

A covariance formula for the number of excursion set components of Gaussian fields and applications

Dmitry Beliaev^{1,a} , Michael McAuley^{2,b}  and Stephen Muirhead^{3,c} 

¹*Mathematical Institute, University of Oxford, belyaev@maths.ox.ac.uk*

²*Department of Mathematics and Statistics, University of Helsinki, m.mcauley@cantab.net*

³*School of Mathematics and Statistics, University of Melbourne, smui@unimelb.edu.au*

Received 1 April 2023; revised 31 August 2023; accepted 5 September 2023

Abstract. We derive a covariance formula for the number of excursion or level set components of a smooth stationary Gaussian field on \mathbb{R}^d contained in compact domains. We also present two applications of this formula: (1) for fields whose correlations are integrable we prove that the variance of the component count in large domains is of volume order and give an expression for the leading constant, and (2) for fields with slower decay of correlation we give an upper bound on the variance which is of optimal order if correlations are regularly varying, and improves on best-known bounds if correlations are oscillating (e.g. monochromatic random waves).

Résumé. Nous démontrons une formule de covariance pour le nombre d'excursions ou de composantes d'ensemble de niveau d'un champ Gaussien stationnaire lisse sur \mathbb{R}^d sur un domaine compact. Nous présentons également deux applications de cette formule : (1) pour les champs dont la corrélation est intégrable nous démontrons que la variance du nombre de composantes dans des grands domaines est de l'ordre de volume du domaine et donnons une expression pour le coefficient constant du terme principale, et (2) pour les champs dont la corrélation décroît lentement nous donnons une majoration de la variance qui est d'ordre optimal si la corrélation varie régulièrement, et qui améliore les meilleures majorations connues si la corrélation oscille (par exemple dans le cas des ondes monochromatiques aléatoires).

MSC2020 subject classifications: 60G15; 60G60

Keywords: Gaussian fields; Excursion set; Level set; Component count; Covariance formula

1. Introduction

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a smooth stationary centred Gaussian field, let $\ell \in \mathbb{R}$, and for a compact domain $D \subset \mathbb{R}^d$ let $N_{\text{ES}}(D)$ and $N_{\text{LS}}(D)$ denote respectively the number of connected components of the excursion set $\{x \in \mathbb{R}^d : f(x) \geq \ell\}$ and the level set $\{x \in \mathbb{R}^d : f(x) = \ell\}$ which are contained in D (i.e. which intersect D but not its boundary). We abbreviate $N_{\star}(D)$ for $\star \in \{\text{ES}, \text{LS}\}$, and refer to these variables generically as the ‘component count’.

Recently, the asymptotic distribution of the component count in large domains has attracted considerable attention, motivated by potential applications (in, e.g., cosmology [2,37], medical imaging [46], quantum chaos [39]), as well as by connections to other areas such as percolation theory [11].

By now the first-order asymptotics of the component count are well-understood: in [31] Nazarov and Sodin established that, under very general conditions, the component count satisfies a strong law of large numbers with volume-order scaling. Although the limiting constant is defined implicitly, its properties have been studied in, e.g., [4,5,24].

A natural next step is to investigate second-order properties. Unlike the law of large numbers, it is expected that the order of the variance depends strongly on the covariance structure of the field. In Section 1.2.1 below we give a brief overview of known and conjectured second-order results.

In this paper we establish an exact covariance formula for the component counts on two compact domains. This formula is inspired by an analogous covariance formula for ‘topological events’ that appeared in [8], and is of a similar ‘interpolation’ type to related covariance formulae for Gaussian vectors [13,14,36] and fields [8,28].

As an application of the covariance formula we prove new results for the variance of the component count. First, assuming correlations are integrable, we prove that the variance has volume-order scaling, and give an expression for

the limiting constant (see Theorem 3). Second, we give a general upper bound on the variance in terms of the covariance kernel of the field which in many cases improves on best-known results, e.g. the important case of monochromatic random waves, and is conjecturally optimal for fields with regularly varying correlations (see Theorem 4 and Corollary 2).

1.1. *A covariance formula for component counts*

For the remainder of the paper we assume the following general conditions on the field f . Let $K(x) = \mathbb{E}[f(0)f(x)]$ be the *covariance kernel* of f . Recall that the *spectral measure* of f is the finite measure μ on \mathbb{R}^d such that K is the Fourier transform of μ . The choice of the normalization in the Fourier transform is unimportant for this paper. We will use the following definition: $K(x) = \mathcal{F}[\mu](x) := \int_{\mathbb{R}^d} e^{it \cdot x} \mu(dt)$.

Assumption 1 (General conditions on the field).

1. f is C^3 -smooth almost surely;
2. As $|x| \rightarrow \infty$, $\max_{|\alpha| \leq 2} \partial^\alpha K(x) \rightarrow 0$;
3. The support of the spectral measure μ contains either (i) an open set, or (ii) a scaled sphere $a\mathbb{S}^{d-1}$ for some $a > 0$.

We actually prove our results using only a weaker version of the third item above – see Remark 14 – however the present statement is convenient and covers all of our examples. The second item of Assumption 1 is used in only one place (in the proof of Lemma 6) and could perhaps also be weakened.

To state the covariance formula we introduce some notation. The formula will apply to the component count on *boxes*, which are sets of the form $B = [a_1, b_1] \times \dots \times [a_d, b_d] \subset \mathbb{R}^d$, for finite $a_i \leq b_i$. Henceforth we fix a level $\ell \in \mathbb{R}$, two boxes B_1 and B_2 , and $\star, \diamond \in \{\text{ES}, \text{LS}\}$.

We first introduce an interpolation between f and an independent copy \tilde{f}

$$(1.1) \quad f^t := tf + \sqrt{1-t^2}\tilde{f}, \quad t \in [0, 1],$$

remarking that f^t is equal in law to f for each $t \in [0, 1]$, and that $f^0 = \tilde{f}$ and $f^1 = f$. We denote by $N_\star^t(D)$ the analogue of $N_\star(D)$ for the field f^t .

We next define a notion of ‘pivotal measures’ for the pair of fields (f, f^t) ; to keep the discussion simple we defer precise definitions to Section 2.2 and instead convey the intuition.

Consider conditioning the field f to have a critical point at $x \in B_1$ with critical value ℓ . Then we can classify this critical point as being either ‘positively pivotal’, ‘negatively pivotal’, or ‘not pivotal’, depending on whether a small local positive perturbation of the field f near the point x increases the component count $N_\star(B_1)$ by one, decreases the component count $N_\star(B_1)$ by one, or keeps the component count the same (see Figure 1 for examples). This classification is well-defined for non-degenerate critical points (see Lemma 2).

In Sections 2.2.2–2.2.3 we will define the ‘two-point density function’ $p_t(x, y)$, which can intuitively be thought of as the density of non-degenerate critical points of (f, f^t) with critical value (ℓ, ℓ) . This function has a similar form to densities that appear in the Kac–Rice formula for the expected number of critical points of a field (see [1, Chapter 11]) but there is no exact correspondence because the expected number of critical points of (f, f^t) with critical value (ℓ, ℓ) is zero. Let $p_t^{++}(x, y)$ be the fraction of this density which corresponds to pairs of critical points which are positively pivotal for $N_\star(B_1, \ell)$ and $N_\diamond(B_2, \ell)$ respectively, and define $p_t^{\pm\pm}(x, y)$ similarly. Although these densities depend on $B_1, B_2 \subset \mathbb{R}^d$ and $\star, \diamond \in \{\text{ES}, \text{LS}\}$ (as well as f and ℓ), to ease notation we drop these dependencies. We also note that, by definition,

$$(1.2) \quad \sum_{i,j \in \{+, -\}} p_t^{ij} \leq p_t.$$

Strictly speaking, in the definition of $p_t(x, y)$ and $p_t^{\pm\pm}(x, y)$ we also consider critical points of (f, f^t) restricted to substrata of the boxes B_1, B_2 (i.e. boundary faces of dimension $0 \leq m \leq d - 1$), so $p_t(x, y)$ and $p_t^{\pm\pm}(x, y)$ are actually measures which have absolutely continuous components with respect to (pairwise products of) the Lebesgue measure on all (pairs of) strata. (See Definition 1 for a precise formulation of these statements.) Hence when we integrate $p_t^{\pm\pm}(x, y)$ over the domain $B_1 \times B_2$, we will denote it as $dp_t^{\pm\pm}(x, y)$ to keep in mind that we integrate over Lebesgue measures on all strata.

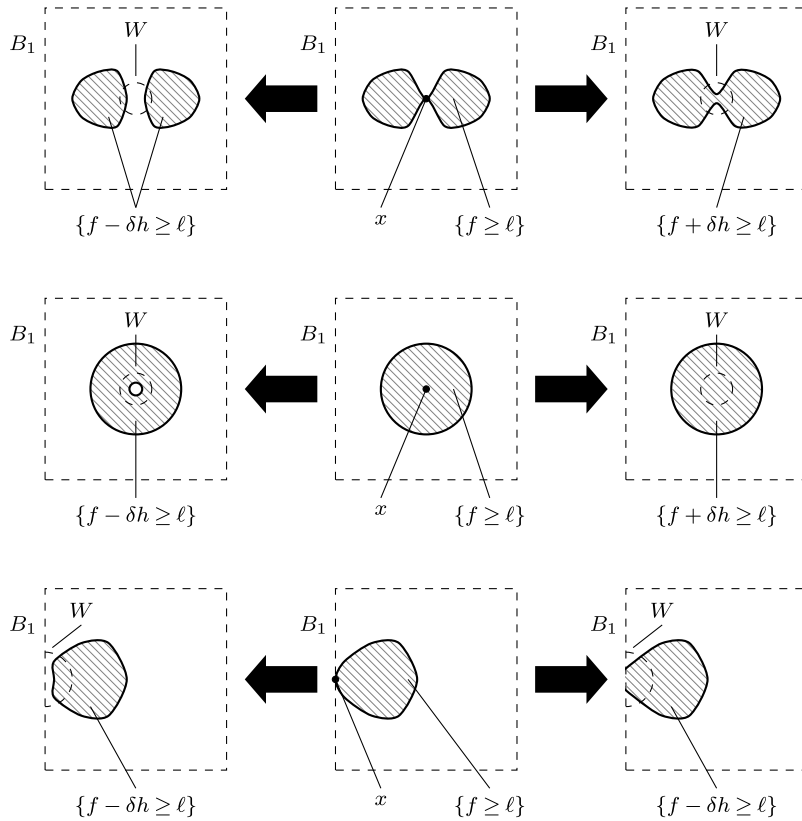


Fig. 1. Top panel: An illustration of a critical point which is ‘negatively pivotal’ for both the excursion and level set component counts at level ℓ . We consider a small perturbation $\pm\delta h$ of the excursion set on a neighbourhood W of the critical point x , where h is a local positive perturbation (precise definitions will be given in Section 2). In this case both excursion and level component counts decrease by one as we pass from the negative perturbation (left) to the positive perturbation (right). Middle panel: A critical point which is ‘negatively pivotal’ for the level set and ‘not pivotal’ for the excursion set. Bottom panel: A critical point of a boundary stratum which is ‘negatively pivotal’ for both the level set and the excursion set. Note that other configurations are also possible.

Theorem 1 (Covariance formula for component counts). *Suppose Assumption 1 holds and let $\ell \in \mathbb{R}$, $B_1, B_2 \subset \mathbb{R}^d$ be boxes, and $\star, \diamond \in \{\text{ES}, \text{LS}\}$. Then*

$$\text{Cov}[N_\star(B_1), N_\diamond(B_2)] = \int_0^1 \int_{B_1 \times B_2} K(x - y) (dp_t^{++}(x, y) + dp_t^{--}(x, y) - dp_t^{+-}(x, y) - dp_t^{-+}(x, y)) dt.$$

As mentioned, Theorem 1 is inspired by an analogous covariance formula for ‘topological events’ that appeared in [8], and in fact we deduce Theorem 1 as a consequence of that formula. In turn, these two formulae can be viewed as generalisations of classical interpolation formulae for finite Gaussian vectors [10,36], and are also related to semigroup interpolation methods developed more recently [13,14,25]. See also [28,42,43] for other recent uses of similar interpolation formulae. For comparison let us state the interpolation formula in the classical case of smooth functions of finite Gaussian vectors.

Proposition 1 (Classical Gaussian interpolation formula). *Let X be a centred Gaussian vector with covariance matrix $(K(i, j))_{1 \leq i, j \leq n}$. Let $g_1, g_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ be absolutely continuous functions such that $\mathbb{E}[g_i(X)^2]$ and $\mathbb{E}[\|\nabla g_i(X)\|_2^2]$ are finite for $i = 1, 2$. Then, for every $t \in [0, 1]$,*

$$\text{Cov}[g_1(X), g_2(X^t)] = \int_0^t \sum_{i,j} K(i, j) \mathbb{E} \left[\frac{\partial g_1(X)}{\partial X_i} \frac{\partial g_2(X^s)}{\partial X_j^s} \right] ds,$$

where

$$X^t = tX + \sqrt{1 - t^2} \tilde{X}, \quad t \in [0, 1],$$

and \tilde{X} is an independent copy of X .

Proof. The proof in the case $g_1 = g_2$ can be found in [13, Lemma 3.4] (note the different parameterisation of the interpolation), and the proof in the general case is identical. \square

Compared to the formula in Proposition 1, the covariance formula in Theorem 1 has two main differences: (i) it concerns functions of a continuous Gaussian field instead of a finite Gaussian vector, and (ii) it concerns integer-valued (and so non-differentiable) functions. Heuristically one handles these differences by interpreting the partial derivatives in Proposition 1 as ‘densities’ in an appropriate sense.

On the other hand, the covariance formula in Proposition 1 does not require stationarity, whereas Theorem 1 is restricted to the stationary setting. We conjecture that Theorem 1 should also be true for non-stationary Gaussian fields; indeed we only use stationarity in one place in the proof – to obtain the integrability of the pivotal measures, see (2.12) – and we expect this to be true under very general conditions.

Remark 1. We also conjecture that the covariance formula in Theorem 1 should extend to component counts on B_1 and B_2 taken with respect to *distinct* levels $\ell_1 \neq \ell_2$ respectively. In fact in that case we expect that the pivotal measures are uniformly bounded, so in theory it should be easier to verify the integrability in (2.12). However this would require extra arguments and is not needed for our applications, so we do not consider it in this work.

1.2. Applications to variance asymptotics

We next present applications of the covariance formula to the asymptotics of the variance of the component count. For brevity we write $N_\star(R) = N_\star(\Lambda_R)$, where $\Lambda_R = [-R/2, R/2]^d$.

1.2.1. Known and conjectured results

As previously mentioned, the second-order properties of the component count are expected to depend on the covariance structure of the field, of which there are three main categories:

1. We say that f is *short-range correlated* if $K \in L^1(\mathbb{R}^d)$; an important example is the *Bargmann–Fock field* with $K(x) = e^{-|x|^2/2}$ (see [3] for background and motivation). By analogy with known results for other ‘local’ geometric statistics of the excursion/level sets (for instance their volume or Euler characteristic [20,22,23,29,30]), as well as with known results for the number of clusters in classical percolation [16,35,47], in this case it is expected that the variance of the component count $N_\star(R)$ has volume-order scaling $\asymp R^d$, and satisfies a central limit theorem (CLT).
2. We say that f is *regularly varying with index $\beta \in (0, d)$ and remainder L* if $K(x) = |x|^{-\beta} L(|x|)$ where L is a slowly varying function (see [9] for a definition); important examples are the *Cauchy fields* with covariance $K(x) = (1 + |x|^2)^{-\beta/2}$. The expected second-order properties are not clear in this case, although by analogy with the zero count for Gaussian processes [41] one might suppose that the variance has super-volume scaling $\asymp R^{2d-\beta} L(R)$ for most levels, with the possibility of a finite number of *anomalous levels* with lower-order variance (although this latter phenomenon is not well-understood, see [6, Section 2.2]). It is unclear whether a CLT holds in this case.
3. We say that f has *oscillating long-range correlations* if it is neither short-range correlated nor positively correlated. An important example is the *monochromatic random wave* for which $K(x) = |x|^{-(d/2-1)} J_{d/2-1}(|x|)$ where J_n is the order- n Bessel function; equivalently K is the Fourier transform of the normalised Lebesgue measure on the sphere \mathbb{S}^{d-1} . In $d = 2$ this is the *random plane wave* (RPW) with $K(x) = J_0(|x|)$. Again by analogy with known results for ‘local’ statistics [26,34], one might expect that for the monochromatic random wave the variance has super-volume scaling $\asymp R^{d+1}$ for most levels, with the possibility of anomalous levels with lower-order variance. Again it is unclear whether a CLT holds in this case.

The predicted behaviour described above has so far only been fully verified for a restricted class of short-range correlated fields. In particular, it was recently shown in [7] that, for a certain subclass of fields whose correlations decay as

$$|K(x)| \leq c(1 + |x|)^{-\beta}$$

for some $\beta > 9d$, $N_\star(R)$ satisfies a CLT with volume-order variance scaling. Although this subclass contains important examples such as the Bargmann–Fock field, the assumption on the decay of K is quite strong, and it appears the method in [7] cannot be extended, without significant new ideas, to cover the entire short-range correlated case (see [7, Remark 3.12]).

Outside of the short-range correlated fields covered by the result in [7], the existing second-order results are limited to variance bounds which are mostly sub-optimal and apply only to planar fields. We now briefly describe these, starting with lower bounds.

In [32] Nazarov and Sodin proved a polynomial lower bound

$$\text{Var}[N_\star(R)] \geq cR^\eta$$

valid for general planar fields with polynomially decaying correlations (to be more precise, they considered families $(f_n)_{n \geq 1}$ of Gaussian fields defined on the sphere \mathbb{S}^2 which converge locally, and only considered the nodal set, but we expect the proof extends, up to boundary effects, to the Euclidean setting we consider here). The exponent $\eta > 0$ was not quantified but is small. In [6] sharper results were proven for planar fields under stronger conditions. More precisely, if f is short-range correlated, and if $\int K(x) dx > 0$ and $\frac{d}{d\ell} \mu_\star(\ell) \neq 0$ where $\mu_\star(\ell)$ is the asymptotic mean number of components per unit volume at level ℓ , then

$$\text{Var}[N_\star(R)] \geq cR^2.$$

Further, in [5] the condition $\frac{d}{d\ell} \mu_\star(\ell) \neq 0$ was shown to hold for a large range of levels (including the zero level for excursion sets). It was also shown in [6] that if f is regularly varying with index $\beta \in (0, 2)$ and remainder L (to be more precise, the condition in [6] was given in terms of the spectral measure, but it is roughly equivalent) then, under the same derivative condition,

$$\text{Var}[N_\star(R)] \geq cR^{2-\beta}L(R),$$

and if f is the RPW then, under the same derivative condition but excluding level $\ell = 0$,

$$\text{Var}[N_\star(R)] \geq cR^3.$$

Note that all these bounds are expected to be of optimal order whenever they apply.

Turning to upper bounds, it is straightforward to establish that, in all dimensions,

$$(1.3) \quad \text{Var}[N_\star(R)] \leq cR^{2d}$$

using a comparison with critical points. More precisely, since each excursion (resp. level) set component contains (resp. surrounds) at least one critical point, the component count in a compact domain is bounded by the number of critical points in the domain. Since the latter quantity has a second moment of order $\asymp R^{2d}$ [18,19], we deduce (1.3). Note that this bound is only expected to be attained for very degenerate Gaussian fields (see [4,6] for examples). For the nodal level $\ell = 0$ of the random plane wave, Priya [38] recently improved the upper bound to $cR^{4-1/8}$. This bound is deduced from the exponential concentration of the nodal component count, and the proof does not extend to general levels/fields. While weaker concentration bounds have also been established for $N_\star(R)$ in general [8,40], these do not lead to improved bounds on the variance.

1.2.2. A formula for the variance: Monotonicity and convexity

We first specialise the covariance formula to the case of the variance, which enjoys additional monotonicity and convexity properties:

Theorem 2 (Variance formula for the component count). *Suppose Assumption 1 holds and $\ell \in \mathbb{R}$, $\star \in \{\text{ES}, \text{LS}\}$, and $R > 0$. Then the map*

$$(1.4) \quad t \mapsto \mathbb{E}[N_\star(R)N_\star'(R)], \quad t \in [0, 1],$$

is non-decreasing, convex, and absolutely continuous with derivative (almost everywhere)

$$(1.5) \quad \int_{\Lambda_R \times \Lambda_R} K(x-y)(dp_t^{++}(x,y) + dp_t^{--}(x,y) - dp_t^{+-}(x,y) - dp_t^{-+}(x,y)),$$

where $dp_t^{\pm\pm}$ are the ‘pivotal measures’ as in Theorem 1 in the case $B_1 = B_2 = \Lambda_R$ and $\star = \diamond$. In particular,

$$(1.6) \quad \text{Var}[N_\star(R)] = \int_0^1 \int_{\Lambda_R \times \Lambda_R} K(x-y)(dp_t^{++}(x,y) + dp_t^{--}(x,y) - dp_t^{+-}(x,y) - dp_t^{-+}(x,y)) dt.$$

Remark 2. It seems plausible that, in fact, $t \mapsto \mathbb{E}[N_\star(R)N_\star^t(R)]$ is continuously differentiable under the conditions of Theorem 2. One possible strategy to prove this statement would be to show that the pivotal intensities $p_t^{\pm\pm}(x, y)$ are continuous functions of $(t, x, y) \in [0, 1) \times B_1 \times B_2$ when restricted to pairs of substrata. We do not pursue such an argument here as it would not directly impact on our applications.

1.2.3. *Exact asymptotics in the short-range correlated case*

In the case $K \in L^1(\mathbb{R}^d)$, we can deduce from Theorem 2 the asymptotic behaviour of the variance:

Theorem 3. *Suppose Assumption 1 holds and let $\star \in \{\text{ES}, \text{LS}\}$ and $\ell \in \mathbb{R}$. If f is short-range correlated (i.e. $K \in L^1(\mathbb{R}^d)$) then as $R \rightarrow \infty$*

$$(1.7) \quad \text{Var}[N_\star(R)] = \sigma^2 R^d + o(R^d),$$

where $\sigma^2 \in [0, \infty)$ is the constant defined as

$$(1.8) \quad \sigma^2 = \int_0^1 \int_{\mathbb{R}^d} K(x) (\hat{I}_t^{++}(0, x) + \hat{I}_t^{--}(0, x) - \hat{I}_t^{+-}(0, x) - \hat{I}_t^{-+}(0, x)) dx dt$$

for the ‘stationary pivotal intensities’ $\hat{I}_t^{\pm\pm}$ defined in Section 2.4.

As mentioned in Section 1.2.1 above, the asymptotic (1.7) was already known for a subset of short-range correlated fields, including the Bargmann–Fock field [7], although using a very different ‘martingale’ approach. This approach led to an alternative representation for σ^2 (see [7, Eq. (3.34)]) which is interesting to compare to (1.8). However (1.7) is new for all short-range fields whose correlations do not decay faster than $|x|^{-9d}$, and the expression (1.8) is new in all cases.

Roughly speaking we deduce Theorem 3 from Theorem 2 by showing that the pivotal measures $dp_t^{\pm\pm}(x, y)$ can be replaced with their stationary counterparts $\hat{I}_t^{\pm\pm}(x, y) dx dy$. More precisely, this involves showing that (i) the contribution from the lower dimensional strata of Λ_R are bounded by $O(R^{d-1})$, and (ii) the dependence of the pivotal measures on the domain Λ_R is negligible for the bulk of points $x, y \in \Lambda_R$.

The previous result determines the asymptotic growth rate of the variance whenever $\sigma^2 > 0$. A natural question is whether this is always true:

Question 1. Is the constant $\sigma^2 \geq 0$ in Theorem 3 strictly positive in general?

At present we only know how to verify $\sigma^2 > 0$ for a restricted class of fields [6,7], which includes the Bargmann–Fock field, although we expect it to be true in general.

As mentioned in Remark 2 above, it may be possible to show that $t \mapsto \mathbb{E}[N_\star(R)N_\star^t(R)]$ is continuously differentiable. Assuming this, using the convexity in Theorem 2 and an argument similar to that in Theorem 3 one would obtain that

$$(1.9) \quad \begin{aligned} \sigma^2 &\geq \int_{\mathbb{R}^d} K(x) (\hat{I}_0^{++}(0, x) + \hat{I}_0^{--}(0, x) - \hat{I}_0^{+-}(0, x) - \hat{I}_0^{-+}(0, x)) dx \\ &= (\hat{I}^+ - \hat{I}^-)^2 \int_{\mathbb{R}^d} K(x) dx, \end{aligned}$$

where $\hat{I}^+ = \hat{I}_0^{++}(0, 0)^{1/2}$ and $\hat{I}^- = \hat{I}_0^{--}(0, 0)^{1/2}$ can be interpreted as ‘one-point stationary pivotal intensities’, and in the last step we used the fact that f and $f^0 = \tilde{f}$ are independent.

Such analysis, if carried out, would show that $\sigma^2 > 0$ as long as

$$(1.10) \quad \int_{\mathbb{R}^d} K(x) dx > 0 \quad \text{and} \quad \hat{I}^+ \neq \hat{I}^-.$$

It is interesting to compare (1.10) to the criteria for $\sigma^2 > 0$ previously established in [6,7]:

- The condition $\int_{\mathbb{R}^d} K(x) dx > 0$ was also required in [6,7], and is quite natural – it implies that the reproducing kernel Hilbert space of the field contains functions which are approximately constant – although we do not believe it to be necessary for $\sigma^2 > 0$.

- The condition $\hat{I}^+ \neq \hat{I}^-$ was also (implicitly) present in [6], although was not needed in [7]. In fact, we previously showed in [6] that, for a wide class of planar fields,

$$\frac{d\mu_\star(\ell)}{d\ell} = \hat{I}^+ - \hat{I}^-,$$

so that the condition $\hat{I}^+ \neq \hat{I}^-$ is equivalent to the condition $\frac{d}{d\ell}\mu_\star(\ell) \neq 0$ mentioned in Section 1.2.1 above. Hence if one were to make the above arguments rigorous, it would recover the bound $\text{Var}[N_\star(R)] \geq cR^2$ proven in [6] for planar fields, using a very different method.

1.2.4. *General variance bounds*

Our next application of Theorem 2 is to deduce a general upper bound on the variance that applies outside the short-range correlated case. By bounding the pivotal intensities with the critical point intensities (see (1.2)) we immediately obtain the following:

Corollary 1. *Suppose Assumption 1 holds and let $\ell \in \mathbb{R}$, $\star \in \{\text{ES}, \text{LS}\}$, and $R > 0$. Then*

$$(1.11) \quad \text{Var}[N_\star(R)] \leq \int_0^1 \int_{\Lambda_R \times \Lambda_R} |K(x - y)| dp_t(x, y) dt.$$

In Section 3 we prove that the critical point intensities $p_t(x, y)$ are integrable on the diagonal (i.e. as $t \rightarrow 1$ and $x - y \rightarrow 0$), from which we deduce the following upper bound:

Theorem 4 (Variance upper bound). *Suppose Assumption 1 holds and let $\ell \in \mathbb{R}$ and $\star \in \{\text{ES}, \text{LS}\}$. Then there exists a $c > 0$ such that, for all $R \geq 1$,*

$$(1.12) \quad \text{Var}[N_\star(R)] \leq cR^d \int_{\Lambda_{2R}} \tilde{K}(x) dx,$$

where $\tilde{K}(x) = \sup_{y:|y-x|\leq 1} |K(y)|$.

Remark 3. The proof of Theorem 4 shows that the constant c could be replaced by $c'(1 + \ell^{2d})$ for a $c' > 0$ depending only on the field f . The term $R^d \int_{\Lambda_{2R}} \tilde{K}(x) dx$ could also be replaced by $\int_{\Lambda_R \times \Lambda_R} |K(x - y)| dx dy$, where this double integral is understood as being over the box Λ_R as well as its boundary faces of all dimensions equipped with their respective Lebesgue measures.

To illustrate the optimality (or lack thereof) of this result, let us consider some examples:

Corollary 2. *Suppose Assumption 1 holds and let $\ell \in \mathbb{R}$ and $\star \in \{\text{ES}, \text{LS}\}$.*

1. *If f is regularly varying with index $\beta \in (0, d)$ and remainder L , there exists a $c > 0$ such that, for all $R \geq 1$,*

$$\text{Var}[N_\star(R)] \leq cR^{2d-\beta} L(R).$$

2. *If f is the monochromatic random wave, there exists a $c > 0$ such that, for all $R \geq 1$,*

$$\text{Var}[N_\star(R)] \leq cR^{(3d+1)/2}.$$

Proof. In the regularly varying case this follows directly from Theorem 4 and the additional fact that, as $R \rightarrow \infty$,

$$\int_{\Lambda_{2R}} \tilde{K}(x) dx \leq \int_{|x|\leq\sqrt{d}R} \tilde{K}(x) dx \sim \int_{|x|\leq\sqrt{d}R} K(x) dx \sim \frac{1}{d-\beta} (\sqrt{d}R)^{d-\beta} L(R)$$

by regular variation and Karamata’s integral theorem [9, Theorem 1.5.11]. In the monochromatic case we additionally use that, for $R \geq 1$,

$$\int_{\Lambda_{2R}} \tilde{K}(x) dx \leq \int_{|x|\leq\sqrt{d}R} \tilde{K}(x) dx = O(R^{(d+1)/2}),$$

since $K(x) = |x|^{-(d/2-1)} J_{d/2-1}(|x|) = O(|x|^{-(d-1)/2})$. □

Comparing with the discussion in the previous section, the bound in Theorem 4 is (conjecturally) of the correct order in the regularly varying case, but not necessarily in the oscillating case due to the fact that we have $\tilde{K} \approx |K|$ instead of K in the right-hand side of (1.12). Nevertheless, even in the oscillating case the attained bounds are best known. For instance, for the nodal set of the RPW the attained bound $cR^{7/2}$ improves the previously best-known bound $cR^{4-1/8}$ [38] (and moreover is valid at all levels).

Remark 4. As in the discussion around (1.9), if one knew the continuous differentiability of $t \mapsto \mathbb{E}[N_\star(R)N_\star^t(R)]$ (see Remark 2) one could use the convexity in Theorem 2 to obtain the alternative upper bound

$$(1.13) \quad \begin{aligned} \text{Var}[N_\star(R)] &\leq \lim_{t \rightarrow 1} \int_{\Lambda_R \times \Lambda_R} K(x-y) (dp_t^{++}(x,y) + dp_t^{--}(x,y) - dp_t^{+-}(x,y) - dp_t^{-+}(x,y)), \\ &\leq \lim_{t \rightarrow 1} \int_{\Lambda_R \times \Lambda_R} |K(x-y)| dp_t(x,y). \end{aligned}$$

This is comparable to the *Gaussian Poincaré inequality*

$$\text{Var}[f(X)] \leq \sum_{i,j} K(i,j) \mathbb{E} \left[\frac{\partial f(X)}{\partial X_i} \frac{\partial f(X)}{\partial X_j} \right] \leq \sum_{i,j} |K(i,j)| \left| \mathbb{E} \left[\frac{\partial f(X)}{\partial X_i} \frac{\partial f(X)}{\partial X_j} \right] \right|$$

for Gaussian vectors X and sufficiently smooth f (see, e.g., [15, Corollary 3.2]). However it is not clear whether the bound (1.13) is effective in our context, since our analysis in Section 3 does *not* show that $\int_{\Lambda_R \times \Lambda_R} dp_t(x,y)$ remains bounded as $t \rightarrow 1$ (c.f. (1.11) where we only require the integrability of $\int_{\Lambda_R \times \Lambda_R} dp_t(x,y)$ as $t \rightarrow 1$).

Remark 5 (General lower bounds). As in the previous remark (see also the discussion around (1.9)), if one knew the continuous differentiability of $t \mapsto \mathbb{E}[N_\star(R)N_\star^t(R)]$ one could use the convexity in Theorem 2 to obtain the general lower bound

$$(1.14) \quad \begin{aligned} \text{Var}[N_\star(R)] &\geq \int_{\Lambda_R \times \Lambda_R} K(x-y) (dp_0^{++}(x,y) + dp_0^{--}(x,y) - dp_0^{+-}(x,y) - dp_0^{-+}(x,y)) dx dy \\ &= \int_{\Lambda_R \times \Lambda_R} K(x-y) (dp^+(x) - dp^-(x))(dp^+(y) - dp^-(y)) dx dy, \end{aligned}$$

where dp^\pm are ‘one-point pivotal measures’ defined similarly to \hat{I}^\pm (we refrain from giving a formal definition), and the last step uses the independence of f and $f^0 = \tilde{f}$. This is similar in spirit to the bound in [12, Proposition 3.7] for smooth functions of finite Gaussian vectors, which is proven using a very different method.

1.3. Outline of the paper

In Section 2 we prove the covariance and variance formulae in Theorems 1 and 2, and also deduce Theorems 3 and 4 as a consequence, subject to some auxiliary results, namely the integrability of the pivotal measures and the topological stability of the component count. The integrability of the pivotal measures is proven in Section 3, and the topological stability is proven in Section 4. Finally in the [Appendix](#) we collect some non-degeneracy properties of smooth Gaussian fields.

2. An exact formula for the covariance and applications

In this section we establish the exact formula for the component count covariance stated in Theorem 1 above, and then deduce the various consequences of this formula. The proof makes use of auxiliary results (integrability of the pivotal measures and topological stability of the component count) which are proven in the following sections.

2.1. Stability of the component count

We begin by stating three deterministic topological stability properties of the component count. Roughly speaking the first states that the component count at level ℓ cannot change under small perturbations if there are no critical points at level ℓ . The second states that the component count can change by at most one if the function is perturbed near a non-degenerate critical point. The third states that the change in the component count when perturbing near a critical point is the same for all sufficiently large boxes. These properties are all consequences of standard results of stratified Morse theory (see [21] for a comprehensive background) as we will explain below. However in the interest of completeness, we present proofs of our results in Section 4.

To formalise the stability properties we require the notion of stratified sets. A box $B = [a_1, b_1] \times \dots \times [a_d, b_d]$ may be considered as a *stratified set* by partitioning it into the finite collection $\mathcal{F} = (F_i)$ of (the interiors of) its faces of all dimensions $0 \leq m \leq d$, which we refer to as the *strata*. We let $\mathcal{F}_0 \subset \mathcal{F}$ be the strata of dimension $m = 0$. A *stratified box* will be any such B considered as a stratified set. We equip each stratum $F \in \mathcal{F} \setminus \mathcal{F}_0$ with its Lebesgue measure dv_F , and each stratum $F \in \mathcal{F}_0$ with a Delta mass, also denoted dv_F .

Let B be a stratified box and let U be an open neighbourhood of B . For $x \in B$ and $g \in C^1(U)$, $\nabla_F g(x)$ denotes the derivative of g restricted to the unique stratum F containing x . A (stratified) *critical point* of g is a point $x \in F$ such that $\nabla_F g(x) = 0$. The *level* of this critical point is the value $g(x)$. The critical point is *non-degenerate* provided that $\det \nabla_F^2 g(x) \neq 0$ and if x is in the closure of a higher dimensional stratum F' , then $\nabla_{F'} g(y) \not\rightarrow 0$ as $y \rightarrow x$, $y \in F'$. We take any point belonging to a stratum of dimension $m = 0$ to be a non-degenerate critical point by definition.

Extending our previous notation, for a function $g \in C^1(U)$ we let $N_\star(B, g, \ell)$ denote the component count of g at level ℓ inside B , i.e. the number of connected components of $\{g \geq \ell\}$ (if $\star = \text{ES}$) or $\{g = \ell\}$ (if $\star = \text{LS}$) which intersect B but not its boundary.

The first two stability properties are the following:

Lemma 1 (Stability away from critical points). *Suppose that $g \in C^1(U)$ has no (stratified) critical points at level $\ell \in \mathbb{R}$. There exists a neighbourhood of g in $C^1(U)$ such that for all g' in the neighbourhood*

$$N_\star(B, g', \ell) = N_\star(B, g, \ell).$$

Lemma 2 (Change near critical points). *Suppose $g \in C^2(U)$ has a non-degenerate stratified critical point at x at level ℓ . Then there exists a unique pair $(k, k') \in \mathbb{N}_0 \times \mathbb{N}_0$, with $|k - k'| \leq 1$, satisfying the following: for any sufficiently small open neighbourhood W of x , any function $h \in C_c^2(W)$ such that $h(x) > 0$, and every sufficiently small $\delta > 0$,*

$$N_\star(B, g - \delta h, \ell) = k \quad \text{and} \quad N_\star(B, g + \delta h, \ell) = k'.$$

Lemma 1 follows quite directly from Thom's first isotopy lemma, as will become apparent from our proof in Section 4. More general forms of this result are proven in Part I, Chapter 4 and Chapter 7.4 of [21]. Lemma 2 follows from the very general 'Main Theorem' (Theorem 3.5.4) of [21], although again we will give a proof in our particular setting in Section 4.

In particular, the second stability property allows us to classify all non-degenerate critical points depending on whether the component count increases by one, decreases by one, or stays the same, after perturbation. More precisely, define the index sets

$$P^+ = \{(k, k') \in \mathbb{N}_0 \times \mathbb{N}_0 : k' - k = 1\}, \quad P^- = \{(k, k') \in \mathbb{N}_0 \times \mathbb{N}_0 : k' - k = -1\},$$

$$P = P^+ \cup P^- \quad \text{and} \quad P^0 = \{(k, k') \in \mathbb{N}_0 \times \mathbb{N}_0 : k' - k = 0\}.$$

We say that a non-degenerate critical point is *positively pivotal* (for the component count) if $(k, k') \in P^+$, *negatively pivotal* if $(k, k') \in P^-$, and *not pivotal* if $(k, k') \in P^0$, where (k, k') is the pair guaranteed to exist by Lemma 2. Note that these definitions depend on the box B . When we wish to emphasise this we will say that the critical point is (Δ, B) -pivotal for $\Delta = +, -, 0$ respectively.

Finally we state our third stability property, which also follows as a consequence of the 'Main Theorem' of [21]:

Lemma 3 (Asymptotic topological stabilisation). *Suppose $g \in C^2(\mathbb{R}^d)$ has a non-degenerate critical point at x at level ℓ . Then there exists $r > 0$ and $\Delta \in \{+, -, 0\}$ such that, for all boxes B containing $x + \Lambda_r$, the critical point x is (Δ, B) -pivotal.*

2.2. Proof of the covariance formula

In this subsection we prove the covariance formula in Theorem 1. Throughout the rest of this section we fix $\ell \in \mathbb{R}$, boxes $B_1, B_2 \subset \mathbb{R}^d$, and $\star, \diamond \in \{\text{ES}, \text{LS}\}$, and drop these variables from the notation. For brevity we write $N = N_\star(B_1)$ and $N^t = N_\diamond^t(B_2)$, recalling that $N^1 = N_\diamond^1(B_2) = N_\diamond(B_2)$.

2.2.1. Gaussian conditioning

Throughout this paper we will make use of conditioning for Gaussian vectors; that is, given two jointly Gaussian vectors X and Y of length m_1 and m_2 respectively we will consider variables of the form $(Y|X = x)$ for some $x \in \mathbb{R}^{m_1}$. Since $\{X = x\}$ may have measure zero, to make sense of this definition we condition on the event $|X - x| < \epsilon$ and take $\epsilon \downarrow 0$. This is sometimes known as ‘vertical window’ conditioning (and is the most common form of conditioning on a continuous variable taking a particular value). This form of conditioning is always well-defined whenever X is non-degenerate. To see this, one can define the variable Z by

$$Y = Z + \Sigma_{YX} \Sigma_X^{-1} X,$$

where $\Sigma_{YX} = \text{Cov}[Y, X]$ and $\Sigma_X = \text{Cov}[X]$. Basic linear algebra shows that Z is uncorrelated with X and hence, by Gaussianity, independent. Then by the above equation $(Y|X = x) = Z + \Sigma_{YX} \Sigma_X^{-1} x$. From this we see that the conditioned variable is also Gaussian and we can explicitly compute its mean and covariance matrix as

$$\mathbb{E}[Y|X = x] = \mathbb{E}[Y] + \Sigma_{YX} \Sigma_X^{-1} (x - \mathbb{E}[X]) \quad \text{and} \quad \text{Cov}[Y|X = x] = \text{Cov}[Y] - \Sigma_{YX} \Sigma_X^{-1} \Sigma_{YX}^T.$$

This relationship between the conditional and unconditional moments is sometimes known as *Gaussian regression*.

2.2.2. The pivotal measures

We begin by defining the pivotal measures $dp_t^{\pm\pm}(x, y)$ that appear in Theorem 1.

Throughout this section we view B_1 and B_2 as stratified boxes in the sense of Section 2.1 and fix open neighbourhoods $U_1, U_2 \subset \mathbb{R}^d$ of B_1 and B_2 respectively. We denote by $\mathcal{F}^{(1)}, \mathcal{F}^{(2)}$ the set of strata of B_1, B_2 respectively, $\mathcal{F}_0^{(i)} \subset \mathcal{F}^{(i)}$ the subset of strata of dimension $m = 0$, and write $\mathcal{F}^{(1,2)} = \mathcal{F}^{(1)} \times \mathcal{F}^{(2)}$. Recall that each stratum $F \in \mathcal{F}^{(i)} \setminus \mathcal{F}_0^{(i)}$ is equipped with its Lebesgue measure dv_F , and each stratum $F \in \mathcal{F}_0^{(i)}$ with a delta mass, also denoted dv_F .

For $x \in B_1$ we define $F_x = F_x(B_1)$ to be the unique stratum of B_1 containing x and for $y \in B_2$ we define $F_y = F_y(B_2)$ similarly. Let us fix momentarily $t \in [0, 1)$, $x \in B_1$ and $y \in B_2$. Consider the fields f and f^t conditioned on

$$(2.1) \quad (f(x), f^t(y), \nabla_{F_x} f(x), \nabla_{F_y} f^t(y)) = (\ell, \ell, 0, 0).$$

This conditioning is well-defined since the vector on the left-hand side of (2.1) is non-degenerate even if $x = y$ (see Lemma 15 for a proof).

For a square matrix M , we write $|M|$ to denote the absolute value of the determinant of M . Using the fact that $(f(x), \nabla f(x), f(y), \nabla f(y))$ is a non-degenerate Gaussian vector (see Remark 14) we see that under the conditioning (2.1) the partial derivatives of f at x in directions normal to F_x must be non-zero (and similarly for f^t at y). Therefore, assuming that $|\nabla_{F_x}^2 f(x)|$ and $|\nabla_{F_y}^2 f^t(y)|$ are non-zero, we let $A(\iota_1, \iota_2)$ denote the event that the pairs (k, k') , guaranteed to exist by Lemma 2 for f and f^t , are given by $\iota_1, \iota_2 \in P^+ \cup P^- \cup P^0$ respectively. In particular we have the decomposition

$$(2.2) \quad |\nabla_{F_x}^2 f(x) \nabla_{F_y}^2 f^t(y)| = |\nabla_{F_x}^2 f(x) \nabla_{F_y}^2 f^t(y)| \sum_{\iota_1, \iota_2 \in P^+ \cup P^- \cup P^0} \mathbb{1}_{A(\iota_1, \iota_2)}.$$

Definition 1. For each $\iota_1, \iota_2 \in P^+ \cup P^- \cup P^0$ we define the associated *two-point intensity*

$$I_t^{\iota_1, \iota_2}(x, y) = \varphi_t(x, y) \tilde{\mathbb{E}}[|\nabla_{F_x}^2 f(x) \nabla_{F_y}^2 f^t(y)| \mathbb{1}_{A(\iota_1, \iota_2)}],$$

where $\varphi_t(x, y)$ denotes the density of the conditioning (2.1) (i.e. the density of the non-degenerate Gaussian vector $(f(x), f^t(y), \nabla_{F_x} f(x), \nabla_{F_y} f^t(y))$ evaluated at $(\ell, \ell, 0, 0)$), and $\tilde{\mathbb{E}}$ denotes expectation under this conditioning. If $F_x \in \mathcal{F}_0^{(1)}$ (the set of strata with dimension 0), then by convention we remove $\nabla_{F_x} f$ from the conditioning (2.1) and corresponding density $\varphi_t(x, y)$ and set $|\nabla_{F_x}^2 f(x)| = 1$, and similarly for $F_y \in \mathcal{F}_0^{(2)}$.

We also define a corresponding measure equal to the product of Lebesgue measures on each pair of strata weighted by this intensity:

$$dp_t^{\iota_1, \iota_2}(x, y) = I_t^{\iota_1, \iota_2}(x, y) dv_{F_x}(x) dv_{F_y}(y).$$

The *pivotal intensities* $I_t^{\pm, \pm}(x, y)$, and their corresponding *pivotal measures* $dp_t^{\pm, \pm}(x, y)$, are then defined as

$$(2.3) \quad I_t^{\pm, \pm}(x, y) = \sum_{\iota_1 \in P^\pm, \iota_2 \in P^\pm} I_t^{\iota_1, \iota_2}(x, y) \quad \text{and} \quad dp_t^{\pm, \pm}(x, y) = I_t^{\pm, \pm}(x, y) dv_{F_x}(x) dv_{F_y}(y).$$

Let us observe that

$$(2.4) \quad I_t^{+,+}(x, y) + I_t^{-,-}(x, y) + I_t^{+,-}(x, y) + I_t^{-,+}(x, y) = \sum_{\iota_1 \in P, \iota_2 \in P} I_t^{\iota_1, \iota_2}(x, y),$$

and

$$(2.5) \quad I_t^{+,+}(x, y) + I_t^{-,-}(x, y) - I_t^{+,-}(x, y) - I_t^{-,+}(x, y) = \sum_{\iota_1 \in P, \iota_2 \in P} \sigma(\iota_1)\sigma(\iota_2)I_t^{\iota_1, \iota_2}(x, y),$$

where $\sigma(\iota) = \pm 1$ if $\iota \in P^\pm$.

2.2.3. The critical point measure

We next define the *critical point intensity and measure* as

$$(2.6) \quad I_t(x, y) = \varphi_t(x, y) \tilde{\mathbb{E}}[|\nabla_{F_x}^2 f(x) \nabla_{F_y}^2 f^t(y)|]$$

and

$$dp_t(x, y) = I_t(x, y) dv_{F_x}(x) dv_{F_y}(y)$$

respectively. Note that (2.2)–(2.4) imply the bound (c.f. (1.2))

$$(2.7) \quad I_t(x, y) = \sum_{\iota_1, \iota_2 \in P \cup P^0} I_t^{\iota_1, \iota_2}(x, y) \geq \sum_{\iota_1, \iota_2 \in P} I_t^{\iota_1, \iota_2}(x, y) = I_t^{+,+}(x, y) + I_t^{-,-}(x, y) + I_t^{+,-}(x, y) + I_t^{-,+}(x, y).$$

Moreover in Section 3 we prove that the pivotal measures are integrable on the diagonal and uniformly bounded away from the diagonal:

Proposition 2 (Integrability and uniform bound for the critical point intensity). *There exists a constant $c > 0$, depending only on the field f and the level $\ell \in \mathbb{R}$, such that, for all $(F_1, F_2) \in \mathcal{F}^{(1,2)}$,*

$$(2.8) \quad \sup_{x \in F_1} \int_0^1 \int_{\substack{y \in F_2 \\ |x-y| \leq 1}} I_t(x, y) dv_{F_y}(y) dt \leq c$$

and

$$(2.9) \quad \sup_{t \in [0, 1]} \sup_{\substack{(x, y) \in F_1 \times F_2 \\ |x-y| \geq 1}} I_t(x, y) \leq c.$$

In the following section we shall need a slight generalisation of this integrability, which we state now. This applies to a ‘stationary’ version of pivotal intensity, i.e. one that does not depend on the boxes B_1 and B_2 . Specifically, for $x, y \in \mathbb{R}^d$, $t \in [0, 1)$, and $k \in \mathbb{N}$, we define

$$(2.10) \quad \hat{I}_t^{(k)}(x, y) = \hat{\varphi}_t(x, y) \widehat{\mathbb{E}}[|\nabla^2 f(x) \nabla^2 f^t(y)|^k]^{1/k},$$

where $\hat{\varphi}_t(x, y)$ denotes the density of the conditioning

$$(2.11) \quad (f(x), f^t(y), \nabla f(x), \nabla f^t(y)) = (\ell, \ell, 0, 0),$$

and $\widehat{\mathbb{E}}$ denotes expectation under this conditioning. Abbreviating $\widehat{I}_t(x, y) := \widehat{I}_t^{(1)}(x, y)$, we note that if x and y lie in the interior of the boxes B_1 and B_2 respectively, then $\widehat{I}_t(x, y) = \widehat{I}_t(0, x - y)$ coincides with the pivotal intensity $I_t(x, y)$ defined in (2.6).

Proposition 3 (Bounds for the stationary critical point intensity). *There exists a constant $c > 0$, depending only on the field f , the level $\ell \in \mathbb{R}$, and $k \in \mathbb{N}$, such that*

$$\int_0^1 \int_{|u| \leq 1} \widehat{I}_t^{(k)}(0, u) du dt \leq c \quad \text{and} \quad \sup_{t \in [0, 1]} \sup_{|u| \geq 1} \widehat{I}_t^{(k)}(0, u) \leq c.$$

Remark 6. We only use this proposition in the case $k = 1, 2$, but formulate its general statement to convey that $k \leq 2$ plays no role in the proof

2.2.4. *The covariance formula for component count exceedances*

As mentioned above, we deduce the variance formula in Theorem 2 as a consequence of a covariance formula for ‘topological events’ established in [8]; in particular we shall apply this formula to the component count tail events $\{N \geq k_1\}$ and $\{N^t \geq k_2\}$.

For $k \in \mathbb{N}$ define

$$P_k^+ = \{(k - 1, k)\} \subset P^+ \quad \text{and} \quad P_k^- = \{(k, k - 1)\} \subset P^-,$$

and $P_k = P_k^+ \cup P_k^-$.

Proposition 4 (Special case of [8, Theorem 2.14]). *For each $t \in [0, 1]$ and $k_1, k_2 \in \mathbb{N}$,*

$$\text{Cov}[\mathbb{1}_{N \geq k_1}, \mathbb{1}_{N^t \geq k_2}] = \int_0^t \int_{\Lambda_R \times \Lambda_R} K(x - y) \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \sigma(\iota_1)\sigma(\iota_2) dp_s^{\iota_1, \iota_2}(x, y) ds,$$

recalling that $\sigma(\iota) = \pm 1$ if $\iota \in P^\pm$.

Remark 7. In Proposition 4 we could have omitted the strata $\mathcal{F}_0^{(1)}, \mathcal{F}_0^{(2)}$ of dimension 0 from the definition of the pivotal measures, because critical points on such strata are never pivotal for the component count (i.e if the field is conditioned to have a critical point on a dimension 0 strata, then almost surely $A(\iota_1, \iota_2)$ occurs for some $\iota_1, \iota_2 \in P^0$). However this is a specific feature of the way we defined the component count (i.e. excluding boundary components) and the convexity of B_1, B_2 , so we prefer not to do this.

Proof. This is [8, Theorem 2.14] applied to the stratified domains B_1, B_2 (which is a ‘tame stratification’ in the sense of [8]) and the topological events $\{N \geq k_1\}$ and $\{N \geq k_2\}$, since one can check (recalling Lemma 2) that the ‘pivotal intensities’ $I_t^+(x, y)$ and $I_t^-(x, y)$ as defined in [8, Theorem 2.14] coincide, respectively, with

$$\sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \mathbb{1}_{\sigma(\iota_1) = \sigma(\iota_2)} I_t^{\iota_1, \iota_2}(x, y) \quad \text{and} \quad \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \mathbb{1}_{\sigma(\iota_1) \neq \sigma(\iota_2)} I_t^{\iota_1, \iota_2}(x, y).$$

Strictly speaking an application of [8, Theorem 2.14] only gives the limiting case $t \uparrow 1$ of this formula, that is,

$$\begin{aligned} \text{Cov}[\mathbb{1}_{N \geq k_1}, \mathbb{1}_{N \geq k_2}] &= \lim_{t \uparrow 1} \text{Cov}[\mathbb{1}_{N \geq k_1}, \mathbb{1}_{N^t \geq k_2}] \\ &= \lim_{t \uparrow 1} \int_0^t \int_{\Lambda_R \times \Lambda_R} K(x - y) \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \sigma(\iota_1)\sigma(\iota_2) dp_s^{\iota_1, \iota_2}(x, y) ds, \end{aligned}$$

but the desired formula for all $t \in [0, 1]$ is established in the final step of the proof of [8, Theorem 2.14] immediately before taking the $t \uparrow 1$ limit. □

2.2.5. Proof of the covariance formula

To deduce Theorem 1 from Proposition 4 we make use of the integrability of the critical point measure

$$(2.12) \quad \int_0^1 \int_{B_1 \times B_2} dp_t(x, y) dt = \int_0^1 \int_{B_1 \times B_2} \sum_{F_1, F_2 \in \mathcal{F}} I_t(x, y) dv_{F_1}(x) dv_{F_2}(y) dt < \infty,$$

which follows from (2.8)–(2.9) above.

We also make use of the first stability property from the previous section applied to the interpolation f^t and the component count on the box B_2 :

Lemma 4. *Let f satisfy Assumption 1, then $N_\diamond^t(B_2) \rightarrow N_\diamond(B_2)$ almost surely as $t \rightarrow 1$.*

Proof. By Lemma 13, $(f(x), \nabla f(x))$ is non-degenerate for any x . Combining this with the fact that f is C^2 -smooth, we may apply Bulinskaya’s lemma ([1, Lemma 11.2.10]), which states that f has no (stratified) critical points on B_2 at level ℓ almost surely. (To be specific, this follows from applying Bulinskaya’s lemma on each stratum of B_2 .) The result then follows from Lemma 1. \square

Proof of Theorem 1. Recalling that $N^1 = N_\diamond(B_2)$, we seek a formula for $\text{Cov}[N, N^1]$.

We first establish that, for every $t \in [0, 1)$,

$$(2.13) \quad \text{Cov}[N, N^t] = \int_0^t \int_{B_1 \times B_2} K(x - y) (dp_s^{++}(x, y) + dp_s^{--}(x, y) - dp_s^{+-}(x, y) - dp_s^{-+}(x, y)) ds.$$

For this we fix a truncation parameter $\tau \in \mathbb{N}$ and consider the variables

$$\min\{N, \tau\} = \sum_{k=1}^\tau \mathbb{1}_{N \geq k} \quad \text{and} \quad \min\{N^t, \tau\} = \sum_{k=1}^\tau \mathbb{1}_{N^t \geq k}.$$

Applying Proposition 4, and interchanging the finite summation with the integrals, we deduce that

$$(2.14) \quad \text{Cov}[\min\{N, \tau\}, \min\{N^t, \tau\}] = \int_0^t \int_{B_1 \times B_2} K(x - y) \sum_{k_1, k_2=1}^\tau \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \sigma(\iota_1)\sigma(\iota_2) dp_s^{\iota_1, \iota_2}(x, y) ds.$$

Now consider a pair of strata $(F_1, F_2) \in \mathcal{F}^{(1,2)}$. For every $(x, y) \in F_1 \times F_2$ and $s \in [0, 1)$, we have

$$\lim_{\tau \rightarrow \infty} \sum_{k_1, k_2=1}^\tau \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \sigma(\iota_1)\sigma(\iota_2) I_s^{\iota_1, \iota_2}(x, y) = \sum_{\iota_1, \iota_2 \in P} \sigma(\iota_1)\sigma(\iota_2) I_s^{\iota_1, \iota_2}(x, y),$$

and also, by (2.7), the uniform bound

$$\left| \sum_{k_1, k_2=1}^\tau \sum_{\iota_1 \in P_{k_1}, \iota_2 \in P_{k_2}} \sigma(\iota_1)\sigma(\iota_2) I_s^{\iota_1, \iota_2}(x, y) \right| \leq I_s(x, y),$$

which is integrable on $(s, x, y) \in [0, t] \times F_1 \times F_2$ by (2.12). Hence by dominated convergence we deduce that the limit of (2.14) as $\tau \rightarrow \infty$ is equal to

$$(2.15) \quad \begin{aligned} & \int_0^t \int_{B_1 \times B_2} K(x - y) \sum_{\iota_1, \iota_2 \in P} \sigma(\iota_1)\sigma(\iota_2) dp_s^{\iota_1, \iota_2}(x, y) ds \\ &= \int_0^t \int_{B_1 \times B_2} K(x - y) (dp_s^{++}(x, y) + dp_s^{--}(x, y) - dp_s^{+-}(x, y) - dp_s^{-+}(x, y)) ds, \end{aligned}$$

where in the last equality we used (2.5). Since $\mathbb{E}[N^2], \mathbb{E}[(N^t)^2] < \infty$, taking $\tau \rightarrow \infty$ on both sides of (2.14) establishes (2.13).

In light of (2.13), it remains to prove that

$$(2.16) \quad \lim_{t \uparrow 1} \text{Cov}[N, N^t] = \text{Cov}[N, N^1].$$

By our argument above,

$$(2.17) \quad \lim_{t \uparrow 1} \text{Cov}[N, N^t] = \lim_{t \uparrow 1} \lim_{\tau \rightarrow \infty} \text{Cov}[\min\{N, \tau\}, \min\{N^t, \tau\}].$$

Consider taking these limits in the opposite order: by Lemma 4 $N^t \rightarrow N^1$ almost surely, then applying the dominated convergence theorem twice (using square integrability of N, N^1) we see that

$$\lim_{\tau \rightarrow \infty} \lim_{t \uparrow 1} \text{Cov}[\min\{N, \tau\}, \min\{N^t, \tau\}] = \lim_{\tau \rightarrow \infty} \text{Cov}[\min\{N, \tau\}, \min\{N^1, \tau\}] = \text{Cov}[N, N^1].$$

Therefore by (2.17), we can establish (2.16) by justifying this interchange of limits. Let us denote

$$g(t, \tau) = \text{Cov}[\min\{N, \tau\}, \min\{N^t, \tau\}],$$

then by the Moore–Osgood theorem we can say that $\lim_{\tau \rightarrow \infty} \lim_{t \uparrow 1} g(t, \tau) = \lim_{t \uparrow 1} \lim_{\tau \rightarrow \infty} g(t, \tau)$ provided that

1. $\lim_{\tau \rightarrow \infty} g(t, \tau)$ exists for every $t \in [0, 1)$, and
2. $\lim_{t \uparrow 1} g(t, \tau)$ exists uniformly over τ .

We have already shown that $\lim_{\tau \rightarrow \infty} g(t, \tau)$ is given by (2.15) for each t , verifying the first requirement. Turning to the second requirement, let $\tau \in \mathbb{N}$ and $0 \leq t_1 < t_2 < 1$, by (2.14) and (2.7)

$$\begin{aligned} |g(t_1, \tau) - g(t_2, \tau)| &\leq \int_{t_1}^{t_2} \int_{B_1 \times B_2} |K(0)| \sum_{k_1, k_2=1}^{\tau} \sum_{t_1 \in P_{k_1}, t_2 \in P_{k_2}} dp_s^{t_1, t_2}(x, y) ds \\ &\leq |K(0)| \int_{t_1}^{t_2} \int_{B_1 \times B_2} \sum_{k_1, k_2=1}^{\infty} \sum_{t_1 \in P_{k_1}, t_2 \in P_{k_2}} dp_s^{t_1, t_2}(x, y) ds \leq |K(0)| \int_{t_1}^{t_2} \int_{B_1 \times B_2} dp_s(x, y) ds. \end{aligned}$$

Hence, by (2.12), $g(t, \tau)$ is uniformly (in τ) Cauchy as $t \uparrow 1$ justifying the Moore–Osgood theorem. Combining (2.13) and (2.16) completes the proof of Theorem 1. □

2.3. Specialisation to the variance and the general upper bound

We next specialise the formula to the case of the variance and deduce the general upper bound, stated in Theorems 2 and 4 respectively.

Compared to the covariance formula, the specialisation to the variance has additional monotonicity and convexity properties. These are essentially due to Chatterjee, who proved this result for functions of finite-dimensional Gaussian vectors [13,14]. Here we state a general version for continuous (not necessarily stationary) Gaussian fields:

Proposition 5 (Monotonicity and convexity of the interpolation). *Let $D \subseteq \mathbb{R}^d$, let f and \tilde{f} be independent copies of a continuous Gaussian field on D , and define the interpolation $f^t = tf + \sqrt{1-t^2}\tilde{f}$, $t \in [0, 1]$. Let $G : C^0(D) \rightarrow \mathbb{R}$ be such that $\mathbb{E}[G(f)^2] < \infty$. Then*

$$t \mapsto \text{Cov}[G(f), G(f^t)], \quad t \in [0, 1],$$

is non-decreasing and convex.

Proof. As mentioned above, the corresponding result in the finite-dimensional setting is due to Chatterjee. More precisely, let X and \tilde{X} be n -dimensional i.i.d. standard Gaussian vectors, define $X^t = tX + \sqrt{1-t^2}\tilde{X}$, $t \in [0, 1]$, and let $G : \mathbb{R}^n \rightarrow \mathbb{R}$ be absolutely continuous with $G(X) \in L^2$ and $\|\nabla G(X)\|_2 \in L^2$. Then [13, Lemma 3.5] states that

$$t \mapsto \text{Cov}[G(X), G(X^t)] \quad \text{and} \quad t \mapsto \text{Cov}[\partial G(X)/\partial X_i, \partial G(X^t)/\partial X_i^t], \quad 1 \leq i \leq n,$$

are non-decreasing on $[0, 1]$ (note that [13, Lemma 3.5] uses a different parameterisation for the interpolation). Recalling that (see Proposition 1)

$$(2.18) \quad \text{Cov}[G(X), G(X^t)] = \int_0^t \sum_{1 \leq i \leq n} \mathbb{E} \left[\frac{\partial G(X)}{\partial X_i} \frac{\partial G(X^s)}{\partial X_i^s} \right] ds$$

this proves the result.

The extension to a general Gaussian vector X follows immediately by writing $X = QZ$, where Q is a square-root of the covariance matrix of X and Z is an i.i.d. Gaussian vector, and applying the result to $G \mapsto G \circ Q$.

It remains to extend the result to arbitrary functions of continuous Gaussian fields. For this we will repeatedly use the fact that the pointwise limit of a sequence of non-decreasing and convex functions on $[0, 1]$ is also non-decreasing and convex.

First we reduce to the case of bounded G by taking the limit as $\tau \rightarrow \infty$ of the maps

$$t \mapsto \text{Cov}[G(f)\mathbb{1}_{|G(f)| \leq \tau}, G(f^t)\mathbb{1}_{|G(f^t)| \leq \tau}], \quad t \in [0, 1],$$

which converge pointwise to $t \mapsto \text{Cov}[G(f), G(f^t)]$ since $G(f), G(f^t) \in L^2$.

Next we reduce to the finite-dimensional case. Let D_n be a finite subset of D such that $D_n \subseteq D_{n+1}$ and $\bigcup_{n \in \mathbb{N}} D_n$ is dense in D . We then define the random variables

$$G_n = \mathbb{E}[G(f) \mid f|_{D_n}] \quad \text{and} \quad G_n^t = \mathbb{E}[G(f^t) \mid f^t|_{D_n}].$$

Since $(G_n)_{n \in \mathbb{N}}$ (resp. $(G_n^t)_{n \in \mathbb{N}}$) is a bounded martingale, by martingale convergence there exists a random variable G_∞ (resp. G_∞^t) such that $G_n \rightarrow G_\infty$ (resp. $G_n^t \rightarrow G_\infty^t$) almost surely and in L^2 . By the continuity of f , $G_\infty = G(f)$ and $G_\infty^t = G(f^t)$ almost surely, and we conclude that as $n \rightarrow \infty$ the maps

$$t \mapsto \text{Cov}[G_n, G_n^t], \quad t \in [0, 1],$$

converge pointwise to $t \mapsto \text{Cov}[G(f), G(f^t)]$.

Finally we reduce to the case of smooth $G : \mathbb{R}^n \rightarrow R$ with bounded derivatives of all orders by defining, for $\varepsilon > 0$, the mollification

$$G_\varepsilon(x) = \mathbb{E}[G(x + \varepsilon \bar{X})] = (G * \varphi_\varepsilon)(x),$$

where \bar{X} is an independent i.i.d. standard Gaussian vector, φ_ε is the density of $\varepsilon \bar{X}$, and $*$ denotes convolution. Since φ_ε is smooth with derivatives in $L^1(\mathbb{R}^n)$ and G is bounded, by Young's convolution inequality G_ε is smooth with bounded derivatives of all order. Moreover, since G is bounded, $G_\varepsilon(x) \rightarrow G(x)$ for almost every x , and also $G_\varepsilon(X) \rightarrow G(X)$ and $G_\varepsilon(X^t) \rightarrow G(X^t)$ in L^2 . Hence we deduce that, as $\varepsilon \rightarrow 0$, the maps

$$t \mapsto \text{Cov}[G_\varepsilon(X), G_\varepsilon(X^t)], \quad t \in [0, 1],$$

converge pointwise to $t \mapsto \text{Cov}[G(X), G(X^t)]$, which completes the proof. □

Remark 8. One can deduce from the formula (2.18) that, in the finite-dimensional case, $t \mapsto \text{Cov}[G(X), G(X^t)]$ is absolutely continuous if G is absolutely continuous with $\|\nabla G(X)\|_2 \in L^2$, although unlike the monotonicity and convexity it is not clear that this property transfers to the continuum limit.

Remark 9. The monotonicity properties in Proposition 5 are special to the case of the variance, i.e. it is not true in general that

$$t \mapsto \text{Cov}[G_1(X), G_2(X^t)], \quad t \in [0, 1],$$

is non-decreasing and convex, even for finite-dimensional X .

Proof of Theorem 2. From the proof of Theorem 1 (specifically (2.13)) in the case that $\star = \diamond$ and $B_1 = B_2 = \Lambda_R$, we see that $t \mapsto \text{Cov}[N, N^t]$ is absolutely continuous with derivative (1.5). By Proposition 5 this map is also convex and non-decreasing, completing the proof of the theorem. □

Proof of Theorem 4. Recall that Corollary 1 states that

$$\text{Var}[N_\star(R)] \leq \int_0^1 \int_{\Lambda_R \times \Lambda_R} |K(x - y)| dp_t(x, y) dt.$$

Summing over pairs of strata $F_1, F_2 \in \mathcal{F}$, partitioning the integral over $(x, y) \in F_1 \times F_2$ according to whether or not $|x - y| \leq 1$, and applying Proposition 2, the above is at most

$$\begin{aligned} & \sum_{F_1, F_2 \in \mathcal{F}} K(0) \int_0^1 \int_{\substack{(x,y) \in F_1 \times F_2 \\ |x-y| \leq 1}} I_t(x, y) dv_{F_1}(x) dv_{F_2}(y) dt + \int_0^1 \int_{\substack{(x,y) \in F_1 \times F_2 \\ |x-y| \geq 1}} |K(x - y)| I_t(x, y) dv_{F_1}(x) dv_{F_2}(y) dt \\ & \leq \sum_{F_1, F_2 \in \mathcal{F}} K(0) \int_{x \in F_1} c dv_{F_1}(x) + \int_{(x,y) \in F_1 \times F_2} c |K(x - y)| dv_{F_1}(x) dv_{F_2}(y) \\ & \leq c_1 \left(\max\{1, R^d\} + \sum_{F_1, F_2 \in \mathcal{F}} \int_{F_1 \times F_2} |K(x - y)| dv_{F_1}(x) dv_{F_2}(y) \right), \end{aligned}$$

for a constant $c_1 > 0$. Moreover there is a $c_d > 0$ depending only on the dimension such that, for all $R \geq 1$,

$$\sum_{F_1, F_2 \in \mathcal{F}} \int_{F_1 \times F_2} |K(x - y)| dv_{F_1}(x) dv_{F_2}(y) \leq c_d \int_{\Lambda_R \times \Lambda_R} |\tilde{K}(x - y)| dx dy \leq c_d R^d \int_{\Lambda_{2R}} |\tilde{K}(x)| dx.$$

Combining the above gives the result. □

2.4. Asymptotic formula

In this subsection we prove the asymptotic formula in Theorem 3. We begin by defining stationary versions $\hat{I}_t^{\pm\pm}$ of the pivotal intensities $I_t^{\pm\pm}$ introduced in (2.3) above; the key difference will be that, unlike the latter, the former will be independent of the box Λ_R .

Recall from Lemma 2 that each critical point in the interior of a box B can be classified as either positively, negatively, or not pivotal, and by Lemma 3 this classification is the same for all sufficiently large boxes. We shall need this statement for the Gaussian fields f and f^t under the conditioning (2.1):

Lemma 5. *Assume that $K \in L^1(\mathbb{R}^d)$. Let $x, y \in \mathbb{R}$ and $t \in [0, 1)$. Then with probability one, for any realisation of*

$$(2.19) \quad (f, f^t \mid f(x) = f^t(y) = \ell, \nabla f(x) = \nabla f^t(y) = 0)$$

the following holds: there exists $r_x, r_y > 0$ and $\Delta_x^t, \Delta_y^t \in \{+, -, 0\}$ such that for all boxes B containing $x + \Lambda_{r_x}$ and $y + \Lambda_{r_y}$, x is (Δ_x^t, B) -pivotal for f and y is (Δ_y^t, B) -pivotal for f^t .

Proof. By Lemma 3 it is enough to show that the critical points of f and f^t at x and y respectively are almost surely non-degenerate. Since $K \in L^1(\mathbb{R}^d)$, the spectral measure μ has a density, and in particular its support contains an open set. Hence, by Lemmas 13 and 15, the Gaussian vector

$$(\nabla^2 f(x) \mid f(x) = f^t(y) = \ell, \nabla f(x) = \nabla f^t(y) = 0)$$

is non-degenerate. (Note that here we view $\nabla^2 f(x)$ as the $d(d + 1)/2$ -dimensional Gaussian vector $(\partial_{ij} f(x) \mid i \leq j)$ so as to avoid the trivial degeneracy $\partial_{ij} f(x) = \partial_{ji} f(x)$. This convention is standard in the literature on smooth Gaussian fields.) The determinant of a symmetric $d \times d$ matrix, viewed as a function of $d(d + 1)/2$ variables, is a polynomial and so its zero-set has Lebesgue measure zero in $\mathbb{R}^{d(d+1)/2}$. Therefore with probability one

$$\mathbb{P}(|\nabla^2 f(x)| = 0 \mid f(x) = f^t(y) = \ell, \nabla f(x) = \nabla f^t(y) = 0) = 0.$$

The same argument applies if we replace $\nabla^2 f(x)$ by $\nabla^2 f^t(y)$, completing the proof of the lemma. □

We define the *stationary pivotal intensities* as

$$\hat{I}_t^{\pm\pm}(x, y) = \hat{\varphi}_t(x, y) \widehat{\mathbb{E}}[|\nabla^2 f(x) \nabla^2 f^t(y)| \mathbb{1}_{(\Delta_x^t, \Delta_y^t) = (\pm\pm)}],$$

recalling that $\hat{\varphi}_t(x, y)$ denotes the density of the conditioning (2.11) and $\widehat{\mathbb{E}}$ denotes expectation under this conditioning.

As already mentioned, in the proof of Theorem 3 we shall compare $\hat{I}_t^{\pm\pm}(x, y)$ and $I_t^{\pm\pm}(x, y)$ for x, y lying in the bulk of Λ_R . To this end, given $s > 0$, $x, y \in \mathbb{R}^d$, a realisation of (2.19), and r_x^*, r_y^* defined as the infima of r_x, r_y in Lemma 5, denote

$$q_t(x, y, s) = \mathbb{P}(\max\{r_x, r_y\} > s).$$

So roughly speaking this is the probability that boxes Λ_s , centred at x and y respectively, are not big enough to capture the change in topology at the pivotal points x, y . Clearly this is non-increasing in s and, by Lemma 5, tends to zero as $s \rightarrow \infty$ for each choice of t, x, y . Also, by stationarity, $q_t(x, y, s) = q_t(0, x - y, s)$.

Let us now observe an equivalent definition of the pivotal intensities $I_t^{\pm\pm}(x, y)$ for x, y in the interior of Λ_R . Given a realisation of (2.19) we write $\Delta_x^t(\Lambda_R) \in \{+, -, 0\}$ for the symbol such that x is a $(\Delta_x^t(\Lambda_R), \Lambda_R)$ -critical point, and we define $\Delta_y^t(\Lambda_R)$ analogously. Then for x, y in the interior of Λ_R we can equivalently define $I_t^{\pm\pm}(x, y)$ as

$$I_t^{\pm\pm}(x, y) = \hat{\varphi}_t(x, y)\widehat{\mathbb{E}}\left[\left|\nabla^2 f(x)\nabla^2 f^t(y)\right| \mathbb{1}_{(\Delta_x^t(\Lambda_R), \Delta_y^t(\Lambda_R)) = (\pm\pm)}\right].$$

Now suppose $r \in (0, R)$ and let $x, y \in \Lambda_{R-r}$. Then, applying the Cauchy–Schwarz inequality to the definitions above,

$$(2.20) \quad \left|I_t^{\pm\pm}(x, y) - \hat{I}_t^{\pm\pm}(x, y)\right| \leq \hat{I}_t^{(2)}(x, y)q_t(x, y, r)^{1/2} = \hat{I}_t^{(2)}(0, x - y)q_t(0, x - y, r)^{1/2},$$

where $\hat{I}_t^{(k)}$ is defined in (2.10).

Proof of Theorem 3. In the proof $c > 0$ will denote constants that are independent of R and may change from line to line.

First we show that the contribution from boundary strata to the variance formula in (1.6) is of order R^{d-1} . Let F_1, F_2 be two strata of Λ_R one of which has dimension at most $d - 1$. Then by the fact that $I_t^{\pm\pm} \leq I_t$ and Proposition 2

$$\begin{aligned} & \left| \int_0^1 \int_{F_1 \times F_2} K(x - y) \sum_{i, j \in \{+, -\}} \sigma(i)\sigma(j)I_t^{ij}(x, y) dv_{F_1}(x) dv_{F_2}(y) dt \right| \\ & \leq 4K(0) \int_0^1 \int_{F_1 \times F_2} I_t(x, y) dv_{F_1}(x) dv_{F_2}(y) dt \leq cR^{d-1}. \end{aligned}$$

Summing over all such pairs of strata, by (1.6) we see that

$$(2.21) \quad \text{Var}[N_\star(R)] = \int_0^1 \int_{\text{int}(\Lambda_R)^2} K(x - y) \sum_{i, j \in \{\pm\}} \sigma(i)\sigma(j)I_t^{ij}(x, y) dx dy dt + O(R^{d-1})$$

as $R \rightarrow \infty$, as claimed.

Next we choose $r = \sqrt{R}$ (in fact any $1 \ll r \ll R$ would do) and argue that, in (2.21), one can replace the domain of integration $\text{int}(\Lambda_R)^2$ by Λ_{R-r}^2 with an $o(R^d)$ error. Indeed using $I_t^{\pm\pm} \leq I_t$ and Proposition 2 once more,

$$(2.22) \quad \begin{aligned} & \left| \int_0^1 \int_{(\text{int}(\Lambda_R) \setminus \Lambda_{R-r}) \times \text{int}(\Lambda_R)} K(x - y) \sum_{i, j \in \{\pm\}} \sigma(i)\sigma(j)I_t^{ij}(x, y) dx dy dt \right| \\ & \leq 4K(0) \int_0^1 \int_{\substack{(\text{int}(\Lambda_R) \setminus \Lambda_{R-r}) \times \text{int}(\Lambda_R) \\ |x-y| \leq 1}} I_t(x, y) dx dy dt + c \int_0^1 \int_{\substack{(\text{int}(\Lambda_R) \setminus \Lambda_{R-r}) \times \text{int}(\Lambda_R) \\ |x-y| \geq 1}} |K(x - y)| dx dy dt. \end{aligned}$$

Note that, in the above expression, the critical point intensity $I_t(x, y)$ is equal to its stationary version $\hat{I}_t(x, y)$, since x, y are in the interior of Λ_R . Therefore by the change of variables $u = x - y$ and stationarity, (2.22) is bounded above by

$$4K(0) \int_{\text{int}(\Lambda_R) \setminus \Lambda_{R-r}} \int_0^1 \int_{|u| \leq 1} \hat{I}_t(0, u) du dt dx + c \int_{\text{int}(\Lambda_R) \setminus \Lambda_{R-r}} \int_{\mathbb{R}^d} |K(u)| du dx \leq crR^{d-1} = o(R^d),$$

where we used Proposition 2 and the fact that $K \in L^1(\mathbb{R}^d)$ for the final inequality. Combining with (2.21) we have shown that, as $R \rightarrow \infty$,

$$(2.23) \quad \text{Var}[N_\star(R)] = \int_0^1 \int_{\Lambda_{R-r}^2} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) I_t^{ij}(x,y) dx dy dt + o(R^d)$$

A similar argument shows that, as $R \rightarrow \infty$,

$$(2.24) \quad \begin{aligned} & \int_0^1 \int_{\Lambda_{R-r} \times \mathbb{R}^d} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) \hat{I}_t^{ij}(x,y) dx dy dt \\ &= \int_0^1 \int_{\Lambda_{R-r}^2} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) \hat{I}_t^{ij}(x,y) dx dy dt + o(R^d). \end{aligned}$$

Indeed using $\hat{I}_t^{\pm\pm} \leq \hat{I}_t$ and Proposition 2 again,

$$(2.25) \quad \begin{aligned} & \left| \int_0^1 \int_{\Lambda_{R-r} \times (\mathbb{R}^d \setminus \Lambda_{R-r})} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) \hat{I}_t^{ij}(x,y) dx dy dt \right| \\ & \leq 4K(0) \int_0^1 \int_{\Lambda_{R-r} \times (\mathbb{R}^d \setminus \Lambda_{R-r})} \hat{I}_t(x,y) dx dy dt + c \int_0^1 \int_{\Lambda_{R-r} \times (\mathbb{R}^d \setminus \Lambda_{R-r})} |K(x-y)| dx dy dt. \end{aligned}$$

By the change of variables $u = x - y$ and stationarity, (2.25) is bounded above by

$$\begin{aligned} & 4K(0) \int_{\Lambda_{R-r} \setminus \Lambda_{R-2r}} \int_0^1 \int_{|u| \leq 1} \hat{I}_t(0,u) du dt dx \\ & + c \int_{(\Lambda_{R-r} \setminus \Lambda_{R-2r}) \times \mathbb{R}^d} |K(x-y)| dx dy + c \int_{\Lambda_{R-2r} \times (\mathbb{R}^d \setminus \Lambda_{R-r})} |K(x-y)| dx dy \\ & \leq crR^{d-1} + cR^d \int_{|u| \geq r} |K(u)| du = o(R^d). \end{aligned}$$

It remains to replace, in (2.23), the pivotal intensities $I_t^{\pm\pm}$ with their stationary versions $\hat{I}_t^{\pm\pm}$ up to $o(R^d)$ error, so as to compare with (2.24). Using (2.20),

$$(2.26) \quad \begin{aligned} & \left| \int_0^1 \int_{\Lambda_{R-r}^2} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) (I_t^{ij}(x,y) - \hat{I}_t^{ij}(x,y)) dx dy dt \right| \\ & \leq \int_0^1 \int_{\Lambda_{R-r}^2} |K(x-y)| \sum_{i,j \in \{\pm\}} |I_t^{ij}(x,y) - \hat{I}_t^{ij}(x,y)| dx dy dt \\ & \leq \int_0^1 \int_{\Lambda_{R-r}^2} |K(x-y)| \hat{I}_t^{(2)}(0, x-y) q_t(0, x-y, r)^{1/2} dx dy dt, \end{aligned}$$

which is bounded above by

$$R^d \int_0^1 \int_{\mathbb{R}^d} |K(u)| \hat{I}_t^{(2)}(0, u) q_t(0, u, r)^{1/2} du dt.$$

We next claim the dominated convergence theorem implies that

$$\int_0^1 \int_{\mathbb{R}^d} |K(u)| \hat{I}_t^{(2)}(u, 0) q_t(u, 0, r)^{1/2} du dt \rightarrow 0$$

as $R \rightarrow \infty$, from which we conclude that (2.26) is $o(R^d)$. Indeed since $q_t(u, 0, r) \leq 1$ and $q_t(u, 0, r) \rightarrow 0$ as $r \rightarrow \infty$, dominated convergence applies since

$$\int_0^1 \int_{\mathbb{R}^d} |K(u)| \hat{I}_t^{(2)}(0, u) \, du \, dt \leq K(0) \int_0^1 \int_{|u| \leq 1} \hat{I}_t^{(2)}(0, u) \, du \, dt + c \int_0^1 \int_{|u| \geq 1} |K(u)| \, du \, dt < \infty,$$

where we used Proposition 3 and the integrability of K .

Combining the fact that (2.26) is $o(R^d)$ with (2.23)–(2.24), and by stationarity, we obtain

$$\begin{aligned} \text{Var}[N_\star(R)] &= \int_0^1 \int_{\Lambda_{R-r} \times \mathbb{R}^d} K(x-y) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) \hat{I}_t^{ij}(x,y) \, dx \, dt + o(R^d) \\ &= R^d(1 - o(1)) \int_0^1 \int_{\mathbb{R}^d} K(x) \sum_{i,j \in \{\pm\}} \sigma(i)\sigma(j) \hat{I}_t^{ij}(0,x) \, dx \, dt + o(R^d), \end{aligned}$$

which completes the proof of (1.7) and (1.8). □

3. Analysis of the critical point intensity

In this section we prove the integrability and boundedness of the critical point intensity I_t , stated in Proposition 2 above, as well as its generalisation in Proposition 3.

3.1. Reduction to pointwise bounds

We first reduce the proof of Propositions 2 and 3 to pointwise bounds on the two components whose product gives the intensity I_t .

Let $\{e_i\}_{1 \leq i \leq d}$ denote the standard basis of \mathbb{R}^d , and for any non-empty $I \subseteq \{1, 2, \dots, d\}$, let ∇_I denote the gradient with respect to the basis directions $\{e_i\}_{i \in I}$. For $I_1, I_2 \subseteq \{1, 2, \dots, d\}$ and $(t, x) \in [0, 1) \times \mathbb{R}^d$ we shall consider the Gaussian vector

$$(3.1) \quad (f(0), f^t(x), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x))$$

and the conditional expectation, for $\ell \in \mathbb{R}$ and $k \in \mathbb{N}$,

$$(3.2) \quad \mathbb{E}[|\nabla_{I_1}^2 f(0) \nabla_{I_2}^2 f^t(x)|^k \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0].$$

If I_1 is empty then by convention we remove the corresponding term ∇_{I_1} from (3.1) and the conditioning in (3.2), and set $|\nabla_{I_1}^2 f(0)| = 1$ in (3.2), and similarly if I_2 is empty.

For a square-integrable random vector X , let $\text{DC}(X) = |\text{Cov}[X]|$ denote the determinant of its covariance matrix.

The pointwise bounds are the following:

Lemma 6. *There exists a $c > 0$ depending only on the field f such that, for every $I_1, I_2 \subseteq \{1, 2, \dots, d\}$ and $(t, x) \in [0, 1) \times \mathbb{R}^d$,*

$$\text{DC}(f(0), f^t(x), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)) \geq \begin{cases} c & \text{if } |x| \geq 1, \\ c \max\{(1-t)^{1/2}, |x|\}^{2d'} (1-t) & \text{if } |x| \leq 1, \end{cases}$$

where $d' = \dim(I_1 \cap I_2)$.

Lemma 7. *Let $k \in \mathbb{N}$. Then there exists $c > 0$ depending only on the field f and k , such that, for every $\ell \in \mathbb{R}$, $I_1, I_2 \subseteq \{1, 2, \dots, d\}$, and $(t, x) \in [0, 1) \times \mathbb{R}^d \setminus \{0\}$,*

$$\mathbb{E}[|\nabla_{I_1}^2 f(0) \nabla_{I_2}^2 f^t(x)|^k \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \leq c(1 + \ell^{2dk}).$$

Remark 10. Lemma 6 is a quantification of our previous claim that

$$(f(x), f^t(y), \nabla_{F_x} f(x), \nabla_{F_y} f^t(y))$$

is non-degenerate for all $t \in [0, 1)$, $x \in B_1$, and $y \in B_2$, and in particular the intensity $I_t(x, y)$ is well-defined. In Lemma 15 below we give a simpler non-quantitative proof of this fact.

Remark 11. Although we only require Lemma 7 for fixed $\ell \in \mathbb{R}$, we include the dependency since we believe it may prove useful in future work.

Let us complete the proofs of Propositions 2 and 3 given these pointwise bounds:

Proof of Proposition 2. Recall that $\ell \in \mathbb{R}$ is fixed. In the proof $c > 0$ will denote constants which depend only on f and ℓ , and may change from line to line.

Let $(F_1, F_2) \in \mathcal{F}^{(1,2)}$ be given, and let I_1 and I_2 denote the indices of the basis directions that span F_1 and F_2 respectively, with $d_i = \dim(I_i)$. Observe that, by stationarity, for every $(t, x, y) \in [0, 1) \times F_1 \times F_2$,

$$(f(x), f^t(y), \nabla_{F_x} f(x), \nabla_{F_y} f^t(y)) \stackrel{d}{=} (f(0), f^t(u), \nabla_{I_1} f(0), \nabla_{I_2} f^t(u)),$$

where $u = y - x$.

Noting the trivial bound

$$\varphi_t(0, u) \leq (2\pi)^{-(d_1+d_2+2)/2} \text{DC}(f(0), f^t(u), \nabla_{I_1} f(0), \nabla_{I_2} f^t(u))^{-1/2}$$

on the Gaussian density, we see that

$$(3.3) \quad \begin{aligned} I_t(x, y) &\leq c \times \text{DC}(f(0), f^t(u), \nabla_{I_1} f(0), \nabla_{I_2} f^t(u))^{-1/2} \\ &\quad \times \mathbb{E}[|\nabla_{I_1}^2 f(0) \nabla_{I_2}^2 f^t(u)| \mid f(0) = f^t(u) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(u) = 0]. \end{aligned}$$

Combining this with Lemmas 6 and 7, we immediately obtain the uniform bound on $I_t(x, y)$ for $|y - x| \geq 1$ stated in (2.9).

It remains to prove the integrability in (2.8). Let L_I denote the linear span of the basis directions $\{e_i\}_{i \in I}$ equipped with its Lebesgue measure dv_I ; if I is empty then we take L_I as the origin and dv_I a Delta mass at the origin. By (3.3), Lemmas 6 and 7, for every $t \in [0, 1)$ and $x \in F_1$ we have

$$(3.4) \quad \begin{aligned} \int_{\substack{y \in F_2 \\ |y-x| \leq 1}} I_t(x, y) dv_{F_y}(y) &\leq c \int_{\substack{y \in F_2 \\ |y-x| \leq 1}} (\max\{(1-t)^{1/2}, |y-x|\}^{2\min\{d_1, d_2\}} (1-t))^{-1/2} dv_{F_y}(y) \\ &\leq c \int_{\substack{u \in L_{I_2} \\ |u| \leq 1}} \max\{(1-t)^{1/2}, |u|\}^{-d_2} (1-t)^{-1/2} dv_{I_2}(u) \end{aligned}$$

where in the last inequality we used the triangle inequality and the fact that the integrand is non-increasing in $|y - x|$. It remains to show that (3.4) is integrable over $t \in [0, 1)$. For this we first make the substitution $t \mapsto 1 - t$, and then split the integral depending on whether $|u| \leq t^{1/2}$ or $|u| \geq t^{1/2}$. In the first case we have

$$\int_{|u| \leq t^{1/2}} \max\{t^{1/2}, |u|\}^{-d_2} t^{-1/2} dv_{I_2}(u) = t^{-d_2/2-1/2} \int_{|u| \leq t^{1/2}} dv_{I_2}(u) = ct^{-1/2},$$

whereas in the second case we have

$$\begin{aligned} \int_{|u| \geq t^{1/2}} \max\{t^{1/2}, |u|\}^{-d_2} t^{-1/2} dv_{I_2}(u) &= t^{-1/2} \int_{|u| \geq t^{1/2}} |u|^{-d_2} dv_{I_2}(u) \\ &= \begin{cases} ct^{-1/2} \int_{t^{1/2}}^1 r^{-1} dr & \text{if } d_2 \geq 1, \\ ct^{-1/2} & \text{if } d_2 = 0, \end{cases} \\ &\leq ct^{-1/2} \log(1/t), \end{aligned}$$

where we integrated in polar coordinates in the second equality in the case $d_2 \geq 1$. Since $t^{-1/2} \log(1/t)$ is integrable over $t \in [0, 1]$, we are done. \square

Proof of Proposition 3. The proof is identical to that of Proposition 2, except that we apply the general case $k \in \mathbb{N}$ of Lemma 7. \square

3.2. Proof of the pointwise bounds

We collect some properties of the DC operator:

Lemma 8. *Let (X, Y) be a Gaussian vector of dimension $m_1 + m_2$.*

1. For $A \in \mathbb{R}^{m_1 \times m_1}$,

$$\text{DC}(AX) = \det(A)^2 \text{DC}(X).$$

In particular, for $a \in \mathbb{R}$ and $B \in \mathbb{R}^{m_1 \times m_2}$,

$$\text{DC}(aX, Y) = a^{2m_1} \text{DC}(X, Y) \quad \text{and} \quad \text{DC}(X + BY, Y) = \text{DC}(X, Y).$$

2. If X_n is a sequence converging to X in L^2 , then $\text{DC}(X_n) \rightarrow \text{DC}(X)$.

3. If X is non-degenerate then $\text{DC}(Y) \geq \text{DC}(Y|X) = \text{DC}(Y, X)/\text{DC}(X)$.

4. Let \tilde{Y} be an independent copy of Y , and define $Y^t = tY + \sqrt{1-t^2}\tilde{Y}$, $t \in [0, 1]$. If X is non-degenerate, then $\text{DC}(X, Y^t) \geq \text{DC}(X, Y^1) = \text{DC}(X, Y)$.

Remark 12. In the proof of this lemma and below, we write $\text{Cov}[Y|X]$ to denote the covariance matrix of Y conditioned on $X = x$ for an arbitrary $x \in \mathbb{R}^{m_1}$. By Gaussian regression (Section 2.2.1) the matrix does not depend on the value of x , which justifies the choice of notation.

Proof. The first item is elementary, the second follows from the continuity of the determinant operator.

Let Σ_X and Σ_Y be the covariance matrices of X and Y respectively, and Σ_{YX} the cross-covariance matrix. By Gaussian regression (Section 2.2.1)

$$\text{Cov}[Y|X] = \Sigma_Y - \Sigma_{YX} \Sigma_X^{-1} \Sigma_{YX}^T.$$

Since the determinant is non-decreasing in the positive definite ordering, the determinant of the above expression is at most $\text{DC}(Y)$. The identity

$$\begin{pmatrix} \Sigma_Y & \Sigma_{YX} \\ \Sigma_{YX}^T & \Sigma_X \end{pmatrix} = \begin{pmatrix} I & \Sigma_{YX} \\ 0 & \Sigma_X \end{pmatrix} \begin{pmatrix} \Sigma_Y - \Sigma_{YX} \Sigma_X^{-1} \Sigma_{YX}^T & 0 \\ \Sigma_X^{-1} \Sigma_{YX}^T & I \end{pmatrix}$$

proves that $\text{DC}(Y, X) = \text{DC}(X)\text{DC}(Y|X)$, verifying the third item.

By the definition of Y^t and the analogue of the previous identity

$$\text{DC}(Y^t, X) = \text{DC} \begin{pmatrix} \Sigma_Y & t \Sigma_{YX} \\ t \Sigma_{YX}^T & \Sigma_X \end{pmatrix} = \det(\Sigma_X) \det(\Sigma_Y - t^2 \Sigma_{YX} \Sigma_X^{-1} \Sigma_{YX}^T).$$

Since the determinant is non-decreasing in the positive-definite ordering,

$$t \mapsto \det(\Sigma_Y - t^2 \Sigma_{YX} \Sigma_X^{-1} \Sigma_{YX}^T)$$

is non-increasing on $t \in [0, 1]$, and hence so is $t \mapsto \text{DC}(Y^t, X)$. \square

Let us establish two particular consequences of Lemma 8. These can be viewed as examples of the ‘divided difference’ method (see [7, Section 4.2] for an overview).

Lemma 9. *Suppose f is a C^3 -smooth stationary Gaussian field, and suppose that, for every $I_1, I_2 \subseteq \{1, 2, \dots, d\}$ with $I_3 = I_1 \cap I_2$ non-empty, and every unit vector $v \in \mathbb{S}^{d-1}$, the Gaussian vector*

$$(\partial_v \nabla_{I_3} f(0), f(0), \nabla_{I_1 \cup I_2} f(0))$$

is non-degenerate, where $\partial_v f$ denotes the partial derivative in the direction v . Then

$$\text{DC}\left(\frac{\nabla_{I_3} f(x) - \nabla_{I_3} f(0)}{|x|} \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x)\right)$$

is uniformly bounded below as $x \rightarrow 0$ (with $x \neq 0$).

Proof. Suppose for contradiction that there exists a sequence $x_n \rightarrow 0$ such that

$$(3.5) \quad \text{DC}\left(\frac{\nabla_{I_3} f(x_n) - \nabla_{I_3} f(0)}{|x_n|} \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x_n)\right) \rightarrow 0.$$

By the third item of Lemma 8 this determinant is equal to

$$(3.6) \quad \text{DC}\left(\frac{\nabla_{I_3} f(x_n) - \nabla_{I_3} f(0)}{|x_n|}, f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x_n)\right) / \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x_n)).$$

Since $f \in C^3$, $\nabla_{I_2 \setminus I_3} f(x_n) \rightarrow \nabla_{I_2 \setminus I_3} f(0)$ almost surely as $n \rightarrow \infty$. Since these variables are jointly Gaussian, convergence occurs in L^2 and so by the second item of Lemma 8

$$(3.7) \quad \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x_n)) \rightarrow \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(0)) > 0,$$

where the final inequality follows by assumption.

By passing to a subsequence, we may assume that $\hat{x}_n := x_n/|x_n| \rightarrow v$ for some $v \in \mathbb{S}^{d-1}$. By a Taylor expansion, for $|x_n| \leq 1$

$$\left| \frac{\nabla_{I_3} f(x_n) - \nabla_{I_3} f(0)}{|x_n|} - \partial_{\hat{x}_n} \nabla_{I_3} f(0) \right| \leq c|x_n| \cdot \|f\|_{C^3(B(1))},$$

where $c > 0$ depends only on the dimension d , and $B(1)$ denotes the Euclidean ball of unit radius. Therefore

$$\frac{\nabla_{I_3} f(x_n) - \nabla_{I_3} f(0)}{|x_n|} \rightarrow \partial_v \nabla_{I_3} f(0)$$

almost surely, and hence in L^2 . By the second item of Lemma 8 the numerator of (3.6) converges to

$$\text{DC}(\partial_v \nabla_{I_3} f(0), f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(0)),$$

which is positive by assumption. Together with (3.7), this contradicts (3.5). \square

Lemma 10. Suppose f is a C^3 -smooth stationary Gaussian field such that, for every unit vector $v \in \mathbb{S}^{d-1}$, the Gaussian vector

$$(f(0), \nabla f(0), \partial_v \nabla f(0), \partial_v^3 f(0))$$

is non-degenerate. Then

$$\text{Var}[f(0) \mid f^t(x) - f(0), \nabla f(0), \nabla f^t(x)]$$

is uniformly bounded below as $(x, t) \rightarrow (0, 1)$ (with $x \neq 0$ and $t \neq 1$).

Proof. Suppose for contradiction that the conditional variance is not bounded below, so there exists a sequence $(x_n, t_n) \rightarrow (0, 1)$ such that, by the third item of Lemma 8,

$$(3.8) \quad \frac{\text{DC}(f(0), f^{t_n}(x_n) - f(0), \nabla f(0), \nabla f^{t_n}(x_n))}{\text{DC}(f^{t_n}(x_n) - f(0), \nabla f(0), \nabla f^{t_n}(x_n))} \rightarrow 0.$$

By passing to a subsequence, we may assume that

$$\hat{x}_n := \frac{x_n}{|x_n|} \rightarrow v \in \mathbb{S}^{d-1}, \quad \frac{|x_n|}{\sqrt{1-t_n^2+|x_n|^2}} \rightarrow \alpha \in [0, 1], \quad \text{and} \quad \frac{|x_n|^2}{\sqrt{1-t_n^2+|x_n|^2}} \rightarrow \beta \in [0, 1].$$

By the first item of Lemma 8, the left hand side of (3.8) is unchanged if we replace $\nabla f^{t_n}(x_n)$ in the numerator and denominator by

$$A_n := \frac{\nabla f^{t_n}(x_n) - t_n \nabla f(0)}{\sqrt{1 - t_n^2} + |x_n|}.$$

By Taylor’s theorem, for $|x_n| \leq 1$,

$$A_n = \frac{\sqrt{1 - t_n^2} \tilde{f}(x_n) + t_n |x_n| \partial_{\hat{x}_n} \nabla f(0)}{\sqrt{1 - t_n^2} + |x_n|} + O_{C^3(B(1))}(|x_n|),$$

where we introduce the notation $U_n = V_n + O_{C^3(B(1))}(a_n)$ to mean that there exists a constant $c > 0$ depending only on the dimension d such that

$$|U_n - V_n| \leq ca_n (\|f\|_{C^3(B(1))} + \|\tilde{f}\|_{C^3(B(1))}).$$

Therefore as $n \rightarrow \infty$

$$(3.9) \quad A_n \rightarrow (1 - \alpha) \nabla \tilde{f}(0) + \alpha \partial_v \nabla f(0)$$

almost surely and hence in L^2 .

We first consider the case that $\beta < 1$. Again by the first item of Lemma 8, the left hand side of (3.8) is unchanged if we replace $f^{t_n}(x_n) - f(0)$ in the numerator and denominator by

$$B_n := \frac{f^{t_n}(x_n) - f(0) - t_n x_n \cdot \nabla f(0)}{\sqrt{1 - t_n^2} + |x_n|^2}.$$

By a Taylor expansion

$$B_n = \frac{\sqrt{1 - t_n^2} \tilde{f}(x_n) - (1 - t_n) f(0) + t_n \frac{|x_n|^2}{2} \partial_{\hat{x}_n}^2 f(0)}{\sqrt{1 - t_n^2} + |x_n|^2} + O_{C^3(B(1))}(|x_n|) \rightarrow (1 - \beta) \tilde{f}(0) + \frac{\beta}{2} \partial_v^2 f(0)$$

almost surely and in L^2 . Combining this with (3.9) and the second item of Lemma 8, we see that the numerator on the left hand side of (3.8) converges to

$$\text{DC} \left(f(0), \nabla f(0), (1 - \alpha) \nabla \tilde{f}(0) + \alpha \partial_v \nabla f(0), (1 - \beta) \tilde{f}(0) + \frac{\beta}{2} \partial_v^2 f(0) \right).$$

This expression is strictly positive since $\beta < 1$ and the Gaussian vector $(f(0), \nabla f(0), \partial_v \nabla f(0))$ is non-degenerate (and independent of \tilde{f}). The denominator of (3.8) converges to the same expression with $f(0)$ removed, which is also positive. Hence (3.8) has a strictly positive limit, yielding a contradiction.

Next we consider the case that $\beta = 1$ (which implies $\alpha = 1$). Note that the limiting DC() terms in the numerator and denominator considered above are both zero, so we must consider a higher order Taylor expansion. By passing to a further subsequence, we assume that

$$\frac{|x_n|^3}{\sqrt{1 - t_n^2} + |x_n|^3} \rightarrow \gamma \in [0, 1].$$

By the first item of Lemma 8 we can replace $f^{t_n}(x_n) - f(0)$ in (3.8) by

$$C_n := \frac{f^{t_n}(x_n) - f(0) - t_n x_n \cdot \nabla f(0) - \frac{x_n}{2} \cdot (\nabla f^{t_n}(x_n) - t_n \nabla f(0))}{\sqrt{1 - t_n^2} + |x_n|^3}.$$

By Taylor’s theorem, there exists $\theta_n, \theta'_n \in [0, 1]$ such that

$$f^{t_n}(x_n) - f(0) = \sqrt{1 - t_n^2} \tilde{f}(x_n) - (1 - t_n) f(0) + t_n |x_n| \partial_{\hat{x}_n} f(0) + t_n \frac{|x_n|^2}{2} \partial_{\hat{x}_n}^2 f(0) + t_n \frac{|x_n|^3}{3!} \partial_{\hat{x}_n}^3 f(\theta_n x_n)$$

and

$$x_n \cdot (\nabla f^{t_n}(x_n) - t_n \nabla f(0)) = \sqrt{1 - t_n^2} |x_n| \partial_{\tilde{x}_n} \tilde{f}(x_n) + t_n |x_n|^2 \partial_{\tilde{x}_n}^2 f(0) + t_n \frac{|x_n|^3}{2} \partial_{\tilde{x}_n}^3 f(\theta'_n x_n).$$

We therefore see that

$$C_n \rightarrow (1 - \gamma) \tilde{f}(0) - \frac{\gamma}{12} \partial_v^3 f(0)$$

almost surely and in L^2 . Hence the numerator of (3.8) converges to

$$\text{DC}\left(f(0), \nabla f(0), \partial_v \nabla f(0), (1 - \gamma) \tilde{f}(0) - \frac{\gamma}{12} \partial_v^3 f(0)\right),$$

which is positive by assumption. The denominator of (3.8) converges to the same expression with $f(0)$ removed, which is also positive and therefore gives the required contradiction. \square

Proof of Lemma 6. Let $I_1, I_2 \subseteq \{1, 2, \dots, d\}$ be given, and abbreviate

$$F(t, x) = \text{DC}(f(0), f^t(x), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)).$$

It suffices to prove the following two statements:

1. For all $\delta \in (0, 1]$ there is a $c > 0$ such that, if either $t \leq 1 - \delta$ or $|x| \geq \delta$,

$$F(t, x) \geq c.$$

2. There are $\delta, c > 0$ such that, for all $t \geq 1 - \delta$ and $|x| \leq \delta$,

$$F(t, x) \geq c \max\{(1 - t)^{1/2}, |x|\}^{2d} (1 - t).$$

We prove these statements in turn. First note that, by the second item of Lemma 8 and the fact that f is C^1 -smooth, $F(t, x)$ is a continuous function of $(t, x) \in [0, 1] \times \mathbb{R}^d$. Now for the first statement we note that, by the third item of Lemma 8,

$$F(t, x) = \text{DC}(f(0), \nabla_{I_1} f(0)) \times \text{DC}(f^t(x), \nabla_{I_2} f^t(x) \mid f(0), \nabla_{I_1} f(0)),$$

and by the first and third items of Lemma 8,

$$\text{DC}(f^t(x), \nabla_{I_2} f^t(x) \mid f(0), \nabla_{I_1} f(0)) \geq \text{DC}(f^t(x), \nabla_{I_2} f^t(x) \mid f) = \text{DC}((1 - t^2)^{1/2} \tilde{f}(x), (1 - t^2)^{1/2} \nabla_{I_2} \tilde{f}(x)).$$

By the first item of Lemma 8, and also by stationarity and the equality in law of f and \tilde{f} , the above is equal to

$$(1 - t^2)^{\dim(I_2)+1} \text{DC}(f(0), \nabla_{I_2} f(0)).$$

Since $(f(0), \nabla f(0))$ is non-degenerate, we conclude that there is a $c_1(\delta) > 0$ such that

$$(3.10) \quad F(t, x) \geq c_1(\delta)$$

for all $t \leq 1 - \delta$.

Moreover, by the fourth item of Lemma 8,

$$F(t, x) \geq F(1, x) = \text{DC}(f(0), \nabla_{I_1} f(0), f(x), \nabla_{I_2} f(x)).$$

By stationarity and since $\partial^\alpha K(x) \rightarrow 0$ as $|x| \rightarrow \infty$ for $|\alpha| \leq 2$,

$$\text{DC}(f(0), \nabla_{I_1} f(0), f(x), \nabla_{I_2} f(x)) \rightarrow \text{DC}(f(0), \nabla_{I_1} f(0)) \text{DC}(f(0), \nabla_{I_2} f(0))$$

as $|x| \rightarrow \infty$. By continuity of $F(t, x)$, and since $(f(0), \nabla f(0))$ is non-degenerate, we conclude that there is a $c_2(\delta) > 0$ such that

$$(3.11) \quad F(t, x) \geq c_2(\delta)$$

for all $|x| \geq \delta$. Combining (3.10) and (3.11) we have proven the first statement.

We turn to the second statement. Abbreviating $I_3 = I_1 \cap I_2$, by repeated application of the third item of Lemma 8 we have

$$F(t, x) = \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f^t(x)) \times \text{DC}(\nabla_{I_3} f^t(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f^t(x)) \\ \times \text{Var}[f^t(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)] =: F_1(t, x) \times F_2(t, x) \times F_3(t, x),$$

where we interpret $F_2(t, x) = 1$ if I_3 is empty. We make the following four claims, which combined give the second statement:

1. $F_1(t, x)$ is uniformly bounded below as $x \rightarrow 0$;
2. $F_2(t, x)(1 - t^2)^{-d'}$ is uniformly bounded below;
3. Restricting to $t \geq 1/2$, $F_2(t, x)|x|^{-2d'}$ is uniformly bounded below as $x \rightarrow 0$;
4. $F_3(t, x)(1 - t^2)^{-1}$ is uniformly bounded below.

Let us establish these claims in turn:

1. By the second and fourth items of Lemma 8, as $x \rightarrow 0$,

$$F_1(t, x) = \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f^t(x)) \geq \text{DC}(f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x)) \rightarrow \text{DC}(f(0), \nabla_{I_1 \cup I_2} f(0)),$$

and the claim follows since $(f(0), \nabla f(0))$ is non-degenerate.

2. We may assume that $d' = \dim(I_3) \geq 1$. By conditioning on f and using the first and fourth items of Lemma 8,

$$F_2(t, x) = \text{DC}(\nabla_{I_3} f^t(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f^t(x)) \\ \geq \text{DC}((1 - t^2)^{1/2} \nabla_{I_3} \tilde{f}(x) \mid \nabla_{I_2 \setminus I_3} \tilde{f}(x)) = (1 - t^2)^{d'} \text{DC}(\nabla_{I_3} f(0) \mid \nabla_{I_2 \setminus I_3} f(0)),$$

where the second step also used stationarity and the equality in law of f and \tilde{f} . The claim follows since $\nabla f(0)$ is non-degenerate.

3. By conditioning on \tilde{f} and by the first and fourth items of Lemma 8,

$$F_2(t, x) = \text{DC}(\nabla_{I_3} f^t(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f^t(x)) \\ \geq \text{DC}(t \nabla_{I_3} f(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x)) \\ \geq 2^{-2d'} |x|^{2d'} \text{DC}\left(\frac{\nabla_{I_3} f(x) - \nabla_{I_3} f(0)}{|x|} \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2 \setminus I_3} f(x)\right),$$

where the last step used that $t \geq 1/2$. Then the claim follows from Lemma 9 above.

4. By again conditioning on f and using the first and fourth items of Lemma 8,

$$F_3(t, x) = \text{Var}[f^t(x) \mid f(0), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)] \geq (1 - t^2) \text{Var}[f(0) \mid \nabla_{I_2} f(0)],$$

and the claim follows since $(f(0), \nabla f(0))$ is non-degenerate. □

Before proving Lemma 7 we establish a uniform bound on the mean of the derivatives of the field under the conditioning (3.1):

Lemma 11. *There exists a $c > 0$ such that, for all $I_1, I_2 \subseteq \{1, 2, \dots, d\}$, $\ell \in \mathbb{R}$, $|\alpha| \leq 2$, and $(t, x) \in [0, 1) \times \mathbb{R}^d \setminus \{0\}$,*

$$\left| \mathbb{E}[\partial^\alpha f(0) \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \right| \leq c|\ell|.$$

Remark 13. This result is only non-trivial in the case $\ell \neq 0$, since if $\ell = 0$ then in fact the left-hand side is zero by Gaussian regression.

Proof of Lemma 11. By first conditioning on $\nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0$ and $f^t(x) - f(0) = 0$, and then on $f(0) = \ell$, we have by Gaussian regression (Section 2.2.1) that

$$\mathbb{E}[\partial^\alpha f(0) \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \\ = \mathbb{E}[\partial^\alpha f(0) \mid f(0) = \ell, f^t(x) - f(0) = 0, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] = \ell s_1 / s_2,$$

where

$$s_1 = \text{Cov}[\partial^\alpha f(0), f(0) \mid f^t(x) - f(0), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)]$$

and

$$s_2 = \text{Var}[f(0) \mid f^t(x) - f(0), \nabla_{I_1} f(0), \nabla_{I_2} f^t(x)].$$

By the Cauchy–Schwarz inequality and the third item of Lemma 8,

$$s_1 \leq \max\{\text{Var}[\partial^\alpha f(0)], \text{Var}[f(0)]\},$$

so it remains to show that s_2 is uniformly bounded away from zero over $(t, x) \in [0, 1] \times \mathbb{R}^d \setminus \{0\}$.

By the third item of Lemma 8

$$(3.12) \quad s_2 \geq \text{Var}[f(0) \mid f^t(x) - f(0), \nabla f(0), \nabla f^t(x)].$$

Arguing as in the proof of Lemma 6, since $\partial^\alpha K(x) \rightarrow 0$ as $|x| \rightarrow \infty$ we deduce that, for any $\delta > 0$, there is a $c > 0$ such that $s_2 \geq c$ if either $t \leq 1 - \delta$ or $|x| \geq \delta$. Lemma 10 and (3.12) then show that s_2 is bounded away from zero for (x, t) close to $(0, 1)$, completing the proof. \square

Proof of Lemma 7. Expanding the determinants, applying Hölder’s inequality, and by stationarity and the equality in law of f and \tilde{f} , it suffices to prove that

$$\mathbb{E}[|\partial_v \partial_{v'} f(0)|^k \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \leq c(1 + |\ell|^k)$$

for all $k \in \mathbb{N}$ and unit vectors $v, v' \in \mathbb{S}^{d-1}$, where the constant $c > 0$ depends only on the field. Since these are Gaussian variables, this moment bound is implied by uniform bounds on the mean

$$\mathbb{E}[|\partial_v \partial_{v'} f(0)| \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \leq c|\ell|,$$

which is a special case of Lemma 11 above, and on the variance

$$\text{Var}[\partial_v \partial_{v'} f(0) \mid f(0) = f^t(x) = \ell, \nabla_{I_1} f(0) = \nabla_{I_2} f^t(x) = 0] \leq \text{Var}[\partial_v \partial_{v'} f(0)] \leq c,$$

where in the first inequality we used the third item of Lemma 8. \square

4. Topological arguments

In this section we prove Lemmas 1, 2 and 3. We start with a fundamental lemma of (stratified) Morse theory, which says roughly that if we smoothly perturb a function then the topology of its level set cannot change unless passing through a critical point. We will deduce Lemmas 1 and 2 as consequences of this result.

Let B be a stratified box (as defined in Section 2.1) with strata \mathcal{F} , let $A \subseteq B$, and let $I \subset \mathbb{R}$ be a compact interval. We say that a continuous map $H : A \times I \rightarrow B$ is a *stratified isotopy of A* if, for each $t \in I$, $H(\cdot, t)$ is a homeomorphism onto its image and, for each $F \in \mathcal{F}$, $H(F \times \{t\}) \subseteq F$. In the case that $A = B$, we also require that $H(B, t) = B$ for every $t \in I$.

Lemma 12. *Let B be a stratified box and U an open neighbourhood of B . Let g and h be functions in class $C^2(U)$, and define $g_t = g + th$ for $t \in [0, 1]$. Suppose that g_t has no (stratified) critical points in B at level ℓ for all $t \in [0, 1]$. Then there exists a stratified isotopy $H : B \times [0, 1] \rightarrow B$ such that, for all $t \in [0, 1]$,*

$$H(\{g = \ell\} \times \{t\}) = \{g_t = \ell\}.$$

In particular $N_\star(B, g_t, \ell)$ is constant over $t \in [0, 1]$ for $\star \in \{\text{ES}, \text{LS}\}$.

Proof. Let us consider the domain $B \times [0, 1]$ and the function $G(x, t) = (g_t(x), t)$. The main idea of the lemma is that the t -sections of the preimage $G^{-1}(\ell, [0, 1])$ are level sets of g_t and this preimage is a trivial (stratified) fibration over the level set of g . Hence the (stratified) topology of level sets of g_t does not change with t .

First, we claim that there exists a neighbourhood W of $G^{-1}(\ell, [0, 1])$ such that the function g_t has no stratified critical points in W . If this were not the case then by compactness we could find a sequence (x_n, t_n) such that x_n is a stratified critical point of g_t and $(x_n, t_n) \rightarrow (x, t)$ where $g_t(x) = \ell$. By taking a further subsequence we may assume that all x_n are in the same stratum F . It may not be the case that $F_x = F$ however we must have $F_x \subseteq \overline{F}$ (since $x_n \rightarrow x$) and therefore by continuity we have $\nabla_{F_x} g(x) = 0$, which contradicts the fact that $\{g_t = \ell\}$ has no stratified critical points. Reducing W further we can assume that W is of the form $\{(x, t) \in B \times [0, 1] : |g_t(x) - \ell| < \epsilon\}$ for some $\epsilon > 0$. We can endow W with a stratification inherited from the stratification of B .

The absence of stratified critical points ensures that G is a submersion when restricted to each stratum. It follows from the continuity of g_t and boundedness of W that G is a proper map (i.e. the inverse image of any compact set is compact). Using these two facts, we may apply Thom’s first isotopy lemma [27, Section 8] to claim that W is a trivial fibration over $G^{-1}((\ell - \epsilon, \ell + \epsilon), 0)$ and $G^{-1}(\ell, [0, 1])$ is a trivial fibration over $G^{-1}(\ell, 0) = g^{-1}(\ell)$. Moreover, these fibrations respect stratification. The proof of Thom’s lemma is based on the construction of a flow x_t which preserves the stratification, trivializes the fibration and along which $g_t(x_t)$ is constant. This immediately implies that there is a stratified isotopy between the level sets of the functions g_t . By the isotopy extension theorem [17, Corollary 1.4] we may extend this to an isotopy of B . □

Proof of Lemma 1. Since g has no stratified critical points in B , there exists $\epsilon > 0$ such that for all $x \in B$

$$\|\nabla_{F_x} g(x)\| + |g(x) - \ell| \geq \epsilon,$$

where we recall that F_x denotes the stratum of B containing x . If

$$\|g - g'\|_{C^1(U)} < \epsilon,$$

then for all $t \in [0, 1]$ we see that $g + t(g' - g)$ has no stratified critical points at level ℓ . Hence by Lemma 12 we conclude that $N_\star(B, g, \ell) = N_\star(B, g', \ell)$, as required. □

Proof of Lemma 2. We first consider the case that $x \in \partial B$. We choose a box $B' \subset U$ which contains x in its interior (so B' is not a subset of B). We assume B' is small enough that it intersects only one component of $\{g \geq \ell\}$ (i.e. the component containing x) and that g has no stratified critical points at level ℓ in B' (this is possible since x is non-degenerate). Let $h \in C_c^2(B')$ such that $h(x) > 0$. Then for any $\delta > 0$

$$\{g + \delta h \geq \ell\} \cap B \setminus B' = \{g - \delta h \geq \ell\} \cap B \setminus B'.$$

Therefore any change in the number of excursion/level set components in B must be due to changes in $B' \cap B$. Now choose $\delta > 0$ sufficiently small that $g + th$ has no stratified critical points at level ℓ in B' for all $|t| \leq \delta$. By Lemma 12 there exists a stratified isotopy between the sets $\{g - \delta h \geq \ell\} \cap B'$ and $\{g + \delta h \geq \ell\} \cap B'$. In particular, both of these sets consist of a single component and have the same intersection with $\partial B'$ (since $h \equiv 0$ on $\partial B'$). This argument shows that the topology of the level set $\{g + th = \ell\}$ on a neighbourhood of B does not change as t varies in $[-\delta, \delta]$. Moreover only one component of this level set varies with t .

Consider the component of $\{g = \ell\}$ containing x ; if this component intersects $\mathbb{R}^d \setminus B$ then the same will be true of the corresponding components of $\{g + th = \ell\}$ and so there will be no change in the number of interior components, i.e. $N_\star(B, g - \delta h, \ell) = N_\star(B, g + \delta h, \ell)$ (see the first two panels of Figure 2). Now assume that the level component containing x does not intersect $\mathbb{R}^d \setminus B$. Then, since B' has no stratified critical points, the component of $\{g \geq \ell\} \cap B'$ containing x must intersect ∂B only at x (i.e. the excursion component touches the boundary at a single point). Since $h(x) > 0$, the set $\{g + \delta h \geq \ell\}$ must contain a neighbourhood of x whereas $\{g - \delta h \geq \ell\}$ does not contain x . Therefore the number of interior components increases by one as we pass from $\{g + \delta h \geq \ell\}$ to $\{g - \delta h \geq \ell\}$ (see the bottom panel of Figure 2). Note that this does not depend on the choice of h . The same conclusion applies to the number of level set components, since these are the boundaries of the excursion sets.

Now let us assume that x is in the interior of B . Let $B' \subset B$ be a box which intersects only one component of $\{g \geq \ell\} \cap B$ and contains x in its interior. Let $h \in C_c^2(B')$ with $h(x) > 0$ and let V be an open ball around x on which $h > 0$. Then for $t \in \mathbb{R}$

$$(4.1) \quad \{g + th \geq \ell\} \cap V = \{G \geq t\} \cap V,$$

where $G := (g - \ell)/h$. Moreover since $\nabla G = h^{-1} \nabla g - h^{-2}(g - \ell) \nabla h$ we see that x is a critical point for G at level 0. Calculating the Hessian of G in a similar way shows that x is a non-degenerate critical point for G . By the Morse lemma

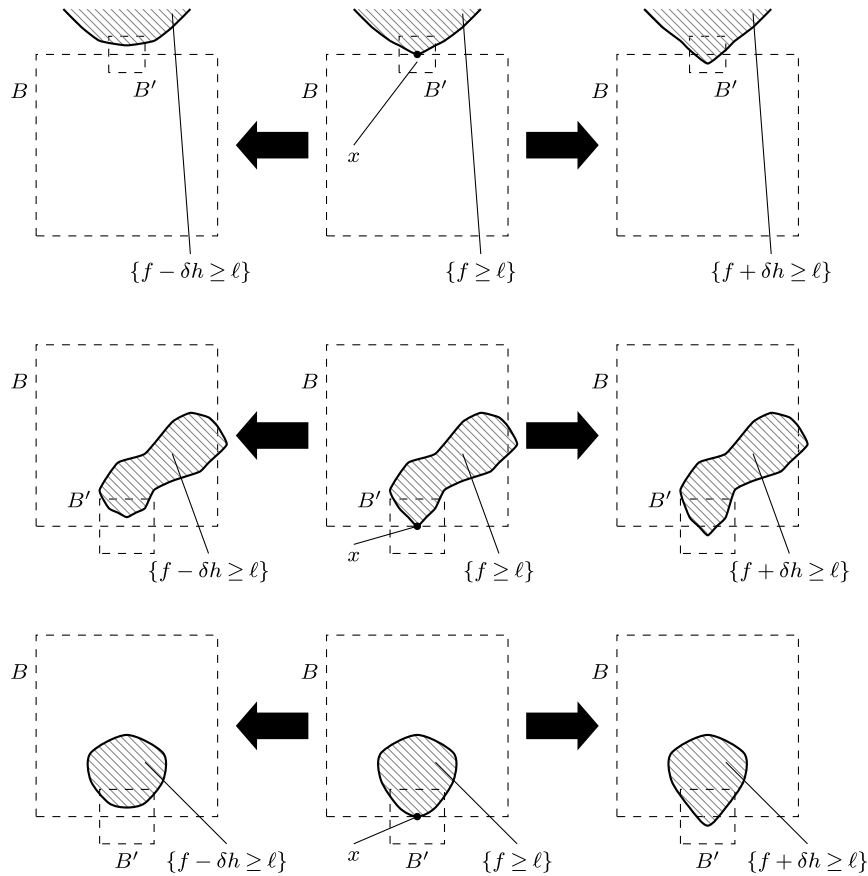


Fig. 2. Stratified critical points on the boundary and the effect of local perturbations on excursion/level sets. The excursion sets $\{f \geq \ell\}$ in the top and middle panels intersect $\mathbb{R}^d \setminus B$ and so they do not contribute to the (interior) component count regardless of the local perturbation around x . The excursion set $\{f \geq \ell\}$ in the bottom panel is interior but touches the boundary at x , therefore the component count is changed by a local perturbation.

[33, Theorem 1.1.12] (and reducing the size of V if necessary) there is a C^2 -chart $\phi : V \rightarrow \mathbb{R}^d$ such that $\phi(x) = 0$ and

$$(4.2) \quad G \circ \phi^{-1}(y_1, \dots, y_d) = -y_1^2 - \dots - y_i^2 + y_{i+1}^2 + \dots + y_d^2,$$

where i is the index of the critical point x . (Note that this index depends on g but not on h .)

We consider how the excursion sets of $g + th$ vary on each of the sets V , $B' \setminus V$ and $B \setminus B'$. Let $\delta > 0$ be small enough that $g + th$ has no stratified critical points at level ℓ in $B' \setminus V$ for all $|t| \leq \delta$. By Lemma 12 there exists a stratified isotopy between the sets

$$\{g + \delta h \geq \ell\} \cap B' \setminus V \quad \text{and} \quad \{g - \delta h \geq \ell\} \cap B' \setminus V.$$

(Strictly speaking we must repeat the proof of Lemma 12 for $B' \setminus V$ since this is not a box, however the arguments are identical.) Since h is supported on B' ,

$$\{g + \delta h \geq \ell\} \cap B \setminus B' = \{g - \delta h \geq \ell\} \cap B \setminus B'.$$

Therefore any change in the topology of excursion/level sets must be captured in V .

By (4.1) and (4.2) we see that if $i = 0$ then as t moves from $-\delta$ to δ the excursion set changes from the empty set to a ball (and the reverse occurs if $i = n$). In either case the number of excursion/level sets either increases or decreases by one. If $i \in \{1, n - 1\}$, then the level set goes from a one-sheet hyperboloid to two-sheet hyperboloid (with a cone at $t = 0$). In this case we gain (or lose) one level set component in V . The number of excursion set components might be constant (if $\lambda = n - 1$) or change by one (if $\lambda = 1$). In all other cases (i.e. $1 < \lambda < n - 1$) the corresponding hyperbolic surface is connected for all t and separates \mathbb{R}^n into two domains. This means that the number of components does not change.

Note that the above argument depends only on the canonical form, which is independent of h , hence the change in the number of components is the same for all functions h . □

Proof of Lemma 3. Let A be the component of $\{g \geq \ell\}$ containing x and let A_1, \dots, A_n be the components of $A \setminus \{x\}$. (The Morse lemma [33, Theorem 1.1.12] implies that there are only finitely many such components.) By reordering, we may assume that A_1, \dots, A_m are bounded and A_{m+1}, \dots, A_n are unbounded for some $m \in \{0, \dots, n\}$.

Now fix an $r > 1$ such that $A_1, \dots, A_m \subset x + \Lambda_r$ (if $m = 0$ then simply choose $r = 1$). We claim that r satisfies the conclusion of the lemma, i.e. there exists a $\Delta \in \{+, -, 0\}$ such that for any box $B \supset x + \Lambda_r$ the point x is (Δ, B) -pivotal. By Lemma 2, there exists $\Delta \in \{+, -, 0\}$ such that x is $(\Delta, x + \Lambda_r)$ -pivotal. By definition, for any sufficiently small neighbourhood W of x , any function $h \in C_c^2(W)$ satisfying $h(x) > 0$ and any δ sufficiently small

$$(4.3) \quad N_\star(x + \Lambda_r, g + \delta h, \ell) - N_\star(x + \Lambda_r, g - \delta h, \ell) = \begin{cases} 1 & \text{if } \Delta = +, \\ -1 & \text{if } \Delta = -, \\ 0 & \text{if } \Delta = 0. \end{cases}$$

If $B \supset (x + \Lambda_r)$ then by definition

$$(4.4) \quad N_\star(B, g - \delta h, \ell) = N_\star(x + \Lambda_r, g - \delta h, \ell) + \tilde{N}_\star(B, x + \Lambda_r, g - \delta h, \ell),$$

where the final term denotes the number of components of $\{g - \delta h \geq \ell\}$ which are contained in B and intersect $B \setminus (x + \Lambda_r)$. Any component contributing to this latter term must be bounded and is not contained in $x + \Lambda_r$, so cannot intersect $A_1 \cup \dots \cup A_m$. Therefore there is a neighbourhood W' of x which does not intersect any component contributing to $\tilde{N}_\star(B, x + \Lambda_r, g - \delta h, \ell)$. These components are disjoint from perturbations on W' and therefore

$$\tilde{N}_\star(B, x + \Lambda_r, g - \delta h, \ell) = \tilde{N}_\star(B, x + \Lambda_r, g + \delta h, \ell)$$

for $h \in C_c^2(W')$. Combining this with (4.3), (4.4), and the corresponding equation with $g + \delta h$ replacing $g - \delta h$ we have

$$N_\star(B, g - \delta h, \ell) - N_\star(B, g + \delta h, \ell) = \begin{cases} 1 & \text{if } \Delta = +, \\ -1 & \text{if } \Delta = -, \\ 0 & \text{if } \Delta = 0, \end{cases}$$

completing the proof. □

Appendix: Non-degeneracy properties of smooth Gaussian fields

In this appendix we collect some basic non-degeneracy properties of smooth Gaussian fields. Throughout we fix $k \in \mathbb{N}$ and let f be a C^k -smooth Gaussian field on \mathbb{R}^d . We define \mathcal{A} to be the collection of multi-indices on d variables with order at most k , and for $A \subseteq \mathcal{A}$ we let

$$\text{deg}(A) = \max\{|\alpha| : \alpha \in A\} \leq k$$

be the largest order among indices in A . For $x \in \mathbb{R}^d$ and $A \subseteq \mathcal{A}$, we define the Gaussian vector $D^A f(x) = (\partial^\alpha f(x))_{\alpha \in A}$.

A.1. Non-degeneracy of the field and its derivatives

Our first lemma concerns non-degeneracy properties of $D^A f(x)$, including when jointly evaluated at multiple points.

We say that a set of multi-indices $A \subseteq \mathcal{A}$ is *spherical* if there exists a multi-index $\alpha' \in A$ such that

$$\{\alpha' + \alpha : \alpha \in \{(0, \dots, 0), (2, 0, \dots, 0), (0, 2, \dots, 0), \dots, (0, 0, \dots, 2)\}\} \subseteq A,$$

and *non-spherical* otherwise. For example, if $d = k = 3$ then

$$\{(1, 0, 0), (3, 0, 0), (1, 2, 0), (1, 0, 2)\} \quad \text{and} \quad \{(1, 0, 1), (3, 0, 0), (1, 2, 0), (1, 0, 2)\}$$

are respectively spherical and non-spherical.

Lemma 13. *Suppose f is stationary and let μ be its spectral measure. Let $x_1, \dots, x_n \in \mathbb{R}^d$ be distinct points, and let $A_1, \dots, A_n \subseteq \mathcal{A}$.*

1. If the support of μ contains an open set, the Gaussian vector $(D^{A_i} f(x_i))_{1 \leq i \leq n}$ is non-degenerate.
2. Suppose each A_i is non-spherical, and either (i) $n = 1$, or (ii) $n = 2$ and $\deg(A_i) \leq 1, i = 1, 2$. Then if the support of μ contains a scaled sphere $a\mathbb{S}^{d-1}, a > 0$, the Gaussian vector $(D^{A_i} f(x_i))_{1 \leq i \leq n}$ is non-degenerate.

Remark 14. In the paper we do not use the full strength of Lemma 13. Indeed we only use the fact that if the support of μ contains either (i) an open set, or (ii) $a\mathbb{S}^{d-1}$, then the following vectors are non-degenerate for $x \neq 0$ and $v \in \mathbb{S}^{d-1}$:

- $(f(0), \nabla f(0), f(x), \nabla f(x))$;
- $(f(0), \nabla f(0), \partial_v \nabla f(0), \partial_v^3 f(0))$;
- (only in case (i)) $(f(0), \nabla f(0), \nabla^2 f(0))$.

Note in particular that the second vector can be written (after a linear transformation of full rank) as $D^A f(0)$ for a non-spherical $A \subset \mathcal{A}$.

Remark 15. We conjecture that the restrictions on n and $\deg(A_i)$ in the second item of Lemma 13 are not necessary for the conclusion of the lemma, but their inclusion simplifies the proof and is sufficient for our needs. On the other hand, the restriction to non-spherical A_i is necessary, for instance, if the support of μ is contained in a sphere $a\mathbb{S}^{d-1}$ then the vector $D^A f(0) = (f(0), \nabla^2 f(0))$, corresponding to a spherical $A \subset \mathcal{A}$, is degenerate.

In the proof of Lemma 13 we use some facts about the zero locus of certain analytic functions. For distinct $x_1, \dots, x_n \in \mathbb{R}^d, A_1, \dots, A_n \subseteq \mathcal{A}$, and real coefficients $c = (c_{i,\alpha})_{1 \leq i \leq n, \alpha \in A_i}$, define the function

$$g = g_{x_i; A_i; c} : \mathbb{R}^d \rightarrow \mathbb{C}^d, \quad g(t) = \sum_{1 \leq i \leq n, \alpha \in A_i} c_{i,\alpha} (-i)^{|\alpha|} t^\alpha e^{-it \cdot x_i},$$

and let $Z = \{t \in \mathbb{R}^d : g(t) = 0\}$ be its zero locus.

Lemma 14. Fix distinct $x_1, \dots, x_n \in \mathbb{R}^d, A_1, \dots, A_n \subseteq \mathcal{A}$, and $c = (c_{i,\alpha})_{1 \leq i \leq n, \alpha \in A_i}$.

1. If Z contains an open set, then $c = 0$.
2. Suppose each A_i is non-spherical, and either (i) $n = 1, x_1 = 0$, or (ii) $n = 2, x_1 = 0$, and $\deg(A_1), \deg(A_2) \leq 1$. If Z contains the sphere \mathbb{S}^{d-1} , then $c = 0$.

Proof. In both statements it suffices to prove that $g = 0$, since this implies that $c = 0$ as follows. Suppose $c \neq 0$, and rewrite g as

$$g(t) = \sum_{\alpha' \in \mathcal{A}} (-i)^{|\alpha'|} t^{\alpha'} h_{\alpha'}(t), \quad h_{\alpha'}(t) = \sum_{\{1 \leq i \leq n, \alpha \in A_i : \alpha = \alpha'\}} c_{i,\alpha} e^{-it \cdot x_i}.$$

Noting that $h_{\alpha'}(t)$ is bounded and non-zero almost everywhere, we may take $|t| \rightarrow \infty$ in such a way that $|g(t)| \rightarrow \infty$, and hence $g \neq 0$. So it remains to prove that g is identically zero.

We separate g into its real and imaginary parts

$$g_1 : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad g_1(t) = \sum_{i=1}^n \sum_{\substack{\alpha \in A_i \\ |\alpha|=2j}} c_{i,\alpha} (-1)^j t^\alpha \cos(t \cdot x_i) + \sum_{\substack{\alpha \in A_i \\ |\alpha|=2j+1}} c_{i,\alpha} (-1)^{j+1} t^\alpha \sin(t \cdot x_i)$$

and

$$g_2 : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad g_2(t) = \sum_{i=1}^n \sum_{\substack{\alpha \in A_i \\ |\alpha|=2j}} c_{i,\alpha} (-1)^{j+1} t^\alpha \sin(t \cdot x_i) + \sum_{\substack{\alpha \in A_i \\ |\alpha|=2j+1}} c_{i,\alpha} (-1)^j t^\alpha \cos(t \cdot x_i),$$

both of whose zero locus contains Z .

For the first statement, since g_1 and g_2 are real-analytic functions whose zero loci contain open sets, g_1 and g_2 are identically zero, hence so is g .

For the second statement, we split into the two cases (i) and (ii). In case (i), since $n = 1$ and $x_1 = 0$, the functions

$$P_1 : \mathbb{C}^d \rightarrow \mathbb{C}, \quad P_1(t) = g_1(t) = \sum_{\substack{\alpha \in A_1 \\ |\alpha|=2j}} c_{1,\alpha} (-1)^j t^\alpha$$

and

$$P_2 : \mathbb{C}^d \rightarrow \mathbb{C}, \quad P_2(t) = g_2(t) = \sum_{\substack{\alpha \in A_1 \\ |\alpha|=2j+1}} c_{1,\alpha} (-1)^{j+1} t^\alpha$$

are polynomials that vanish on the real sphere \mathbb{S}^{d-1} , and hence also on the complex sphere since the former is Zariski dense in the latter. Therefore, by Hilbert's Nullstellensatz, both $P_1(t)$ and $P_2(t)$ have the polynomial $Q(t) = 1 - \sum_{1 \leq i \leq d} t_i^2$ as a divisor, which contradicts the fact that A_1 is assumed non-spherical unless $g = 0$. In case (ii), without loss of generality we may assume x_2 lies on the d^{th} coordinate axis. Then considering the restriction of g_1 and g_2 to $\{t_d = 0\}$, we see that

$$P'_1 : \mathbb{C}^{d-1} \rightarrow \mathbb{C}, \quad P'_1(t) = g_1(t, 0) = c_{1,0} + c_{2,0}$$

and

$$P'_2 : \mathbb{C}^{d-1} \rightarrow \mathbb{C}, \quad P'_2(t) = g_2(t, 0) = (-1) \sum_{1 \leq j \leq d-1} (c_{1,j} + c_{2,j}) t_j$$

are affine functions that vanish on the real sphere \mathbb{S}^{d-2} , and therefore have null coefficients. Thus viewed as a function on \mathbb{R}^d , $g(t)$ depends only on the coordinate t_d . Since it vanishes on \mathbb{S}^{d-1} , Z contains an open set, and $g = 0$. \square

Proof of Lemma 13. In the case $A_i = \{(0, \dots, 0)\}$ this is a well-known result (see [45, Theorem 6.8]), and Lemma 14 allows us to extend the argument to the general case.

Suppose for the sake of contradiction that the vector $(D^{A_i} f(x_i))_{1 \leq i \leq n}$ is degenerate. By definition this means that there exists non-zero $c = (c_{i,\alpha})_{1 \leq i \leq n, \alpha \in A_i}$ such that

$$0 = \text{Var} \left(\sum_{1 \leq i \leq n, \alpha \in A_i} c_{i,\alpha} \partial^\alpha f(x_i) \right) = \int_{\mathbb{R}^d} \left| \sum_{1 \leq i \leq n, \alpha \in A_i} c_{i,\alpha} (-i)^{|\alpha|} t^\alpha e^{-it \cdot x_i} \right|^2 d\mu(t).$$

Therefore the function

$$g(t) = \sum_{1 \leq i \leq n, \alpha \in A_i} c_{i,\alpha} (-i)^{|\alpha|} t^\alpha e^{-it \cdot x_i}$$

vanishes on the support of μ , and the desired contradiction follows from Lemma 14 (after transformation and scaling). \square

A.2. Non-degeneracy under the interpolation

Let \tilde{f} be an independent copy of the field f , and for $t \in [0, 1]$ define $f^t = tf + \sqrt{1-t^2} \tilde{f}$.

Our next lemma states that, for $t \in [0, 1)$, the pair of fields (f, f^t) inherits the non-degeneracy properties of f . In particular this shows the non-degeneracy of the conditioning (2.1) in the definition of the pivotal measures in Theorem 1.

Lemma 15. *Let $x_1, x_2 \in \mathbb{R}^d$ be points, not necessarily distinct, and let $A_1, A_2 \subseteq \mathcal{A}$. Suppose the Gaussian vectors $D^{A_1} f(x_1)$ and $D^{A_2} f(x_1)$ are both non-degenerate. Then the Gaussian vector $(D^{A_1} f(x_1), D^{A_2} f^t(x_2))$ is non-degenerate for every $t \in [0, 1)$.*

Proof. By the third item of Lemma 8

$$\begin{aligned} \text{DC}(D^{A_1} f(x_1), D^{A_2} f^t(x_2)) &= \text{DC}(D^{A_1} f(x_1)) \text{DC}(tD^{A_2} f(x_2) + \sqrt{1-t^2} D^{A_2} \tilde{f}(x_2) \mid D^{A_1} f(x_1)) \\ &\geq \text{DC}(D^{A_1} f(x_1)) \text{DC}(tD^{A_2} f(x_2) + \sqrt{1-t^2} D^{A_2} \tilde{f}(x_2) \mid f). \end{aligned}$$

By the first item of Lemma 8, and the equality in law of f and \tilde{f} , the above is equal to

$$\mathrm{DC}(D^{A_1} f(x_1))\mathrm{DC}(\sqrt{1-t^2} D^{A_2} \tilde{f}(x_2)) = (1-t^2)^{|A_2|} \mathrm{DC}(D^{A_1} f(x_1))\mathrm{DC}(D^{A_2} f(x_2)) > 0. \quad \square$$

Acknowledgments

The authors are grateful to Hugo Vanneville for pointing out in [44] that interpolation methods can be used to study the variance of the component count. We also thank an anonymous referee for detailed references to results in [21], as well as general comments and corrections which improved the presentation of this work.

Funding

The second author was supported by the European Research Council (ERC) Advanced Grant QFPROBA (grant number 741487).

The third author was supported by the Australian Research Council (ARC) Discovery Early Career Researcher Award DE200101467, and acknowledges the hospitality of the Statistical Laboratory, University of Cambridge, where part of this work was carried out.

References

- [1] R. J. Adler and J. E. Taylor. *Random Fields and Geometry. Springer Monographs in Mathematics*. Springer, New York, 2007. MR2319516
- [2] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay. The statistics of peaks of Gaussian random fields. *Astrophys. J.* **304** (1986) 15–61.
- [3] V. Beffara and D. Gayet. Percolation of random nodal lines. *Publ. Math. Inst. Hautes Études Sci.* **126** (2017) 131–176. MR3735866 <https://doi.org/10.1007/s10240-017-0093-0>
- [4] D. Beliaev, M. McAuley and S. Muirhead. On the number of excursion sets of planar Gaussian fields. *Probab. Theory Related Fields* **178** (2020) 655–698. MR4168385 <https://doi.org/10.1007/s00440-020-00984-9>
- [5] D. Beliaev, M. McAuley and S. Muirhead. Smoothness and monotonicity of the excursion set density of planar Gaussian fields. *Electron. J. Probab.* **25** (2020) 93. MR4136473 <https://doi.org/10.1214/20-ejp470>
- [6] D. Beliaev, M. McAuley and S. Muirhead. Fluctuations of the number of excursion sets of planar Gaussian fields. *Probab. Math. Phys.* **3** (2022) 105–144. MR4420297 <https://doi.org/10.2140/pmp.2022.3.105>
- [7] D. Beliaev, M. McAuley and S. Muirhead. A central limit theorem for the number of excursion set components of Gaussian fields, *Ann. Probab.* To appear, 2023. MR4168385 <https://doi.org/10.1007/s00440-020-00984-9>
- [8] D. Beliaev, S. Muirhead and A. Rivera. A covariance formula for topological events of smooth Gaussian fields. *Ann. Probab.* **48** (2020) 2845–2893. MR4164455 <https://doi.org/10.1214/20-AOP1438>
- [9] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge University Press, Cambridge, 1987. MR0898871 <https://doi.org/10.1017/CBO9780511721434>
- [10] S. G. Bobkov, F. Götze and C. Houdré. On Gaussian and Bernoulli covariance representations. *Bernoulli* **7** (2001) 439–451. MR1836739 <https://doi.org/10.2307/3318495>
- [11] E. Bogomolny and C. Schmit. Percolation model for nodal domains of chaotic wave functions. *Phys. Rev. Lett.* **88** (2002) 114102.
- [12] T. Cacoullos. On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.* **10** (1982) 799–809. MR0659549
- [13] S. Chatterjee Chaos, concentration, and multiple valleys, 2008. arXiv preprint. Available at [arXiv:0810.4221](https://arxiv.org/abs/0810.4221).
- [14] S. Chatterjee. *Superconcentration and Related Topics. Springer Monographs in Mathematics*. Springer, Cham, 2014. MR3157205 <https://doi.org/10.1007/978-3-319-03886-5>
- [15] L. H. Y. Chen. An inequality for the multivariate normal distribution. *J. Multivariate Anal.* **12** (1982) 306–315. MR0661566 [https://doi.org/10.1016/0047-259X\(82\)90022-7](https://doi.org/10.1016/0047-259X(82)90022-7)
- [16] J. T. Cox and G. Grimmett. Central limit theorems for associated random variables and the percolation model. *Ann. Probab.* **12** (1984) 514–528. MR0735851
- [17] R. D. Edwards and R. C. Kirby. Deformations of spaces of imbeddings. *Ann. of Math. (2)* **93** (1971) 63–88. MR0283802 <https://doi.org/10.2307/1970753>
- [18] A. I. Elizarov. On the variance of the number of stationary points of a homogeneous Gaussian field. *Theory Probab. Appl.* **29** (1985) 569–570. MR0761142
- [19] A. Estrade and J. Fournier. Number of critical points of a Gaussian random field: Condition for a finite variance. *Statist. Probab. Lett.* **118** (2016) 94–99. MR3531487 <https://doi.org/10.1016/j.spl.2016.06.018>
- [20] A. Estrade and J. R. León. A central limit theorem for the Euler characteristic of a Gaussian excursion set. *Ann. Probab.* **44** (2016) 3849–3878. MR3572325 <https://doi.org/10.1214/15-AOP1062>
- [21] M. Goresky and R. MacPherson. *Stratified Morse Theory. Ergebnisse der Mathematik und Ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **14**. Springer-Verlag, Berlin, 1988. MR0932724 <https://doi.org/10.1007/978-3-642-71714-7>
- [22] M. Kratz and S. Vadlamani. Central limit theorem for Lipschitz–Killing curvatures of excursion sets of Gaussian random fields. *J. Theoret. Probab.* **31** (2018) 1729–1758. MR3842168 <https://doi.org/10.1007/s10959-017-0760-6>
- [23] M. F. Kratz and J. R. León. Central limit theorems for level functionals of stationary Gaussian processes and fields. *J. Theoret. Probab.* **14** (2001) 639–672. MR1860517 <https://doi.org/10.1023/A:1017588905727>

- [24] P. Kurlberg and I. Wigman. Variation of the Nazarov–Sodin constant for random plane waves and arithmetic random waves. *Adv. Math.* **330** (2018) 516–552. MR3787552 <https://doi.org/10.1016/j.aim.2018.03.026>
- [25] M. Ledoux. L’algèbre de Lie des gradients itérés d’un générateur Markovien – développements de moyennes et entropies. *Ann. Sci. Éc. Norm. Supér.* **28** (1995) 435–460. MR1334608
- [26] D. Marinucci, M. Rossi and I. Wigman. The asymptotic equivalence of the sample trispectrum and the nodal length for random spherical harmonics. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** (2020) 374–390. MR4058991 <https://doi.org/10.1214/19-AIHP964>
- [27] J. N. Mather. Stratifications and mappings. In (*Proc. Sympos., Univ. Bahia, Salvador, 1971*) 195–232. Academic Press, New York-London, 1973. MR0368064
- [28] S. Muirhead, A. Rivera and H. Vanneuville. The phase transition for planar Gaussian percolation models without FKG. *Ann. Probab.* **51** (2023) 1785–1829. With an appendix by L. Köhler-Schindler. MR4642224 <https://doi.org/10.1214/23-aop1633>
- [29] D. Müller. A central limit theorem for Lipschitz–Killing curvatures of Gaussian excursions. *J. Math. Anal. Appl.* **452** (2017) 1040–1081. MR3632692 <https://doi.org/10.1016/j.jmaa.2017.03.036>
- [30] G. Naitzat and R. J. Adler. A central limit theorem for the Euler integral of a Gaussian random field. *Stochastic Process. Appl.* **127** (2017) 2036–2067. MR3646439 <https://doi.org/10.1016/j.spa.2016.09.007>
- [31] F. Nazarov and M. Sodin. Asymptotic laws for the spatial distribution and the number of connected components of zero sets of Gaussian random functions. *J. Math. Phys. Anal. Geom.* **12** (2016) 205–278. MR3522141 <https://doi.org/10.15407/mag12.03.205>
- [32] F. Nazarov and M. Sodin. Fluctuations in the number of nodal domains. *J. Math. Phys.* **61** (2020) 123302. MR4191640 <https://doi.org/10.1063/5.0018588>
- [33] L. Nicolaescu. *An Invitation to Morse Theory*, 2nd edition. *Universitext*. Springer, New York, 2011. MR2883440 <https://doi.org/10.1007/978-1-4614-1105-5>
- [34] I. Nourdin, G. Peccati and M. Rossi. Nodal statistics of planar random waves. *Comm. Math. Phys.* **369** (2019) 99–151. MR3959555 <https://doi.org/10.1007/s00220-019-03432-5>
- [35] M. D. Penrose. A central limit theorem with applications to percolation, epidemics and Boolean models. *Ann. Probab.* **29** (2001) 1515–1546. MR1880230 <https://doi.org/10.1214/aop/1015345760>
- [36] V. I. Piterbarg. *Asymptotic Methods in the Theory of Gaussian Processes and Fields. Translations of Mathematical Monographs* **148**. American Mathematical Society, Providence, RI, 1996. Translated from the Russian by V. V. Piterbarg, Revised by the author. MR1361884 <https://doi.org/10.1090/mmono/148>
- [37] P. Pranav et al. Unexpected topology of the temperature fluctuations in the cosmic microwave background. *Astron. Astrophys.* **627** (2019) A163.
- [38] L. Priya. Concentration for nodal component count of Gaussian Laplace eigenfunctions, 2020. arXiv preprint. Available at [arXiv:2012.10302](https://arxiv.org/abs/2012.10302).
- [39] J. S. Ranjan and R. Samajdar. Nodal portraits of quantum billiards: Domains, lines, and statistics. *Rev. Modern Phys.* **89** (2017) 045005. MR3746353 <https://doi.org/10.1103/RevModPhys.89.045005>
- [40] A. Rivera and H. Vanneuville. Quasi-independence for nodal lines. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (2019) 1679–1711. MR4010948 <https://doi.org/10.1214/18-aihp931>
- [41] E. V. Slud. MWI representation of the number of curve-crossings by a differentiable Gaussian process, with applications. *Ann. Probab.* **22** (1994) 1355–1380. MR1303648
- [42] K. Tanguy. Some superconcentration inequalities for extrema of stationary Gaussian processes. *Statist. Probab. Lett.* **106** (2015) 239–246. MR3389997 <https://doi.org/10.1016/j.spl.2015.07.028>
- [43] V. Tassion and H. Vanneuville. Noise sensitivity of percolation via differential inequalities. *Proc. Lond. Math. Soc. (3)* **126** (2023) 1063–1091. MR4574826
- [44] H. Vanneuville. Reading group on random nodal lines, 2019. Available at at ETH Zurich. https://metaphor.ethz.ch/x/2019/hs/401-4600-69L/Reading_group_nodal_1_2.pdf. Accessed on 25/01/23.
- [45] H. Wendland. Scattered data approximation. In *Cambridge Monographs on Applied and Computational Mathematics* **17**. Cambridge University Press, Cambridge, 2005. MR2131724
- [46] K. J. Worsley et al. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* **4** (1996) 58–73.
- [47] Y. Zhang. A martingale approach in the study of percolation clusters on the \mathbb{Z}^d lattice. *J. Theoret. Probab.* **14** (2001) 165–187. MR1822899 <https://doi.org/10.1023/A:1007877216583>