, New York. MR0900810

[9] VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2008). Reproducing kernel Hilbert spaces of Gaussian priors. In *Pushing the Limits of Contemporary Statistics*: *Contributions in Honor of Jayanta K*. *Ghosh*. *Inst*. *Math*. *Stat*. *Collect*. **3** 200–222. IMS, Beachwood, OH. MR2459226

R. J. ADLER
E. MOLDAVSKAYA
ELECTRICAL ENGINEERING
TECHNION, HAIFA
ISRAEL 32000
E-MAIL: robert@ee.technion.ac.il
elinamoldavskaya@gmail.com
URL: webee.technion.ac.il/people/adler

G. SAMORODNITSKY
ORIE
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
USA
E-MAIL: gs18@cornell.edu
URL: www.orie.cornell.edu/~gennady/