A PHASE TRANSITION IN ZERO COUNT PROBABILITY FOR STATIONARY GAUSSIAN PROCESSES

NAOMI DVORA FELDHEIM, OHAD NOY FELDHEIM, AND LAKSHMI PRIYA M.E.

ABSTRACT. We study the probability that a real stationary Gaussian process has at least ηT zeros in $[0, T]$ (overcrowding), or at most this number (undercrowding). We show that if the spectral measure of the process is supported on $\pm[B, A]$, overcrowding probability transitions from exponential decay to Gaussian decay at $\eta = \frac{A}{\pi}$, while undercrowding probability undergoes the reverse transition at $\eta = \frac{B}{\pi}$.

1. INTRODUCTION

Let $F : \mathbb{R} \to \mathbb{R}$ be a continuous centered stationary Gaussian process (SGP), that is, a shift-invariant random function whose finite marginal distributions are multi-normal with zero mean. Such a process is characterized by its *covariance kernel*, $k(t) = \mathbb{E}[F(0)F(t)]$, or, equivalently, by its *spectral measure* μ, which is the finite, non-negative, symmetric measure obtained as the inverse Fourier transform of k, so that,

$$k(t) = \mathcal{F}[\mu](t) = \int_{\mathbb{R}} e^{-i\lambda t} d\mu(\lambda).$$

Throughout we assume $k(0) = 1$, so that μ is normalized to be a probability measure. The counting measure for the zero point process associated with F is generated by

$$N_F(T) = N(T) = \#\{t \in [0, T] : f(t) = 0\}.$$

The zeros of stationary Gaussian functions form a family of well-studied stochastic point processes with numerous applications [1, 7, 12, 24]. The expectation of $N_F(T)$ is given by the Kac-Rice formula [23, 33],

$$\mathbb{E}[N_F(T)] = \frac{\gamma}{\pi} T, \quad \text{where} \quad \gamma^2 = -k''(0) = \int_{\mathbb{R}} \lambda^2 d\mu(\lambda).$$

Variance and higher moments of $N(T)$ have also been studied [4, 37], along with results about clustering and limit theorems [2, 25]. Nevertheless, rare events involving such processes are generally not well-understood (see discussion in [8] and [38, Q. 7]).

Here we study linear deviations of $N_F(T)$ as T tends to infinity. These consist of η-*overcrowding* events, given by $\{N_F(T) \geq \eta T\}$ for $\eta > \mathbb{E}[N_F(1)]$, and η-*undercrowding* events, given by $\{N_F(T) \leq \eta T\}$ for $\eta < \mathbb{E}[N_F(1)]$.

Basu, Dembo, Zeitouni and the first author showed in [8] that, for well-behaved SGPs and any $\eta > \mathbb{E}[N_F(1)]$, the probability of η-overcrowding decays at most exponentially with T. In contrast, the third author showed in [32] that when the spectral measure is compactly supported and η is sufficiently large, the η-overcrowding probability decays in a Gaussian fashion or faster.

Here we establish a sharp phase transition between these two decay profiles, occurring at $\eta = \frac{A}{\pi}$ where A is the supremum of the support of the spectrum.

Theorem 1. Suppose that μ is compactly supported with $\mu_{ac} \not\equiv 0$. Let A be the smallest positive number such that $\mathrm{sprt}(\mu) \subseteq [-A, A]$. Then:

(1) There exist $c, d_1, d_2 > 0$ such that, for every $T \geq e^8$ and $\varepsilon \in \left[d_1 \frac{\log T}{\sqrt{T}}, d_2 \right]$, we have

$$\mathbb{P} \left(N_F(T) \geq \frac{A}{\pi} T + \varepsilon T \right) \leq \exp \left(-c(\varepsilon / \log \varepsilon)^4 T^2 \right).$$

(2) For every $\varepsilon > 0$ there exists $c > 0$ such that, for all $T \geq 1$, we have

$$\mathbb{P} \left(N_F(T) \geq \frac{A}{\pi} T - \varepsilon T \right) \geq \exp(-cT).$$

An analogous phase transition is established for undercrowding, in case μ has a spectral gap.

Theorem 2. Suppose that μ is compactly supported with $\mu_{ac} \not\equiv 0$. Let B be the largest and A the smallest non-negative numbers such that $\mathrm{sprt}(\mu) \subseteq [-A, -B] \cup [B, A]$. Then:

(1) There exist $c, d_1, d_2 > 0$ such that, for every $T \geq e^8$ and $\varepsilon \in \left[d_1 \frac{\log T}{\sqrt{T}}, d_2 \right]$, we have

$$\mathbb{P} \left(N_F(T) \leq \frac{B}{\pi} T - \varepsilon T \right) \leq \exp(-c(\varepsilon / \log \varepsilon)^4 T^2).$$

(2) For every $\varepsilon > 0$, there exists $c > 0$ such that, for all $T \geq 1$, we have

$$\mathbb{P} \left(N_F(T) \leq \frac{B}{\pi} T + \varepsilon T \right) \geq \exp(-cT).$$

The remainder of the paper is structured as follows. In Sections 1.1 and 1.2 we describe the main steps of the proofs of the lower and upper bounds, respectively. In Section 1.3 we discuss tightness of the results. In Section 1.4 we discuss their role in the broader context of the study of zeros of SGPs. In Section 1.5 we mention deterministic analogues which inspired the proof. Section 1.6 is dedicated to open problems and future research. Section 2 is a preliminary section containing useful results about entire functions and Gaussian processes. In Section 3 we prove a result (Proposition 1.1 below) which implies the lower bounds in both theorems. Section 4 contains the proof of the upper bound of Theorem 1, and Section 5 contains a reduction of the upper bound of Theorem 2 to that of Theorem 1.

1.1. **Main steps of the proof: lower bound.** The lower bound in both theorems readily follows from the following proposition, which is concerned with obtaining tight control over $\frac{N(T)}{T}$ and may be of independent interest.

Proposition 1.1. Suppose that $\int \lambda^2 \log^2 \lambda \, d\mu(\lambda) < \infty$ and $\mu([X - \varepsilon, X + \varepsilon]) > 0$ for some $X \in (0, \infty)$ and $\varepsilon > 0$. Then there exists $c > 0$ such that

$$\mathbb{P} \left(\frac{N(T)}{T} \in \left[\frac{X - 2\varepsilon}{\pi}, \frac{X + 2\varepsilon}{\pi} \right] \right) \geq \exp(-cT), \quad \text{for sufficiently large } T. \tag{1}$$

Moreover, if $\limsup_{\delta \to 0} \frac{\mu([X - \delta, X + \delta])}{2\delta} = \infty$, then (1) holds for any choice of $c > 0$.

This is obtained by estimating the probability that F closely imitates a pure wave. The lower bounds in Theorems 1 and 2 are then obtained by the choice of $X = A$ and $X = B$ respectively.

1.2. **Main steps of the proof: upper bound.** The proof of the upper bound given in Theorem 1 uses classical tools from complex analysis. In fact, rather than providing a direct upper bound for the number of real zeros F has in $[0, T]$, we shall provide a bound for the number of complex zeros of the analytic extension of F in a neighbourhood of this interval. A similar idea was used in [8].

An entire function $f : \mathbb{C} \to \mathbb{C}$ is of *exponential type*, if there exist $\sigma, C > 0$ such that

$$|f(z)| \leq C \exp(\sigma|z|), \quad \forall z \in \mathbb{C}. \tag{2}$$

The *exponential type of f* is then defined as the infimum over all σ for which this inequality is satisfied for some value of C. An entire function $f : \mathbb{C} \to \mathbb{C}$ is said to belong to the *Cartwright's class $\mathcal{C}$* if the following two conditions hold.

(1) f is of exponential type σ for some $\sigma \in [0, \infty)$.
(2) f satisfies the integral condition:

$$\int_{\mathbb{R}} \frac{\log^+ |f(t)|}{1 + t^2} dt < \infty. \tag{3}$$

Zeros of functions from Cartwright's class have special properties, see [29, §16.1] and [18] for more details.

The first step in our proof is to show that F is almost surely in Cartwright's class, and of exponential type at most A (Proposition 4.1 below). This guarantees a tight upper bound on $\log |F|$ (Corollary 2.3). The celebrated Jensen's formula (Theorem 2.1 below) is then used in order to relate weighted zero counts of F to evaluation and integration of $\log |F|$ around a point. To establish the upper bound we further average this formula over a carefully selected set of points, so that the total weight given to each real zero in $[0, T]$ will be roughly the same. The challenge then lies in providing probabilistic bounds on $\log |F|$ and on the event of existence of suitable points.

The upper bound of Theorem 2 is obtained by applying a non-trivial coupling between F and another SGP G (Claim 5.1 below), under which undercrowding for F corresponds to over-crowding for G. Applying Theorem 1 to G completes the argument. This useful coupling was suggested by Eremenko and Novikov [17].

1.3. **Tightness of the bounds.** The one-sided bounds of Theorems 1 and 2 often capture the exact probability decay above and below the transition point, up to a constant in the exponent. Here we survey results concerning the conditions for this to hold.

Exponential upper bounds. Results of Basu, Dembo, Zeitouni and the second author in [8] imply that for SGPs with compactly supported spectral density and absolutely integrable covariance kernel, any (upper or lower) linear deviation of $N_F(T)$ from its mean indeed decays exponentially or faster[1]. We conjecture that this property should hold whenever the spectral density is bounded, and has some finite exponential moment (see Conjectures 3 and 4 below).

Gaussian lower bounds. In [19], Jay, Nazarov and Nitzan together with the first two authors, showed that the probability of persistence, namely 0-undercrowding, is of probability at least $\exp(-CT^2)$ if the spectral measure is compactly supported, having a non-trivial absolutely continuous component. By monotonicity, this implies a Gaussian lower bound on η-undercrowding probability for all η. The third author [32] showed an analogous Gaussian lower

---

[1]in fact, in [8] this is shown for a slightly larger family of SGPs.

bound on η-overcrowding, for spectral measures with finite moments and non-trivial absolutely continuous component. Hence the first part of both Theorems 1 and 2 is tight up to a constant in the exponent.

1.4. **Background and related literature.** For many years there has been little progress in studying large deviations of the zero count. It has been anticipated that, similar to many classical models for random sums, the probability of such a large deviation should be asymptotically exponential in T. Until recently, such exponential concentration was not established even for a single non-trivial example of SGPs (as mentioned in lecture notes by Tsirelson [40] and Sodin [38]); even though such results were obtained for related models, including complex zeros of the planar and hyperbolic Gaussian analytic functions [27, 39] and nodal component count of Gaussian Laplace eigenfunctions [30, 31, 35]. Some exponential bounds on deviations were also obtained for zeros of high-degree Kostlan polynomials and other Gaussian models on manifolds [22], as well as counts of nodal components and other topological events in high dimensional smooth Gaussian fields with fast covariance decay [9].

The extremely fast decay of η-overcrowding for sufficiently large values of η, has been pointed out in [32]. There, larger than linear overcrowding has also been considered, and in the case of compactly supported spectral measures with absolutely continuous component, it was shown that $\log \mathbb{P}(N_F(T) \geq n) \asymp -n^2 \log(n/T)$ for $n \geq CT$, with some constant $C > 0$.

The particular case of 0-undercrowding, also known as the *persistence event*, gained substantial attention by both physicists [10] and mathematicians [6]. In 1962 Slepian introduced his famous comparison inequality in order to bound these events [36], but this was not enough to determine the rate of decay for many examples of interest (see e.g. [15]). It was only in the last decade that major advances were made in understanding persistence and estimating its probability (e.g. [5, 13, 14, 26]). Using new spectral techniques, the decay rate of persistence probabilities was shown to depend mainly on the behaviour of the spectral measure near the origin [21], with asymptotically exponential behaviour corresponding to existence of a spectral density at the origin [20]. The methods developed in these works play a major role in the proof of Proposition 1.1.

1.5. **Deterministic analogues.** The number of zeros of Cartwright's class functions in the complex plane inside a ball around the origin, is asymptotically restricted by their exponential growth. This is shown by the following theorem attributed to Levinson and Cartwright, which is analogous to Theorem 1.

Theorem 1.2 ([29, Ch. 17, Thm. 1]). *Let $f \in \mathcal{C}$ be of exponential type $A > 0$. Then we have the following asymptotic zero density:*

$$\lim_{R \to \infty} \frac{n_f^+(R)}{R} = \lim_{R \to \infty} \frac{n_f^-(R)}{R} = \frac{A}{\pi},$$

where $n_f^{\pm}(R)$ denotes the zero count of f in $\{z \in \mathbb{C} : |z| < R \text{ and } Re(z) \gtrless 0\}$.

This has been used by Eremenko and Novikov to study the asymptotic zero density of functions with a spectral gap at the origin, a result analogous to Theorem 2.

Theorem 1.3 ([17, Prop. 1]). *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function whose spectrum is supported on $\pm[B, A]$, for some $A, B > 0$, for which $\int_{\mathbb{R}} \frac{\log^+ |f(t)|}{1+t^2} dt < \infty$. Then*

$$\liminf_{R \to \infty} \frac{N_f(R)}{R} \geq \frac{B}{\pi}.$$

N.b. that for Theorem 1.3 to hold in such generality, one needs to generalise the notion of a Fourier transform, see [17] for more details.

For an SGP F whose spectral measure is supported on $\pm[B, A]$, these results tell us that, asymptotically, $\frac{N_F(T)}{T} \in [\frac{B}{\pi}, \frac{A}{\pi}]$. The purpose of this paper is to provide a probabilistic estimate for the rate of convergence. While we do not use Theorems 1.2 and 1.3 themselves, ideas from the proofs of these results will be used in proving Theorems 1 and 2 respectively.

1.6. **Future Work.** Here we survey several directions for future work concerning overcrowding and undercrowding events.

Transition window. While the results of Theorems 1 and 2 demonstrate the existence of a phase transition in the overcrowding and undercrowding probabilities, they leave to be desired in understanding the fine nature of this transition and recovering the transition window. Theorem 1 implies that for $\beta \in (\frac{1}{2}, 1]$ we have

$$\mathbb{P}\left(N(T) \geq \frac{A}{\pi}T + T^\beta\right) \leq \exp(-T^{4\beta-2-\varepsilon})$$

for any $\varepsilon > 0$, while Theorem 2 implies an analogous bound for undercrowding. For $\beta \in [\frac{1}{2}, \frac{3}{4}]$ this bound is worse than the bound obtained in [8], and is unlikely to be tight. Our first question is concerned with improving this bound and finding a matching lower bound.

Problem 1. Obtain tight upper and lower bounds for $\mathbb{P}\left(N(T) \geq \frac{A}{\pi}T + T^\beta\right)$ for $\beta \in [0, 1]$.

Undercrowding for non-compactly supported spectrum. It seems natural that sub-Gaussian undercrowding probability decay should occur in all processes demonstrating a spectral gap, and that the condition of compact spectral support is not needed. This restriction in Theorem 2 stems only from our use of Theorem 1 in obtaining it.

Problem 2. Remove the condition that μ is compactly supported in Theorem 2.

General exponential upper bounds. Proposition 1.1 indicates that the exponential upper bounds of [8] do not hold for spectral measures with unbounded spectral density, or a singular spectral component, away from $\pm\pi\mathbb{E}[N_F(1)]$. Exponential upper bounds also require the measure to have some exponential moment, as otherwise $N_F(T)$ cannot decay exponentially. We conjecture that these are the true obstructions for such a bound.

Conjecture 3. If μ has bounded density, and some finite exponential moment, then for any $\alpha > 0$ there exists $c > 0$ such that

$$\mathbb{P}(|N_F(T) - \mathbb{E}N_F(T)| \geq \alpha T) \leq e^{-cT}.$$

In fact, we conjecture the following stronger statement.

Conjecture 4. If $a > \pi\mathbb{E}[N_F(1)]$ is such that $\mu|_{[a,\infty)}$ has bounded density and some finite exponential moment, then there exists $c > 0$ such that $\mathbb{P}\left(N_F(T) > \frac{a}{\pi}T\right) \leq e^{-cT}$.

A similar property should hold for undercrowding.

Exponents and conditional behaviour. Lastly, we state two additional problems which aim to establish a large-deviations theory for overcrowding and undercrowding events.

Problem 5. Describe conditions under which the limit $\lim\limits_{T\to\infty} \frac{1}{T}\log \mathbb{P}\left(N_F(T) \geq \eta T\right)$ exists.

Problem 6. Describe the process F conditioned on the event $\{N_F(T) \geq \eta T\}$. In particular, describe the empirical power spectrum $\lim\limits_{T\to\infty} \frac{1}{T}\left|\mathcal{F}^{-1}\left[F\mathbb{1}_{[0,T]}\right]\right|^2$, where F is the conditioned process.

Analogous problems could be stated concerning undercrowding events. A possible approach to these problems may be to refine known large deviation principles for empirical measures of SGPs, established by Donsker-Varadhan [16] and Bryc-Dembo [11], whose current conditions are too restrictive to be applied directly.

2. Preliminaries

2.1. **Theory of entire functions.** We require a few classical results about entire functions. For $z \in \mathbb{C}$ and $r > 0$, denote by $B(z,r) := \{w \in \mathbb{C} : \ |w - z| < r\} \subseteq \mathbb{C}$ the disc of radius r around z, and by $n_f(z,r)$ the number of zeros, counted with multiplicity, of an entire function $f : \mathbb{C} \to \mathbb{C}$ in that disc. The following is the famous zero-counting formula by Jensen.

Theorem 2.1 (Jensen [29, §2.3]). *Let $f : \mathbb{C} \to \mathbb{C}$ be an entire function such that $f(0) \neq 0$. Suppose that f has no zeros on $\partial B(0,R)$. Then*

$$\int_0^R \frac{n_f(0,t)}{t}dt = \frac{1}{2\pi}\int_0^{2\pi} \log|f(Re^{i\theta})|d\theta - \log|f(0)|.$$

We shall also require the Phragmén–Lindelöf theorem. Let $\mathbb{U} := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ denote the upper half plane, and $\overline{\mathbb{U}}$ denote its closure. The notion of functions of exponential type σ naturally extends to holomorphic functions on $\mathbb{U}$.

Theorem 2.2 (Phragmén–Lindelöf [29, §6.1, Theorem 3]). *Let $f : \overline{\mathbb{U}} \to \mathbb{C}$ be a continuous function, holomorphic on $\mathbb{U}$, of exponential type σ, which satisfies $|f| \leq M$ on $\mathbb{R}$. Then for every $z = x + iy \in \overline{\mathbb{U}}$,*

$$|f(x+iy)| \leq Me^{\sigma y}.$$

From this we draw the following corollary, concerning functions in the Cartwright's class, which rather than being bounded, satisfies the slow growth assumption (3). This result is obtained in the course of the proof of [29, §16.1, Theorem 1] and brought here for completeness.

Corollary 2.3. *Let f be a Cartwright's class function which is of exponential type $\sigma > 0$. Then for every $z = x + iy \in \mathbb{C} \setminus \mathbb{R}$, we have*

$$\log|f(z)| \leq \frac{|y|}{\pi}\int_{\mathbb{R}} \frac{\log^+|f(t)|}{|t-z|^2}dt + \sigma|y|.$$

Proof. Assume, without loss of generality that $z \in \mathbb{U}$. Denote

$$g(z) = \frac{y}{\pi}\int_{\mathbb{R}} \frac{\log^+|f(t)|}{|t-z|^2}dt,$$

which converges by (3). This is a harmonic function on $\mathbb{U}$ with boundary values on $\mathbb{R}$ coinciding with $\log^+ |f|$. Denote the harmonic conjugate of g by $\tilde{g}$ and write

$$F(z) = f(z)e^{-g(z)-i\tilde{g}(z)}.$$

Since $g \geq 0$, we have $|F(z)| \leq |f(z)| \exp(-g(z)) \leq |f(z)|$, so that F is analytic of exponential type at most σ on $\mathbb{U}$. Moreover $|F(t)| \leq 1$ for all $t \in \mathbb{R}$, so that we may apply Theorem 2.2 to obtain $|F(x+iy)| \leq \exp(\sigma y)$. Taking logarithm on both sides, the corollary follows. $\qquad\square$

2.2. **Tools from the theory of SGPs.** Throughout this section, let F be a fixed centred SGP with spectral measure μ, the moments of which are all finite. It follows that F is smooth and its higher derivatives $F^{(n)}$ are centred SGPs with spectral measure μ_n given by $d\mu_n(x) = x^{2n}d\mu(x)$, and covariance kernel given by

$$\mathbb{E}[F^{(n)}(x) \cdot F^{(n)}(0)] = (-1)^n \, k^{(2n)}(x), \text{ for } x \in \mathbb{R}, \tag{4}$$

where k is the covariance kernel of F. Define the space of symmetric square integrable functions with respect to μ by

$$\mathcal{L}^2_{\mathrm{symm}}(\mu) := \left\{ \varphi : \mathbb{R} \to \mathbb{C} : \int_{\mathbb{R}} |\varphi|^2 d\mu < \infty, \ \varphi(-\lambda) = \overline{\varphi(\lambda)} \ \forall \lambda \in \mathbb{R} \right\}.$$

F boasts the following orthonormal representation.

Lemma 2.4 ([21, Lem. 3.7]). Let $\{\varphi_n\}$ be an orthonormal basis of $\mathcal{L}^2_{\mathrm{symm}}(\mu)$ and write $\psi_n(t) = \int_{\mathbb{R}} e^{-i\lambda t}\varphi_n(\lambda)d\mu(\lambda)$. Then

$$F \overset{d}{=} \sum_n \zeta_n \psi_n, \quad \zeta_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}_{\mathbb{R}}(0,1).$$

The following lemma shows that F grows at most linearly on the real line, providing a probabilistic bound on its growth rate.

Lemma 2.5. There exists $c > 0$ such that for every $M > 0$ sufficiently large

$$\mathbb{P}\Big(\forall s \in \mathbb{R} \, : \, |F(s)| \leq M(1+|s|)\Big) \geq 1 - e^{-cM^2}.$$

Proof of Lemma 2.5. By a simple application of Dudley's bound [1, Thm. 1.3.3] and the Borell-TIS inequality [1, Thm. 2.1.1] (see [32, Lem. 3.10] for a full proof), there exist $a, c \in (0,1)$ such that for every M sufficiently large we have

$$\mathbb{P}\left(\sup_{[0,aM]} |F| \geq M\right) \leq \exp(-cM^2).$$

By the stationarity of F, the events

$$E_{k,M} := \left\{ \sup\{|F(t)| : t \in [akM, a(k+1)M]\} > \frac{(|k|+1)M}{2} \right\}$$

satisfy $\mathbb{P}(E_{k,M}) \leq \exp(-ck^2M^2/4)$. Taking a union bound, we obtain, for sufficiently large M,

$$\mathbb{P}(\cup_{k\in\mathbb{Z}}E_{k,M}) \leq \exp(-cM^2/5).$$

Observing that $\{\exists s \in \mathbb{R} \, : \, |F(s)| > M(1+|s|)\} \subset \cup_{k\in\mathbb{Z}}E_{k,M}$, the lemma follows. $\qquad\square$

We also need bounds on *ball event* probabilities.

Lemma 2.6 (Application of [32, Lemma 3.4]). *Suppose that $\mu_{\mathrm{ac}} \neq 0$. Then there exist $\alpha, \beta > 0$ such that for all $T \geq 1$ and $\ell > 0$ we have:*

$$\mathbb{P}\left(\sup_{[0,T]} |F| \leq \ell\right) \leq e^{\beta T^2} \ell^{\alpha T}.$$

Lemma 2.7 ([21, Lemma 3.17]). *Suppose that $\int |\lambda|^{\delta} d\mu(\lambda) < \infty$ for some $\delta > 0$. Then there exist $\ell_0 > 0$ and $c > 1$ such that for all $T \geq 1$ and $\ell > \ell_0$ we have:*

$$\mathbb{P}\left(\sup_{[0,T]} |F| \leq \ell\right) \geq \mathbb{P}\left(c|F(0)| \leq \ell\right)^T.$$

The following famous comparison inequality is due to Anderson.

Lemma 2.8 ([3]). *Let X and Y be independent centered Gaussian processes on I. Then for any $\ell > 0$,*

$$\mathbb{P}\left(\sup_{I} |X + Y| \leq \ell\right) \leq \mathbb{P}\left(\sup_{I} |X| \leq \ell\right).$$

We shall also need the celebrated Gaussian correlation inequality.

Lemma 2.9 ([28, 34]). *Let X be a centred Gaussian vector in $\mathbb{R}^d$. Then, for any convex sets $K, L \subset \mathbb{R}^d$, symmetric around 0, we have*

$$\mathbb{P}(X \in (K \cap L)) \geq \mathbb{P}(X \in K)\,\mathbb{P}(X \in L).$$

Lastly, we state standard bounds on the tail of the Gaussian distribution.

Lemma 2.10 ([1, Chapter 1.2]). *Let $Z \sim \mathcal{N}_{\mathbb{R}}(0, 1)$. Then for any $x > 0$:*

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x} - \frac{1}{x^3}\right) e^{-x^2/2} \leq \mathbb{P}(Z > x) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{x} e^{-x^2/2}.$$

In particular, for $x \geq 2$ it holds that $e^{-x^2} \leq \mathbb{P}(Z > x) \leq e^{-x^2/2}$.

3. Lower bounds

This section is dedicated to the proof of Proposition 1.1. Assume, without loss of generality, that $X = 1$, $0 < \varepsilon < \frac{1}{10}$ and $T > 1$. Observe that there must exist $a \in (1 - \varepsilon, 1 + \varepsilon)$ such that

$$\mu\left(\left[a - \tfrac{1}{10T}, a + \tfrac{1}{10T}\right]\right) \geq \tfrac{1}{10\varepsilon T}\mu([1 - \varepsilon, 1 + \varepsilon]). \tag{5}$$

Fix such a and denote

$$I = \left[a - \tfrac{1}{10T}, a + \tfrac{1}{10T}\right] \text{ and } \pm I = I \cup -I.$$

Take an orthonormal basis $\{f_n\}_{n=1}^{\infty}$ for $\mathcal{L}^2_{\mathrm{symm}}(\mu)$, with $f_1 = (\mu(\pm I))^{1/2}\mathbb{1}_{\pm I}$. From Lemma 2.4 we obtain that

$$F(x) = \sum_{n=1}^{\infty} \xi_n \widehat{f}_n(x), \text{ where } \xi_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1).$$

Denote $G := F - \xi_1\widehat{f}_1$ and observe that G is independent of ξ_1. We compute,

$$\sqrt{2\mu(I)}\widehat{f}_1(x) = 2\int_I \cos(\lambda x)d\mu(\lambda) = 2\mu(I)\cos(ax) + 2\int_I (\cos(\lambda x) - \cos(ax))\,d\mu(\lambda),$$

$$-\sqrt{2\mu(I)}\widehat{f}_1{}'(x) = 2\int_I \lambda\sin(\lambda x)d\mu(\lambda) = 2a\mu(I)\sin(ax) + 2\int_I (\lambda\sin(\lambda x) - a\sin(ax))\,d\mu(\lambda).$$

Since both sin and cos are 1-Lipschitz, we conclude that for any $|x| \leq T$,

$$\left| \frac{\widehat{f}_1(x)}{\sqrt{2\mu(I)}} - \cos(ax) \right| \leq \frac{1}{\mu(I)} \int_I |\lambda - a||x|d\mu(\lambda) < \frac{1}{10}, \quad \text{and}$$

$$\left| \frac{\widehat{f}_1'(x)}{a\sqrt{2\mu(I)}} + \sin(ax) \right| \leq \frac{1}{a\mu(I)} \int_I |\lambda - a|(|ax| + 1)d\mu(\lambda) \leq \frac{1}{10}\left(1 + \frac{2}{1-\varepsilon}\right) < \frac{4}{10},$$

where we used the fact that $|\lambda - a| \leq \frac{1}{10T}$ for all $\lambda \in I$. Observe that G is almost surely continuously differentiable and that G' is a centred Gaussian process, by our assumption $\int \lambda^2 \log^2 \lambda d\mu(\lambda) < \infty$. Given $L > 0$, define the events

$$E_1 = \left\{ \sqrt{2\mu(I)}\xi_1 \geq 10L \right\}, \quad E_2 = \left\{ \sup_{[0,T]} |G| \leq L \right\}, \quad E_3 = \left\{ \sup_{[0,T]} |G'| \leq aL \right\},$$

and observe that E_1 is independent from E_2 and E_3. We deduce that on the event $E_1 \cap E_2 \cap E_3$, we have

$$\left| \frac{F(x)}{\sqrt{2\mu(I)}\xi_1} - \cos(ax) \right| = \left| \frac{\widehat{f}_1(x)\xi_1}{\sqrt{2\mu(I)}\xi_1} + \frac{G(x)}{\sqrt{2\mu(I)}\xi_1} - \cos(ax) \right|$$

$$\leq \left| \frac{\widehat{f}_1(x)}{\sqrt{2\mu(I)}} - \cos(ax) \right| + \left| \frac{G(x)}{\sqrt{2\mu(I)}\xi_1} \right| < \frac{1}{2}.$$

$$\left| \frac{F'(x)}{a\sqrt{2\mu(I)}\xi_1} + \sin(ax) \right| = \left| \frac{\widehat{f}_1'(x)\xi_1}{a\sqrt{2\mu(I)}\xi_1} + \frac{G'(x)}{a\sqrt{2\mu(I)}\xi_1} + \sin(ax) \right|$$

$$\leq \left| \frac{\widehat{f}_1'(x)}{a\sqrt{2\mu(I)}} + \sin(ax) \right| + \left| \frac{G'(x)}{a\sqrt{2\mu(I)}\xi_1} \right| < \frac{1}{2}.$$

Denote $S_1 = \left(\frac{\pi}{a}\mathbb{Z} + \left[-\frac{\pi}{4a}, \frac{\pi}{4a}\right]\right) \cap [0,T]$ and $S_2 = \left(\frac{\pi}{a}\mathbb{Z} + \left[\frac{\pi}{4a}, \frac{3\pi}{4a}\right]\right) \cap [0,T]$, and observe that $S_1 \cup S_2 = [0,T]$. We conclude that on the event $E_1 \cap E_2 \cap E_3$ we have

$$\forall x \in S_1 : \quad \left| \frac{F(x)}{\sqrt{2\mu(I)}\xi_1} \right| > \cos(\tfrac{\pi}{4}) - \tfrac{1}{2} \tag{6}$$

$$\forall x \in S_2 : \quad \left| \frac{F'(x)}{a\sqrt{2\mu(I)}\xi_1} \right| > \sin(\tfrac{\pi}{4}) - \tfrac{1}{2} \tag{7}$$

$$\forall k \in \mathbb{Z} \cap [0,T] : \quad \text{sgn}\left(F(\tfrac{k\pi}{a})\right) = (-1)^k. \tag{8}$$

By (6), F has no zeros in S_1. By (8), it has at least one zero in each interval of S_2 except the last, while by (7), it cannot have more than one zero in each interval of S_2. Since there are $\lfloor \frac{aT}{\pi} \rfloor$ such intervals, and we deduce that $\frac{N(T)}{T} \in \left[\frac{1-2\varepsilon}{\pi}, \frac{1+2\varepsilon}{\pi}\right]$ for sufficiently large T. Therefore

$$\mathbb{P}\left(\frac{N(T)}{T} \in \left[\frac{1-2\varepsilon}{\pi}, \frac{1+2\varepsilon}{\pi} \right] \right) \geq \mathbb{P}(E_1 \cap E_2 \cap E_3) = \mathbb{P}(E_1)\mathbb{P}(E_2 \cap E_3). \tag{9}$$

We are thus left with choosing L and obtaining a bound on $\mathbb{P}(E_1)$ and $\mathbb{P}(E_2 \cap E_3)$.

Writing $\kappa = \frac{\varepsilon}{10\mu([1-\varepsilon,1+\varepsilon])}$ and using (5) together with the Gaussian tail bound in Lemma 2.10, we get, for sufficiently large T,

$$\mathbb{P}(E_1) = \mathbb{P}\left(\xi_1 \geq 5\sqrt{\tfrac{2}{\mu(I)}}L \right) \geq \mathbb{P}\left(\xi_1 \geq 5\sqrt{2\kappa T}L \right) \geq e^{-50L^2\kappa T}.$$

In addition, we observe that there exists some $c_{1,L} > 0$ which depends on μ, for which

$$\mathbb{P}(E_2) = \mathbb{P}\left(\sup_{[0,T]}|G| \leq L\right) \geq \mathbb{P}\left(\sup_{[0,T]}|F| \leq L\right) \geq e^{-c_{1,L}T},$$

where the first inequality follows from Anderson's inequality (Lemma 2.8) and the second one follows from the ball estimate of Lemma 2.7. From the latter lemma and from the fact that $\lim_{\ell \to \infty} \mathbb{P}(c|F(0)| \leq \ell) = 1$ we also deduce that $\lim_{L \to \infty} c_{1,L} = 0$.

Similarly we obtain the existence of $c_{2,L}$ which satisfies $\lim_{L \to \infty} c_{2,L} = 0$, for which

$$\mathbb{P}(E_3) = \mathbb{P}\left(\sup_{[0,T]}|G'| \leq aL\right) \geq \mathbb{P}\left(\sup_{[0,T]}|F'| \leq aL\right) \geq e^{-c_{2,L}T}.$$

Since G, G' are almost surely continuous, we may apply the Gaussian correlation inequality (Lemma 2.9) to obtain that $\mathbb{P}(E_2 \cap E_3) \geq \mathbb{P}(E_2)\mathbb{P}(E_3)$. Setting $L = \kappa^{-1/3}$ and plugging the estimates we obtained into (9) we get

$$\mathbb{P}\left(\frac{N(T)}{T} \in \left[\frac{1-2\varepsilon}{\pi}, \frac{1+2\varepsilon}{\pi}\right]\right) \geq e^{-(50L^{-1}+c_{1,L}+c_{2,L})T}.$$

To obtain the moreover part, apply this for arbitrarily small $\varepsilon = \delta$, and notice that $L = L(\delta)$ is then arbitrarily large. □

4. Upper bound on overcrowding probabilities

In this section we prove the first part of Theorem 1, which gives a Gaussian upper bound on certain overcrowding probabilities. Note that the second part of Theorem 1 follows readily from Proposition 1.1.

Throughout the remainder of the paper, we let F be a fixed centred SGP with spectral measure μ supported on $[-A, A]$, which satisfies $\mu_{\mathrm{ac}} \not\equiv 0$. Firstly we make the observation that, since μ is compactly supported, F is almost surely real analytic, so that for all $t \in \mathbb{R}$ we have

$$F(t) = \sum_{k=0}^{\infty} \frac{F^{(k)}(0)t^k}{k!}. \tag{10}$$

This formula naturally extends F to an entire function on $\mathbb{C}$. By a slight abuse of notation, we denote this extended function also by F.

In order to apply the Jensen's formula to F, we require the following proposition.

Proposition 4.1. F is almost surely in Cartwright's class and of exponential type at most A.

Using this proposition we develop the following probabilistic estimate.

Proposition 4.2. There exists $c > 0$ such that for all $T > e^8$,

$$\mathbb{P}\left(\forall a \in [0,T], \ r \in (0,T): \ \frac{1}{2\pi}\int_0^{2\pi} \log|F(a + re^{i\theta})|d\theta \leq \frac{A}{\pi} \cdot 2r + 6\log T\right) \geq 1 - e^{-cT^2}.$$

The proof of both propositions is provided in Section 4.2 below. Equipped with them we are now ready to prove the theorem itself.

4.1. **Proof of the upper bound in Theorem 1.** Let $T \geq e^8$. Given $\varepsilon \in \left[\frac{1}{\sqrt{T}}, e^{-3}\right]$, write

$$\Delta = \varepsilon^2 T, \ \ r = \varepsilon T \text{ and } n = \left\lceil \frac{T + 2r}{\Delta} \right\rceil = \left\lceil \frac{1}{\varepsilon^2} + \frac{2}{\varepsilon} \right\rceil. \tag{11}$$

Let $x_1, \dots, x_n \in [0, T]$ be such that

$$x_k \in I_k := [(k-1)\Delta - r, k\Delta - r]$$

and assume that $F(x_k) \neq 0$ for all $k \in [n]$. By Jensen's formula (Theorem 2.1), for all $r > 0$ we have

$$\sum_{k \in [n]} \int_0^r \frac{n_F(x_k, t)}{t} dt = \sum_{k \in [n]} \frac{1}{2\pi} \int_0^{2\pi} \log |F(re^{i\theta} + x_k)| d\theta - \sum_{k \in [n]} \log |F(x_k)|. \qquad (12)$$

We proceed in three steps. Firstly, we relate the left-hand-side of (12) to $N_F(T)$. Next, we use Proposition 4.2 to replace the first term in the right-hand-side with the desired bound and show that with high probability there exists a good collection of points $x_1, \dots, x_n$ for which the second term of the right-hand-side is negligible. Finally, we combine these to obtain the desired upper bound.

Step 1: Relating the left-hand-side of (12) and the number of real zeros. Observe that for all $x \in [0, T]$, $t > 0$ we have

$$|\{k \in [n] : \ x \in [x_k - t, x_k + t]\}| \geq \left\lceil \frac{2t}{\Delta} \right\rceil - 2.$$

In particular, this holds for values of x which are real zeros of F. Denoting the set of real zeros of F by $\mathcal{Z}$, we therefore have

$$\sum_{k \in [n]} n_F(x_k, t) \geq \sum_{k \in [n]} \sum_{w \in \mathcal{Z}} \mathbb{1}_{(x_k - t, x_k + t)}(w) \geq \left(\left\lceil \frac{2t}{\Delta} \right\rceil - 2 \right) N_F(T)$$

Integrating both sides against $\frac{1}{t}$ yields

$$\sum_{k \in [n]} \int_0^r \frac{n_F(x_k, t)}{t} dt \geq N_F(T) \int_\Delta^r \frac{1}{t} \left(\left\lceil \frac{2t}{\Delta} \right\rceil - 2 \right) dt$$

$$\geq 2 N_F(T) \left(\frac{r - \Delta}{\Delta} - \log \left(\frac{r}{\Delta} \right) \right)$$

$$= 2 N_F(T) \left(\frac{1}{\varepsilon} - 1 - \log \left(\frac{1}{\varepsilon} \right) \right)$$

$$\geq 2 N_F(T) \left(\frac{1}{\varepsilon} - 2 \log \left(\frac{1}{\varepsilon} \right) \right), \qquad (13)$$

where we used (11) and the fact that $\varepsilon < e$.

Step 2: Bounding the right-hand-side of (12). By Proposition 4.2, there exists $c_1 > 0$ such that the event

$$E_1 := \left\{ \sum_{k \in [n]} \frac{1}{2\pi} \int_0^{2\pi} \log |F(re^{i\theta} + x_k)| d\theta \leq \frac{A}{\pi} \cdot 2rn + 6n \log T \right\}$$

satisfies

$$\mathbb{P}(E_1) \geq 1 - e^{-c_1 T^2}. \qquad (14)$$

By Lemma 2.6, there exist $C, c_2 > 0$ such that if $\varepsilon^2 T > 1$ we have

$$\mathbb{P} \left(\sup_{I_k} |F| \leq e^{-C\varepsilon^2 T} \right) = \mathbb{P} \left(\sup_{[0, \varepsilon^2 T]} |F| \leq e^{-C\varepsilon^2 T} \right) \leq e^{-c_2 \varepsilon^4 T^2}.$$

Writing

$$E_2 := \left\{ \forall k \in [n] : \sup_{I_k} |F| \geq e^{-C\varepsilon^2 T} \right\},$$

taking a union bound over $k \in [n]$, and using the fact that $n \leq \frac{1}{\varepsilon^2} + \frac{2}{\varepsilon} + 1 \leq \left(\sqrt{T} + 1 \right)^2 \leq 2T$, we obtain

$$\mathbb{P}(E_2) \geq 1 - ne^{-c_2 \varepsilon^4 T^2} \geq 1 - 2Te^{-c_2 \varepsilon^4 T^2}. \tag{15}$$

On the event E_2, we can select the points $\{x_k\}_{k \in [n]}$ which satisfy $|F(x_k)| \geq e^{-C\varepsilon^2 T}$. Thus, on the event $E_1 \cap E_2$, we can bound the right-hand-side of (12) by

$$\sum_{k \in [n]} \frac{1}{2\pi} \int_0^{2\pi} \log |F(re^{i\theta} + x_k)| d\theta - \sum_{k \in [n]} \log |F(x_k)| \leq \frac{A}{\pi} \cdot 2rn + 6n \log T + Cn\varepsilon^2 T. \tag{16}$$

Step 3: Obtaining the probabilistic bound. Plugging (13) and (16) into (12) we obtain that, on the event $E_1 \cap E_2$, we have

$$2 \left(\frac{1}{\varepsilon} - 2 \log(\frac{1}{\varepsilon}) \right) N_F(T) \leq \frac{A}{\pi} \cdot 2rn + 6n \log T + Cn\varepsilon^2 T.$$

Recalling (11) and the fact that $n\varepsilon^2 \leq 1 + 3\varepsilon \leq 2$, we conclude that on $E_1 \cap E_2$,

$$(1 - 2\varepsilon \log(\tfrac{1}{\varepsilon})) N_F(T) \leq \frac{A}{\pi} T(1 + 3\varepsilon) + 3(\tfrac{1}{\varepsilon} + 3) \log T + C\varepsilon T$$

$$\leq \frac{A}{\pi} T + D\varepsilon T + D\tfrac{1}{\varepsilon} \log T,$$

where $D > 0$ is some constant. In what follows we use D to denote constants which may change from line to line. Since $(1 - 2\varepsilon \log(\tfrac{1}{\varepsilon}))^{-1} \leq 1 + 3\varepsilon \log(\tfrac{1}{\varepsilon})$ for $\varepsilon < e^{-3}$, we get that on $E_1 \cap E_2$,

$$N_F(T) \leq \frac{A}{\pi} T + D\varepsilon \log \left(\tfrac{1}{\varepsilon} \right) T + D\tfrac{1}{\varepsilon} \log T$$

$$\leq \frac{A}{\pi} T + D\varepsilon \log \left(\tfrac{1}{\varepsilon} \right) T,$$

where in the last inequality we used that $\varepsilon > \frac{1}{\sqrt{T}}$. By (14) and (15) we have $\mathbb{P}(E_1 \cap E_2) \geq 1 - e^{-c\varepsilon^4 T^2}$ for some $c > 0$. Changing variables by defining $\delta = D\varepsilon \log \left(\tfrac{1}{\varepsilon} \right)$, we conclude that

$$\mathbb{P} \left(N_F(T) \leq \tfrac{A}{\pi} T + \delta T \right) \geq \mathbb{P}(E_1 \cap E_2) \geq 1 - \exp \left(-c \left(\delta / \log \delta \right)^4 T^2 \right),$$

where $\delta \in [d_1 \log T / \sqrt{T}, d_2]$, for some constants $d_1, d_2 > 0$. $\qquad \square$

4.2. **Proof of Propositions 4.1 and 4.2.** Here we establish Proposition 4.1 and use it together with tools from Sections 2.1 and 2.2 to establish Proposition 4.2.

Proof of Proposition 4.1. We note that the integral condition (3) holds by Lemma 2.5, which ensures that, almost surely, F grows at most linearly on $\mathbb{R}$. It remains to show that the exponential type of F is at most A. Recall that F has the power series expansion (10):

$$F(z) = \sum_{k=0}^{\infty} \frac{F^{(k)}(0)}{k!} z^k, \quad \forall z \in \mathbb{C}.$$

Note that, for $k \in \mathbb{N}$, we have $F^{(k)}(0) \sim \mathcal{N}(0, C_{2k})$, where $C_{2k} := \int_{\mathbb{R}} |\lambda|^{2k} d\mu(\lambda)$. Using the fact that $\mathrm{sprt}(\mu) \subseteq [-A, A]$, we obtain $C_{2k} \leq A^{2k}$. An application of Lemma 2.10 yields for $k \geq 2$:

$$\mathbb{P}(|F^{(k)}(0)| \geq kA^k) \leq \mathbb{P}(|F^{(k)}(0)| \geq k\sqrt{C_{2k}}) \leq \exp(-k^2/2).$$

Hence, almost surely, there exists $m \in \mathbb{N}$ such that $|F^{(k)}(0)| \leq kA^k$ for every $k \geq m+1$. Denoting $C = \max_{k \leq m} |F^{(k)}(0)/A^k|$, we obtain

$$
|F(z)| \leq \sum_{k=0}^{m} |F^{(k)}(0)| \frac{|z|^k}{k!} + \sum_{k=m+1}^{\infty} kA^k \frac{|z|^k}{k!}
$$
$$
\leq C \sum_{k=0}^{m} A^k \frac{|z|^k}{k!} + A|z| \sum_{k=m}^{\infty} A^k \frac{|z|^k}{k!}
$$
$$
\leq (C + A|z|) \exp(A|z|).
$$

For any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that $C + A|z| \leq C_\varepsilon e^{\varepsilon|z|}$ for all $z \in \mathbb{C}$, which, in turn, implies $|F(z)| \leq C_\varepsilon e^{(A+\varepsilon)|z|}$. Recalling the definitions in Section 1.2, we deduce that F is a.s. of exponential type at most A. $\square$

Proof of Proposition 4.2. With Proposition 4.1 at hand, apply Corollary 2.3 to get that for every $z = x + iy \in \mathbb{C}$, we have

$$
\log |F(z)| \leq \frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log^+ |F(t)|}{|t-z|^2} dt + A|y|. \tag{17}
$$

Write

$$
E_M = \{\forall t \in \mathbb{R} : |F(t)| \leq M(1+|t|)\}.
$$

By Lemma 2.5, there exist $c > 0$ such that

$$
\mathbb{P}(E_M) \geq 1 - e^{-cM^2}, \tag{18}
$$

for sufficiently large values of M. Using the fact that $\frac{y}{\pi} \int_{\mathbb{R}} \frac{1}{|t-z|^2} dt = 1$, we obtain from (17) that on the event E_M, we have

$$
\log |F(z)| \leq \log M + \frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log(1+|t|)}{|t-z|^2} dt + A|y|, \quad \forall z \in \mathbb{C}. \tag{19}
$$

For $y \neq 0$ and $|z| \leq 2T$, we obtain

$$
\frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log(1+|t|)}{(t-x)^2 + y^2} dt = \frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log(1+|x+s|)}{s^2 + y^2} ds
$$
$$
\leq \frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log(1+2T+|s|)}{s^2 + y^2} ds
$$
$$
\leq 2\log(1+2T) + \frac{2|y|}{\pi} \int_{\mathbb{R}} \frac{\log^+ |s|}{s^2 + y^2} ds
$$
$$
\leq 2\log(1+2T) + 2\log^+ |y| + 2 \leq 4(2 + \log T), \tag{20}
$$

where for the second inequality we use the fact that $\log(\alpha + \beta) \leq 2(\log \alpha + \log^+ \beta)$ for $\alpha \geq 1$ and $\beta > 0$, and for the third we use the bound

$$
\frac{|y|}{\pi} \int_{\mathbb{R}} \frac{\log^+ |s|}{s^2 + y^2} ds = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\log^+ |ty|}{1+t^2} dt = \log^+ |y| + \frac{1}{\pi} \int_{\mathbb{R}} \frac{\log^+ |t|}{1+t^2} dt \leq \log^+ |y| + 1.
$$

Plugging (20) into (19), we conclude that on the event E_M, for $|z| \leq 2T$ with $\text{Im}(z) \neq 0$,

$$
\log |F(z)| \leq A|y| + \log M + 4(2 + \log T). \tag{21}
$$

Thus on the event E_M, for all $a \in [0, T]$, $r \in (0, T)$, and $\theta \in (0, \pi) \cup (\pi, 2\pi)$ we have

$$
\log |F(a + re^{i\theta})| \leq Ar|\sin \theta| + \log M + 4(2 + \log T). \tag{22}
$$

Set $M = T$. By integrating on $\theta \in (0, 2\pi)$, we obtain that on the event E_T,

$$\forall a \in [0, T],\ r \in (0, T): \quad \frac{1}{2\pi} \int_0^{2\pi} \log |F(a + re^{i\theta})| \le \frac{2}{\pi} Ar + 6 \log T,$$

as long as $T \ge e^8$. Recalling (18), this completes the proof. $\qquad \square$

5. Upper bound on undercrowding probabilities

This section consists of the proof Theorem 2, by a reduction to Theorem 1. Let μ, A and B be as in the statement of Theorem 2. We define a new measure ν as follows: for any Borel measurable $S \subseteq \mathbb{R}$,

$$\nu(S) := \mu_+(S + A) + \mu_-(S - A),$$

where $\mu_+(S) = \mu(S \cap [0, \infty))$ and $\mu_-(S) = \mu(S \cap (-\infty, 0))$. Observe that ν is a symmetric Borel probability measure on $\mathbb{R}$ supported on $[B - A, A - B]$, so that it is the spectral measure of a centered SGP on $\mathbb{R}$ which we denote by G.

Claim 5.1. *The process G has the representation*

$$G(x) = \cos(Ax)F(x) + \sin(Ax)H(x), \tag{23}$$

where H has the same distribution as F (but is dependent of it).

First, we use this to prove Theorem 2 and then establish the claim itself. Note that, by (23),

$$G(k\pi/A) = (-1)^k F(k\pi/A), \quad \text{for any } k \in \mathbb{Z}.$$

By continuity, we deduce that if G has no zeros in the interval $[k_0 \frac{\pi}{A}, (k_0 + 1)\frac{\pi}{A}]$ for some $k_0 \in \mathbb{Z}$, then F must have at least one zero in that interval. Since both F and G are non-constant SGPs, their zero set is almost surely disjoint from the lattice $(\pi/A)\mathbb{Z}$. We conclude that

$$N_F(T) \ge \left\lfloor \frac{AT}{\pi} \right\rfloor - N_G(T). \tag{24}$$

Since G satisfies all the assumptions of Theorem 1, we may use it to obtain the existence of $c, d_1, d_2 > 0$ such that, for every $T > e^8$ and $\varepsilon \in \left[d_1 \frac{\log T}{\sqrt{T}}, d_2\right]$,, we have

$$\mathbb{P}\left(N_G(T) \ge \left\lceil \frac{A - B}{\pi} T \right\rceil + \varepsilon T\right) \le \exp(-c(\varepsilon/\log\varepsilon)^4 T^2).$$

This together with (24) yields the upper bound in Theorem 2. $\qquad \square$

Proof of Claim 5.1. Let $\{f_n\}_{n=1}^\infty$ be an orthonormal basis for $\mathcal{L}^2_{\text{symm}}(\mu)$ as per Lemma 2.4. For every $n \in \mathbb{N}$, consider the function g_n defined as follows:

$$g_n(t) := \begin{cases} f_n(t + A), & \text{if } t < 0, \\ f_n(t - A), & \text{if } t \ge 0. \end{cases}$$

Note that $\{g_n\}_{n=1}^\infty$ form an orthonormal basis for $\mathcal{L}^2_{\text{symm}}(\nu)$. Lemma 2.4 thus implies

$$F \stackrel{d}{=} \sum_{n=1}^\infty \xi_n \widehat{f_n}, \quad \text{and} \quad G \stackrel{d}{=} \sum_{n=1}^\infty \xi_n \widehat{g_n}, \tag{25}$$

where $\xi_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$, where

$$\widehat{f}_n(x) = \int_{\mathbb{R}} f_n(\lambda)e^{-i\lambda x}d\mu(\lambda), \quad \widehat{g}_n(x) = \int_{\mathbb{R}} g_n(\lambda)e^{-i\lambda x}d\nu(\lambda).$$

Using the relation between g_n and f_n, we compute:

$$\begin{aligned}
\widehat{g}_n(x) &= \int_{\mathbb{R}} g_n(\lambda)e^{-i\lambda x}d\nu(\lambda) \\
&= \int_{B-A}^{0} f_n(\lambda + A)e^{-i\lambda x}d\nu(\lambda) + \int_{0}^{A-B} f_n(\lambda - A)e^{-i\lambda x}d\nu(\lambda) \\
&= e^{iAx}\int_{B}^{A} f_n(\tau)e^{-i\tau x}d\mu(\tau) + e^{-iAx}\int_{-A}^{-B} f_n(\tau)e^{-i\tau x}d\mu(\tau) \\
&= (\cos(Ax) + i\sin(Ax))\int_{B}^{A} f_n(\tau)e^{-i\tau x}d\mu(\tau) + (\cos(Ax) - i\sin(Ax))\int_{-A}^{-B} f_n(\tau)e^{-i\tau x}d\mu(\tau) \\
&= \cos(Ax)\cdot\widehat{f}_n(x) + \sin(Ax)\cdot h_n(x) \tag{26}
\end{aligned}$$

where

$$\begin{aligned}
h_n &= i\left(\int_{B}^{A} f_n(\tau)e^{-i\tau x}d\mu(\tau) - \int_{-A}^{-B} f_n(\tau)e^{-i\tau x}d\mu(\tau)\right) \\
&= i\int_{\mathbb{R}} \left(f_n\mathbb{1}_{[0,\infty)}(\tau) - f_n\mathbb{1}_{(-\infty,0)}(\tau)\right)e^{-i\tau x}d\mu(\tau).
\end{aligned}$$

Note that $\{if_n\mathbb{1}_{[0,\infty)} - if_n\mathbb{1}_{(-\infty,0)}\}$ also forms a basis to $\mathcal{L}^2_{\text{symm}}(\mu)$, so that the random series $H := \sum_{n=1}^{\infty} \xi_n h_n$ is a SGP with spectral measure μ (by means of Lemma 2.4). Plugging (26) into (25), we obtain

$$\begin{aligned}
G(x) &= \sum_{n=1}^{\infty} \xi_n\widehat{g}_n(x) = \cos(Ax)\sum_{n=1}^{\infty} \xi_n\widehat{f}_n(x) + \sin(Ax)\sum_{n=1}^{\infty} \xi_n h_n(x), \\
&= \cos(Ax)F(x) + \sin(Ax)H(x),
\end{aligned}$$

which completes the proof of Claim 5.1. $\square$

Acknowledgements

References

[1] Robert J. Adler and Jonathan E. Taylor. *Random fields and geometry*, volume 80 of *Springer monographs in mathematics*. Springer, 2007.

[2] Michele Ancona and Thomas Letendre. Zeros of smooth stationary Gaussian processes. *Electronic Journal of Probability*, 26:1–81, 2021.

[3] Theodore W Anderson. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proceedings of the American Mathematical Society*, 6(2):170–176, 1955.

[4] Eran Assaf, Jeremiah Buckley, and Naomi D. Feldheim. An asymptotic formula for the variance of the number of zeroes of a stationary Gaussian process. *arXiv preprint arXiv:2101.04052*, 2021.

[5] Frank Aurzada, Sumit Mukherjee, and Ofer Zeitouni. Persistence exponents in markov chains. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(3):1411–1441, 2021.

[6] Frank Aurzada and Thomas Simon. Persistence probabilities and exponents. In *Lévy matters V*, pages 183–224. Springer, 2015.

[7] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[8] Riddhipratim Basu, Amir Dembo, Naomi Feldheim, and Ofer Zeitouni. Exponential Concentration for Zeroes of Stationary Gaussian Processes. *International Mathematics Research Notices*, 2020(23):9769–9796, 12 2018.

[9] Dmitry Beliaev, Stephen Muirhead, and Alejandro Rivera. A covariance formula for topological events of smooth gaussian fields. *Annals of Probability*, 48:2845–2893, 2020.

[10] Alan J. Bray, Satya N. Majumdar, and Grégory Schehr. Persistence and first-passage properties in nonequilibrium systems. *Advances in Physics*, 62(3):225–361, 2013.

[11] Włodzimierz Bryc and Amir Dembo. On large deviations of empirical measures for stationary Gaussian processes. *Stochastic processes and their applications*, 58(1):23–34, 1995.

[12] Harald Cramér and M. R. Leadbetter. *Stationary and related stochastic processes. Sample function properties and their applications.* John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[13] Amir Dembo and Sumit Mukherjee. No zero-crossings for random polynomials and the heat equation. *The Annals of Probability*, 43(1):85–118, 2015.

[14] Amir Dembo and Sumit Mukherjee. Persistence of Gaussian processes: non-summable correlations. *Probability Theory and Related Fields*, 169(3):1007–1039, 2017.

[15] Bernard Derrida, Vincent Hakim, and Reuven Zeitak. Persistent spins in the linear diffusion approximation of phase ordering and zeros of stationary Gaussian processes. *Physical review letters*, 77(14):2871, 1996.

[16] Monroe D. Donsker and Srinivasa R.S. Varadhan. Large deviations for stationary Gaussian processes. *Communications in Mathematical Physics*, 97(1):187–210, 1985.

[17] A. Eremenko and D. Novikov. Oscillation of Fourier integrals with a spectral gap. *J. Math. Pures Appl. (9)*, 83(3):313–365, 2004.

[18] S. Yu. Favorov. Zero sets of entire functions of exponential type with additional conditions on the real line. *Algebra i Analiz*, 20(1):138–145, 2008.

[19] Naomi D. Feldheim, Ohad N. Feldheim, Benjamin Jaye, Fedor Nazarov, and Shahaf Nitzan. On the probability that a stationary Gaussian process with spectral gap remains non-negative on a long interval. *International Mathematics Research Notices*, 2020(23):9210–9227, 2020.

[20] Naomi D. Feldheim, Ohad N. Feldheim, and Sumit Mukherjee. Persistence and ball exponents for Gaussian stationary processes. *arXiv preprint arXiv:2112.04820*, 2021.

[21] Naomi D. Feldheim, Ohad N. Feldheim, and Shahaf Nitzan. Persistence of Gaussian stationary processes: a spectral perspective. *The Annals of Probability*, 49(3):1067–1096, 2021.

[22] Damien Gayet and Jean-Yves Welschinger. Exponential rarefaction of real curves with many components. *Publications mathématiques de l'IHÉS*, 113:69–96, 2011.

[23] M. Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 49:314–320, 1943.

[24] Marie F. Kratz. Level crossings and other level functionals of stationary Gaussian processes. *Probab. Surv.*, 3:230–288, 2006.

[25] Marie F. Kratz and José R. León. Central limit theorems for level functionals of stationary Gaussian processes and fields. *Journal of Theoretical Probability*, 14(3):639–672, 2001.

[26] Maddaly Krishna and Manjunath Krishnapur. Persistence probabilities in centered, stationary, Gaussian processes in discrete time. *Indian Journal of Pure and Applied Mathematics*, 47(2):183–194, 2016.

[27] Manjunath Krishnapur. Overcrowding estimates for zeroes of planar and hyperbolic Gaussian analytic functions. *Journal of statistical physics*, 124(6):1399–1423, 2006.

[28] Rafał Latała and Dariusz Matlak. Royen's proof of the gaussian correlation inequality. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2014–2016*, pages 265–275. Springer, 2017.

[29] B. Ya. Levin. *Lectures on entire functions*, volume 150 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1996. In collaboration with and with a preface by Yu. Lyubarskii, M. Sodin and V. Tkachenko, Translated from the Russian manuscript by Tkachenko.

[30] Fedor Nazarov and Mikhail Sodin. On the number of nodal domains of random spherical harmonics. *American Journal of Mathematics*, 131(5):1337–1357, 2009.

[31] Lakshmi Priya. Concentration for nodal component count of Gaussian Laplace eigenfunctions. *arXiv e-prints*, page arXiv:2012.10302, December 2020.

[32] Lakshmi Priya. Overcrowding Estimates for the Nodal Volume of Stationary Gaussian Processes. *International Mathematics Research Notices*, 02 2022. rnac035.

[33] S. O. Rice. Mathematical analysis of random noise. *Bell System Tech. J.*, 24:46–156, 1945.

[34] Thomas Royen. A simple proof of the gaussian correlation conjecture extended to multivariate gamma distributions. *Far East J. Theor. Stat.*, 48:139–145, 2014.

[35] Yoni Rozenshein. The number of nodal components of arithmetic random waves. *Int. Math. Res. Not. IMRN*, (22):6990–7027, 2017.

[36] David Slepian. The one-sided barrier problem for Gaussian noise. *Bell System Technical Journal*, 41(2):463–501, 1962.

[37] Eric V. Slud. Mwi expansions for functionals related to level-crossing counts. *Chaos Expansions, Multiple Wiener-Ito Integrals and their Applications by Christian Houdré and Victor Pérez-Abreu, Probability and Stochastics Series*, pages 125–143, 1994.

[38] Mikhail Sodin. Lectures on random nodal portraits. In *Probability and statistical physics in St. Petersburg*, volume 91 of *Proc. Sympos. Pure Math.*, pages 395–422. Amer. Math. Soc., Providence, RI, 2016.

[39] Mikhail Sodin and Boris Tsirelson. Random complex zeroes, iii. decay of the hole probability. *Israel Journal of Mathematics*, 147(1):371–379, 2005.

[40] Tsirelson, Boris. Gaussian measures (gradute course lecture notes). http://www.tau.ac.il/~tsirel/Courses/Gauss3/main.html, 2010.

NAOMI DVORA FELDHEIM

DEPARTMENT OF MATHEMATICS, BAR-ILAN UNIVERSITY

Email address: naomi.feldheim@biu.ac.il

OHAD NOY FELDHEIM

EINSTEIN INSTITUTE OF MATHEMATICS, HEBREW UNIVESITY OF JERUSALEM

Email address: ohad.feldheim@mail.huji.ac.il

LAKSHMI PRIYA M.E.

SCHOOL OF MATHEMATICAL SCIENCES, TEL-AVIV UNIVERSITY

Email address: lpriyame@gmail.com