A COVARIANCE FORMULA FOR TOPOLOGICAL EVENTS OF SMOOTH GAUSSIAN FIELDS

DMITRY BELIAEV[1], STEPHEN MUIRHEAD[2], AND ALEJANDRO RIVERA[3]

ABSTRACT. We derive a covariance formula for the class of 'topological events' of smooth Gaussian fields on manifolds; these are events that depend only on the topology of the level sets of the field, for example (i) crossing events for level or excursion sets, (ii) events measurable with respect to the number of connected components of level or excursion sets of a given diffeomorphism class, and (iii) persistence events. As an application of the covariance formula, we derive strong mixing bounds for topological events, as well as lower concentration inequalities for additive topological functionals (e.g. the number of connected components) of the level sets that satisfy a law of large numbers. The covariance formula also gives an alternate justification of the Harris criterion, which conjecturally describes the boundary of the percolation university class for level sets of stationary Gaussian fields. Our work is inspired by [44], in which a correlation inequality was derived for certain topological events on the plane, as well as by [39], in which a similar covariance formula was established for finite-dimensional Gaussian vectors.

1. INTRODUCTION

In recent years there has been much progress in the study of the topology of level sets of smooth Gaussian fields. Techniques have been developed to estimate their homology (see [36, 37], and also [10, 14, 20, 30, 46]), and also their large scale connectivity properties (see [1, 5], and also [9, 34, 35, 43]) using ideas from Bernoulli percolation. When studying the topology of level sets, one often has to estimate quantities such as

$$\mathrm{Cov}(A_1, A_2) := \mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2],$$

where A_1 and A_2 are events of topological nature. Since the events A_1 and A_2 in general do not admit explicit integral representations, the quantity $\mathrm{Cov}(A_1, A_2)$ is often estimated indirectly, leading to inequalities of varying precision. In the present work we prove an *exact* formula for $\mathrm{Cov}(A_1, A_2)$, where A_1 and A_2 belong to a large class of 'topological events'.

Let us illustrate our formula with a simple example. Let f be an a.s. C^2 centred Gaussian field on $\mathbb{R}^2$, with covariance $K(x, y) := \mathrm{Cov}(f(x), f(y))$, such that, for each distinct $x, y \in \mathbb{R}^2$,

$(f(x), \nabla f(x), f(y), \nabla f(y))$ is a non-degenerate Gaussian vector. Let B_1 and B_2 be two boxes on the plane $\mathbb{R}^2$, not necessarily disjoint, each with two opposite sides distinguished (we call these 'left' and 'right', with the remaining sides being 'top' and 'bottom'). For each $i \in \{1, 2\}$, consider the event A_i that there exists a continuous path in $B_i \cap \{f \geq 0\}$ joining the 'left' and 'right' sides. This is known as a 'crossing event' for the excursion set $\{f \geq 0\}$, and is of fundamental importance in the study of the connectivity of the level sets [5]. As a corollary of our general covariance formula, we establish the following exact formula for $\mathrm{Cov}(A_1, A_2)$:

Corollary 1.1. *The quantity* $\mathrm{Cov}(A_1, A_2)$ *is equal to*

$$\sum_{j_1, j_2 = 0, 1, 2, 3, 4} \int_{F_{j_1}^1 \times F_{j_2}^2} K(x_1, x_2) \int_0^1 \gamma_{t; x_1, x_2}(0) \, \mathbb{E}_{t; x_1, x_2} \left[\prod_{i=1,2} |\det(H_{x_i}^{j_i} f_t^i)| \mathbb{1}_{\mathrm{Piv}_{x_i}^{t,i}(A_i)} \right] dt \, \mathrm{dv}_{F_{j_1}^1} \mathrm{dv}_{F_{j_2}^2},$$

where:

- *For each* $i \in \{1, 2\}$, $F_0^i := \mathring{B}_i$ *denotes the interior of* B_i, *equipped with its two-dimensional Lebesgue measure* $\mathrm{dv}_{F_0^i}$, *and* $(F_j^i)_{j=1,2,3,4}$ *denote the sides of* B_i, *equipped with their natural length measure* $\mathrm{dv}_{F_j^i}$; *the* F_j^i *are therefore disjoint.*
- *For each* $t \in [0, 1]$, $f_t = (f_t^1, f_t^2) = (f^1, tf^1 + \sqrt{1 - t^2} f^2)$ *denotes a Gaussian field on* $\mathbb{R}^2 \times \mathbb{R}^2$ *that interpolates between* (f^1, f^1) *and* (f^1, f^2), *where* f^1 *and* f^2 *are independent copies of* f. *For each distinct* $x_1 \in B_1$ *and* $x_2 \in B_2$, $\gamma_{t; x_1, x_2}(0)$ *denotes the density at* 0 *of the Gaussian vector*

$$(1.1) \qquad (f_t^1(x_1), \nabla f_t^1|_{F_{j_1}^1}(x_1), f_t^2(x_2), \nabla f_t^2|_{F_{j_2}^2}(x_2)) \in \mathbb{R} \times T_{x_1} F_{j_1}^1 \times \mathbb{R} \times T_{x_2} F_{j_2}^2,$$

  where $F_{j_i}^i$ *denotes the unique face/interior that contains* x_i; *moreover,* $\mathbb{E}_{t; x_1, x_2}[\cdot]$ *denotes expectation conditional on the vector* (1.1) *vanishing, and* $H_{x_i}^{j_i} f_t^i = \nabla^2 f_t^i|_{F_{j_i}}(x_i)$ *denotes the Hessian at the point* x_i *of* f_t^i *restricted to the face* $F_{j_i}^i$.
- *For each* $i \in \{1, 2\}$, $t \in [0, 1]$ *and* $x \in B_i$, $\mathrm{Piv}_x^{t,i}(A_i)$ *denotes the event that there exists a continuous path in* $B_i \cap \{f_t^i \geq 0\}$ *joining the 'left' and 'right' sides, and a continuous path in* $B_i \cap \{f_t^i \leq 0\}$ *joining the 'top' and 'bottom' sides, both of which pass through* x *(see Figure 1; central panels). This is a natural analogue of a 'pivotal event' in Bernoulli percolation (see [12, 23]).*

Let us make three observations concerning the formula in Corollary 1.1:

- If K is non-negative then so is the integrand in the formula, and we deduce that

$$\mathbb{P}[A_1 \cap A_2] \geq \mathbb{P}[A_1] \mathbb{P}[A_2].$$

  This is the analogue of the Fortuyn-Kasteleyn-Ginibre (FKG) inequality (see [12, 23]), originally proven in the Gaussian setting by Pitt [40].
- Assume that f is stationary, let $\kappa(x) = K(0, x)$ and denote $\overline{\kappa}(r) = \sup_{|x| \geq r} |\kappa(x)|$. Since f is Gaussian, the Hessians $H_{x_i}^{j_i} f_t^i$ have finite moments and so, by stationarity, the conditional expectation in the formula is bounded. Thus, if B_1 and B_2 have sides of length $O(R)$ and are at distance of order at least R, we deduce a 'strong mixing' bound for crossing events, namely that

$$(1.2) \qquad |\mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1] \mathbb{P}[A_2]| = O(R^4 \overline{\kappa}(R)).$$

  In particular, as long as $\kappa(R) = o(R^{-4})$, the crossing events A_1 and A_2 are asymptotically independent, recovering the recent result of Rivera and Vanneuville [44].
- Setting $B_1 = B_2$ (and so $A_1 = A_2$), Corollary 1.1 also yields a formula for the *variance* of (the indicator function of) the crossing event A_i.
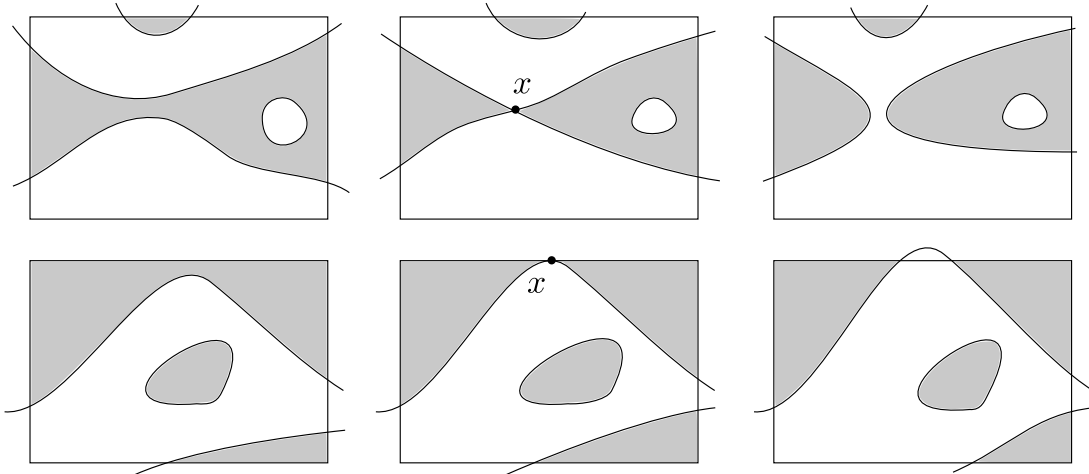
FIGURE 1. An illustration of the crossing events A_i and the pivotal events $\mathrm{Piv}_x(A_i)$ that appear in the covariance formula in Corollary 1.1. Left panels: Two realisations of a field f which exhibit the left-right crossing event for $\{f > 0\}$ in the rectangle B (shown in grey). Right panels: After a small perturbation of f (compared to the left panel), the left-right crossing event no longer occurs. Central panels: The 'pivotal event' at which the crossing event first fails in this perturbation; this event can be of two possible types, either involving a level-0 critical point x of f in the interior of B (top figure), or involving a level-0 critical point x of f restricted to the top side of B (bottom figure).

The main result of this paper (see Theorem 2.14) consists of a vast generalisation of Corollary 1.1 to the class of topological events of smooth Gaussian fields on manifolds of any dimension. In particular, this permits a generalisation of the mixing bound (1.2) to arbitrary topological events on manifolds (see Corollary 1.2 for the Euclidean case and Theorem 2.15 for the general case). Since the statement of Theorem 2.14 requires several preliminary definitions, in this introduction we instead focus on applications of this formula, including (i) the aforementioned strong mixing bounds, and (ii) lower concentration inequalities for additive topological functionals of the level sets, such as such as the number of connected components contained in a given domain.

Our work was largely inspired by [44] in which the mixing bound (1.2) was first established, improving similar bounds that had previously appeared in [5, 8]. Here we extend the techniques and results in [44] to arbitrary topological events and to higher dimensions; the key difference in our approach is that we work *directly in the continuum*, rather than with discretisations of the field as in [5, 8, 44].

1.1. **Topological events.** We begin by describing the class of topological events to which our results apply. Broadly speaking, we study events that depend only on the topology of the level sets $\{f = \ell\}$ (or excursion sets $\{f > \ell\}$) of a Gaussian field f restricted to reasonable bounded domains $B \subset \mathbb{R}^d$. One might think that it would therefore be enough to study homeomorphism classes of pairs $(\{f > \ell\} \cap B, B)$, however, this would in fact *not* identify crossing events, which distinguish marked sides of the reference domain B. Moreover, as in the case of a product of homeomorphic sets, one might wish to distinguish between factors. For these reasons, we work instead with equivalence classes induced by isotopies that preserve certain subsets of B, using the formalism of *stratifications*.

An *affine stratified set* in $\mathbb{R}^d$ is a compact subset $B \subset \mathbb{R}^d$ equipped with a finite partition $B = \sqcup_{F \in \mathcal{F}} F$ into open connected subsets of affine subspaces of $\mathbb{R}^d$, such that for each $F, F' \in \mathcal{F}$, $F \cap \overline{F'} \neq \emptyset \Rightarrow F \subset \overline{F'}$. The partition $\mathcal{F}$ is called a *stratification* of B. When there is no risk of

ambiguity, we will often refer to B itself as an affine stratified set. For example, a closed cube in $\mathbb{R}^d$, equipped with the collection of the interiors of its faces of all dimensions, is an affine stratified set.

Given an affine stratified set $(B, \mathcal{F})$ of $\mathbb{R}^d$ and a continuous map $H : B \times [0, 1] \to B$, we say that H is a *stratified isotopy if* for each $t \in [0, 1]$, $H(\cdot, t)$ is a homeomorphism such that for each $F \in \mathcal{F}$, $H(F \times \{t\}) = F$. The *stratified isotopy class* of a subset $E \subset B$, denoted by $[E]_B$, is the set of $H(E \times \{1\})$ where $H : B \times [0, 1] \to B$ ranges over the set of stratified isotopies of B with $H(\cdot, 0) = id_B$. We consider the stratified isotopy class $[\{f > 0\}]_B$ of the excursion set $\{f > 0\}$, which captures what we mean by the 'topology' of the level set $\{f = 0\}$ restricted to B. As we verify in Corollary 5.8, under mild conditions on f the stratified istotopy class $[\{f > 0\}]_B$ is measurable with respect to f.

A *topological event* in B is an event measurable with respect to $[\{f > 0\}]_B$. Important examples include:

- As in Corollary 1.1, crossing events for level or excursion sets inside a box B, e.g. the event that a connected component of $\{f = 0\} \cap B$ or $\{f > 0\} \cap B$ intersects opposite $(d-1)$-dimensional faces of B (Corollary 1.1 concerned the case $d = 2$).
- Events that depend on the number of the connected components of a level or excursion set inside a polytope B, or more generally the number of such components of a given diffeomorphism class (see, e.g., [14, 20, 36, 37, 46]).
- The 'persistence' event that $\{f|_B > 0\}$ (see, e.g., [2, 15, 19, 42]).

We write $\sigma_{\text{top}}(B)$ to denote the σ-algebra of topological events on B.

1.2. **Strong mixing in the Euclidean setting.** The *strong mixing* of a random field is defined via the decay, for domains B_1 and B_2 that are well-separated in space, of the α-mixing coefficient

$$(1.3) \qquad \alpha(B_1, B_2) = \sup_{A_1 \in \sigma(B_1),\, A_2 \in \sigma(B_2)} |\mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2]|,$$

where $\sigma(B)$ denotes the sub-σ-algebra generated by the restriction of f to the domain B. Strong mixing is a classical notion in probability theory with important connections to laws of large numbers, central limit theorems, and extreme value theory (see, e.g., [16, 31, 32, 45]) among other topics. While for general continuous processes there is a rich literature on strong mixing (see [13] for a review), in the study of *smooth* random fields the concept of strong mixing is often far too restrictive. For example, if the spectral density of a stationary Gaussian process decays exponentially (which implies the real analyticity of the covariance kernel and the corresponding sample paths), then by [27] there is no strong mixing regardless of how rapidly correlations decay, unless one restricts the class of events that are controlled by the α-mixing coefficient. As a first application of our covariance formula we derive conditions that guarantee the strong mixing of the class of topological events.

Let f be an a.s. C^2 stationary Gaussian field on $\mathbb{R}^d$ with covariance $\kappa(x) = \text{Cov}(f(0), f(x))$, and suppose that, for each distinct $x, y \in \mathbb{R}^d$, $(f(x), \nabla f(x), f(y), \nabla f(y))$ is a non-degenerate Gaussian vector. These conditions ensure that κ is C^4, and that the level set $\{f = 0\}$ is a C^2-smooth hypersurface. For each pair of affine stratified sets $B_1, B_2 \subset \mathbb{R}^d$, define the 'topological' α-mixing coefficient

$$\alpha_{\text{top}}(B_1, B_2) = \sup_{A_1 \in \sigma_{\text{top}}(B_1),\, A_2 \in \sigma_{\text{top}}(B_2)} |\mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2]|.$$

Corollary 1.2 (Strong mixing for topological events)**.** *There exist $c_1, c_2 > 0$ such that, for every pair of affine stratified sets $(B_1, \mathcal{F}_1)$ and $(B_2, \mathcal{F}_2)$ in $\mathbb{R}^d$ satisfying*

$$\max_{\alpha \in \mathbb{N}^d : |\alpha| \le 2} \sup_{x_1 \in B_1, x_2 \in B_2} |\partial^\alpha \kappa(x_1 - x_2)| < c_1,$$

it holds that

$$\alpha_{\text{top}}(B_1, B_2) \leq c_2 \, |\mathcal{F}_1||\mathcal{F}_2| \max_{F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2} \int_{F_1 \times F_2} |\kappa(x_1 - x_2)| \, \mathrm{dv}_{F_1}(x_1) \, \mathrm{dv}_{F_2}(x_2).$$

In particular, recalling that $\bar{\kappa}(s) = \sup_{|x| \geq s} |\kappa(x)|$, if

$$(1.4) \qquad \lim_{|x| \to \infty} |\partial^\alpha \kappa(x)| = 0 \,, \quad \text{for all } \alpha \in \mathbb{N}^d \text{ such that } |\alpha| \leq 2,$$

then for every pair of disjoint affine stratified sets $B_1, B_2 \subset \mathbb{R}^d$ there exist $c_3, c_4 > 0$ such that

$$(1.5) \qquad \alpha_{\text{top}}(sB_1, sB_2) \leq c_3 s^{2d} \, \bar{\kappa}(c_4 s) \quad \text{for all } s \geq 1.$$

Corollary 1.2 demonstrates that topological events on well-separated boxes $B_1, B_2 \subset \mathbb{R}^d$ are independent up to an additive error that depends (up to a constant) solely on the double integral of the absolute value of the covariance kernel on the boxes; we expect this result to have many applications. Later we present a generalisation of Corollary 1.2 to Gaussian fields on general manifolds (see Theorem 2.15). The proof of Corollary 1.2 is given in Section 6.

Remark 1.3. The constant c_1 in Corollary 1.2 can be chosen in a way that depends only on the dimension d, on $\kappa(0)$, and on the Hessian of κ at 0, whereas the constant c_2 can be chosen in a way that depends, in addition to these, also on $\max_j (\partial^4 \kappa(0)/\partial x_j^4)$.

Remark 1.4. We do not assume that the field f is centred. Since adding a constant does not change the covariance kernel, Corollary 1.2 also bounds the strong mixing of topological events that are defined in terms of non-zero levels. Notably, neither c_1 nor c_2 depends on the mean value of the field.

Remark 1.5. As explained above, the mixing bound in Corollary 1.2 was already known in two dimensions, at least in the case of crossing events [44] (see also (1.2)); our results extends this mixing bound to arbitrary dimensions and arbitrary topological events. Note also that an analogue of (1.5) was recently established [35] for a version of the α-mixing coefficient that controls all events (not necessarily topological) that depend *monotonically* on f (this includes, for instance, crossing events for $\{f > 0\}$); in this case the factor s^{2d} can be improved to s^d.

1.3. **Application to lower concentration for topological counts.** We next present a simple application of Corollary 1.2 to give a taste of the utility of mixing bounds. A *topological count* is a set of integer-valued random variables $N = N(B)$, indexed by affine stratified sets $B \subset \mathbb{R}^d$, each of which is measurable with respect to the corresponding σ-algebra $\sigma_{\text{top}}(B)$. We call a topological count *super-additive* if, for every affine stratified set B and every collection of disjoint affine stratified sets $(B_i)_{i \leq k}$ contained in B,

$$(1.6) \qquad N(B) \geq \sum_{i \leq k} N(B_i).$$

Examples of super-additive topological counts include the number of connected components of level or excursion sets that are fully contained in a set [37], or more generally the number of connected components of these sets that have a certain diffeomorphism class [14, 20, 46]. In one dimension, topological counts reduce to the number of solutions to $\{f = 0\}$ in intervals, a quantity studied extensively since the works of Kac and Rice in the 1940s [25, 41]. We say that a topological count N satisfies a *law of large numbers* if there exists a $c_N > 0$ such that, for every affine stratified set $B \subset \mathbb{R}^d$, as $s \to \infty$

$$(1.7) \qquad \frac{N(sB)}{s^d \operatorname{Vol}(B)} \to c_N \quad \text{in probability.}$$

Nazarov–Sodin have shown [36, 37] (see also [7, 30]) that if f is ergodic (and under certain mild extra conditions) the number of connected components of level or excursion sets satisfies

a law of large numbers, and in fact, (1.7) converges a.s. and in mean; the same result was later shown to be true also for the number of connected components of a given diffeomorphism type [10, 14, 46] (in the one dimensional case this follows immediately from the ergodic theorem). As was shown in [44], quantitative mixing bounds can be used to deduce the lower concentration of super-additive topological counts:

Corollary 1.6 (Lower concentration for topological counts)**.** *Let N denote a super-additive topological count that satisfies a law of large numbers (1.7) with limiting constant $c_N > 0$. Assume that (1.4) holds. Then for every affine stratified set $B \subset \mathbb{R}^d$ and constants $\varepsilon, C > 0$, there exist $c_1, c_B > 0$ such that, for every $s \geq 1$,*

$$(1.8) \qquad \mathbb{P}\left[\frac{N(sB)}{s^d \operatorname{Vol}(B)} \leq c_N - \varepsilon\right] \leq c_1 \inf_{r \in [1,s]} \left(e^{-C(s/r)^d} + e^{c_B(s/r)^d}(rs)^d \bar{\kappa}(r)\right),$$

where the constant $c_B > 0$ depends only on the stratified set B. In particular, if there exist $c_2, \alpha > 0$ such that $\kappa(x) \leq c_2|x|^{-\alpha}$ for every $|x| \geq 1$, then for every $\varepsilon, \delta > 0$ we can set $r = c_3 s/(\log s)^{1/d}$ for a sufficiently large choice of $c_3 > 0$ (depending on c_B, α and δ) and apply (1.8) for $C > 0$ sufficiently large (depending on c_3 and δ) to deduce the existence of a $c_4 > 0$ such that, for every $s \geq 1$,

$$\mathbb{P}\left[\frac{N(sB)}{s^d \operatorname{Vol}(B)} \leq c_N - \varepsilon\right] \leq c_4 s^{2d - \alpha + \delta}.$$

Similarly, if there exist $c_2, \alpha, \beta > 0$ such that $\kappa(x) \leq c_2 e^{-\beta|x|^\alpha}$ for every $|x| \geq 1$, then setting $r = c_3 s^{d/(d+\alpha)}$ for a sufficiently large choice of $c_3 > 0$ and then choosing C sufficiently large we deduce that for every $\gamma > 0$ there is $c_4 > 0$ such that, for every $s \geq 1$,

$$\mathbb{P}\left[\frac{N(sB)}{s^d \operatorname{Vol}(B)} \leq c_N - \varepsilon\right] \leq c_4 \exp\left(-\gamma s^{d\alpha/(d+\alpha)}\right).$$

Remark 1.7. As for Corollary 1.2, Corollary 1.6 was also already known in two dimensions (at least in the case of the number of connected components of level sets [44]) but not in higher dimensions. A stronger version of Corollary 1.6 was also recently established in the one dimensional case (i.e. for the number of zeros of a one-dimensional stationary Gaussian process [4]), and also for the number of connected components of the zero level set of random spherical harmonics (RSHs) [36]; the results in [4, 36] are proven using very different techniques to ours, and in the latter case relies heavily on the specific structure of the RSHs.

2. A covariance formula for topological events

In this section we present our covariance formula in the general setting of smooth Gaussian fields on smooth manifolds. We also discuss further applications of the formula beyond those we gave in Section 1, and give a sketch of its proof.

2.1. **The covariance formula.** We begin by fixing definitions, starting with the 'stratified sets' on which we work; our main reference is [22]. Let (M, g) be a smooth Riemannian manifold of dimension d.

Definition 2.1 (Stratified set)**.** Let $B \subset M$ be a compact subset. Assume there is a partition of B into a finite collection $\mathcal{F}$ of smooth locally closed submanifolds, called *strata*, satisfying the following additional properties:
- The strata cover B, i.e. $B = \coprod_{F \in \mathcal{F}} F$.
- Any two strata F_1 and F_2 satisfy $F_1 \cap \overline{F_2} \neq \emptyset \Leftrightarrow F_1 \subset \overline{F_2}$. This allows us to equip $\mathcal{F}$ with the partial order $<$ defined such that, for any two strata F_1 and F_2,

$$F_1 \cap \overline{F_2} \neq \emptyset \Leftrightarrow F_1 = F_2 \text{ or } F_1 < F_2.$$

- For each $F_1 < F_2$ the following is true. Consider any embedding of M in Euclidean space, and let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be sequences of points satisfying (i) for each $k \in \mathbb{N}$, $x_k \in F_2$ and $y_k \in F_1$, (ii) x_k and y_k converge to a common point $y \in F_1$, (iii) the tangent planes $T_{x_k}F_2$ converge to a limit τ, and (iv) the lines λ_k generated by the vectors $x_k - y_k$ converge to a limit λ. Then it holds that $\lambda \subset \tau$. Equivalently, it is enough that this condition be fulfilled for one fixed embedding of M in Euclidean space. Limits τ of this kind are called *generalised tangent spaces at y*.
- For each $F_1, F_2 \in \mathcal{F}$ such that $F_1 < F_2$, there exists a smooth sub-bundle $TF_2|_{F_1}$ of $TM|_{F_2}$, whose rank is the dimension of F_2, that contains TF_1 as a sub-bundle, and such that (i) the map $y \mapsto T_yF_2$, with values in the adequate Grassmannian bundle defined on F_2, extends by continuity to F_1 together with all of its derivatives, and (ii) for each sequence of points $x_k \in F_2$ converging to a limit $x \in F_1$, $\lim_{k\to+\infty} T_{x_1}F_2 = T_xF_2|_{F_1}$. We call $TF_2|_{F_1}$ the *generalised tangent bundle of F_2 over F_1* (see Figure 2).

The collection $\mathcal{F}$ is called a *tame stratification* of B. A *stratified set of M* is a pair $(B, \mathcal{F})$ consisting of a compact subset $B \subset M$ and a tame stratification $\mathcal{F}$ of B. When there is no risk of ambiguity, we will often write that $B \subset M$ is a stratified set without explicit mention of its tame stratification $\mathcal{F}$.
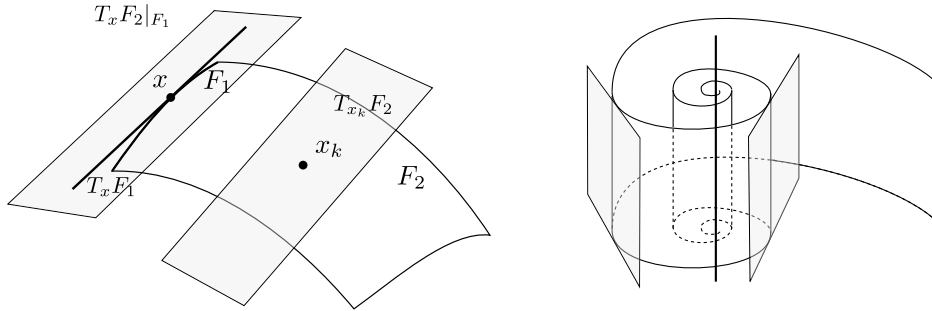

FIGURE 2. Left: An example of a tame stratification $\mathcal{F} = \{F_1, F_2\}$ of a compact set B. Here the generalised tangent bundle $T_xF_2|_{F_1}$ is well-defined since, as the points x_k converge to $x \in F_1$, the respective tangent planes also converge. Right: A rough depiction of the 'rapid spiral sheet', which is an example of a set that cannot be tamely stratified (see Example 2.6); here tangent planes do not converge, and so the generalised tangent bundle is not well-defined.

Remark 2.2. A partition $\mathcal{F}$ of a compact subset B satisfying the first three properties required in Definition 2.1 is called a Whitney stratification (see for instance Part I, Section 1.2 of [22]); indeed, the third property is known as 'Whitney's condition (b)'. While Whitney stratifications have many interesting properties, sometimes the structure of a stratification can force functions on it to have degenerate stratified critical points (see Example 2.6). To avoid such pathologies, we add the additional fourth condition which is satisfied in most natural examples. In fact, this additional 'tameness' property is only used at a single place in the proof of the covariance formula, namely, to prove Claim 4.6.

Let us present several important examples (and one non-example) of stratified sets, beginning with the trivial stratification:

Example 2.3 (Trivial stratification). Let M be a compact manifold without boundary. Then $\mathcal{F} = \{M\}$ is a tame stratification of M. Moreover, let $\Omega \subset M$ be a compact subset with smooth boundary $\partial\Omega$. Then $\mathcal{F} = \{\mathring{\Omega}, \partial\Omega\}$ is a tame stratification of Ω.

In the case that $M = \mathbb{R}^d$, by gluing boxes and other polytopes together one obtains sets equipped with a natural stratification that will, in most case, be tame. Our definition of 'affine stratified set', introduced in Section 1, covers all such examples:

Example 2.4 (Affine stratified sets)**.** The affine stratified sets introduced in Section 1 are stratified sets of $M = \mathbb{R}^d$.

One can also consider individual 'polytopes', such as the boxes in Corollary 1.1, to be stratified sets of $M = \mathbb{R}^d$:

Example 2.5 (Polytopes)**.** A polytope in $\mathbb{R}^d$ is naturally equipped with a stratification whose strata are the faces of the polytope of all dimensions. Though to our knowledge there is no consensus on the definition of a polytope in $\mathbb{R}^d$, it is easy to check whether or not a specific example satisfies Definition 2.1.

We also present one non-example, in the form of the 'rapid spiral':

Example 2.6 (Rapid spiral)**.** The rapid spiral $B = \{r = e^{-\theta^2}\}$ (see Figure 2) admits a natural partition that satisfies all the conditions of a tame stratification except the last; in particular, this partition is a Whitney stratification. The rapid spiral B exhibits certain pathologies that result from the lack of tameness, for instance, there are no stratified Morse functions on B (see [22, Part I, Example 2.2.2]).

We next extend the definition of topological events given in Section 1 to the general setting of stratified sets. Let f be a continuous Gaussian field on M, defined on a probability space Ω. Let $\mu : M \to \mathbb{R}$ and $K : M \times M \to \mathbb{R}$ denote respectively the mean and covariance kernel of f. Assume that f satisfies the following condition (generalising the conditions in Section 1):

Condition 2.7. The field f is a.s. C^2. Moreover, for each distinct $x, y \in M$, the Gaussian vector
$$(f(x), d_x f, f(y), d_y f) \in \mathbb{R} \times T_x^* M \times \mathbb{R} \times T_y^* M$$
is non-degenerate.

This condition ensures that μ is C^2 and that K is of class $C^{2,2}$. Let us now define the class of *topological events* on a stratified set B.

Definition 2.8 (Topological events)**.** Let $(B, \mathcal{F})$ be a stratified set of M. A *stratified homeomorphism* of B is a homeomorphism $h : B \to B$ such that for each $F \in \mathcal{F}$, $h(F) = F$. A *stratified isotopy* of B is a continuous map $H : B \times [0,1] \to B$ such that for each $t \in [0,1]$, $H(\cdot, t) : B \to B$ is a stratified homeomorphism of B. We say that two stratified homeomorphisms $h_0, h_1 : B \to B$ are $\mathcal{F}$-*isotopic* if there exists a stratified isotopy H such that $H(\cdot, 0) = h_0$ and $H(\cdot, 1) = h_1$.

Let $\mathcal{D}$ denote the excursion set $\{f > 0\}$. The *stratified isotopy class* of $\mathcal{D}$ in B, denoted $[\mathcal{D}]_B$, is the set of $h(\mathcal{D} \cap B)$ where h ranges over all stratified homeomorphisms of B that are $\mathcal{F}$-isotopic to the identity. As we establish in Corollary 5.8, under Condition 2.7 there are a countable number of stratified isotopy classes, and we equip the set of classes with its maximal σ-algebra. We will also verify in Corollary 5.8 that the map $[\mathcal{D}]_B$ from the probability space Ω into the set of stratified isotopy classes is measurable. A *topological event on B* is an event $A \subset \Omega$ measurable with respect to the random variable $[\mathcal{D}]_B$.

Henceforth we fix two stratified sets $(B_1, \mathcal{F}_1)$ and $(B_2, \mathcal{F}_2)$ of M (not necessarily disjoint). Our main formula expresses the covariance between topological events on B_1 and B_2 in terms of an integral over the 'pivotal measure' of the events. This measure is defined in terms of (i) 'pivotal points', and (ii) a certain interpolation between f and an independent copy of itself; we introduce these concepts now. Our definition of 'pivotal points' is related to the notion of 'pivotal sites' in percolation theory (see [23, Section 2.4]), whereas the interpolation is based on the classical interpolation argument of Piterbarg [39].

Definition 2.9 (Pivotal points). Fix $\hat{A} \subset C^1(M)$. For every $u \in C^1(M)$, we say that $x \in M$ is *pivotal for u (with respect to $\hat{A}$)* if, for any open neighbourhood W of x in M, there exists a function $h \in C^2_c(W)$ such that for every sufficiently small $\delta > 0$, $u + \delta h \in \hat{A}$ and $u - \delta h \notin \hat{A}$. Such a function u is described as having a *pivotal point at $x \in M$*, and we denote by $\mathrm{Piv}_x(\hat{A}) \subset C^1(M)$ the set of all such u's. If h can be chosen so that $h \geq 0$, we say that x is *positively pivotal for u*, and we denote by $\mathrm{Piv}^+_x(\hat{A}) \subset C^1(M)$ the set of such u's. Similarly, x is *negatively pivotal for u* if h can be chosen so that $h \leq 0$, and we denote $\mathrm{Piv}^-_x(\hat{A}) \subset C^1(M)$ the set of such u's.

Definition 2.10 (Interpolation). Let $\tilde{f}$ be an independent copy of f. For each $t \in [0,1]$, define the Gaussian field on $M \times M$

$$(2.1) \qquad f_t(x) = (f^1_t(x), f^2_t(x)) := (f(x), t(f(x) - \mu(x)) + \sqrt{1-t^2}(\tilde{f}(x) - \mu(x)) + \mu(x)).$$

Observe that f^1_t and f^2_t have the same law as f, and $\mathrm{Cov}(f^1_t(x_1), f^2_t(x_2)) = tK(x_1, x_2)$; in particular, f^1_0 and f^2_0 are independent, while $f^1_1 = f^2_1$. Also, observe that f^1_t and f^2_t both satisfy Condition 2.7. For each $x_1 \in F_1 \in \mathcal{F}_1$ and $x_2 \in F_2 \in \mathcal{F}_2$, denote by $\gamma_{t;x_1,x_2}(0)$ the density at zero of the Gaussian vector

$$(2.2) \qquad (f^1_t(x_1), d_{x_1}f^1_t|_{F_1}, f^2_t(x_2), d_{x_2}f^2_t|_{F_2})$$

in orthonormal coordinates of $\mathbb{R} \times T^*_{x_1}F_1 \times \mathbb{R} \times T^*_{x_2}F_2$, and denote by $\mathbb{E}_{t;x_1,x_2}[\cdot]$ expectation conditional on the vector (2.2) vanishing; this conditional expectation is well defined and described by the usual Gaussian regression formula ([3, Proposition 1.2]) since the vector (2.2) is non-degenerate. Note that, since x_1 and x_2 correspond to unique strata F_1 and F_2, to ease notation we have dropped the explicit dependence of $\gamma_{t;x_1,x_2}(0)$ and $\mathbb{E}_{t;x_1,x_2}[\cdot]$ on F_1 and F_2.

We are now ready to define the pivotal measure, or more precisely, two 'signed' pivotal measures. Fix topological events A_1 and A_2 on B_1 and B_2 respectively. Denote by $\widetilde{A}_1, \widetilde{A}_2$ the measurable sets of stratified isotopy classes in B_1 and B_2 respectively that define these topological events, and let $\hat{A}_1$ (resp. $\hat{A}_2$) be the set of functions $u \in C^1(M)$ such that $[\{u > 0\}]_{B_1} \in \widetilde{A}_1$ (resp. $[\{u > 0\}]_{B_2} \in \widetilde{A}_2$).

Denote by dv_g the Riemannian volume measure on M. Similarly, for each stratum $F \in \mathcal{F}_1 \cup \mathcal{F}_2$, denote by dv_F the Riemannian volume measure induced by g_F, the restriction of g to F. If $u \in C^2(M)$ and x is a critical point of u, we denote by $H_x u$ the Hessian of u at x (which is well defined since x is a critical point of u; see for instance [38, Chapter 1]). More generally, if $F \subset M$ is a smooth submanifold of M and $d_x u|_F = 0$, then let $H^F_x u$ be the Hessian of $u|_F$ at x.

Definition 2.11 (Pivotal measures). For each $t \in [0,1]$ and $\sigma \in \{-,+\}$, define the *signed pivotal intensity function $I^\sigma_t(x_1, x_2)$ on $B_1 \times B_2$* to be

$$(2.3) \qquad \sum_{\substack{\sigma_1, \sigma_2 \in \{-,+\}, \\ \sigma_1 \sigma_2 = \sigma}} \gamma_{t;x_1,x_2}(0)\, \mathbb{E}_{t;x_1,x_2}\left[|\det(H^{F_1}_{x_1} f^1_t) \det(H^{F_2}_{x_2} f^2_t)|;\ f^1_t \in \mathrm{Piv}^{\sigma_1}_{x_1}(\hat{A}_1), f^2_t \in \mathrm{Piv}^{\sigma_2}_{x_2}(\hat{A}_2) \right]$$

where F_1 and F_2 denote the (unique) strata in $\mathcal{F}_1$ and $\mathcal{F}_2$ that contain x_1 and x_2 respectively, and the determinants are taken with respect to orthonormal bases of $T_{x_i}F_i$. The *signed pivotal measures $\mathrm{d}\pi^\sigma(x_1, x_2)$ on $B_1 \times B_2$* are defined, for $\sigma \in \{-,+\}$, as

$$\mathrm{d}\pi^\sigma(x_1, x_2) = \left(\int_0^1 I^\sigma_t(x_1, x_2)\, \mathrm{d}t \right) \mathrm{dv}_{F_1}(x_1) \mathrm{dv}_{F_2}(x_2).$$

We emphasise that, although the 'pivotal measures' depend on both (i) the stratified sets B_i, and (ii) the topological events A_i, to ease notation we have left these dependencies implicit. Observe also that $\mathrm{d}\pi^\sigma$ is a sum of measures of different dimensions that are supported on pairs of strata $(F_1, F_2) \in \mathcal{F}_1 \times \mathcal{F}_2$. On each such pair, the measures $\mathrm{d}\pi^\pm$ are singular with respect to each other and mutually continuous with respect to the product of Riemannian volume measures.

Remark 2.12. By definition, the Hessian on 0-dimensional strata is always equal to zero. This implies that, when at least one of x_i belongs to a 0-dimensional stratum, the corresponding term is zero independently of how we interpret $\mathrm{d}v_F$ for 0-dimensional F. Hence in (2.3), as well as in all subsequent formulae of similar type, we can discard the contribution from 0-dimensional strata.

Remark 2.13. If A_1 and A_2 are both increasing events (meaning that, for $i \in \{1, 2\}$, if $u \in \hat{A}_i$ and h is a non-negative function, then $u + h \in \hat{A}_i$), then the negative pivotal measure $\mathrm{d}\pi^-$ is identically zero since $\mathrm{Piv}_{x_i}^-(\hat{A}_i)$ is empty by definition. The same is true if A_1 and A_2 are both decreasing events, since then $\mathrm{Piv}_{x_i}^+(\hat{A}_i)$ is empty. Similarly, if A_1 is increasing and A_2 is decreasing, then $\mathrm{d}\pi^+$ is identically zero.

We are now ready to present our covariance formula in full generality:

Theorem 2.14 (Covariance formula for topological events). *Let $(B_1, \mathcal{F}_1)$ and $(B_2, \mathcal{F}_2)$ be stratified sets of M. Let f be a Gaussian field on M satisfying Condition 2.7. Then the covariance of topological events A_1 and A_2 on B_1 and B_2 respectively can be expressed as*

$$\mathbb{P}\left[A_1 \cap A_2\right] - \mathbb{P}\left[A_1\right]\mathbb{P}\left[A_2\right] = \int_{B_1 \times B_2} K(x, y)\left(\mathrm{d}\pi^+(x, y) - \mathrm{d}\pi^-(x, y)\right),$$

where $\mathrm{d}\pi^+$ and $\mathrm{d}\pi^-$ denote the pivotal measures introduced in Definition 2.11.

Let us offer some intuition behind the covariance formula in Theorem 2.14. The starting point of our analysis is the observation that

$$\mathbb{P}\left[A_1 \cap A_2\right] = \mathbb{P}[f_1 \in \hat{A}_1 \times \hat{A}_2] \quad \text{and} \quad \mathbb{P}\left[A_1\right]\mathbb{P}\left[A_2\right] = \mathbb{P}[f_0 \in \hat{A}_1 \times \hat{A}_2],$$

and hence

$$\mathbb{P}\left[A_1 \cap A_2\right] - \mathbb{P}\left[A_1\right]\mathbb{P}\left[A_2\right] = \int_0^1 \frac{d}{dt}\mathbb{P}[f_t \in \hat{A}_1 \times \hat{A}_2]\,\mathrm{d}t.$$

As we explain in Section 2.3, the structure of the Gaussian measure allows us to express

$$\frac{d}{dt}\mathbb{P}[f_t \in \hat{A}_1 \times \hat{A}_2]$$

as an integral, over pairs of strata $(F_1, F_2) \in \mathcal{F}_1 \times \mathcal{F}_2$, of the (signed) two-point intensity functions I_t^σ of critical points that are 'pivotal' for the events A_1 and A_2 respectively, weighted by a term that is the inner product of the outward normal vectors at the boundary of the events A_1 and A_2; by the properties of the Gaussian measure (in particular, the reproducing property of the covariance kernel), this inner product is just (a normalisation of) the covariance kernel K.

To understand the form of the intensity functions I_t^σ, notice that pivotal points are necessarily critical points at the zero level. Hence we can understand I_t^σ as a restriction to pivotal points of the standard two-point intensity function for critical points of f_t on (F_1, F_2) at the zero level, which by the well-known Kac-Rice formula (see [3, Chapter 6]) is given by

$$\gamma_{t;x_1,x_2}(0)\,\mathbb{E}_{t;x_1,x_2}\left[|\det(H_{x_1}^{F_1}f_t^1)\det(H_{x_2}^{F_2}f_t^2)|\right].$$

Note that our intensity functions are signed; this is because we must distinguish pairs of pivotal points that are pivotal 'in the same direction', in the sense that a local increase in f causes the events A_1 and A_2 to *both* occur or to *both not* occur, from those that are pivotal 'in opposite directions'.

It is possible that some variant of Theorem 2.14 remains true for a wider class of smooth random fields. The Kac-Rice formula applies far beyond the Gaussian setting, and in principle one can also express the intensity of pivotal points for non-Gaussian fields. As for the initial interpolation step, by formulating it using the Ornstein-Uhlenbeck semigroup (as in, say, [**?**] or as suggested in [**?**]) the setting could perhaps be extended to measures related to other Markov semigroups. We leave this for future investigation.

2.2. **Applications.** We next present applications of the covariance formula in Theorem 2.14; some of these have already been discussed (see Corollaries 1.2 and 1.6), but here we give extensions to more general settings. The proofs will be deferred to Section 6. Throughout this section we assume that f satisfies Condition 2.7.

2.2.1. *Strong mixing for topological events.* Our first application generalises the strong mixing statement in Corollary 1.2 to the set-up in Section 2.1. For a stratified set $B \subset M$, let $\sigma_{\text{top}}(B)$ denote the σ-algebra consisting of topological events in B, and for a pair of stratified sets $B_1, B_2 \subset M$, define the corresponding 'topological' α-mixing coefficient

$$(2.4) \qquad \alpha_{\text{top}}(B_1, B_2) = \sup_{A_1 \in \sigma_{\text{top}}(B_1),\, A_2 \in \sigma_{\text{top}}(B_2)} |\mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2]|.$$

Theorem 2.15 (Strong mixing for topological events)**.** *There exists a constant $c_d > 0$, depending only on the dimension of the manifold M, such that for every pair of stratified sets $(B_1, \mathcal{F}_1)$ and $(B_2, \mathcal{F}_2)$ of M,*

$$\alpha_{top}(B_1, B_2) \leq c_d \sum_{F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2} c_{F_1, F_2} \int_{F_1 \times F_2} |K(x_1, x_2)| \, dv_{F_1}(x_1) \, dv_{F_2}(x_2),$$

where c_{F_1, F_2} is equal to the maximum, over $i, j, k \in \{1, 2\}$, of

$$\sup_{x_1 \in F_1, x_2 \in F_2} \frac{\left(\mathbb{E}\left[\|H_{x_i}^{F_i} f\|_{op}^2 \,|\, d_{x_i} f|_{F_i} = 0\right]\right)^{d_i}}{\sqrt{\det(\Delta(x_1, x_2))}} \max\left\{1, \left(\frac{K(x_j, x_j) \det(d_{x_k} \otimes d_{x_k} K|_{F_i \times F_i})}{\sqrt{\det(\Delta(x_1, x_2))}}\right)^{2d_i}\right\},$$

and where $\|\cdot\|_{op}$ denotes the (L^2-)operator norm, $d_i = \dim(F_i)$, and $\Delta(x_1, x_2)$ is the covariance matrix, in orthonormal coordinates, of the (non-degenerate) Gaussian vector

$$(f(x_1), d_{x_1} f|_{F_1}, f(x_2), d_{x_2} f|_{F_2}).$$

Remark 2.16. All the terms in the definition of c_{F_1, F_2} can be written as a quotient of powers of polynomials of partial derivatives of K of order at most $(2, 2)$. This means that (i) c_{F_1, F_2} depends continuously on the $C^{2,2}$ norm of K, and (ii) c_{F_1, F_2} is homogeneous in K (the degree of homogeneity is easily seen to be -1, which compensate the presence of $K(x_1, x_2)$ in the integral).

2.2.2. *Sequences of fields: The Kostlan ensemble.* In Corollary 1.2 we stated a quantitative mixing bound for rescaled (affine) stratified sets sB_1 and sB_2 as $s \to \infty$. In the setting of compact manifolds M, it is often more appropriate to work with a *sequence* of Gaussian fields on M that converge to a local limit, and consider the topological mixing between *fixed* disjoint stratified sets $B_1, B_2 \subset M$ (in fact, this includes the setting in Corollary 1.2 as a special case, by rescaling the field rather than the sets).

Rather than work in full generality, here we work only with the *Kostlan ensemble*, which is the sequence $(f_n)_{n \in \mathbb{N}}$ of smooth centred isotropic Gaussian fields on $\mathbb{S}^d$ with covariance kernels

$$K(x, y) = \cos^n(d_{\mathbb{S}^d}(x, y)) = \langle x, y \rangle^n,$$

where $d_{\mathbb{S}^d}(\cdot, \cdot)$ denotes the spherical distance; it is easy to check that each f_n satisfies Condition 2.7. The sequence f_n converges to a local limit on the scale $s_n = 1/\sqrt{n}$, in the sense that for any $x_0 \in \mathbb{S}^d$ the rescaled field

$$(2.5) \qquad f(\exp_{x_0}(x/\sqrt{n})) \,, \quad x \in \mathbb{R}^d$$

converges on compact sets to the smooth stationary Gaussian field on $\mathbb{R}^d$ with covariance $\kappa(x) = e^{-\|x-y\|^2/2}$; here $\exp_{x_0} : \mathbb{R}^d \to \mathbb{S}^d$ denotes the exponential map based at x_0. The Kostlan ensemble is a natural model for random homogeneous polynomials (see [28, 29]), and its level sets have been the focus of recent study [9]. Its local limit is known as the *Bargmann-Fock* field.

Corollary 2.17 (Strong mixing for the Kostlan ensemble)**.** *For each pair of disjoint stratified sets $B_1, B_2 \subset \mathbb{S}^d$ that are contained in an open hemisphere, there exist $c_1, c_2 > 0$ such that, for each $n \geq 1$,*

$$\alpha_{n;top}(B_1, B_2) \leq c_1 e^{-c_2 n},$$

where $\alpha_{n;top}$ denotes the 'topological' mixing coefficient (2.4) for the field f_n.

Remark 2.18. Since f_n are homogeneous polynomials, they are naturally defined on the real projective space rather than the sphere, which makes it natural to restrict B_1 and B_2 to be contained in an open hemisphere. Indeed, f_n is degenerate at antipodal points.

The lower concentration result in Corollary 1.6 can also be generalised to the setting of sequences of Gaussian fields on manifolds; again we focus just on the Kostlan ensemble $(f_n)_{n \in \mathbb{N}}$ on $\mathbb{S}^d$. We define a *topological count* $N = N_n(B)$ for f_n analogously to in Section 1, after substituting affine stratified sets $B \subset \mathbb{R}^d$ with general stratified sets $B \subset \mathbb{S}^d$; these counts are now indexed by $B \subset \mathbb{S}^d$ and $n \in \mathbb{N}$. A topological count N is called *super-additive* if (1.6) holds for each N_n. We say that a topological count N satisfies a *law of large numbers* if there exists a $c_N > 0$ such that, for every stratified set $B \subset \mathbb{S}^d$, as $n \to \infty$,

$$(2.6) \qquad \frac{N_n(B)}{n^{d/2} \operatorname{Vol}(B)} \to c_N \quad \text{in probability;}$$

the scale $n^{d/2}$ can be understood as the natural volume scaling induced by the rate $s_n = 1/\sqrt{n}$ at which the Kostlan ensemble converges to a local limit in (2.5).

Corollary 2.19 (Lower concentration for topological counts of the Kostlan ensemble)**.** *Let N denote a super-additive topological count that satisfies a law of large numbers (2.6) with limiting constant $c_N > 0$. Then for every stratified set $B \subset \mathbb{S}^d$ and every $\varepsilon > 0$, there exist $c_1, c_2 > 0$ such that, for every $n \geq 1$,*

$$(2.7) \qquad \mathbb{P}\left[\frac{N_n(B)}{n^{d/2} \operatorname{Vol}(B)} \leq c_N - \varepsilon\right] \leq c_1 e^{-c_2 n^{d/(d+2)}}.$$

In particular, taking $B = \mathbb{S}^d$ with its trivial stratification $\mathcal{F} = \{\mathbb{S}^d\}$, the conclusion of Corollary 2.19 is true for N_n the number of connected components of $\{f_n > 0\}$ or $\{f_n = 0\}$ on the sphere $\mathbb{S}^d$ (see [37] for a proof of the law of large numbers for N_n).

2.2.3. *Decorrelation for topological counts.* In the classical theory of strong mixing, a major application of mixing bounds is to prove central limit theorems (CLTs) (see, e.g., [16, 31, 45]). Although establishing CLTs for topological counts is beyond the scope of this work, we illustrate here how mixing bounds can be used to deduce the 'decorrelation' of topological counts, a key intermediate step in proving a CLT.

For simplicity we return to the Euclidean setting of Section 1. We say that a topological count N has a *finite two-plus-delta moment* on an affine stratified set $B \subset \mathbb{R}^d$ if there exist $\delta, c > 0$ such that

$$(2.8) \qquad \mathbb{E}[N(B)^{2+\delta}] < c < \infty.$$

Although the finiteness of two-plus-delta moments is not known for the topological counts discussed in Section 1 (except in the one-dimensional case), in principle one can bound (2.8) by the purely local quantity

$$\mathbb{E}[(\# \text{ of critical points of } f \text{ in } B)^{2+\delta}],$$

which we suspect is finite in great generality.

Corollary 2.20 (Decorrelation for topological counts). *Fix affine stratified sets $B_1, B_2 \subset \mathbb{R}^d$ and suppose that N_1 and N_2 are topological counts that have finite two-plus-delta moments (2.8) on B_1 and B_2 with constants $\delta, c > 0$. Then*

$$(2.9) \qquad \mathrm{Cov}\big(N_1(B_1), N_2(B_2)\big) \leq 8 \, c^{2/(2+\delta)} \, \alpha_{top}(B_1, B_2)^{\delta/(2+\delta)}.$$

We expect that standard methods (i.e. [18, 45]) should allow one to deduce, from Corollary 2.20, a CLT for rescaled topological counts that satisfy a law of large numbers whenever strong enough two-plus-delta moment bounds can be established, at least as long as $\kappa(x)$ decays at a high enough polynomial rate (with the polynomial exponent depending on δ).

2.2.4. *Positive association for increasing topological events.* Recall that a random vector is said to be 'positively associated' if increasing events (or equivalently decreasing events) are positively correlated. To state an analogous property for continuous random fields some care must be taken to specify an appropriate class of increasing events, and here we restrict the discussion to topological events. An important example of topological events that are increasing are crossing events for the excursion set $\{f > 0\}$ (but *not* crossing events for the level set $\{f = 0\}$), and the fact that crossing events are positively correlated is crucial in the analysis of level set percolation [5, 9, 35, 43].

In the setting of Gaussian fields, it is known that the class of increasing topological events on a stratified set are positively correlated if and only if the covariance kernel K is positive. The standard approach is to invoke a classical result that (finite-dimensional) Gaussian vectors are positively associated if and only if they are positively correlated [40], and then to apply an approximation argument (see [44]). Here we deduce, directly from our exact formula, a quantitative version of this result, whose proof is immediate from Theorem 2.14 and the observation in Remark 2.13.

Corollary 2.21 (Positive associations). *Let A_1 and A_2 be topological events on stratified sets B_1 and B_2, and suppose that A_1 and A_2 are both increasing. Then*

$$(2.10) \qquad \mathbb{P}\,[A_1 \cap A_2] - \mathbb{P}\,[A_1]\,\mathbb{P}\,[A_2] = \int_{B_1 \times B_2} K(x, y)\, \mathrm{d}\pi^+(x, y),$$

where $\mathrm{d}\pi^+$ is the measure defined in Definition 2.11. In particular, A_1 and A_2 are positively correlated if $K|_{B_1 \times B_2} \geq 0$.

The fact that positive associations fails in general if a Gaussian field is not positively correlated is a serious limitation to many applications; for example, the current theory of level set percolation for Gaussian fields fails more or less completely unless $K \geq 0$ (see however [6] for recent progress in this direction). One advantage of (2.10) is that the failure of positive associations can be *quantified*, which gives hope that the errors that arise might be controllable.

2.2.5. *The Harris criterion.* Lastly, we present an informal discussion of the 'Harris criterion' (HC), demonstrating in particular that Theorem 2.14 can be used to give an alternative derivation of this criterion.

In its original formulation (see, e.g., [47]), the HC was a heuristic to determine whether long-range correlations influence the large-scale connectivity of discrete critical percolation models. Translated to the setting of Gaussian fields on $\mathbb{R}^d$ (see [11]), the HC claims that the connectivity of the level set of smooth centred Gaussian fields will, at the *critical level* $\ell_c \leq 0$ (known to be zero if $d = 2$, but believed to be strictly negative if $d \geq 3$), be well-described on large scales by critical (Bernoulli) percolation (the 'percolation hypothesis') if and only if

$$(2.11) \qquad s^{2/\nu - 2d} \int_{B_s \times B_s} \kappa(x - y)\, dx dy \to 0 \quad \text{as } s \to \infty,$$

where B_s denotes the ball of radius s centred at the origin, and ν is the *correlation length exponent* of critical percolation, widely believed to be universal and satisfy

$$\nu = \begin{cases} 4/3, & d = 2, \\ \in (1/2, 1), & d = 3, 4, 5, \\ 1/2, & d \geq 6. \end{cases}$$

In the positively-correlated case $\kappa \geq 0$, (2.11) is roughly equivalent to demanding that κ has polynomial decay with exponent at least $2/\nu$. The original argument of Harris (as translated to our setting in [11]) goes as follows. Define

$$m_s = \frac{1}{|B_s|} \int_{B_s} f(x)\, dx$$

to be the average value of f on the ball B_s. The fluctuations of m_s are of order

$$(2.12) \qquad \sqrt{\mathbb{E}[m_s^2]} = \frac{1}{|B_s|} \left(\int_{B_s \times B_s} \kappa(x - y)\, dxdy \right)^{1/2}.$$

Recall now that the behaviour of critical (and near-critical) percolation follows a set of power-laws with certain *universal exponents*, one of which is the *correlation length exponent* ν. Roughly speaking, this claims that the connectivity of percolation with probability $p \in [0, 1]$ closely approximates the connectivity of critical percolation on the ball B_s as long as $|p - p_c| \ll s^{-1/\nu}$, where p_c is the critical probability. Under the assumption that f can be replaced by $m_s + f$ on B_s, the 'percolation hypothesis' therefore generates a contradiction unless $m_s \ll s^{-1/\nu}$, and combining with (2.12) gives (2.11). Note that the HC should really be understood as a *necessary condition* for the 'percolation hypothesis', since the argument assumes the 'percolation hypothesis' and derives a contradiction.

We now demonstrate that Theorem 2.14 yields an alternative criterion, more or less equivalent to (2.11), that we claim is also a necessary condition for the 'percolation hypothesis'. Fix a pair of disjoint boxes $B_1, B_2 \subset \mathbb{R}^d$ and, for each $s \geq 1$ and $i \in \{1, 2\}$, let A_i^s denote the crossing events for the critical level set in sB_i. Note that pivotal points for crossing events roughly correspond to *four-arm saddles at distance s*, i.e. saddle points x such that all four arms of the level set $\{f = f(x)\}$ hit the ball of radius s around x. Putting this approximation into Theorem 2.14, we deduce that

$$\mathbb{P}\left[A_1^s \cap A_2^s\right] - \mathbb{P}\left[A_1^s\right] \mathbb{P}\left[A_2^s\right] \approx c_\kappa \int_{sB_1 \times sB_2} \kappa(x - y) I_s(x, y)\, dxdy$$

$$\approx c_\kappa\, I_s(0)^2 \int_{sB_1 \times sB_2} \kappa(x - y)\, dxdy.$$

where I_s denotes the intensity of four-arm saddles at distance s, and where in the last step we used stationarity and an (unjustified) factorisation of this intensity. Consider now the universal exponent ζ_4 that is believed to describe the decay of the probability of critical 'four-arm' events for all percolation models. If the 'percolation hypothesis' is true, then $I_s(0) \approx s^{-\zeta_4}$, and since under the 'percolation hypothesis' the events A_1^s and A_2^s decorrelate, we end up with the following criterion for this hypothesis:

$$(2.13) \qquad s^{-2\zeta_4} \int_{sB_1 \times sB_2} \kappa(x - y)\, dxdy \to 0 \quad \text{as } s \to \infty.$$

To compare to (2.11), recall that by the 'Kesten scaling relations' [26] $\zeta_4 = d - 1/\nu$, and so the exponents $2/\nu - 2d$ and $-2\zeta_4$ in (2.11) and (2.13) match. The only difference is the domain of integration, but as $s \to \infty$ this difference is negligible under mild assumptions on the decay of covariance.

2.3. **Proof sketch.** Theorem 2.14 can be considered as a generalisation to topological events of a simple formula, essentially due to Piterbarg [39], that gives a covariance formula for finite-dimensional Gaussian vectors. This lemma is both the inspiration for Theorem 2.14, and also one of the key ingredients in the proof. We state Piterbarg's formula in the simplest case of standard Gaussian vectors, since this is all that we need, but a similar statement exists for general non-degenerate Gaussian vectors; for completeness, we give the proof in Appendix B.

Lemma 2.22 (Piterbarg's formula; see [39, Theorem 1.4]). *For each $t \in [0,1]$, let X_t and Y_t be jointly Gaussian vectors in $\mathbb{R}^m$, not necessarily centred, whose covariance matrix is*

$$\begin{pmatrix} I & tI \\ tI & I \end{pmatrix};$$

that is, $\mathrm{Cov}(X_{t,i}, X_{t,j}) = \mathrm{Cov}(Y_{t,i}, Y_{t,j}) = \delta_{i,j}$ and $\mathrm{Cov}(X_{t,i}, Y_{t,j}) = t\delta_{i,j}$. Let $\gamma_t(x,y)$ denote the density of $Z_t = (X_t, Y_t) \in \mathbb{R}^{2m}$. Let A and B be domains in $\mathbb{R}^m$ whose boundaries are piecewise smooth, and which have surface areas, inside the ball of radius R, that grow at most polynomially in R. Denote by ν_A and ν_B the outward unit normal vectors on the boundaries of A and B respectively. Then $\mathbb{P}[Z_t \in A \times B]$ is differentiable in $t \in (0,1)$, and

$$\frac{d}{dt}\mathbb{P}[Z_t \in A \times B] = \int_{\partial A \times \partial B} \langle \nu_A(x), \nu_B(y) \rangle \gamma_t(x,y) \, \mathrm{d}x \, \mathrm{d}y,$$

where by $\int \mathrm{d}x \, \mathrm{d}y$ we understand integration with respect to the natural $m-1$ dimensional measures on ∂A and ∂B respectively.

In particular, if X denotes an arbitrary translation of a standard Gaussian vector in $\mathbb{R}^m$, then

$$\mathbb{P}[X \in A \cap B] - \mathbb{P}[X \in A]\mathbb{P}[X \in B] = \int_0^1 \int_{\partial A \times \partial B} \langle \nu_A(x), \nu_B(y) \rangle \gamma_t(x,y) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t;$$

the integral converges since the integral $\int_0^s \mathrm{d}t$ on the right-hand side exists for all $s < 1$, and converges as $s \to 1$ to the left-hand side.

Let us now give a brief sketch of the proof of Theorem 2.14, showing how Piterbarg's formula plays an essential role. We begin by considering the case of finite-dimensional Gaussian fields, i.e. the case in which f is a Gaussian vector in a finite-dimensional space of continuous functions V (see Proposition 3.9). More precisely, we fix $\langle \cdot, \cdot \rangle$ a scalar product on V and take f to be a translation of the standard Gaussian vector in V. The scalar product also induces a volume measure du on V, and allows us to identify V with $\mathbb{R}^{\dim(V)}$ up to isometries. Hence, we can apply Piterbarg's formula in V and deduce that

$$(2.14) \qquad \frac{d}{dt}\mathbb{P}\left[f_t \in \hat{A}_1 \times \hat{A}_2 \right] = \int_{\partial \hat{A}_1 \times \partial \hat{A}_2} \langle \nu_{A_1}(u_1), \nu_{A_2}(u_2) \rangle \gamma_t(u_1, u_2) \, \mathrm{d}u_1 \mathrm{d}u_2 \,,$$

where $f_t = (f_t^1, f_t^2)$ has covariance $\begin{pmatrix} I & tI \\ tI & I \end{pmatrix}$ in orthonormal coordinates of $V \times V$ equipped with the product scalar product (this coincides with the definition of f_t in (2.1)).

The next step is to analyse the boundaries of $\hat{A}_i$. The path $(f_t^i)_{t \in [0,1]}$ is a generic deformation of f. By standard arguments in Morse theory, along this deformation the topology of the set $\{f_t^i \geq 0\}$ changes only when f_t^i passes through a non-degenerate critical point at level 0 (which can cause f_t^i to either enter or exit $\hat{A}_i$); if such a change in topology occurs we say that this critical point is *pivotal* for the event $\hat{A}_i$ and the function f_t^i. We will see (in Lemma 3.10) that, if we exclude a subset $E \subset \partial \hat{A}_i$ of positive codimension containing the functions with multiple stratified critical points at level 0, we can define a surjection

$$\Xi : \partial \hat{A}_i \setminus E \twoheadrightarrow B_i,$$

that induces submersions on each stratum of B_i, by associating to each $u_i \in \partial \hat{A}_i \setminus E$ its unique critical point at level 0. The fibre $\Xi^{-1}(x_i)$ is an open subset of the subspace of functions for which x_i is a stratified critical point at level 0. We will see that it is equal to $\mathrm{Piv}_{x_i}(\hat{A}_i)$ up to a negligible set.

Using the map Ξ, the coarea formula allows us to switch from an integral over $\partial \hat{A}_1 \times \partial \hat{A}_2 \subset V$ to a sum of integrals over pairs of faces of B_1 and B_2. We obtain that (2.14) is equal to

$$
\sum_{F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2} \int_{F_1 \times F_2} \left(\int_{\Xi^{-1}(x_1) \times \Xi^{-1}(x_2)} \frac{\langle \nu_{\partial \hat{A}_1}(u_1), \nu_{\partial \hat{A}_2}(u_2) \rangle}{\mathrm{Jac}^{\perp}_{F_1}(u_1) \mathrm{Jac}^{\perp}_{F_2}(u_2)} \gamma_t(u_1, u_2)\, du_1 du_2 \right) \mathrm{dv}_{F_1}(x_1) \mathrm{dv}_{F_2}(x_2)
$$

where, in the inner integral, the measures du_i are the natural volume measures on the fibres of $\Xi^{-1}(x_i)$ and the terms $\mathrm{Jac}^{\perp}_{F_i}(u_i)$ are the normal Jacobians of Ξ at u_i.

We then turn our attention to the unit normal vectors in the integrand (see Lemma 3.12). Consider $u_i \in \partial \hat{A}_i \setminus E$ such that $\Xi(u_i) = x_i$. Since x_i is the only place at which the topology of $\{u_i \geq 0\}$ can change by infinitesimal perturbations, $T_{u_i} \partial \hat{A}_i$ is the subspace of functions $v \in V$ such that $v(x_i) = 0$. Since K is the reproducing kernel of V, $K(x_i, \cdot)$ is orthogonal to $T_{u_i} \partial \hat{A}_i$. Thus,

$$
\langle \nu_{\partial \hat{A}_1}(u_1), \nu_{\partial \hat{A}_2}(u_2) \rangle = \pm \frac{K(x_1, x_2)}{\sqrt{K(x_1, x_1) K(x_2, x_2)}},
$$

where the sign depends on whether a small positive perturbation of u_i at x_i makes u_i enter or exit $\hat{A}_i$.

Finally, in Lemmas 3.11 and 3.14 we (i) compute the Jacobian of Ξ at $u_i \in \Xi^{-1}(x_i)$ and (ii) reinterpret the integral over $\Xi(x_1)^{-1} \times \Xi^{-1}(x_2)$ as an expectation in $f_t = (f_t^1, f_t^2)$ conditioned on the fact that for $i = 1, 2$, x_i is a critical point of f_t^i at level 0, containing the indicators that the x_i are pivotal for f_t^i. This process involves some standard computations of Jacobians of evaluation maps and a careful study of the relations between the different metrics on the spaces $V \times V$ and $F_1 \times F_2$. As a result, we get exactly the term which appears in the definition of the pivotal intensity functions (see (2.3)), namely

$$
\frac{\langle \nu_{\partial \hat{A}_1}(u_1), \nu_{\partial \hat{A}_2}(u_2) \rangle}{\mathrm{Jac}^{\perp}_{F_1}(u_1) \mathrm{Jac}^{\perp}_{F_2}(u_2)} = \pm K(x_1, x_2) \prod_{i=1,2} \left| \det \left(H^{F_i}_{x_i} u_i \right) \right|,
$$

which completes the proof in the finite-dimensional case.

To extend Theorem 2.14 to the general case, it remains only to argue that f can always be approximated by finite-dimensional fields and that we can successfully pass to the limit in the covariance formula. This latter step is mainly technical, and requires us to show, among other things, that the boundary of $\Xi^{-1}(x_i) = \mathrm{Piv}_{x_i}(\hat{A}_i)$ is a null set for the field f conditioned on the existence of critical points at x_1 and x_2.

3. Heart of the proof: the finite-dimensional case

In this section we state and prove a reinterpretation of our covariance formula in the case where the space V is finite-dimensional (see Proposition 3.9). As discussed in the proof sketch above, we prove this proposition by applying Piterbarg's formula (Lemma 2.22) and then obtaining a rather explicit description of boundaries of topological events (see Lemma 3.10).

Throughout this section, and indeed for the remainder of the paper, $(B, \mathcal{F})$ denotes an arbitrary stratified set of M.

3.1. **Restating the formula in terms of the discriminant.** In this subsection we state the finite-dimensional version of the formula (Proposition 3.9). For this we introduce an alternative notion of 'pivotal sets' defined in terms of the 'discriminant'.[1]

Definition 3.1 (Critical points)**.** Let $u \in C^1(M)$. A *stratified critical point* of u in B is a point $x \in B$ such that $d_x u|_F = 0$, where $F \in \mathcal{F}$ is the (unique) stratum containing x. When there is no ambiguity we will refer to stratified critical points as critical points for brevity. The *level* of a critical point refers to its critical value.

Assume now that $u \in C^2(M)$. A stratified critical point x of u is said to be a *non-degenerate* if (i) $H_x^F u$ is non-degenerate, and (ii) for each $F' \in \mathcal{F}$ such that $F' > F$, $d_x u$ does not vanish on $T_x F'|_F$ (see Definition 2.1). Roughly speaking (ii) means that $d_x u$ vanishes on $T_x F$ but not on tangent spaces to higher dimensional strata. Note that we define non-degeneracy in terms of the generalised tangent bundle $T_x F'|_F$; this is since all strata are open and disjoint, so $T_x F'$ is not defined.

In the following definitions $V \subset C^2(M)$ denotes an arbitrary linear subspace (i.e. not necessarily finite-dimensional). To define the discriminant, it will be convenient to introduce the following subsets of V:

Notation 3.2. For each $x \in M$, $V_x \subset V$ denotes the linear subspace of $u \in V$ such that $u(x) = 0$. Moreover, V_x' denotes the linear subspace of V_x such that also $d_x u|_F = 0$, where F is the (unique) stratum containing x; in other words, V_x' contains the functions that possess a stratified critical point at $x \in B$ at level 0. Similarly, for each $F \in \mathcal{F}$, $V_F' = \cup_{x \in F} V_x'$ denotes the functions that possess a stratified critical point on F at level 0.

Definition 3.3 (Discriminant)**.** The *discriminant* associated to B in V is the set $\mathfrak{D}_B(V) = \cup_{F \in \mathcal{F}} V_F'$, that is, the set of functions that possess a stratified critical point in B at level 0. For each $u \in V \setminus \mathfrak{D}_B(V)$, the *B-discriminant class* of u (in V), written as $[u]_{(B,V)}$ is the connected component of $V \setminus \mathfrak{D}_B(V)$ containing u. By Lemma C.1 the discriminant is closed; since $C^1(M)$ is separable, the number of classes is therefore at most countable. We will denote by $\widetilde{\sigma}_{\mathrm{discr}}(B, V)$ the complete σ-algebra of all collections of B-discriminant classes.

Before defining the alternate notion of 'pivotal sets' in terms of the discriminant, we introduce further subsets of V_x' and V_F' defined above; as we verify later (see Proposition 4.1), these subsets are of full measure:

Notation 3.4. For each $x \in M$, $\widetilde{V}_x' \subset V_x'$ denotes the set of $u \in V$ such that x is a non-degenerate stratified critical point at level 0 and there are no other stratified critical points in B at this level. Similarly, for each $F \in \mathcal{F}$, $\widetilde{V}_F' = \cup_{x \in F} \widetilde{V}_x'$ denotes the subset of V_F' consisting of functions that have a non-degenerate stratified critical point on stratum F at level 0, and no other stratified critical points in B at this level.

Definition 3.5 ('Pivotal sets' in terms of the discriminant)**.** Let $\tilde{A}$ be an element of $\widetilde{\sigma}_{\mathrm{discr}}(B, V)$. We denote by $\hat{A} \subset V$ the set of functions whose B-discriminant class is in $\tilde{A}$, and by $\hat{\sigma}_{\mathrm{discr}}(B, V)$ the σ-algebra of all possible sets $\hat{A}$ of this type. For $\hat{A} \in \hat{\sigma}_{\mathrm{discr}}(B, V)$, we define the 'pivotal sets' $\widetilde{\mathrm{Piv}}_x(\hat{A}) = \partial \hat{A} \cap \widetilde{V}_x'$ and $\widetilde{\mathrm{Piv}}_F(\hat{A}) = \partial \hat{A} \cap \widetilde{V}_F'$; note that these are subsets of the discriminant $\mathfrak{D}_B(V)$. For each $\sigma \in \{+, -\}$, let $\widetilde{\mathrm{Piv}}_x^\sigma(\hat{A})$ be the set of $u \in \widetilde{\mathrm{Piv}}_x(\hat{A})$ such that there exists $h \in V$ with $h(x) > 0$ such that, for all small enough values of $\eta > 0$, $u + \sigma \eta h \in \hat{A}$.

Finally, we introduce the key conditions on the space V:

---

[1]In fact, in the cases that matter to us, this alternative notion of 'pivotal sets' coincides with that of Definition 2.9 up to null sets. See Remark 5.3.

Condition 3.6. For each distinct $x, y \in M$, let $V'_{x,y} \subset V$ denotes the set of $u \in V$ such that $(u(x), d_x u, u(y), d_y u)$ vanishes. Then the following map is surjective:

$$V'_{x,y} \to \mathrm{Sym}^2 \left(T_y^* M \right)$$
$$u \mapsto H_y u \,.$$

Condition 3.7. For each distinct $x, y \in M$, the following map is surjective:

$$V \to \mathbb{R} \times T_x^* M \times \mathbb{R} \times T_y^* M$$
$$u \mapsto (u(x), d_x u, u(y), d_y u) \,.$$

Remark 3.8. For every smooth M there exists a finite-dimensional subspace $V \subset C^\infty(M)$ satisfying Conditions 3.6 and 3.7. Indeed, given a smooth mapping $G : M \to \mathbb{R}^N$ for some $N \in \mathbb{N}$, the coordinates of G generate an N-dimensional subspace of $C^\infty(M)$ which we denote by V^G. For any distinct $x, y \in M$, the set of G such that V^G does not satisfy Conditions 3.6 and 3.7 at x and y has codimension arbitrarily large as $N \to \infty$. Therefore, by the multijet transversality theorem (see Theorem 4.13, Chapter II of [21]), applied to the multijet $(x, y) \mapsto (j^1 G(x), j^2 G(y))$, the set of G such that V^G satisfies Conditions 3.6 and 3.7 is a residual subset of $C^\infty(M, \mathbb{R}^N)$ for sufficiently large N. In particular, such spaces exist.

We are now ready to present our finite-dimensional restatement of the covariance formula:

Proposition 3.9. *Recall the notation introduced in Section 2.1. Let V be a finite-dimensional subspace of $C^2(M)$ that satisfies Conditions 3.6 and 3.7, and assume that the support of f is exactly V, so that f is a non-degenerate Gaussian vector in V. Let $(B_1, \mathcal{F}_1)$ and $(B_2, \mathcal{F}_2)$ be stratified sets of M. For each $i \in \{1, 2\}$, let $\hat{A}_i \in \hat{\sigma}_{\mathrm{discr}}(B_i, V)$ and let A_i be the event $\{f \in \hat{A}_i\}$. Then, for each $t \in [0, 1)$,*

$$\frac{d}{dt} \mathbb{P}\left[f_t \in \hat{A}_1 \times \hat{A}_2 \right] = \sum_{\sigma_1, \sigma_2 \in \{-,+\}} \sum_{F_1 \in \mathcal{F}_1, \ F_2 \in \mathcal{F}_2} \int_{F_1 \times F_2} K(x_1, x_2) \times \gamma_{t;x_1,x_2}(0)$$

$$\times \sigma_1 \sigma_2 \, \mathbb{E}_{t;x_1,x_2}\left[\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}^{\sigma_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}^{\sigma_2}(\hat{A}_2)}(f_t^1, f_t^2) \left| \det \left(H_{x_1}^{F_1} f_t^1 \right) \right| \left| \det \left(H_{x_2}^{F_2} f_t^2 \right) \right| \right] \mathrm{dv}_{F_1}(x_1) \mathrm{dv}_{F_2}(x_2).$$

3.2. **Proof of Proposition 3.9.** Throughout this section we assume that V, f and $\hat{A}_i$ are as in the statement of Proposition 3.9, in particular V is finite-dimensional and satisfies Conditions 3.6 and 3.7 (although all the notation that is introduced applies equally to arbitrary linear subspaces V of $C^2(M)$). We continue to use $(B, \mathcal{F})$ to denote an arbitrary stratified set of M, and we also define an arbitrary $\hat{A} \in \hat{\sigma}_{\mathrm{discr}}(B, V)$. We rely on four technical lemmas (namely Lemmas 3.10–3.12 and 3.14), whose proofs are deferred to Section 4.

The starting point of the proof is to apply Piterbarg's formula to the events $\hat{A}_1$ and $\hat{A}_2$; for this we need to study the regularity of their boundaries. The structure of $\partial \hat{A}_i$ is described by the following lemma:

Lemma 3.10. *For each $F \in \mathcal{F}$, the set $\widetilde{\mathrm{Piv}}_F(\hat{A})$ (from Definition 3.5) is a smooth (immersed) conical hypersurface of V. If x is the unique level-0 stratified critical point of some $u \in \widetilde{\mathrm{Piv}}_F(\hat{A})$, then $T_u \widetilde{\mathrm{Piv}}_F(\hat{A}) = V_x$ (see Notation 3.4). Moreover, there exists a subset $E \subset \partial \hat{A}$ of zero $N - 1$ dimensional Hausdorff measure such that*

$$\partial \hat{A} = E \cup \bigsqcup_{F \in \mathcal{F}} \widetilde{\mathrm{Piv}}_F(\hat{A}).$$

By Lemma 3.10, the boundaries of the sets $\hat{A}_1$ and $\hat{A}_2$ are smooth up to null sets, which implies that their $N - 1$ dimensional volume inside any finite ball is finite. Since they are conical, the volume of the boundary inside a ball of radius R is of order R^{N-1}, ensuring that Piterbarg's formula applies to these sets.
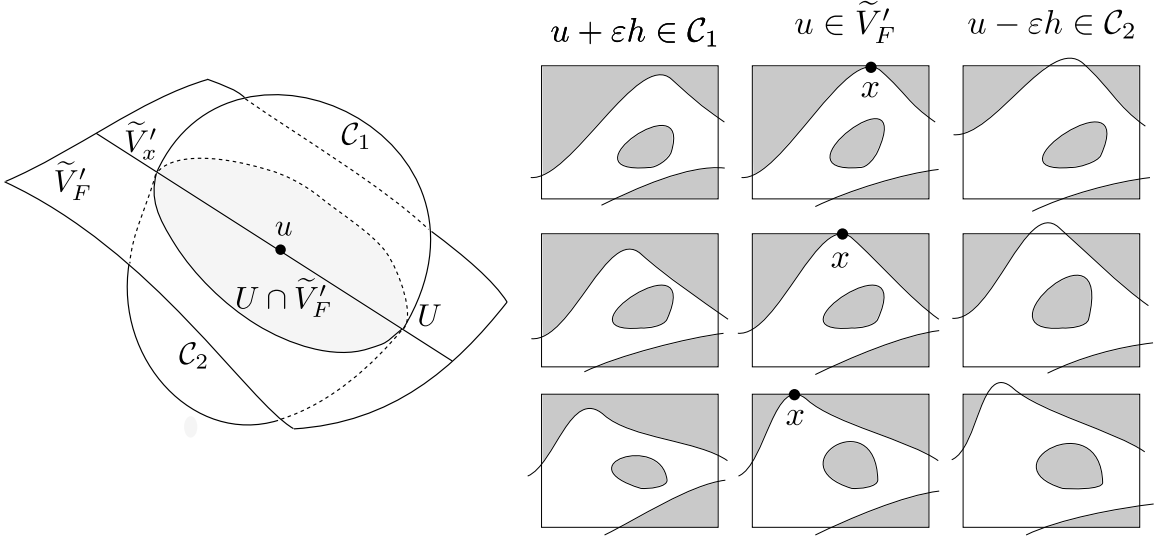
FIGURE 3. Outside of a null set, the boundary of $\hat{A}$ is a hypersurface $\widetilde{V}'_F$ which is covered by the disjoint union over $x \in F$ of the $\widetilde{V}'_x$. Left (functional view): A small neighbourhood U of u in V is split by $\widetilde{V}'_F$ into two parts, $\mathcal{C}_1$ and $\mathcal{C}_2$, which are inside two different topological classes (one of them belongs to $\hat{A}$ and one does not). Right (spatial view): When u changes continuously within $U \cap \widetilde{V}'_F$, the corresponding level-0 stratified critical point x changes continuously within F. Central panels shows three functions in $\widetilde{V}'_F$ and their critical points. Small perturbations of these functions all belong to the same topological class, for perturbations positive near the critical point they belong to $\mathcal{C}_1$ (right panels) and for negative perturbations to $\mathcal{C}_2$ (left panels).

Now, consider a coordinate system orthonormal with respect to the scalar product $\langle \cdot, \cdot \rangle$ induced by f. We typically denote $u = (u^1, \ldots, u^N)$ to be the set of coordinates of an element of V. For each $t \in [0, 1)$, let $\gamma_t : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be the density of the Gaussian vector with covariance

$$\begin{pmatrix} I_N & t I_N \\ t I_N & I_N \end{pmatrix}$$

and mean (μ, μ), where $\mu \in \mathbb{R}^N$ is such that $\mathbb{E}[f] = \sum \mu_i u^i$. This density gives the distribution of f_t as defined in (2.1). Piterbarg's formula (Lemma 2.22) implies that

$$(3.1) \qquad \frac{d}{dt}\mathbb{P}\left[f_t \in \hat{A}_1 \times \hat{A}_2 \right] = \int_{\partial \hat{A}_1 \times \partial \hat{A}_2} \langle \nu_{\hat{A}_1}(u_1), \nu_{\hat{A}_2}(u_2) \rangle \gamma_t(v_1, v_2)\, d\mathcal{H}^{N-1}(u_1) d\mathcal{H}^{N-1}(u_2),$$

where the integral is taken on the product of the smooth part of the boundaries of $\hat{A}_1$ and $\hat{A}_2$, which are seen as subsets of $\mathbb{R}^N$ through the coordinate system fixed above, and where $\nu_{\hat{A}_1}$ (resp. $\nu_{\hat{A}_2}$) is the outward unit normal vector to $\hat{A}_1$ (resp. $\hat{A}_2$) defined on the smooth part of its boundary. Applying the expression for the smooth part of the boundary of $\hat{A}_1$ and $\hat{A}_2$ in Lemma 3.10, we have

$$(3.2) \qquad \begin{aligned} &\frac{d}{dt}\mathbb{P}\left[f_t \in \hat{A}_1 \times \hat{A}_2 \right] \\ &= \sum_{F_1 \in \mathcal{F}_1,\, F_2 \in \mathcal{F}_2} \int_{\widetilde{\mathrm{Piv}}_{F_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{F_2}(\hat{A}_2)} \langle \nu_{\hat{A}_1}(u_1), \nu_{\hat{A}_2}(u_2) \rangle \gamma_t(u_1, u_2)\, d\mathcal{H}^{N-1}(u_1) \mathcal{H}^{N-1}(u_2)\,. \end{aligned}$$

The next step is to to apply the coarea formula to the integrals in (3.2). For each $F \in \mathcal{F}$, let Ξ_F denote the function $\widetilde{V}'_F \to F$ which maps $u \in \widetilde{V}'_F$ to the unique $x \in F$ such that u has a stratified critical point on F at level 0. We note that

$$\widetilde{\mathrm{Piv}}_x(\hat{A}) = (\Xi_F)^{-1}(x) \cap \widetilde{\mathrm{Piv}}_F(\hat{A}),$$

which means that we can parametrise $\widetilde{V}'_F$ by pairs (x, u) where $x \in F$ and $u \in \widetilde{\mathrm{Piv}}_x(\hat{A})$. The next lemma shows that Ξ_F is a submersion and gives an expression for its normal Jacobian:

Lemma 3.11. *For each $F \in \mathcal{F}$, the map Ξ_F is a submersion. Moreover, for each $u \in \widetilde{\mathrm{Piv}}_F(\hat{A})$, if $x := \Xi_F(u)$ then the normal Jacobian of Ξ_F at u is*

$$J_F(u) := \mathrm{Jac}^{\perp}\left[\Xi_F\right](u) = \frac{\mathrm{Jac}^{\perp}(L_x)}{|\det(H_x^F u)|},$$

where $L_x : V_x \to T_x^* F$ denotes the linear operator $u \mapsto d_x u|_F$, and where the determinant is taken in orthonormal coordinates of $T_x F$.

Using Lemma 3.11 we can apply the coarea formula to the integrals in (3.2), converting them from integrals over part of the boundary of the events to integrals over the faces of the stratified sets. As a result, each integral in (3.2) can be written as

$$(3.3) \qquad \int\limits_{F_1 \times F_2} \Gamma(t; x_1, x_2) \, \mathrm{dv}_{F_1}(x_1) \mathrm{dv}_{F_2}(x_2),$$

where, for each $x_1 \in F_1$, $x_2 \in F_2$ and $t \in [0, 1)$,

$$(3.4) \qquad \Gamma(t; x_1, x_2) = \int\limits_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)} \frac{\langle \nu_{\hat{A}_1}(u_1), \nu_{\hat{A}_2}(u_2) \rangle \gamma_t(u_1, u_2)}{J_{F_1}(u_1) J_{F_2}(u_2)} \, \mathrm{dv}_{V'_{x_1}}(u_1) \mathrm{dv}_{V'_{x_2}}(u_2),$$

and where $\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)$ is viewed as an open subset of $V'_{x_1} \times V'_{x_2}$. Here we have identified the spaces V'_{x_i} with their images in $\mathbb{R}^N$ in the coordinate system fixed previously. The measures $\mathrm{dv}_{V'_{x_i}}$ are defined as the canonical $N - \dim(F_i) - 1$ dimensional volume measures on the affine spaces V'_{x_i} of $\mathbb{R}^N$.

We next interpret the normal vectors in (3.4) in more tractable terms (using the sets from Definition 3.5):

Lemma 3.12. *The fibre $\widetilde{\mathrm{Piv}}_x(\hat{A})$ is the disjoint union of the two subsets $\widetilde{\mathrm{Piv}}_x^+(\hat{A})$ and $\widetilde{\mathrm{Piv}}_x^-(\hat{A})$. Moreover, for each $\sigma \in \{+, -\}$ and each $u \in \widetilde{\mathrm{Piv}}_x^\sigma(\hat{A})$, the outward unit normal vector of $\hat{A}$ at u is*

$$\nu_{\hat{A}}(u) = -\sigma \frac{K(x, \cdot)}{\|K(x, \cdot)\|} = -\sigma \frac{K(x, \cdot)}{\sqrt{K(x, x)}} \, .$$

Remark 3.13. Since K is the reproducing kernel in V, the evaluation map Ev_x defined by $v \mapsto v(x)$ is equal to the map $v \mapsto \langle v, K(x, \cdot) \rangle$. Hence $K(x, \cdot)$ is orthogonal to V_x, and so $\|K(x, \cdot)\|$ can also be interpreted as $\mathrm{Jac}^{\perp}(\mathrm{Ev}_x)$, the normal Jacobian of the evaluation operator.

Since K is the reproducing kernel in V, it satisfies $\langle K(x_1, \cdot), K(x_2, \cdot) \rangle = K(x_1, x_2)$. Hence

$$(3.5) \quad \langle \nu_{\hat{A}_1}(u_1), \nu_{\hat{A}_2}(u_2) \rangle = \sigma(u_1, u_2) \frac{K(x_1, x_2)}{\|K(x_1, \cdot)\| \|K(x_2, \cdot)\|} = \sigma(u_1, u_2) \frac{K(x_1, x_2)}{\sqrt{K(x_1, x_1) K(x_2, x_2)}},$$

where $\sigma(u_1, u_2) = +$ if either $(u_1, u_2) \in \widetilde{\mathrm{Piv}}_{x_1}^+(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}^+(\hat{A}_2)$ or $(u_1, u_2) \in \widetilde{\mathrm{Piv}}_{x_1}^-(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}^-(\hat{A}_2)$, and $\sigma(u_1, u_2) = -$ otherwise. Thus, by Lemma 3.11 and (3.5),

$$(3.6) \qquad \Gamma(t; x_1, x_2) = \int_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)} \Upsilon_{x_1, x_2}(u_1, u_2) \gamma_t(u_1, u_2) \, \mathrm{dv}_{V'_{x_1}}(u_1) \mathrm{dv}_{V'_{x_2}}(u_1),$$

where

$$(3.7) \qquad \Upsilon_{x_1,x_2}(u_1,u_2) = \frac{\sigma(u_1,u_2)K(x_1,x_2)}{\sqrt{K(x_1,x_1)K(x_2,x_2)}} \times \frac{\left|\det\left(H^{F_1}_{x_1}u_1\right)\right|\left|\det\left(H^{F_2}_{x_2}u_2\right)\right|}{\mathrm{Jac}^\perp(L_{x_1})\mathrm{Jac}^\perp(L_{x_2})},$$

and where $L_{x_i} : V_x \to T^*_{x_i}F_i$ denotes the linear operator $u \mapsto d_{x_i}u|_{F_i}$. The integral in the definition of Γ can be interpreted as a conditional expectation:

Lemma 3.14. *For each $t \in [0,1)$ and each distinct $x_1 \in F_1$ and $x_2 \in F_2$,*

$$\Gamma(t;x_1,x_2) = K(x_1,x_2)\,\gamma_{t;x_1,x_2}(0)\times$$
$$\mathbb{E}_{t;x_1,x_2}\left[\sigma(f^1_t,f^2_t)\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1)\times\widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(f^1_t,f^2_t)\left|\det\left(H^{F_1}_{x_1}f^1_t\right)\right|\left|\det\left(H^{F_2}_{x_2}f^2_t\right)\right|\right].$$

Combining (3.2), (3.3) and Lemma 3.14 yields the formula in Proposition 3.9.

4. Proof of the auxiliary lemmas

In this subsection we prove the auxiliary lemmas from Section 3, namely Lemmas 3.10–3.12, and Lemma 3.14. While we make use of the notation from Section 3, we do not rely on results from that section.

4.1. **Differential topology in the space of functions: Proof of Lemmas 3.10–3.12.** Throughout this section V denotes a linear subspace of $C^2(M)$; moreover, with the exception of the statement of Proposition 4.1, we will assume that V is finite-dimensional and satisfies Conditions 3.6 and 3.7. Again we fix an arbitrary stratified set $(B,\mathcal{F})$ in M and $\hat{A} \in \hat{\sigma}_{\mathrm{discr}}(B,V)$.

We begin with a couple of definitions; for the time being we work independently of the choice of $\hat{A}$. Let $F \in \mathcal{F}$, and recall from Section 3 the subsets $\widetilde{V}'_F \subset V'_F \subset V$ and the map $\Xi_F(u)$ which sends $u \in \widetilde{V}'_F$ to its unique non-degenerate stratified critical point at level 0. Let $\mathcal{I}_F$ be the set of pairs $(u,x) \in V \times F$ such that x is a stratified critical point of u at level 0 (so that in fact $u \in V'_F$), and let $\widetilde{\mathcal{I}}_F$ be the set of pairs $(u,x) \in \mathcal{I}_F$ such that x is the unique non-degenerate stratified critical point of u at level 0 (so that $u \in \widetilde{V}'_F$). By Condition 3.7, the map $(u,x) \mapsto (u(x),d_xu)$ is a submersion on $V \times F$, and so $\mathcal{I}_F$ is a smooth submanifold of $V \times F$ whose codimension is one plus the dimension of F. Moreover, for each $(u,x) \in \mathcal{I}_F$,

$$(4.1) \qquad T_{(u,x)}\mathcal{I}_F = \left\{(v,\tau) \in V \times T_xF \;:\; v(x) = 0,\; d_xv|_F + H^F_xu(\tau,\cdot) = 0\right\}.$$

Let $\mathrm{pr}^1_F : \mathcal{I}_F \to V$ and $\mathrm{pr}^2_F : \mathcal{I}_F \to F$ be the projections onto the first and second coordinates. Note that $V'_F = \mathrm{pr}^1_F(\mathcal{I}_F)$ and $\widetilde{V}'_F = \mathrm{pr}^1_F(\widetilde{\mathcal{I}}_F)$, and observe also that the map $\Xi_F(u)$ completes the following commutative diagram:

$$(4.2) \qquad \begin{array}{ccc} & \widetilde{\mathcal{I}}_F & \\ {\scriptstyle \mathrm{pr}^1_F}\swarrow & & \searrow{\scriptstyle \mathrm{pr}^2_F} \\ \widetilde{V}'_F & \xrightarrow{\;\Xi_F\;} & F \end{array}$$

Lemmas 3.10 and 3.11 both pertain to elements of this diagram: for Lemma 3.11 this is explicitly so, whereas for Lemma 3.10 it is since, as we shall see, $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is an open subset of $\widetilde{V}'_F$. In the proof of Lemmas 3.10 and 3.11, we use the following proposition (whose proof is postponed until the very end of the subsection):

Proposition 4.1. *Let $F \in \mathcal{F}$. Then the set $\widetilde{\mathcal{I}}_F$ is open in $\mathcal{I}_F$ and the set $\widetilde{V}'_F$ is open in $\mathfrak{D}_B$. Moreover, if V has finite dimension $N \in \mathbb{N}$ and satisfies Conditions 3.6 and 3.7, then*

$$\mathcal{H}^{N-1}\left(V'_F \setminus \widetilde{V}'_F\right) = 0.$$

Remark 4.2. Although we only apply Proposition 4.1 to finite-dimensional V, we state it in full generality so as to clarify which tools are used to prove each point.

Remark 4.3. Roughly speaking, Proposition 4.1 ensures that if the field f is conditioned to have a stratified critical point at level 0, then a.s. this critical point is non-degenerate, and there are no other stratified critical points at level 0.

Proof of Lemma 3.10. To show that $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is a smooth (immersed) conical hypersurface of V, we first show that $\widetilde{V}'_F$ is a smooth immersed (although maybe not embedded) hypersurface of V. By Proposition 4.1, $\widetilde{\mathcal{I}}_F$ is a smooth submanifold of $V \times F$ with the same tangent space as $\mathcal{I}_F$ at each point. The mapping $\mathrm{pr}^1_F : \widetilde{\mathcal{I}}_F \to V$ is one-to-one, and we claim that it has constant rank. To see this, let us take $(u,x) \in \widetilde{\mathcal{I}}_F$ and check that

$$d_{(u,x)}\mathrm{pr}^1_F \left(T_{(u,x)}\mathcal{I}_F \right) = V_x.$$

The inclusion $\subset$ is clear by (4.1). For the reverse inclusion, let $v \in V_x$ and define $\lambda = -d_x v|_F$. Since $(u,x) \in \widetilde{I}_F$, $H^F_x u$ is non-degenerate, and so there exists $\tau \in T_x F$ such that $H^F_x u(\tau, \cdot) = \lambda$. Therefore, $(v,\tau) \in T_{(u,x)}\mathcal{I}_F$ and $d_{(u,x)}\mathrm{pr}^1_F(v,\tau) = v$, which proves the reverse inclusion. To sum up, pr^1_F is a mapping of corank one on $\widetilde{\mathcal{I}}_F$, and so its image $\widetilde{V}'_F$ is a smooth immersed (although maybe not embedded) hypersurface of V with the tangent space

$$(4.3) \qquad T_u\widetilde{V}'_F = d_{(u,x)}\mathrm{pr}^1_F \left(T_{(u,x)}\mathcal{I}_F \right) = V_x.$$

Next, we show that $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is open in $\widetilde{V}'_F$. Indeed, by Proposition 4.1, $\widetilde{V}'_F$ is open in $\mathfrak{D}_B$. Moreover, $\widetilde{V}'_F$ is a smooth submanifold of V, which implies that, for each $u \in \widetilde{V}'_F$, there exists $U \subset V$ containing u such that $(u, U \cap \widetilde{V}'_F, U) \simeq (0, \mathbb{R}^{N-1} \times \{0\}, \mathbb{R}^N)$ and such that $U \cap \mathfrak{D}_B = U \cap \widetilde{V}'_F$. Hence there exist exactly two B-discriminant classes $\mathcal{C}_1, \mathcal{C}_2$ that intersect U and

$$(4.4) \qquad \overline{\mathcal{C}_1} \cap \widetilde{V}'_F \cap U = \overline{\mathcal{C}_2} \cap \widetilde{V}'_F \cap U = \widetilde{V}'_F \cap U$$

as illustrated in Figure 3. In particular, if $u \in \widetilde{\mathrm{Piv}}_F(\hat{A})$ then $\widetilde{V}'_F \cap U \subset \widetilde{\mathrm{Piv}}_F(\hat{A})$, and so $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is an open subset of $\widetilde{V}'_F$.

To sum up, since $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is open in $\widetilde{V}'_F$ and since $\widetilde{V}'_F$ is a smooth (immersed) hypersurface of V, $\widetilde{\mathrm{Piv}}_F(\hat{A})$ is also a smooth (immersed) hypersurface of V. Noting also that $\hat{A}$ is conical hence so is $\partial\hat{A}$, and observing moreover that, by (4.3), $T_u\widetilde{\mathrm{Piv}}_F(\hat{A}) = V_x$ for every $(u,x) \in \widetilde{\mathcal{I}}_F$, we complete the proof of the first two statements of the lemma.

For the third statement of the lemma, we define

$$E = \partial\hat{A} \setminus \left(\bigsqcup_{F \in \mathcal{F}} \widetilde{V}'_F \right).$$

By the definition of $\widetilde{\mathrm{Piv}}_F(\hat{A}) = \partial\hat{A} \cap \widetilde{V}'_F$, we have

$$(4.5) \qquad \partial\hat{A} = E \cup \bigsqcup_{F \in \mathcal{F}} \widetilde{\mathrm{Piv}}_F(\hat{A}).$$

Moreover, we claim that $\mathcal{H}^{N-1}(E) = 0$. To see this, observe that $\partial\hat{A} \subset \mathfrak{D}_B := \bigcup_{F \in \mathcal{F}} V'_F$. Indeed, since the discriminant $\mathfrak{D}_B$ is closed (see Lemma C.1), the B-discriminant class of any $u \in V \setminus \mathfrak{D}_B$ forms a neighbourhood of u; in particular, $u \notin \partial\hat{A}$. Hence we have an alternate expression for E:

$$E = \bigsqcup_{F \in \mathcal{F}} \partial\hat{A} \cap \left(V'_F \setminus \widetilde{V}'_F \right).$$

Since by Proposition 4.1 the $N-1$ dimensional Hausdorff measure of each term of the union on the right-hand side vanishes, it follows that $\mathcal{H}^{N-1}(E) = 0$. $\qquad \square$

Proof of Lemma 3.11. We first show that Ξ_F is a submersion. Let $(u,v) \in T\widetilde{V}'_F$, so that there exist $x \in F$ and $\tau \in T_x F$ such that $((u,x),(v,\tau)) \in T\widetilde{\mathcal{I}}_F$. In particular, by (4.1) we have $d_x v|_F + H^F_x u(\tau,\cdot) = 0$. Since $H^F_x u$ is non-degenerate, τ is uniquely determined by v. More precisely, let $\check{H}^F_x u$ be the image of $H^F_x u$ by the canonical isomorphism $(T^*F)^{\otimes 2} \simeq \mathrm{Hom}(T^*F, TF)$. Then $\tau = - \left(\check{H}^F_x u\right)^{-1} (d_x v|_F)$. Since the diagram (4.2) commutes, we have proven that

$$d_u \Xi_F(v) = - \left(\check{H}^F_x u\right)^{-1} (d_x v).$$

By Condition 3.7, the map $v \mapsto d_x v$ is surjective when restricted to V_x. Hence Ξ_F is a submersion, which proves the first statement of the lemma.

Let us now show that the Jacobian of Ξ_F is as claimed in the lemma. Let g_F^{-1} be the metric induced on T^*F by the metric g_F on TF. Since $(\check{H}^F_x u)^{-1}$ is an isomorphism $(T^*_x F, g^{-1}_{F,x}) \to (T_x F, g_{F,x})$, the normal Jacobian of Ξ_F is the product of the Jacobian of $(\check{H}^F_x u)^{-1}$ and of the normal Jacobian of the map $L_x : (V_x, \langle \cdot, \cdot \rangle) \to (T^*_x F, g^{-1}_{F,x})$, defined in the statement of the lemma to be $L_x(v) = d_x v$. Since the first Jacobian is the absolute value of the inverse of $\det(H^F_x u)$, i.e. the determinant of the matrix of the bilinear form $H^F_x u$ in a $g^{-1}_{F,x}$-orthonormal basis of $T_x F$, the proof is complete. $\square$

Remark 4.4. Although for our purposes we do not need to compute $\mathrm{Jac}^\perp(L_x)$ explicitly (since it eventually cancels out in the main formula), for completeness we have

$$\mathrm{Jac}^\perp(L_x) = \sqrt{\det\left(L_x L^*_x\right)} = \sqrt{\det\left(d_x \otimes d_x K_x|_{F,F}\right)},$$

where $K_x(y_1, y_2) = K(y_1, y_2) - K(x, y_2) K(y_1, x)/K(x,x)$ is the covariance kernel of f conditioned on $f(x) = 0$ or, equivalently, of the orthogonal projection of f onto V_x; this follows from the same routine computation as in Remark 3.13. More generally, if $L : V \to \mathbb{R}^k$ is a linear operator, the orthogonal Jacobian of Lf is the square root of the determinant of the covariance of Lf.

Let us now complete the proof of Lemma 3.12; for this we rely on elements from the proof of Lemma (3.10):

Proof of Lemma 3.12. Let $u \in \widetilde{\mathrm{Piv}}_F(\hat{A})$, $x = \Xi_F(u)$, and take U, $\mathcal{C}_1$ and $\mathcal{C}_2$ as in (4.4). By Lemma 3.10, we have $T_u \widetilde{\mathrm{Piv}}_F(\hat{A}) = V_x$. In particular, for any such v, $\langle K(x,\cdot), v \rangle = v(x) = 0$, so $K(x,\cdot)$ is orthogonal to $T_x \widetilde{\mathrm{Piv}}_F(\hat{A})$. Moreover, $\langle K(x,\cdot), K(x,\cdot) \rangle = K(x,x)$, which must be positive (otherwise all functions in V vanish at x which contradicts Condition 3.7). Therefore, the outward unit normal vector $\nu_{\hat{A}}(u)$ to A at u is plus or minus

$$(4.6) \qquad\qquad v_x := \frac{K(x,\cdot)}{\sqrt{K(x,x)}}.$$

The sign of this vector depends on which of the $\mathcal{C}_i$ belongs to $\hat{A}$. More precisely, a perturbation $u + \eta h$ (with $\eta \ll 1$) enters $\hat{A}$ whenever $\langle v_x, h \rangle = h(x)$ has the right sign. In particular, this shows that the sets $\widetilde{\mathrm{Piv}}^+_x(\hat{A})$ and $\widetilde{\mathrm{Piv}}^-_x(\hat{A})$ form a partition of $\widetilde{\mathrm{Piv}}_x(\hat{A})$ and that, for each $\sigma \in \{+,-\}$ and each $u \in \widetilde{\mathrm{Piv}}^\sigma_x(\hat{A})$, $\nu_{\hat{A}}(u) = -\sigma \frac{K(x,\cdot)}{\sqrt{K(x,x)}}$. $\square$

Finally, we prove Proposition 4.1. For this we use the following standard fact which we state without proof:

Lemma 4.5. *Let $h : M \to M'$ be a Lipschitz map and let $S \subset M$ be a k-dimensional submanifold of M. Then the Hausdorff dimension of $h(S)$ is at most k. In particular, $\mathcal{H}^d(h(S)) = 0$ for every $d > k$.*

Proof of Proposition 4.1. Let us first give some intuition. The set $\mathfrak{D}_B \setminus \widetilde{V}'_F$ consists of functions which, in addition to having a level-0 critical point on F, are degenerate in some way. We express the five different cases of degeneracy as the vanishing of five explicit smooth functionals of pairs $(u, x) \in V \times F$ or triplets $(u, x, y) \in V \times F \times F_2$ for some $F_2 \in \mathcal{F}$. From this we deduce both that $\widetilde{\mathcal{I}}_F$ is open in $\mathcal{I}_F$ and that its complement has positive codimension. We then conclude by projecting the vanishing loci onto V.

Recall that $V \subset C^2(M)$ is a linear space. Let d_1 denote the dimension of F, and let $F_2, F_3 \in \mathcal{F}$ be strata of dimensions d_2 and d_3 respectively. We consider the following five subsets:

(1) If $F < F_2$, let $\mathcal{I}^1_{F, F_2}$ be the set of pairs $(u, x) \in \mathcal{I}_F$ such that $d_x u \in T^*_{F_2} M|_x$.
(2) Let $\mathcal{I}^2_F$ be the set of pairs $(u, x) \in \mathcal{I}_F$ such that $H_x u$ is singular.
(3) If $F_2 < F_3$, let $\mathcal{I}^3_{F, F_2, F_3}$ be the set of triplets $(u, x, y) \in \mathcal{I}_F \times F_2$ such that x and y are distinct, y is also a stratified critical point of u and $d_y u \in T^*_{F_3} M|_y$.
(4) Let $\mathcal{I}^4_{F, F_2}$ be the set of triplets $(u, x, y) \in \mathcal{I}_F \times F_2$ such that x and y are distinct, y is also a stratified critical point of u and $H^{F_2}_y u$ is singular.
(5) Let $\mathcal{I}^5_{F, F_2}$ be the set of triplets $(u, x, y) \in \mathcal{I}_F \times F_2$ such that x and y are distinct and y is also a stratified critical point of u with critical value 0.

Claim 4.6. *Each of the five subsets defined above is a closed subset of $V \times F$ (resp. $V \times F \times F_2$, as appropriate). Moreover, if we assume in addition that V has finite dimension $N \in \mathbb{N}$ and satisfies Conditions 3.6 and 3.7, then each of these subsets is a finite union of submanifolds of codimension at least $N + d_1 + 1$ (resp. $N + d_1 + d_2 + 1$).*

Remark 4.7. The proof of Claim 4.6 is the only place in the paper where we use the fact that $\mathcal{F}$ is a tame stratification of B, rather than merely a Whitney stratification.

Proof. We begin with a couple of definitions. For each $F \in \mathcal{F}$, let $T^*_F M$ be the *conormal bundle* to F, that is, for each $x \in F$, $T^*_F M|_x$ is the set of $\xi \in T^*_x M$ such that $\xi|_{T_x F} = 0$. This is a smooth vector bundle whose rank is exactly the codimension of F in M. In particular, $T^*_F M$ has codimension d in $T^* M$. Recall from the definition of a tame stratification that, given $F_1, F_2 \in \mathcal{F}$ such that $F_1 < F_2$, the set of limit points of $T F_2$ with basepoints on F_1 defines a vector bundle over F_1, denoted by $T F_2|_{F_1}$, which we call the generalised tangent bundle of F_2 over F_1. This allows us to extend the definition of conormal bundle as follows: the *conormal bundle to F_2 over F_1*, denoted $T^*_{F_2} M|_{F_1}$, is the set of $(x, \xi) \in T^* M|_{F_1}$ such that ξ vanishes on $T^* F_2|_x$. This defines a smooth vector bundle over F_1 whose rank is the codimension of F_2 in M. Thus, a point $x \in F_1$ is a non-degenerate stratified critical point of some $u \in C^2(M)$ if and only if it is a non-degenerate critical point of $u|_{F_1}$ and for each $F_2 \in \mathcal{F}$ such that $F_1 < F_2$, $(x, d_x u) \notin T^*_{F_2} M|_{F_1}$.

Now, assume first that V has finite dimension $N \in \mathbb{N}$ and satisfies Conditions 3.6 and 3.7. Since the proofs all follow the same structure, we cover in detail only the case of $\mathcal{I}^1_{F, F_2}$, and then indicate what changes need to be made in the other cases.

Consider the map $\Phi_1 : \mathcal{I}_F \to T^* M|_F$ defined by $(u, x) \mapsto d_x u$, and recall the expression of the tangent spaces of $\mathcal{I}_F$ given in (4.1). By Condition 3.7, the map Φ_1 is a submersion. Moreover, the set $T^*_{F_2} M|_F$ is a smooth submanifold of $T^* M|_F$ of codimension $1 + d_2$ that is also a closed subset, and therefore $\mathcal{I}^1_{F, F_2} = \Phi_1^{-1}\left(T^*_{F_2} M|_F\right)$ is a smooth submanifold of $\mathcal{I}_F$ of codimension d_1 as well as a closed subset of this space. We have thus covered the case of $\mathcal{I}^1_{F, F_2}$.

For $\mathcal{I}^2_F$ we consider the map $\Phi_2 : \mathcal{I}_F \to \mathrm{Sym}^2\left(T^* F\right)$ defined by $(u, x) \mapsto H^F_x u$, which is a submersion by Condition 3.6. Instead of $T^*_{F_2} M|_F$, we consider the zero set of the determinant map $\det : \mathrm{Sym}^2\left(T^* F\right) \to \mathbb{R}$ induced by some auxiliary metric. Its zero set $\mathcal{W}_F$ is closed and can be partitioned into the spaces of matrices of fixed rank in $\{0, \ldots, d_1 - 1\}$ so it is a finite union of smooth submanifolds of positive codimension. Since $\mathcal{I}^2_F = \Phi_2^{-1}\left(\mathcal{W}_F\right)$, we are done.

The cases $\mathcal{I}^3_{F,F_2,F_3}$ and $\mathcal{I}^4_{F,F_2}$ are analogous to the first two cases. The maps Φ_1 and Φ_2 should be replaced by maps Φ_3 and Φ_4 defined on $\mathcal{I}_{F,F_2} = \{(u,x,y) \in V \times F \times F_2 \ : \ u(x) = 0, \ d_x u|_F = 0, \ d_y u|_{F_2} = 0\}$ which is a smooth submanifold of $\mathcal{I}_F \times F_2$ of codimension d_2 and whose tangent space at (u,x,y) is

$$\{(v,\tau_1,\tau_2) \in V \times T_x F \times T_y F_2 \ : \ v(x) = 0, \ d_x v|_F + H_x u(\tau_1, \cdot) = 0, \ d_y v|_{F_2} + H_y u(\tau_2, \cdot) = 0\} \ .$$

They should be defined as follows: $\Phi_3 : (u,x,y) \mapsto (u(x), d_x u, d_y u)$ and $\Phi_4(u,x,y) \mapsto H_y u$. As for $\mathcal{I}^2_F$, Condition 3.7 should be replaced by Condition 3.6 in the case of $\mathcal{I}^4_{F,F_2}$.

Finally, for $\mathcal{I}^5_{F,F_2}$ we can consider the map $\Phi_5 : \mathcal{I}_{F,F_2} \to \mathbb{R}$ that maps each triple (u,x,y) to $u(y)$. This map is a submersion by Condition 3.7. The conclusion follows accordingly.

This ends the proof of the finite-dimensional part of the claim. Consider now the general case. Observe that we still have $\mathcal{I}^1_{F,F_2} = \Phi_1^{-1}\left(T^*_{F_2} M|_F\right)$, which is the preimage of a closed subset by a continuous map; in particular it is also closed. Since the same argument works with the four other cases, we also deduce the infinite-dimensional case of the claim. □

Let us now use Claim 4.6 to prove that $\widetilde{\mathcal{I}}_F$ is open in $\mathcal{I}_F$. Consider $(u_k, x_k) \in (\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F)^{\mathbb{N}}$ that converges in $\mathcal{I}_F$; we claim its limit (u,x) belongs to $\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$. Observe that $\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$ is the union of the following sets:

(1) The union over the $\{F_2 \in \mathcal{F} : F < F_2\}$ of the sets $\mathcal{I}^1_{F,F_2}$.
(2) The set $\mathcal{I}^2_F$.
(3) The union over $\{F_2, F_3 \in \mathcal{F} : F_2 < F_3\}$ of the images of the projections $\mathcal{I}^3_{F,F_2,F_3} \to \mathcal{I}_F$.
(4) The union over $F_2 \in \mathcal{F}$ of the images of the projections $\mathcal{I}^4_{F,F_2} \to \mathcal{I}_F$.
(5) The union over $F_2 \in \mathcal{F}$ of the images of the projections $\mathcal{I}^5_{F,F_2} \to \mathcal{I}_F$.

Since the above union is over a finite set, one of them contains an infinite number of terms of the sequence $(u_k, x_k)_{k \in \mathbb{N}}$. We can and will thus assume, up to extraction, that the sequence (u_k, x_k) belongs to one of the sets just described. We now describe what happens in each case:

(1) By Claim 4.6, $\mathcal{I}^1_{F,F_2}$ is closed in $\mathcal{I}_F$, so $(u,x) \in \mathcal{I}^1_{F,F_2} \subset \mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$.
(2) We reason likewise.
(3) By construction, for each $k \in \mathbb{N}$, (u_k, x_k) is the projection of a triplet in $\mathcal{I}^3_{F,F_2,F_3}$. By compactness of B, we can extract a subsequence for which the third coordinate of the triplet converges in $\overline{F_2}$. Since the subsequence must have the same limit in the projection as the full sequence, we just denote it by $(u_k, x_k, y_k)_{k \in \mathbb{N}} \in (\mathcal{I}^3_{F,F_2,F_3})^{\mathbb{N}}$ so that the third coordinate converges to some $y \in \overline{F_2}$. If $y \in F_2$ then $(u,x,y) \in \mathcal{I}_F \times F_2$. Then by Claim 4.6, $\mathcal{I}^3_{F,F_2,F_3}$ is closed in $\mathcal{I}_F \times F_2$ so $(u,x,y) \in \mathcal{I}^3_{F_2,F_3}$, which implies that (u,x) belongs to its projection onto $\mathcal{I}_F$. If, on the other hand, $y \notin F_2$, (by Definition 2.1), y must belong to some F_4 such that $F_4 \in \mathcal{F}$ such that $F_4 < F_2$. Then $d_y u \in T^*_{F_2} M|_y$ (actually we even have $d_y u \in T^*_{F_3} M|_y$). If $y \neq x$, we must then have $(u,x,y) \in \mathcal{I}^3_{F,F_4,F_2}$ so (u,x) belongs to its projection onto $\mathcal{I}_F$. Otherwise, if $y = x$, then, $(u,x) = (u,y) \in \mathcal{I}^1_{F,F_4}$.
(4) We reason as in the third case. As before, up to extraction, we can find $(y_k)_{k \in \mathbb{N}} \in F_2$ converging to some $y \in \overline{F_2}$ such that for each $k \in \mathbb{N}$, $(u_k, x_k, y_k) \in \mathcal{I}^4_{F,F_2}$. Again, as before, if y belongs to some face $F_3 < F_2$, we have $d_y u \in T^*_{F_2} M|_y$ so $(u,x,y) \in \mathcal{I}^4_{F,F_3}$. Otherwise, if $y \in F_2$, using Claim 4.6 we deduce that $(u,x,y) \in \mathcal{I}^4_{F,F_2}$.
(5) We reason as in the fourth case.

We have therefore proven that $\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$ is closed in $\mathcal{I}_F$.

Next, we show that $\widetilde{V}'_F$ is open in $\mathfrak{D}_B$. By construction, $\mathfrak{D}_B$ is the union of the projections onto the first coordinates of the sets $\mathcal{I}_{F_1}$ for $F_1 \neq F$ and of the sets $\mathcal{I}^1_F$, $\mathcal{I}^2_F$, $\mathcal{I}^3_{F,F_2,F_3}$, $\mathcal{I}^4_{F,F_2}$ and $\mathcal{I}^5_{F,F_2}$ defined above, taken over all the adequate F_2 and F_3. As before we take $(u_k)_{k \in \mathbb{N}} \in$

$(\mathfrak{D}_B \setminus \widetilde{V}'_F)^{\mathbb{N}}$ converging to some $u \in V'_F$ and, up to extraction, there exist two strata $F_1 \leq F_2$ and a sequence $(x_k)_{k \in \mathbb{N}} \in F_2^{\mathbb{N}}$ and $x \in F_1$ such that for each $k \in \mathbb{N}$, $(u_k, x_k) \in \mathcal{I}_F \setminus \widetilde{\mathcal{I}}_{F_2}$ and $\lim_{k \to \infty}(u_k, x_k) = (u, x)$. By Lemma C.1, x is a stratified critical point of u. Let us prove that $u \in \mathfrak{D}_B \setminus \widetilde{V}'_F$. From now on, the reasoning is analogous to that used for $\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$.

(1) If $F_1 \neq F$, then, $(u, x) \in \mathcal{I}_{F_1}$ so $u \notin \widetilde{V}'_F$.
(2) If $F_2 > F_1 = F$, then, as before $d_x u \in T^*|_{F_2} M|_x$ and so $(u, x) \in \mathcal{I}^1_{F, F_2}$ and $u \notin \widetilde{V}'_F$.
(3) If $F_2 = F_1 = F$ then for each $k \in \mathbb{N}$, (u_k, x_k) belongs to $\mathcal{I}_F \setminus \widetilde{\mathcal{I}}_F$ which is closed in $\mathcal{I}_F$ so that $(u, x) \notin \widetilde{\mathcal{I}}_F$ and so $u \notin \widetilde{V}'_F$.

This proves that $\mathfrak{D}_B \setminus \widetilde{V}'_F$ is closed in $\mathfrak{D}_B$ as announced.

To finish, assume that V has finite dimension $N \in \mathbb{N}$ and satisfies Conditions 3.6 and 3.7. By (the finite-dimensional case of) Claim 4.6, $V'_F \setminus \widetilde{V}'_F$ is a finite union of projections of submanifolds of $V \times F$ and $V \times F \times F_2$ for $F_2 \in \mathcal{F}$ of codimensions at least $\dim(F) + 2$ and $\dim(F) + \dim(F_2) + 2$ respectively. By Lemma 4.5, we must therefore have $\mathcal{H}^{N-1}(V'_F \setminus \widetilde{V}'_F) = 0$. □

4.2. Conditional expectation computation: Proof of Lemma 3.14.

In this section we prove Lemma 3.14, that is, we rewrite the function Γ defined by (3.4) (see also (3.6)) in terms of a conditional expectation.

Fix $t \in [0, 1)$ and distinct $x_1 \in F_1$ and $x_2 \in F_2$. In the first part of the proof the exact expression of Υ_{x_1, x_2}, defined by (3.7), will not play any role except through the fact that it is bounded by a polynomial in u_1, u_2. Let P_{x_1, x_2} be the orthogonal projector in $V \times V$ (equipped with the product metric) onto the subspace $V'_{x_1} \times V'_{x_2}$, and let $P^\perp_{x_1, x_2} = I - P_{x_1, x_2}$ be the complementary orthogonal operator onto the orthogonal complement, which we denote by $(V'_{x_1} \times V'_{x_2})^\perp$. We write $(u_1, u_2) = w + w^\perp$ where $w = P_{x_1, x_2}(u_1, u_2)$ and $w^\perp = P^\perp_{x_1, x_2}(u_1, u_2)$. Let us define

$$j_{x_1, x_2} : V \times V \to \mathbb{R} \times T^*_{x_1} F_1 \times \mathbb{R} \times T^*_{x_2} F_2$$

by

$$j_{x_1, x_2}(u_1, u_2) = (u_1(x_1), d_{x_1} u_1|_{F_1}, u_2(x_2), d_{x_2} u_2|_{F_2}).$$

Note that the space $V'_{x_1} \times V'_{x_2}$ is exactly the kernel of j_{x_1, x_2}, hence j_{x_1, x_2} is a linear isomorphism from $(V'_{x_1} \times V'_{x_2})^\perp$ onto $\mathbb{R} \times T^*_{x_1} F_1 \times \mathbb{R} \times T^*_{x_2} F_2$. With this notation we can rewrite the integral in (3.6) as

$$(4.7) \qquad \Gamma(t; x_1, x_2) = \int_{V'_{x_1} \times V'_{x_2}} \mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(w) \Upsilon_{x_1, x_2}(w) \gamma_t(w) \, dw,$$

where $dw = \mathrm{dv}_{V'_{x_1}} \mathrm{dv}_{V'_{x_2}}$. In the same spirit we write $g_t = P_{x_1, x_2} f_t$ and $g^\perp_t = P^\perp_{x_1, x_2} f_t$ so that $f_t = g_t + g^\perp_t$. The density of f_t, conditioned on $g^\perp_t = 0$, at $w = (u_1, u_2) \in V'_{x_1} \times V'_{x_2}$ is given by

$$\gamma_{f_t | g^\perp_t = 0}(w) = \frac{\gamma_t(u_1, u_2)}{\gamma_{g^\perp_t}(0)},$$

where $\gamma_{g^\perp_t}(0)$ is the density of $g^\perp_t$ evaluated at 0. Notice that for f_t, conditioning on $g^\perp_t = 0$ is the same as conditioning on $(f^1_t(x_1), d_{x_1} f^1_t|_{F_1}, f^2_t(x_2), d_{x_2} f^2_t|_{F_2}) = 0$. Since by definition $(u_1, u_2) = w$ on $V'_{x_1} \times V'_{x_2}$, (4.7) becomes

$$(4.8) \qquad \Gamma(t; x_1, x_2) = \gamma_{g^\perp_t}(0) \int_{V'_{x_1} \times V'_{x_2}} \mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(w) \Upsilon_{x_1, x_2}(w) \gamma_{f_t | g^\perp_t = 0}(w) \, dw$$

$$= \gamma_{g^\perp_t}(0) \, \mathbb{E}_{t; x_1, x_2} \left[\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(f_t) \Upsilon_{x_1, x_2}(f^1_t, f^2_t) \right].$$

In the above expression, the density $\gamma_{g_t^\perp}(0)$ is with respect to the orthogonal coordinates in $(V'_{x_1} \times V'_{x_2})^\perp$, and we need to express it in terms of K. Let $\widetilde{Q}_{t;x_1,x_2}$ be the covariance matrix of $g_t^\perp$ in some orthonormal system of coordinates in $(V'_{x_1} \times V'_{x_2})^\perp$. Let $Q_{t;x_1,x_2}$ be the covariance of

$$(f_t^1(x_1), d_{x_1}f_t^1|_{F_1}, f_t^2(x_2), d_{x_2}f_t^2|_{F_2}) = j_{x_1,x_2}(f_t) = j_{x_1,x_2}(g_t)$$

in any orthonormal coordinate system of $\mathbb{R} \times T_{x_1}^*F_1 \times \mathbb{R} \times T_{x_2}^*F_2$ equipped with the product metric. Treating j_{x_1,x_2} as an isomorphism from $(V'_{x_1} \times V'_{x_2})^\perp$ onto $\mathbb{R} \times T_{x_1}^*F_1 \times \mathbb{R} \times T_{x_2}^*F_2$ we see that the covariances $\widetilde{Q}_{t;x_1,x_2}$ and $Q_{t;x_1,x_2}$ are linked by the following relation

$$\widetilde{Q}_{t;x_1,x_2} = \left(j_{x_1,x_2}^*\right)^{-1} Q_{t;x_1,x_2} j_{x_1,x_2}^{-1}.$$

In particular, $\det(\widetilde{Q}_{t;x_1,x_2}) = \det(Q_{t;x_1,x_2})/\det(j_{x_1,x_2}j_{x_1,x_2}^*)^{-1}$. Recalling that $\gamma_{t;x_1,x_2}(0)$ is the density of $j_{x_1,x_2}(f_t)$ at 0, we have

$$\gamma_{g_t^\perp}(0) = \gamma_{t;x_1,x_2}(0)\sqrt{\det\left(j_{x_1,x_2}j_{x_1,x_2}^*\right)}.$$

It remains to compute $\sqrt{\det\left(j_{x_1,x_2}j_{x_1,x_2}^*\right)}$. Notice first that j_{x_1,x_2} factors as the direct product of the two linear maps $j_{x_i} : V'^\perp_{x_i} \to \mathbb{R} \times T_{x_i}^*F_i$ for $i \in \{1,2\}$ defined as $j_{x_i}(u) = (u(x_i), d_{x_i}u|_{F_i})$,

$$(4.9) \qquad \det\left(j_{x_1,x_2}j_{x_1,x_2}^*\right) = \det\left(j_{x_1}j_{x_1}^*\right)\det\left(j_{x_2}j_{x_2}^*\right).$$

To compute $\det\left(j_{x_i}j_{x_i}^*\right)$ note that, since j_{x_i} is 0 on V'_{x_i}, this determinant does not depend on whether j_{x_i} acts on $V'^\perp_{x_i}$ or the entire V; we treat it as an operator on V. Next, we write V as orthogonal sum of V_{x_i} which is the space of functions such that $v(x_i) = 0$ and its orthogonal complement which is spanned by $K(x_i, \cdot)$ (see the discussion preceding (4.6)). Let us choose orthonormal coordinates in V that are adopted to this decomposition, that is $K(x_i, \cdot)/\|K(x_i, \cdot)\|$ must be one of the basis vectors. In this coordinates j_{x_i} factors as $u \mapsto u(x_i)$ acting on the span of $K(x_i, \cdot)$ (this is the operator Ev_{x_i} from Remark 3.13) and $u \mapsto d_{x_i}u|_{F_i}$ on V_x (which is the operator L_{x_i}). The factorisation implies that

$$\sqrt{\det\left(j_{x_i}j_{x_i}^*\right)} = \sqrt{\det\left(\mathrm{Ev}_{x_i}\mathrm{Ev}_{x_i}^*\right)\det\left(L_{x_i}L_{x_i}^*\right)} = \mathrm{Jac}^\perp(\mathrm{Ev}_{x_i})\mathrm{Jac}^\perp(L_{x_i}).$$

Plugging this computation into (4.8) we see that Γ is equal to

$$\gamma_t(x_1,x_2)\sqrt{K(x_1,x_1)K(x_2,x_2)}\mathrm{Jac}^\perp(L_{x_1})\mathrm{Jac}^\perp(L_{x_2})$$

$$\times \mathbb{E}_{t;x_1,x_2}\left[\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1)\times\widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(w)\Upsilon_{x_1,x_2}(f_t^1,f_t^2)\right].$$

Recalling the definition of $\Upsilon_{x_1,x_2}(u_1,u_2)$, and in particular pulling the terms $\sqrt{K(x_i,x_i)}$ and $\mathrm{Jac}^\perp(L_{x_i})$ from this definition out of the expectation (since they do not depend on u_i) so that they cancel with those already present, we deduce the result.

Remark 4.8. The cancellations in the above derivation are not so mysterious, since the relevant terms are Jacobians of evaluations of f and its differential and they appear, first, when we switch from space coordinates to functional coordinates, and then once again when we move back.

5. Proof of the main theorem: from the finite to the infinite-dimensional case

In this section we complete the proof of the covariance formula in Theorem 2.14. The basic idea is to (i) reinterpret topological events in terms of the discriminant, (ii) approximate the field f by a sequence of fields f_k taking values in a finite-dimensional spaces V_k, and then (iii) pass to the limit in the formula of Proposition 3.9.

In Section 5.1 we show that the boundary of pivotal events is well behaved, which will allow us to take limits of the expectations in the right-hand side of Proposition 3.9. In Section 5.2 we verify that topological events are encoded by the discriminant. Next, in Section 5.3 we construct

the finite-dimensional approximation and state an abstract continuity lemma for expectations that we use in the proof. Finally in Section 5.4 we assemble these elements into a proof of Theorem 2.14.

At the end of the section we also verify that Corollary 1.1 is indeed a special case of Theorem 2.14, as claimed in Section 1.

5.1. **On the boundary of pivotal events.** In this section we compare pivotal events in different subspaces of $C^2(M)$, link the two distinct notions of pivotal events we have introduced, and study the boundary of pivotal events.

Recall that $(B, \mathcal{F})$ denotes an arbitrary stratified set of M. Fix a linear subspace $V \subset C^2(M)$, not necessarily finite-dimensional. Also fix $\widetilde{A} \in \widetilde{\sigma}_{\mathrm{discr}}(B, C^2(M))$, and let $\hat{A} \in \hat{\sigma}_{\mathrm{discr}}(B, C^2(M))$ be the set of $u \in C^2(M)$ whose discriminant class (in $C^2(M)$) belongs to $\widetilde{A}$. Observe that the set $\hat{A}_V = \hat{A} \cap V$ belongs to $\hat{\sigma}_{\mathrm{discr}}(B, V)$, i.e. it is encoded by the V-discriminant. Indeed, it is the set of functions $u \in V$ whose discriminant class in $C^2(M)$ belongs to $\widetilde{A}$. Recall also the definition, for $x \in B$ and $\hat{A} \in \hat{\sigma}_{\mathrm{discr}}(B, V)$, of the sets $\widetilde{\mathrm{Piv}}_x(\hat{A})$ and $\widetilde{\mathrm{Piv}}_x^{\sigma}(\hat{A})$ from Definition 3.1.

The main result of this section is the following:

Lemma 5.1 (On pivotal events). *Suppose that V contains the constant functions on M. Then*

(1) $\widetilde{\mathrm{Piv}}_x(\hat{A}_V) = \widetilde{\mathrm{Piv}}_x(\hat{A}) \cap V$ *and, for each $\sigma \in \{+, -\}$, $\widetilde{\mathrm{Piv}}_x^{\sigma}(\hat{A}_V) = \widetilde{\mathrm{Piv}}_x^{\sigma}(\hat{A}) \cap V$.*

Moreover, let f be a Gaussian field on M satisfying Condition 2.7. Then, conditionally on x being a stratified critical point of f with $f(x) = 0$, a.s.

(2) $f \in \widetilde{\mathrm{Piv}}_x(\hat{A})$ *if and only if (i) $f \in \mathrm{Piv}_x(\hat{A})$ and (ii) $H_x^F f$ is a non-degenerate bilinear form. Moreover, for each $\sigma \in \{+, -\}$, the same is true if we replace $\widetilde{\mathrm{Piv}}_x(\hat{A})$ by $\widetilde{\mathrm{Piv}}_x^{\sigma}(\hat{A})$ and $\mathrm{Piv}_x(\hat{A})$ by $\mathrm{Piv}_x^{\sigma}(\hat{A})$.*

(3) *If x is a non-degenerate critical point then $f \notin \partial\widetilde{\mathrm{Piv}}_x(\hat{A})$, where $\widetilde{\mathrm{Piv}}_x(\hat{A})$ is seen as a subset of the space V_x'.*

Remark 5.2. Note that we only apply Lemma 5.1 to approximations of the field f (as opposed to f itself), so it is irrelevant that the constant functions will not belong to the Cameron-Martin space of f in general.

Remark 5.3. If we had been willing to impose a non-degeneracy condition on the Hessian of f, we could have concluded from Lemma 5.1 that, conditionally on x being a level-0 stratified critical point of f, a.s. $f \in \widetilde{\mathrm{Piv}}_x(\hat{A})$ if and only if $f \in \mathrm{Piv}_x(\hat{A})$, and this is the sense in which we think of $\widetilde{\mathrm{Piv}}_x(\hat{A})$ and $\mathrm{Piv}_x(\hat{A})$ as equal up to null sets. Since non-degeneracy of the Hessian is unnecessary for the result to hold, we do not do this.

In order to prove Lemma 5.1, we use the following result:

Lemma 5.4. *Let $u \in C^2(M)$ be such that u has a unique non-degenerate stratified critical point x at level 0 (c.f. the set $\cup_{F \in \mathcal{F}} \widetilde{V}_F'$). Then*

(1) *For $\varepsilon > 0$ small enough, neither $u + \varepsilon$ nor $u - \varepsilon$ have a stratified critical point at level 0. Moreover, let $\mathcal{C}_1 = \mathcal{C}_1(u)$ and $\mathcal{C}_2 = \mathcal{C}_2(u)$ be the connected components in $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$ of $u + \varepsilon$ and $u - \varepsilon$ respectively. Then, $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ (resp. $\overline{\mathcal{C}_1}$, $\overline{\mathcal{C}_2}$) is a neighbourhood of u in $C^2(M)$ (resp. in the set of functions $u' \in C^2(M)$ such that $u(x) \geq 0$, in the set of functions $u' \in C^2(M)$ such that $u(x) \leq 0$).*

(2) *Let $U \subset M$ be a neighbourhood of x. Then there is a neighbourhood $\mathcal{U}$ of u in the discriminant $\mathfrak{D}_B(C^2(M))$ such that, for each $u' \in \mathcal{U}$, u' has exactly one stratified critical point at level 0, which is non-degenerate and belongs to U. Moreover, we have $\mathcal{C}_1(u') = \mathcal{C}_1(u)$ and $\mathcal{C}_2(u') = \mathcal{C}_2(u)$ (defined as in (1)).*

Remark 5.5. If we were working in a finite-dimensional space, we could think of u as belonging to the smooth part of the discriminant. Since this discriminant is a hypersurface, this would mean that in a small neighbourhood of u the discriminant would be diffeomorphic to a hyperplane, separating the ambient space into two connected components $\mathcal{C}_1$ and $\mathcal{C}_2$, and moreover small perturbations of u would yield the same $\mathcal{C}_1$ and $\mathcal{C}_2$. Lemma 5.4 encodes (part of) this intuition.

Remark 5.6. The first point of Lemma 5.4 implies that the definition of $\widetilde{\mathrm{Piv}}_x^\sigma(\hat{A})$ does not change if one takes, in the definition of this set, $h \in C^2(M)$ instead of merely $h \in V$. Similarly, the definition does not change if one requires h to be a positive constant (which assists in showing a function does not belong to $\widetilde{\mathrm{Piv}}_x^\sigma(\hat{A})$).

Proof of Lemma 5.4. We start by showing that the property that u has a stratified critical point near x is stable under C^2 perturbations; this involves isolating x as a critical point in a uniform way. Fix a neighbourhood U of x. Since x is a non-degenerate critical point, it is isolated in the set of critical points of u (see Lemma C.2), which is compact by Lemma C.1. Thus, the critical value $u(x)$ is isolated in the set of critical values of u. In particular, for each $\varepsilon > 0$, both $u + \varepsilon$ and $u - \varepsilon$ belong to $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$, which justifies the existence of $\mathcal{C}_1$ and $\mathcal{C}_2$. Let us show that $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ is a neighbourhood of u in $C^2(M)$. Let $F \in \mathcal{F}$ be the stratum containing x. For each $r > 0$, let B_r be the Riemmanian ball of radius $r > 0$ in F centred at x. Since x is a non-degenerate critical point at x, the section du, which is C^1, vanishes transversally at x on the stratum F and stays bounded from below on the higher strata near x. Therefore, there exist $r = r(u) > 0$ and $\eta = \eta(u) > 0$ such that for each $w \in C^2(M)$ such that $\|w\|_{C^2(B)} \leq \eta$, the following holds:

- The ball B_r is included in U;
- The section $d(u + w)|_F$ vanishes exactly once on B_r;
- For any $F' \neq F$, $d(u + w)|_{F'}$ does not vanish on B_r;
- $u + w$ has no stratified critical points with critical value in $[-10\eta, 10\eta]$ outside of B_r;
- If moreover $\|w\|_{C^2(B)} \leq \eta/8$ then $|u + w| \leq \eta/4$ on B_r.

In particular, for each $s \in (0, \eta]$, $u \pm s$ does not belong to the discriminant. Let $w \in C^2(M)$ be such that $\|w\|_{C^2(B)} \leq \eta/8$. Let us show that $u + w \in \overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$, and that if $u + w \in \mathfrak{D}_B(C^2(M))$ then $u + w$ has a unique stratified critical point at level 0 which belongs to B_r. To this end, we will first consider a path $(v_t)_t$ from u to $u + w$ where w is a small perturbation. Along this path we will find further perturbations $v_{t,s} = v_t + s$ of v_t, for suitable choices of s, that do not belong to the discriminant and that belong to the two connected components $\mathcal{C}_1$ and $\mathcal{C}_2$.

More precisely, for each $t \in [0, 1]$ and each $s \in [-\eta/2, \eta/2]$, let $v_{t,s} = u + tw + s$. Then, for each $t \in [0, 1]$ and each $s \in [-\eta/2, \eta/2]$, $\|v_{t,s} - u\|_{C^2(B)} \leq \eta$ and $\|v_{t,0} - u\|_{C^2(M)} \leq \eta/8$. In particular, $v_{t,s}$ has a unique stratified critical point in B_r, which we call y_t (since it does not depend on s) and no other stratified critical points with critical value in $[-9\eta, 9\eta]$. Moreover, $\sup_{B_r} |v_{t,0}| \leq \eta/4$ so that $\min_{B_r} v_{t,\eta/2} \geq \eta/4$ and $\max_{B_r} v_{t,-\eta/2} \leq -\eta/4$. In particular, for each $t \in [0, 1]$, $v_{t,\pm\eta/2} \notin \mathfrak{D}_B(C^2(M))$. Thus, $v_{1,\pm\eta/2} \in \mathcal{C}_1 \cup \mathcal{C}_2$. Now, for each $s \in [-\eta/2, \eta/2]$, $v_{1,s} = u + w + s$. In particular, if $u(y_1) + w(y_1) \geq 0$, $v_{1,s}$ does not belong to the discriminant for $s \in (0, \eta/2]$ and converges to $u + w$ as $s \to 0$. If on the other hand, $u(y_1) + w(y_1) \leq 0$, the same approximation holds by taking $s \in [-\eta/2, 0)$ and $s \to 0$. In any case, by construction of this approximation $v_{1,s} \in \mathcal{C}_1 \cup \mathcal{C}_2$ as long as $s \neq 0$ so that $u + w \in \overline{\mathcal{C}_1 \cup \mathcal{C}_2} = \overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$. This shows that $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ is a neighbourhood of u in $C^2(M)$ and that for each $v \in (\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}) \cap \mathfrak{D}_B(C^2(M))$, v has a unique stratified critical point at level 0, which is in $B_r \subset U$.

Next notice that $u + w$ is in the same connected component of the complement of the discriminant as $u + s$ for $s \ll 1$ if $w(x) < 0$, and is in the same connected component as $u - s$ for $s \ll 1$ if $w(x) > 0$, which proves that $\overline{\mathcal{C}_1}$ (resp. $\overline{\mathcal{C}_2}$) is a neighbourhood of u in the set of functions taking non-negative (resp. non-positive) values at x. Finally, if $u(y_t) + tw(y_t) = 0$, by

construction $u + w + \varepsilon$ (resp. $u + w - \varepsilon$) is in the same connected component as $u + \varepsilon$ (resp. $u - w$) for $\varepsilon > 0$ small enough. In other words, $\mathcal{C}_i(u) = \mathcal{C}_i(u + w)$ for $i \in \{1, 2\}$. This ends the proof of the lemma.   □

Proof of Lemma 5.1. We prove the three statements in the lemma sequentially:

(1). In order to prove that $\widetilde{\mathrm{Piv}}_x(\hat{A}_V) = \widetilde{\mathrm{Piv}}_x(\hat{A}) \cap V$ and $\widetilde{\mathrm{Piv}}_x^\sigma(\hat{A}_V) = \widetilde{\mathrm{Piv}}_x^\sigma(\hat{A}) \cap V$, it is enough that $\partial \hat{A}_V = \partial \hat{A} \cap V$, where $\partial \hat{A}_V$ is the boundary of $\hat{A}_V$ in V. Clearly, $\partial \hat{A}_V \subset \partial \hat{A} \cap V$. On the other hand, let $u \in \partial \hat{A} \cap V$. Then, by Lemma 5.4, there exist $\mathcal{C}_1$ and $\mathcal{C}_2$ two connected components of $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$ such that for $\varepsilon > 0$ small enough, $u + \varepsilon \in \mathcal{C}_1$, $u - \varepsilon \in \mathcal{C}_2$ and $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ is a neighbourhood of u in $C^2(M)$. Let us assume that $\mathcal{C}_1 \subset \hat{A}$ and $\mathcal{C}_2 \subset C^2(M) \setminus \hat{A}$ since $u \in \partial \hat{A}$, and the case $\mathcal{C}_2 \subset \hat{A}$ and $\mathcal{C}_1 \subset C^2(M) \setminus \hat{A}$ follows by exchanging $\hat{A}$ and its complement. Since $u \in V$ and the constant functions belong to V, $u \pm \varepsilon \in V$. In particular, letting $\varepsilon \to 0$, we deduce that $u \in \overline{\hat{A} \cap V} \cap \overline{V \setminus \hat{A}} = \partial \hat{A}_V$, from which it follows that $\partial \hat{A}_V = \partial \hat{A} \cap V$ as announced.

(2). Let f_x denote the field f conditioned on $f(x) = 0$ and on x being a stratified critical point of f. Then, f_x is a.s. C^2. Assume now that $f_x \in \mathrm{Piv}_x^+(\hat{A})$. Then, there exists a (random) $h \in C^2(M)$ satisfying $h \geq 0$ such that, for small enough values of $\delta > 0$, $f_x + \delta h \in \hat{A}$ and $f_x - \delta h \notin \hat{A}$. In particular, $f_x \in \partial \hat{A}$. Moreover, since f satisfies Condition 2.7, by the regression formula $(f_x(y), d_y f_x)$ is non-degenerate for $y \neq x$, and so by Bulinskaya's lemma ([3, Proposition 1.20]) a.s. f_x has no other stratified critical points at level 0. If we also assume that $H_x^F f_x$ is non-degenerate, then by Remark 5.6 $f_x \in \widetilde{\mathrm{Piv}}_x^+(\hat{A})$. Thus, we have shown that if $f_x \in \mathrm{Piv}_x^+(\hat{A})$ and $H_x^F f_x$ is non-degenerate then a.s. $f_x \in \widetilde{\mathrm{Piv}}_x^+(\hat{A})$.

Conversely, assume that $f_x \in \widetilde{\mathrm{Piv}}_x^+(\hat{A})$ and let $U \subset M$ be a neighbourhood of x in M. Then, $H_x^F f_x$ is non-degenerate. Since f satisfies Condition 2.7, as before by Bulinskaya's lemma a.s. f_x has no other critical points at level 0. We may thus apply Lemma 5.4 to f_x, which implies that there exists a neighbourhood $\mathcal{U}$ of f_x in $C^2(M)$, two connected components $\mathcal{C}_1$ and $\mathcal{C}_2$ of $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$, and a geodesic ball $B_r \subset U$ of radius $r > 0$ centred at x, such that the following holds:

- For all small enough $\delta > 0$, $f_x + \delta \in \mathcal{C}_1$ and $f_x - \delta \in \mathcal{C}_2$.
- The union $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ covers $\mathcal{U}$.
- Each $v \in \mathcal{U} \cap \mathfrak{D}_B(C^2(M))$ has a unique stratified critical point in B_r and no stratified critical points at level 0 outside of B_r.

Since $f_x \in \widetilde{\mathrm{Piv}}_x^+(\hat{A})$, we have $\mathcal{C}_1 \subset \hat{A}$ and $\mathcal{C}_2 \subset \hat{A}^c$. Let $h \in C_c^2(W)$ be equal to 1 on B_r. Then, for all small enough $\delta > 0$, $f_x \pm h \in \mathcal{U}$, so f_x has no stratified critical points at level 0 outside of B_r. Inside B_r it coincides with f_x up to a constant $\pm \delta$. In particular, if $\delta \neq 0$, $f_x \pm h \in \mathcal{C}_1 \cup \mathcal{C}_2$. Moreover, by considering the path $(f_x \pm (\delta h + s(1 - h)))_{s \in [0, \delta]}$, we conclude that $f_x + h \in \mathcal{C}_1 \subset \hat{A}$ and $f_x - h \in \mathcal{C}_2 \subset \hat{A}^c$. But h is supported arbitrarily close to x. Thus, $f_x \in \mathrm{Piv}_x^+(\hat{A})$. Reasoning symmetrically, we get the same statement with the $+$ exponent replaced by $-$, and combining the two results we get the same property for $\mathrm{Piv}_x^\sigma(\hat{A})$ replaced by $\mathrm{Piv}_x(\hat{A})$.

(3). Assume that x is a non-degenerate critical point of f_x. Then, Lemma 5.4 applies so that there are two discriminant classes $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $\overline{\mathcal{C}_1} \cup \overline{\mathcal{C}_2}$ is a neighbourhood of f_x in $C^2(M)$ and there is a neighbourhood W of f_x in the discriminant such that for each $v \in W$, and each small enough $\varepsilon > 0$, $v + \varepsilon \in \mathcal{C}_1$ and $v - \varepsilon \in \mathcal{C}_2$. If $f_x \in \widetilde{\mathrm{Piv}}_x(\hat{A})$ then exactly one of the two classes belongs to $\hat{A}$, and hence the elements of W will all belong to the boundary of $\hat{A}$. Therefore, f_x belongs to the interior of $\widetilde{\mathrm{Piv}}_x(\hat{A})$ in the space V_x' of functions in $C^2(M)$ with a stratified critical point at x at level 0. Similarly, if $f_x \notin \widetilde{\mathrm{Piv}}_x(\hat{A})$ then either both $\mathcal{C}_1$ and $\mathcal{C}_2$ are subsets of $\hat{A}$ or neither of them are. So then, as before, the elements of W cannot belong to the boundary of

$\hat{A}$ so that f_x is in the interior of $V'_x \setminus \widetilde{\mathrm{Piv}}_x(\hat{A})$. In both cases, $f_x \notin \partial \widetilde{\mathrm{Piv}}_x(\hat{A})$, which proves the last part of the proposition. $\qquad \square$

5.2. **The topological class is encoded by the discriminant.** In this subsection we verify that topological events are encoded by the discriminant, making explicit the link between the events that appear in Theorem 2.14 and the events that appear in Proposition 3.9; in passing, we also prove the measurability of the stratified isotopy classes.

In this section $(B, \mathcal{F})$ again denotes an arbitrary stratified set of M; nevertheless, here we prefer to view $\mathcal{F}$ as a general Whitney stratification (see Remark 2.2) since we make use of the standard theory of Whitney stratifications. Recall the definition of B-discriminant classes from Definition 3.3, as well as the definition of the stratified isotopy class from Definition 2.8.

Lemma 5.7 (Topological class is encoded by the discriminant). *Suppose that $u, v \in C^2(M)$ have the same B-discriminant class in $C^2(M)$. Then their excursion sets $\{u > 0\}$ and $\{v > 0\}$ have the same stratified isotopy class, i.e., $[\{u > 0\}]_B = [\{v > 0\}]_B$.*

Since the discriminant classes are C^1-open (see Lemma C.1), there are at most countably many of them. This immediately implies the following:

Corollary 5.8. *There are at most countably many stratified isotopy classes of subsets of B. Moreover, the map $[\mathcal{D}]_B$ from the probability space Ω into the set of stratified isotopy classes is measurable.*

Before proving Lemma 5.7, let us recall some standard facts about Whitney stratifications; they can all be easily checked from the definitions of the objects they involve:

- If $I \subset \mathbb{R}$ is an open interval, then the collection $\mathcal{F}_I = (F \times I)_{F \in \mathcal{F}}$ is a Whitney stratification of $B \times I$.
- For each open subset $W \subset M$, $\mathcal{F}_W = (F \cap W)_{F \in \mathcal{F}}$ is a Whitney stratification of $B \cap W$.
- Consider $f : M \to N$ a smooth map between two Riemannian manifolds. Assume that $f|_B$ is proper and that for each $F \in \mathcal{F}$, $f|_F : F \to N$ is a submersion. Then, for each $y \in N$, the preimage $f^{-1}(y) \cap B$ is naturally equipped with a Whitney stratification $\mathcal{F}_y$ whose strata are the intersections $F \cap f^{-1}(y)$ where $F \in \mathcal{F}$ (see Definition 1.3.1 of Part I of [22]).

The proof of Lemma 5.7 is a standard application of Thom's first isotopy lemma (see (8.1) of [33]) and the isotopy extension theorem (see [17]). In fact, the only place we use C^2 regularity in this proof is when we apply Thom's first isotopy lemma.

Proof of Lemma 5.7. Let $u \in C^2(M) \setminus \mathfrak{D}_B(C^2(M))$, i.e. u has no stratified critical points in B at level 0. Recall that, by Lemma C.1 and since B is compact, the set of critical points of u is compact. In particular, this set is at positive distance from the zero set of u and there exists a bounded open neighbourhood $W \subset M$ of $u^{-1}(0)$ in M and a convex neighbourhood $\mathcal{U}$ of u in $C^1(M)$ such that for each $v \in \mathcal{U}$ and each face $F \in \mathcal{F}$, $dv|_F \neq 0$ in $\overline{W}$ and $v \neq 0$ on $B \setminus W$. We will prove that for each $v \in \mathcal{U} \cap C^2(M)$, $[\{v > 0\}]_B = [\{u > 0\}]_B$. To do so, notice that since $\mathcal{U}$ is open and convex, there exists I an open interval containing $[0, 1]$ such that for each $t \in I$, $u_t = tv + (1 - t)u \in \mathcal{U} \cap C^2(M)$. The family $\mathcal{F}_{W,I} = ((F \cap W) \times I)_{F \in \mathcal{F}}$ defines a Whitney stratification of $(B \cap W) \times I$ in $W \times I$. Moreover, since for each $t \in I$, u_t has no critical points on any face of $\mathcal{F}$ inside W, the map

$$(W \cap B) \times I \to \mathbb{R} \times I , \quad U : (x, t) \mapsto (u_t(x), t)$$

is a submersion when restricted to any face of $\mathcal{F}_{W,I}$. It is proper since B is compact and $id_I : I \to I$ is proper. In particular, by Thom's first isotopy lemma, since U is C^2 there exists a stratified homeomorphism $h : W \times I \to (W \times I) \cap U^{-1}(0) \times \mathbb{R} \times I$ (where $U^{-1}(0) \cap B$ is equipped with the preimage Whitney stratification that exists since F is transverse to $\{0\}$ in $\mathbb{R} \times I$) such

that $U \circ h^{-1}$ is the projection on the last two factors. Note that $U^{-1}(0) = \{(x,t) \in M \times I :$ $(u_t(x), t) = (0,0)\} = u_0^{-1}(0) \times \{0\}$. In particular, the map

$$(B \cap u_0^{-1}(0)) \times I \to B \times I , \quad (z,t) \mapsto (f_t(x), t) := h^{-1}((x,0), 0, t)$$

defines an isotopy of $u_0^{-1}(0)$ in B such that for each $t \in I$, $f_t(u_0^{-1}(0) \cap B) = u_t^{-1}(0) \cap B$. Since it is constructed from h it extends to an isotopy of a tubular neighbourhood of $u_0^{-1}(0)$ in B that preserves strata of $\mathcal{F}$. By Corollary 1.4 of [17] (and its extension provided in Section 7 of the same article), there exists a continuous isotopy $B \times I \to B \times I$ $(x,t) \to (\Phi_t(x), t)$ such that for each $t \in I$, Φ_t is a stratified homeomorphism of B and $\Phi_t \circ f_0 = f_t$. In particular, $\Phi_t(u_0^{-1}(0) \cap B) = u_t^{-1}(0) \cap B$ for each $t \in I$. Since $\Phi_0 = id$, and Φ_t is continuous in t, we also have $\Phi_1(\{u_0 > 0\} \cap B) = \{u_1 > 0\} \cap B$ and so $[\{u > 0\}]_B = [\{v > 0\}]_B$. Given that this is true for all $v \in \mathcal{U} \cap C^2(M)$, we have shown that equivalence classes for the equivalence relation generated by the map $u \mapsto [\{u > 0\}]_B$ are $C^2(M)$-open. In particular, since $\mathfrak{D}_B(C^1(M))$ is C^1-closed (by Lemma C.1 or just C^2-closed by the present argument) each topological class in $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$ must be a union of connected components of $C^2(M) \setminus \mathfrak{D}_B(C^2(M))$ and the proof is over. $\qquad \square$

5.3. **Approximation results.** To deduce Theorem 2.14 from Proposition 3.9, we approximate the field f by a sequence of fields $(f_k)_{k \in \mathbb{N}}$ taking values in finite-dimensional subspaces $(V_k)_{k \in \mathbb{N}}$ of $C^2(M)$. Then, we integrate the result of Proposition 3.9 and pass to the limit.

In this subsection, we first show the existence of an approximating sequence in a general setting (see Lemma 5.9), and then state the abstract continuity lemma for expectations (see Lemma 5.10) which we use to show the convergence of the terms in Proposition 3.9.

Lemma 5.9 (Existence of finite-dimensional approximations). *Fix $l \in \mathbb{N}$ and let f be an a.s. C^l Gaussian field on a smooth manifold M of dimension d. Let $V \subset C^l(M)$ be a linear subspace of $C^l(M)$ such that f belongs a.s. to V. Then the following holds:*

(1) *There exists a sequence $(V_k)_{k \in \mathbb{N}}$ of finite-dimensional linear subspaces of V and a sequence of Gaussian fields $(f_k)_{k \in \mathbb{N}}$, all defined in the same probability space as f, that converges in probability to f in the topology of uniform C^l convergence on compact subsets of M, and such that for each $k \in \mathbb{N}$, $f_k \in V_k$ a.s. and f_k defines a non-degenerate Gaussian vector in V_k. If f is centred then the f_k can also be chosen to be centred.*

(2) *Moreover, let $W \subset V$ be a finite-dimensional subspace. Then, we may find sequences $(V_k)_{k \in \mathbb{N}}$ and $(f_k)_{k \in \mathbb{N}}$ as in (1) such that $W \subset V_k$.*

Proof. Consider a countable atlas $(U_j, \phi_j)_{j \in \mathbb{N}}$ of M. Let $J \subset \mathbb{N}$, let $\eta > 0$ be a parameter to be fixed later, and let $I \subset M$ be a locally finite set such that, for each $j \in J$ and $z \in \phi_j(U_j)$, there exists $x \in \phi_j(I \cap U_j)$ for which $|z - x| \leq \eta$. Let $\varepsilon > 0$, fix $j \in J$ and let $B \subset U_j$ be a compact subset. Let us prove that there exists $\eta_0 = \eta_0(j, B, \varepsilon) > 0$ such that for all $\eta \leq \eta_0$, the field $f_I := \mathbb{E}[f \mid f_I]$ satisfies

$$(5.1) \qquad \qquad \mathbb{P}\left[\|f \circ \phi_j^{-1} - f_I \circ \phi_j^{-1}\|_{C^l(B)} > \varepsilon\right] < \varepsilon.$$

Since the seminorms $\|\cdot\|_{C^l(B)}$, where $j \in \mathbb{N}$ and $B \subset U_j$ ranges over the compact subsets of U_j, generate the topology of $C^l(M)$, repeating the above construction for a sequence $(\varepsilon_k)_{k \in \mathbb{N}} \to 0$ yields a sequence $(f_k)_{k \in \mathbb{N}}$ of fields satisfying (5.1) for $\varepsilon = \varepsilon_k$ which proves the first point of the lemma.

To prove (5.1), fix $\alpha \in \mathbb{N}^d$ and let $g = \partial^\alpha(f \circ \phi_j^{-1})$ and $g_I = \partial^\alpha(f_I \circ \phi_j^{-1})$. Also, let $N = N(d, l) \in \mathbb{N}$ be the number of multi-indices $\alpha \in \mathbb{N}^d$ such that $|\alpha| \leq l$. Observe that, for each $z \in \phi_j(B)$, there exists $z \in \phi_j(U_j \cap I)$ such that $|x - z| < \eta$ so that

$$|g(z) - g_I(z)| \leq |g(z) - g(x)| + \mathbb{E}[|g(x) - \mathbb{E}[g(z) \mid f|_I]|] \leq |g(z) - g(x)| + \mathbb{E}[|g(z) - g(x)|].$$

In particular,

$$\mathbb{E}\big[\sup_{z\in B}|g(z)-g_I(z)|\big] \leq 2\mathbb{E}\Big[\sup_{x,y\in B,\ |x-y|\leq\eta}|g(x)-g(y)|\Big] =: 2\mathbb{E}\left[X_\eta\right].$$

Now, since g is a continuous Gaussian field on B (which is compact), it is a.s. bounded and the family $(X_\eta)_{\eta>0}$ is uniformly L^1 (see for instance [3, Theorem 2.9]). Moreover, since g is continuous, X_η converges a.s. to 0 as $\eta \to 0$. In particular, $\lim_{\eta\to 0}\mathbb{E}[X_\eta] = 0$. Thus, there exists $\eta_\alpha = \eta_\alpha(\varepsilon,d,l) > 0$ such that for each $\eta \leq \eta_\alpha$, $\mathbb{P}\left[\sup_{z\in B}|g(z)-g_I(z)| > \varepsilon\right] \leq \varepsilon/C$. The estimate (5.1) follows by taking a union bound of the probability of the events $\sup_{z\in B}|g(z)-g_I(z)|$ where $\alpha \in \mathbb{N}^d$ ranges over all the multiindices such that $|\alpha| \leq l$, and will be valid for $\eta \leq \eta_0 = \min_{|\alpha|\leq l}\eta_\alpha$.

Consider now a sequence $(\eta_k)_{k\in\mathbb{N}}$ of positive real numbers converging to 0 and $(I_k)_{k\in\mathbb{N}}$ an increasing sequence of finite subsets of M, such that for each $k \in \mathbb{N}$ and $j \leq k$, $\phi_j(I_k \cap U_j)$ is an η_k-net of $\phi_j(U_j)$. For each $k \in \mathbb{N}$, let $f_k = f_{I_k}$ defined as above (with $J = [0,k]$). Then, for each $j \in \mathbb{N}$ and each compact subset $B \subset$, by (5.1), $\lim_{k\to\infty}\|f\circ\phi_j^{-1} - f_k\circ\phi_j^{-1}\|_{C^l(B)} = 0$ in probability so $(f_k)_{k\in\mathbb{N}}$ converges to f in probability.

We claim that each f_k belongs to a finite-dimensional subspace V_k. Indeed, let K be the covariance of f. By the regression formula (Proposition 1.2 of [3]), for each $k \in \mathbb{N}$, f_k is a random linear combination of the functions $K(\cdot,x)$ for $x \in I_k$ and of $\mathbb{E}[f]$ the mean of f. Hence it belongs to the finite-dimensional subspace V_k generated by these functions. Moreover, since f_k is the mean of a random variable with values in V, we have $V_k \subset V$ and if f is centred, by construction, f_k is centred. This concludes the proof of the first statement.

For the second statement, take $(f_k)_{k\in\mathbb{N}}$ as above, let $(h_1,\ldots,h_m)$ be a basis of W, and let $\xi_1,\ldots,\xi_m$ be independent standard normals. Then, clearly, $\frac{1}{k}(\xi_1 h_1 + \cdots + \xi_m h_m)$ converges to 0 in probability in $C^l(M)$, so that replacing f_k by

$$f_k + \frac{1}{k}(\xi_1 h_1 + \cdots + \xi_m h_m)$$

yields the required result. $\qquad\square$

Next we state without proof an abstract continuity lemma for expectations; this can be considered a simple variant of the standard Portemanteau lemma.

Lemma 5.10. *Let (X,Y) and $(X_k,Y_k)_{k\in\mathbb{N}}$ be random variables with values in $\mathbb{R}\times E$, where E is a Polish space. Assume that the sequence $(X_k,Y_k)_{k\in\mathbb{N}}$ converges in law towards (X,Y), and that the sequence $(X_k)_{k\in\mathbb{N}}$ is uniformly integrable. Let $A \subset E$ and assume $\mathbb{P}[X \neq 0,\ Y \in \partial A] = 0$. Then*

$$\lim_{k\to\infty}\mathbb{E}\left[X_k \mathbb{1}_{[Y_k\in A]}\right] = \mathbb{E}\left[X\mathbb{1}_{[Y\in A]}\right].$$

5.4. **Completing the proof of Theorem 2.14.** To complete the proof of Theorem 2.14 we assemble the previous elements together, namely we:

- Approximate f by a sequence of finite-dimensional fields $(f_k)_{k\in\mathbb{N}}$ with nice regularity and non-degeneracy properties constructed using Lemma 5.9.
- Use Lemma 5.7 to encode the topological events A_1 and A_2 via the discriminant.
- Apply Proposition 3.9 to the fields f_k and the events encoded by the discriminant.
- Pass to the limit in each term of the formula given by Proposition 3.9, using Lemmas 5.1 and 5.10.
- Show that the two definitions of pivotal events coincide using Lemma 5.1.

Proof of Theorem 2.14. Recall that $(B_1,\mathcal{F}_1)$ and $(B_2,\mathcal{F}_2)$ are stratified sets of M, and A_1 and A_2 are topological events on B_1 and B_2 respectively. By Lemma 5.7, for each $i \in \{1,2\}$ there exists $\hat{A}_i \in \hat{\sigma}_{\text{discr}}(B_i, C^2(M))$ such that $\mathbb{P}[f \in \hat{A}_i \triangle A_i] = 0$, and so it will be sufficient to work with the events $\hat{A}_i \in \hat{\sigma}_{\text{discr}}(B_i, C^2(M))$.

Let us first define the approximating sequence of fields. By Remark 3.8, there exists a finite-dimensional subspace $W \subset C^2(M)$ satisfying Conditions 3.6 and 3.7 that contains the constant functions. Hence we may define a sequence of Gaussian fields $(f_k)_{k \in \mathbb{N}}$, taking values in a sequence of finite-dimensional linear subspaces $W \subset V_k \subset C^2(M)$, that satisfy all the properties guaranteed by Lemma 5.9 (setting $\ell = 2$, and so in particular the f_k converge in probability in the topology of uniform C^2 on compact sets). Since W satisfies Conditions 3.6 and 3.7 so does each V_k, and so Proposition 3.9 applies to the sets $\hat{A}_{i,k} = \hat{A}_i \cap V$ for $i \in \{1, 2\}$ and the field f_k.

Next, recall that f^1 and f^2 denote independent copies of f, and let $(f_k^1)_{k \in \mathbb{N}}$ and $(f_k^2)_{k \in \mathbb{N}}$ be independent copies of $(f_k)_{k \in \mathbb{N}}$, with f_k^1 converging to f_1 and f_k^2 converging to f_2 (i.e. in C^2). Similarly, recall that f_t denotes the interpolation $(f_t^1, f_t^2) = (f^1, t(f^1 - \mu) + \sqrt{1 - t^2}(f^2 - \mu) + \mu)$, and define for each $k \in \mathbb{N}$ the interpolation $f_{t,k} = (f_{t,k}^1, f_{t,k}^2)$ analogously. Applying Proposition 3.9 we have, for each $k \in \mathbb{N}$ and $t \in [0, 1)$,

$$(5.2) \quad \mathbb{P}\left[f_{t,k} \in \hat{A}_{1,k} \times \hat{A}_{2,k}\right] - \mathbb{P}\left[f_{0,k} \in \hat{A}_{1,k} \times \hat{A}_{2,k}\right]$$

$$= \sum_{F_1 \in \mathcal{F}_1, \ F_2 \in \mathcal{F}_2} \int_0^t \int_{F_1 \times F_2} K_k(x_1, x_2) \times \Lambda_k(s; x_1, x_2) \gamma_{s,k;x_1,x_2}(0) \, \mathrm{dv}_{F_1}(x_1) \mathrm{dv}_{F_2}(x_2) \mathrm{d}s,$$

where $\Lambda_k(s; x_1, x_2)$ equals
$$(5.3)$$
$$\mathbb{E}_{s;x_1,x_2}\left[\sigma(f_{s,k}^1, f_{s,k}^2)\mathbb{1}_{\widetilde{\mathrm{Piv}_{x_1}}(\hat{A}_{1,k}) \times \widetilde{\mathrm{Piv}_{x_2}}(\hat{A}_{2,k})}(f_{s,k}^1, f_{s,k}^2)|\det\left(H_{x_1}f_{s,k}^1|_{F_1}\right)||\det\left(H_{x_2}f_{s,k}^2|_{F_2}\right)|\right],$$

and where K_k is the covariance of f_k, and $\gamma_{s,k;x_1,x_2}$ is the density of

$$(f_{s,k}^1(x_1), d_{x_1}f_{s,k}^1|_{F_1}, f_{s,k}^2(x_2), d_{x_2}f_{s,k}^2|_{F_2})$$

in orthonormal coordinates (recall that the subscript $s; x_1, x_2$ in the expectation denotes conditioning on this vector vanishing).

Let us compute the limits of both sides of (5.2) as $k \to \infty$, beginning with the left-hand side. Notice that for each $t \in [0, 1)$ and $k \in \mathbb{N}$, $f_k^t \in \hat{A}_{1,k} \times \hat{A}_{2,k}$ if and only if $f_k^t \in \hat{A}_1 \times \hat{A}_2$. Since $\partial(\hat{A}_1 \times \hat{A}_2) \subset \mathfrak{D}_B(C^2(M)) \times C^2(M) \cup C^2(M) \times \mathfrak{D}_B(C^2(M))$, and since f satisfies Condition 2.7, by Bulinskaya's lemma (Proposition 1.20 of [3])

$$\mathbb{P}[(f_t^1, f_t^2) \in \partial(\hat{A}_1 \times \hat{A}_2)] = 0.$$

Thus, by Lemma 5.10 (setting $X = 1$, $Y = (f_t^1, f_t^2)$, $E = C^2(M) \times C^2(M)$ and $A = \hat{A}_1 \times \hat{A}_2$), we have

$$(5.4) \qquad\qquad \lim_{k \to \infty} \mathbb{P}[f_{t,k} \in \hat{A}_{1,k} \times \hat{A}_{2,k}] = \mathbb{P}[f_t \in \hat{A}_2 \times \hat{A}_2].$$

We turn now to the right-hand side of (5.2); we begin by computing the pointwise limit of the integrand, and then apply the dominated convergence theorem. Fix $F_1 \in \mathcal{F}_1$, $F_2 \in \mathcal{F}_2$, $x_1 \in F_1$, $x_2 \in F_2$ and $s \in [0, t]$ (so that $s < 1$). Since the Hessians restricted to zero-dimensional faces have vanishing determinants, we may assume that $\dim(F_1), \dim(F_2) > 0$. This allows us to assume that $x_1 \neq x_2$ by removing a set of measure zero from the integral in (x_1, x_2). Now, since f_k^s converges in probability to f^s as $k \to \infty$, it also converges in law. In particular $K_k(x_1, x_2)$ converges to $K(x_1, x_2)$ and $\gamma_{s,k;x_1,x_2}(0)$ converges to $\gamma_{s;x_1,x_2}(0)$. To deal with $\Lambda_k(s; x_1, x_2)$, we note that $f_{s,k}$ converges in law to f_s in $C^2(M)$ as $k \to \infty$ so that K_k converges to K in $C^2(M \times M)$. Now, since the vector $(f_s^1(x_1), d_{x_1}f_s^1|_{F_1}, f_s^2(x_2), d_{x_2}f_s^2|_{F_2})$ is non-degenerate, by the regression formula (Proposition 1.2 of [3]), the law of f conditioned on this vector vanishing is well-defined and depends continuously in K. Since the covariance of a field determines its law, we deduce that the sequence of fields $(f_{s,k})_{k \in \mathbb{N}}$ conditioned on

$(f^1_{s,k}(x_1), d_{x_1}f^1_{s,k}|_{F_1}, f^2_{s,k}(x_2), d_{x_2}f^2_{s,k}|_{F_2}) = 0$ converges in law to f with the above conditioning. We denote the conditional law of these fields by $\mathbb{P}_{s;x_1,x_2}[\dots]$. By the first statement of Lemma 5.1

$$\mathbb{P}_{s;x_1,x_2}\left[f_{t,k} \in \widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_{1,k}) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_{2,k}) \triangle \widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)\right] = 0$$

so, if we temporarily set

$$A = \widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2), \quad X_s = X_s(x_1, x_2) = \sigma(f_s)|\det(H^{F_1}_{x_1}f^1_s)||\det(H^{F_2}_{x_2}f^2_s)|$$

and

$$X_{s,k} = X_{s,k}(x_1, x_2) = \sigma(f_{s,k})|\det(H^{F_1}_{x_1}f^1_{s,k})||\det(H^{F_2}_{x_2}f^2_{s,k})|,$$

we have

$$\Lambda_k(s; x_1, x_2) = \mathbb{E}_{s;x_1,x_2}\left[\mathbb{1}_A(f_{s,k})X_{s,k}(x_1, x_2)\right].$$

Since under the conditioning f_k converges in law to f, and since these are Gaussian fields, the sequence $(X_{s,k}(x_1, x_2))_{k \in \mathbb{N}}$ is uniformly integrable (though the bound may depend on s, x_1 and x_2). On the other hand, the random variables $f_{s,k}$ and f_s take values in the Polish space $C^2(M)$. By the third point of Lemma 5.1, a.s. either the Hessian of one of the f^i_s's is degenerate, which implies that either $X_s = 0$ or $f_t \notin \partial A$. Moreover, the pair $(X_{s,k}, f_{s,k})$ converges in probability to the pair (X_s, f_s). By Lemma 5.10, we have $\lim_{k \to \infty} \Lambda_k(s; x_1, x_2) = \Lambda(s; x_1, x_2)$ which is equal to

$$(5.5) \qquad \mathbb{E}_{s;x_1,x_2}\left[\sigma(f^1_s, f^2_s)\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_1}(\hat{A}_1) \times \widetilde{\mathrm{Piv}}_{x_2}(\hat{A}_2)}(f^1_s, f^2_s)|\det(H^{F_1}_{x_1}f^1_s)||\det(H^{F_2}_{x_2}f^2_s)|\right].$$

In summary, the integrand of the right hand side of (5.2) converges pointwise to the same quantity with $f_{s,k}$ replaced by f_s everywhere.

To apply the dominated convergence theorem to the right hand side of (5.2), we must find a uniform L^1 bound on the integrand. To bound $\Lambda_k(s; x_1, x_2) \times \gamma_{s,k;x_1,x_2}(0)$ we use Lemma A.4 with $X^i_{s,k} = H_{x_i}f^i_{s,k}$, $Y_{s,k} = Y_{s,k}(x_1, x_2) = (f^1_{s,k}(x_1), d_{x_1}f^1_{s,k}|_{T_{x_1}F_1}, f^2_{s,k}(x_2), d_{x_2}f^2_{s,k}|_{T_{x_2}F_2})$ for each $i \in \{1, 2\}$ and $k \in \mathbb{N}$. For any finite-dimensional Gaussian vector X in a space equipped with a scalar product, let $\mathrm{DC}(X)$ be the determinant of the covariance of X in orthonormal coordinates. The covariances of the coordinates of the $X^i_{s,k}$ are bounded in terms of derivatives up to order two in each variable of the covariances K_k; since these are uniformly bounded, there exists a constant $C < \infty$ for which, for each $k \in \mathbb{N}$,

$$|\Lambda_k(s; x_1, x_2)\gamma_{s,k;x_1,x_2}(0)| \leq \frac{C}{\sqrt{\mathrm{DC}(Y_{s,k})}}.$$

Next, for $i, j \in \{1, 2\}$ let $Y^{ij}_k(x_j) = (f^i_k(x_j), d_{x_j}f^i_k|_{T_{x_j}F_j})$ so that, for any $j_1, j_2 \in \{1, 2\}$, $Y^{1j_1}_k$ is independent from $Y^{2j_2}_k$ and

$$Y_{s,k} = (Y^{11}_k, s(Y^{12}_k - \mathbb{E}[Y^{12}_k]) + \sqrt{1-s^2}(Y^{22}_k - \mathbb{E}[Y^{22}_k] + \mathbb{E}[Y^{12}_k])).$$

Then, by Lemma A.3, for each $s \in [0, t]$,

$$\mathrm{DC}(Y_{s,k}) = \mathrm{DC}(Y^{11}_k, sY^{12}_k + \sqrt{1-s^2}Y^{22}_k)$$
$$\geq \mathrm{DC}(Y^{11}_k, sY^{12}_k) + \mathrm{DC}(Y^{11}_k)\mathrm{DC}(\sqrt{1-s^2}Y^{22}_k)$$
$$\geq (1-s^2)^{\dim(F_2)+1}\mathrm{DC}(Y^{11}_k)\mathrm{DC}(Y^{22}_k)$$
$$= (1-t^2)^{n+1}\mathrm{DC}(f_k(x_1), d_{x_1}f_k|_{T_{x_1}F_1})\mathrm{DC}(f_k(x_2), d_{x_2}f_k|_{T_{x_2}F_2}).$$

Since $(f_k)_{k \in \mathbb{N}}$ converges in law to f, and since $(f(x), d_x f)$ is non-degenerate for each $x \in M$ (and since B is compact), there exist $k_0 \in \mathbb{N}$ and a constant $c > 0$ such that for each $t \in [0, 1)$, $s \in [0, t]$, $x_1 \in F_1$ and $x_2 \in F_2$, as long as $k \geq k_0$,

$$\mathrm{DC}(Y_{s,k}) \geq (1-t^2)^{n+1}c.$$

In particular $|\Lambda_k(s; x_1, x_2)\gamma_{s,k;x_1,x_2}(0)| \leq Cc^{-1/2}(1 - t^2)^{-(n+1)/2}$. Hence the integrand in the right hand side of (5.2) is uniformly integrable, so the dominated convergence theorem applies.

All in all, letting $k \to \infty$ in both sides of (5.2) yields

$$(5.6) \quad \mathbb{P}[f_t \in \hat{A}_1 \times \hat{A}_2] - \mathbb{P}[f_0 \in \hat{A}_1 \times \hat{A}_2]$$

$$= \sum_{F_1 \in \mathcal{F}_1, \ F_2 \in \mathcal{F}_2} \int_0^t \int_{F_1 \times F_2} K(x_1, x_2) \times \Lambda(s; x_1, x_2)\gamma_{s;x_1,x_2}(0)\mathrm{dv}_{F_1}(x_1)\,\mathrm{dv}_{F_2}(x_2)\mathrm{d}s,$$

where Λ is defined in (5.5).

Let us now finish off the proof. Fix $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$. By the second point of Lemma 5.1, for each $i \in \{1, 2\}$, each $s \in [0, t]$ and each $x_i \in F_i$, a.s. (under the conditioning present in Λ)

$$\det(H_{x_i}^{F_i} f_s^i)\mathbb{1}_{\widetilde{\mathrm{Piv}}_{x_i}(\hat{A}_i)}(f_s^i) = \det(H_{x_i}^{F_i} f_s^i)\mathbb{1}_{\mathrm{Piv}_{x_i}(\hat{A}_i)}(f_s^i)$$

since the only place at which the two pivotal events do not coincide is where the determinant of the Hessian vanishes. Moreover, $\mathrm{Piv}_{x_i}(\hat{A}_i)$ splits as the disjoint union of $\mathrm{Piv}_{x_i}^+(\hat{A}_i)$ and $\mathrm{Piv}_{x_i}^-(\hat{A}_i)$. All in all, a.s. under the conditioning used in Λ,

$$\sigma(f_s^1, f_s^2)\mathbb{1}_{\mathrm{Piv}_{x_1}(\hat{A}_1) \times \mathrm{Piv}_{x_2}(\hat{A}_2)}(f_s^1, f_s^2)$$

$$= \left(\mathbb{1}_{\mathrm{Piv}_{x_1}^+(\hat{A}_1) \times \mathrm{Piv}_{x_2}^+(\hat{A}_2)}(f_s^1, f_s^2) + \mathbb{1}_{\mathrm{Piv}_{x_1}^-(\hat{A}_1) \times \mathrm{Piv}_{x_2}^-(\hat{A}_2)}(f_s^1, f_s^2)\right)$$

$$- \left(\mathbb{1}_{\mathrm{Piv}_{x_1}^+(\hat{A}_1) \times \mathrm{Piv}_{x_2}^-(\hat{A}_2)}(f_s^1, f_s^2) + \mathbb{1}_{\mathrm{Piv}_{x_1}^-(\hat{A}_1) \times \mathrm{Piv}_{x_2}^+(\hat{A}_2)}(f_s^1, f_s^2)\right).$$

In particular,

$$\Lambda(s; x_1, x_2)\gamma_{s;x_1,x_2}(0) = I_s^+(x_1, x_2) - I_s^-(x_1, x_2),$$

where $I_s^{\pm}$ are the signed pivotal intensity functions from Definition 2.11. Thus, (5.6) yields

$$\mathbb{P}[f_1 \in \hat{A}_1 \times \hat{A}_2] - \mathbb{P}[f_0 \in \hat{A}_1 \times \hat{A}_2]$$

$$= \sum_{F_1 \in \mathcal{F}_1, \ F_2 \in \mathcal{F}_2} \int_{F_1 \times F_2} K(x_1, x_2) \times \int_0^t \left(I_s^+(x_1, x_2) - I_s^-(x_1, x_2)\right) \mathrm{d}s\,\mathrm{dv}_{F_1}(x_1)\mathrm{dv}_{F_2}(x_2).$$

By the definition of $\hat{A}_1$ and $\hat{A}_2$ as well as $(f_t)_{t \in [0,1]}$,

$$\mathbb{P}[f_1 \in \hat{A}_1 \times \hat{A}_2] - \mathbb{P}[f_0 \in \hat{A}_1 \times \hat{A}_2] = \mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2],$$

so letting $t \to 1$ yields

$$\mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1]\mathbb{P}[A_2] = \int_{B_1 \times B_2} K(x_1, x_2)\left(\mathrm{d}\pi^+(x_1, x_2) - \mathrm{d}\pi^-(x_1, x_2)\right),$$

where $d\pi^\sigma(x_1, x_2)$, for $\sigma \in \{+, -\}$, are the signed pivotal measures from Definition 2.11. $\square$

To complete the section, we verify that Corollary 1.1 is indeed a special case of Theorem 2.14:

Proof of Corollary 1.1. Recall that B_1 and B_2 are closed boxes, F_0^1 and F_0^2 are their interiors, and F_j^i, for $j \in \{1, 2, 3, 4\}$ and $i \in \{1, 2\}$, are their four sides. Together with the corners of the boxes, which we denote F_j^i for $j \in \{5, 6, 7, 8\}$ and $i \in \{1, 2\}$, the set of F_j^i form a tame (and affine) stratification of B_1. Moreover, the events A_1 and A_2 are indeed topological events since stratified isotopies preserve crossings, and so Theorem 2.14 applies to these events, yielding an exact formula for $\mathrm{Cov}[A_1, A_2]$.

Let us next analyse the terms in this formula. By Remark 2.12, the corners do not contribute to the sum over strata. Further, by Remark 2.13, since A_i are both increasing events, the sets $\mathrm{Piv}_{x_i}^-(A_i)$ are empty and so the pivotal measure only contains positively pivotal events. Finally notice that, for each $i = 1, 2$, $t \in (0, 1)$ and $(x_1, x_2) \in (B_1, B_2)$, under the conditioning that x_1 and x_2 are stratified critical points at level 0 of f_t^1 and f_t^2 respectively, the fields f_t^1 and f_t^2

have a.s. no other critical points at level 0 (by the non-degeneracy assumption). Moreover, if the Hessians of f_t^i at x_i do not degenerate, the x_i are non-degenerate stratified critical points of f_t^i. The pivotal event for A_i is then equivalent to the existence of a path in $\{f_t^i \geq 0\}$ joining 'left' to 'right' and a path in $\{f_t^i \leq 0\}$ joining 'top' to 'bottom', both passing through x_i.   □

6. Proofs of the applications

In this section we give proofs for the applications that are discussed in Sections 1 and 2, in particular Theorem 2.15 and Corollaries 1.2, 1.6, 2.17, and 2.19.

6.1. Strong mixing for topological events.

Proof of Theorem 2.15. Let c denote a constant, that can change line-to-line, that depends only on d. By Theorem 2.14 and the definition of α_{top}, after replacing K with its absolute value, and dropping the condition that f_t^1 and f_t^2 lie in the pivotal sets, it suffices to show that, for all $t \in [0,1]$,

$$\gamma_t(x_1, x_2) \, \mathbb{E}_{t;x_1,x_2} \left[|\det(H_{x_1}^{F_1} f_t^1) \det(H_{x_2}^{F_2} f_t^2)| \right] \tag{6.1}$$

is bounded above by the maximum, over $i, j, k \in \{1, 2\}$, of

$$c \, \frac{\mathbb{E}\left[\|H_{x_i}^{F_i} f\|_{\text{op}}^2 \mid d_{x_i} f|_{F_i} = 0\right]^{d_i}}{\sqrt{\det(\Delta(x_1, x_2))}} \max\left\{1, \left(\frac{K(x_j, x_j) \det(d_{x_k} \otimes d_{x_k} K|_{F_i \times F_i})}{\sqrt{\det(\Delta(x_1, x_2))}}\right)^{2d_i}\right\}.$$

Let $\Delta_t(x_1, x_2)$ denote the covariance matrix of (2.2) and let $\Omega_t(x_1, x_2)$ denote the covariance matrix of $(d_{x_1} f_t^1|_{F_1}, d_{x_2} f_t^2|_{F_2})$. Applying Lemma A.4 to the matrices $X_1 = H_{x_1}^{F_1} f_t^1$ and $X_2 = H_{x_2}^{F_2} f_t^2$ and the vectors $Y = (f_t^1(x_1), f_t^2(x_2))$ and $Z = (d_{x_1} f_t^1|_{F_1}, d_{x_2} f_t^2|_{F_2})$, (6.1) is bounded above by the maximum, over $i, j, k \in \{1, 2\}$, of

$$c \, \frac{\mathbb{E}\left[\|H_{x_i}^{F_i} f_t^i\|_{\text{op}}^2 \mid d_{x_i} f_t^i|_{F_i} = 0\right]^{d_i}}{\sqrt{\det(\Delta_t(x_1, x_2))}} \max\left\{1, \left(\frac{\text{Var}[f_t^j(x_j)] \sqrt{\det(\Omega_t(x_1, x_2))}}{\sqrt{\det(\Delta_t(x_1, x_2))}}\right)^{2d_i}\right\}.$$

Recall that f_t^i is equal in law to f, and that, by Lemma A.5,

$$\det(\Delta_t(x_1, x_2)) \geq \det(\Delta_1(x_1, x_2)) = \det(\Delta(x_1, x_2))$$

and

$$\det(\Omega_t(x_1, x_2)) \leq \det(\Omega_0(x_1, x_2)) = \Pi_{k=1,2} \text{DC}[d_{x_k} f|_{F_k}] \leq \max_{k=1,2} \text{DC}[d_{x_k} f|_{F_k}]^2,$$

where $\text{DC}(X)$ is the determinant of the covariance of X in orthonormal coordinates. Since

$$\text{Var}[f(x_j)] = K(x_j, x_j) \quad \text{and} \quad \text{DC}[d_{x_k} f|_{F_k}] = \det(d_{x_k} \otimes d_{x_k} K|_{F_i \times F_i})^2,$$

we have the desired result.   □

Proof of Corollary 1.2. The non-degeneracy condition in the statement of Corollary 1.2 is equivalent to Condition 2.7, and so we are in the setting of Theorem 2.15. First we argue that there exists constants $c_1, c_2 > 0$, depending only on d, $\kappa(0)$, and the Hessian of κ at 0, such that, if

$$\max_{\alpha \in \mathbb{N}^d : |\alpha| \leq 2} \sup_{x_1 \in B_1, x_2 \in B_2} |\partial^\alpha \kappa(x_1 - x_2)| < c_1,$$

then for any affine sets F_1 and F_2 the covariance matrix of the Gaussian vector

$$(f(x_1), \nabla f|_{F_1}(x_1), f(x_2), \nabla f|_{F_2}(x_2)) \tag{6.2}$$

has a determinant bounded below by c_2. Let $\Sigma_1(x_1, x_2)$ denote this matrix, and observe that

$$\Sigma_1(x_1, x_2) = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}$$

where, by stationarity,

$$M_{ii} = \left[\begin{array}{cc} \kappa(0) & 0 \\ 0 & H_0^{F_i}\kappa \end{array} \right],$$

and M_{12} depends only on the value and second derivatives of κ at x_1 and x_2 (here H_0^F denotes the Hessian at the point x in an orthonormal basis of the linear span of F). The result then follows by the continuity, on the set of strictly positive-definite matrices, of the determinant with respect to the entry-wise sup-norm.

Combined with the stationarity of f, under the assumption that

$$\max_{\alpha\in\mathbb{N}^d:|\alpha|\leq 2} \sup_{x_1\in B_1,x_2\in B_2} |\partial^\alpha \kappa(x_1-x_2)| < c_1,$$

the quantity $c_{F_1,F_2}(x_1,x_2)$ in Theorem 2.15 can be bound above by

$$c \max_i \left\{ \mathbb{E}\big[\|H_0^{F_i}f\|_{\mathrm{op}}^2 \,|\, \nabla f|_{F_i}(0)=0\big]^{\dim(F_i)} \right\},$$

for some $c > 0$. Since this is a finite quantity, we have proved the result.

To verify the observation in Remark 1.3 note that we have already established that c_1 depends only on d, $\kappa(0)$, and the Hessian of κ at 0. Next, since all norms on $\mathbb{R}^d$ are equivalent,

$$\mathbb{E}\big[\|H_0^{F_i}f\|_{\mathrm{op}}^2 \,|\, \nabla f|_{F_i}(0)=0\big] \leq c_d \max_{j_1,j_2} \mathbb{E}\big[(H_0^{F_1,F_2}f)_{j_1,j_2}^2 \,|\, \nabla f|_{F_i}(0)=0\big],$$

where c_d is a constant depending only on the dimension d. By stationarity, and since conditioning on part of a Gaussian vector reduces the variance of all coordinates, this is at most

$$c_d \max_{j_1,j_2} \mathbb{E}\big[(H_0^{F_1,F_2}f)_{j_1,j_2}^2\big] \leq c_d \max_{j_1,j_2} \frac{\partial^4 \kappa(0)}{\partial x_{j_1}^2 \partial x_{j_2}^2}.$$

Finally, applying the Cauchy–Schwartz inequality in Fourier space,

$$\max_{j_1,j_2} \frac{\partial^4 \kappa(0)}{\partial x_{j_1}^2 \partial x_{j_2}^2} \leq \max_j \frac{\partial^4 \kappa(0)}{\partial x_j^4}$$

and we have the result. $\qquad\square$

Proof of Corollary 2.17. Observe that each of $f_n|_{B_1\cup B_2}$ satisfies Condition 2.7, since $B_1\cup B_2$ does not include antipodal points. Note also that a condition analogous to (1.4) holds; more precisely, as $n\to\infty$,

$$\sup_{x_1\in B_1,x_2\in B_2} \|(K_n(x_1,x_2),d_{x_1}K_n(x_1,x_2),d_{x_2}K_n(x_1,x_2),d_{x_1}\otimes d_{x_2}K_n(x_1,x_2))\|_\infty \to 0,$$

which, as in the proof of Theorem 2.15, implies that, as $n\to\infty$,

$$\sup_{x_1\in F_1,x_2\in F_2} \left| \frac{\det(\Delta_1(x_1,x_2))}{\det(\Delta_0(x_1,x_2))} - 1 \right| \to 0,$$

where $\Delta_t(x_1,x_2)$ is the covariance matrix, for the field f_n, that is defined in the proof of Theorem 2.15 (note that we have omitted the dependence on n in the notation). Observe also that the scale $s_n = 1/\sqrt{n}$ at which the Kostlan ensemble converges to a local limit is a polynomial, and so all derivatives of K_n on the diagonal (x,x) grow at most polynomially, uniformly over x. Along with the discussion in Remark 2.16, we deduce that

$$\sup_{F_1\in\mathcal{F}_1,F_2\in\mathcal{F}_2} \sup_{x_1\in F_1,x_2\in F_2} c_{F_1,F_2}(x_1,x_2)$$

grows at most polynomially as a function of n, where $c_{F_1,F_2}(x_1,x_2)$ is the constant appearing in Theorem 2.15 applied to f_n (again we omit the dependence on n in the notation). Since on the other hand

$$\sup_{x_1\in B_1,x_2\in B_2} |K_n(x_1,x_2)|$$

decays exponentially in n (recall that B_1 and B_2 are contained within an open hemisphere), we deduce the result from Theorem 2.15. $\qquad\square$

6.2. **Lower concentration for topological counts.** Our proof of the lower concentration results in Corollaries 1.6 and 2.19 essentially follows the proof of [44, Theorem 1.4]. We make use of the following simple lemma:

Lemma 6.1. *Let $B_i \subset M$ be a sequence of disjoint stratified sets, and let $A_i \in \sigma_{top}(B_i)$ be topological events such that $\sup_i \mathbb{P}[A_i] < h < 1$. Then*

$$|\mathbb{P}[\cap_i A_i] - \Pi_i \mathbb{P}[A_i]| \leq \frac{1}{1-h} \sup_{n \in \mathbb{N}} \alpha_{top}(B_n, \cup_{j>n} B_j).$$

Proof. By the definition of α_{top}

$$|\mathbb{P}(A_n \cap (\cap_{j>n} A_j)) - \mathbb{P}(A_n)\mathbb{P}(\cap_{j>n} A_j)| \leq \alpha_{\text{top}}(B_n, \cup_{j>n} B_j).$$

Iterating this inequality for $n = 1, 2, \ldots$ and using that $\mathbb{P}(A_n) < h$ we get the upper bound

$$(1 + h + h^2 + \ldots) \times \sup_{n \in \mathbb{N}} \alpha_{\text{top}}(B_n, \cup_{j>n} B_j),$$

which is equal to the desired upped bound. $\qquad\square$

Proof of Corollary 1.6. First observe that we may assume the infimum in (1.8) is eventually attained in the set $r \in [g_s, s/g_s]$ for some function $g_s \to \infty$ as $s \to \infty$, since otherwise the right-hand side of (1.8) is bounded from below, and we may then choose $c_1 > 0$ large enough so that (1.8) holds trivially.

Fix $\varepsilon > 0$, a function $g_s \in (0, \sqrt{s})$ such that $g_s \to \infty$, and a mesoscopic parameter $g_s < r < s/g_s$. Let B_0 denote the unit cube, considered as a stratified set via its collection of generalised faces of all dimensions. Consider placing $(\text{Vol}(B) + o(1))(s/r)^d$ disjoint translations of rB_0 inside sB, and let M_i denote these mesoscopic cubes. By the super-additivity of N, if $N(sB)/s^d \leq (c_N - \varepsilon)\text{Vol}(B)$, then there exist at least $\varepsilon\text{Vol}(B)/(2c_N - \varepsilon) \times (s/r)^d$ mesoscopic cubes M_i such that $N(M_i)/r^d \leq c_N - \varepsilon/2$. If this holds, then we can also find at least $\nu(s/r)^d$ mesoscopic cubes M_i with this property that are separated by a distance r, where

$$\nu = c_d \times \frac{\varepsilon\text{Vol}(B)}{2c_N - \varepsilon},$$

and $c_d > 0$ is a constant that depends only on the dimension.

By stationarity and the law of large numbers (1.7), for each $h \in (0,1)$

$$\mathbb{P}[N(M_i)/r^d \leq c_N - \varepsilon/2] < h$$

eventually as $s \to \infty$. Hence, by Lemma 6.1, for every choice of $\nu(s/r)^d$ mesoscopic cubes M_i which are r-separated, the probability that $N(M_i)/r^d \leq c_N - \varepsilon/2$ for all of them is at most

$$h^{\nu(s/r)^d} + \frac{1}{1-h}\alpha_{r,s},$$

where $\alpha_{r,s} = \max_n \alpha_{\text{top}}(M_n, \cup_{i>n} M_i)$. There are at most $2^{(\text{Vol}(B)+o(1))(s/r)^d}$ ways to choose cubes M_i, hence by the union bound

$$\mathbb{P}\big[N(sB)/s^d \leq (c_N - \varepsilon)\text{Vol}(B)\big] \leq 2^{(\text{Vol}(B)+o(1))(s/r)^d}\Big(h^{\nu(s/r)^d} + \frac{1}{1-h}\alpha_{r,s}\Big).$$

Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be the standard stratifications of M_n and $\cup_{i>n} M_i$. Clearly, there is a constant $c > 0$ which depends on the dimension only, such that $|\mathcal{F}_1| \leq c$ and $|\mathcal{F}_2| \leq c\nu(s/r)^d$. In both stratifications the strata with the largest volume are interiors of mesoscopic cubes that have volume r^d. By Corollary 1.2 this implies that there is a constant $c_1 > 1$ such that

$$\alpha_{r,s} \leq c_1 r^d s^d \bar{\kappa}(r).$$

Combining these estimates, taking $c_B = 2 \log 2 \mathrm{Vol}(B)$ and choosing h small enough, we see that for every $C > 0$

$$\mathbb{P}\big[N(sB)/s^d \leq (c_N - \varepsilon)\mathrm{Vol}(B)\big] \leq c_1\big(e^{-C(s/r)^d} + e^{c_B(s/r)^d}r^d s^d \bar{\kappa}(r)\big),$$

provided s is large enough. This proves the result for $r \in [g_s, s/g_s]$. As mentioned in the very beginning of the proof, by choosing sufficiently large c_1 we can extend the estimate to all $r \in [0, s]$. $\qquad\square$

Proof of Corollary 2.19. This follows closely the proof of Corollary 1.6. We first treat the case that B is contained in an open hemisphere. Equip $\mathbb{S}^d$ with a marked pole x_0, and for $r \in (0, 1]$ let rB_0 denote the symmetric spherical cap centred at x_0 with volume r^d, considered as a stratified set via the stratification $\mathcal{F}_B = \{\mathrm{int}(B_0), \partial B_0\}$. Fix $\varepsilon > 0$ and a function g_n such that $g_n \to \infty$ as $n \to \infty$. Define a mesoscopic scale $g_n/\sqrt{n} < r < 1/g_n$ and consider placing $(\mathrm{Vol}(B) + o(1))/r^d$ disjoint copies of rB_0 inside B . Following exactly the proof of Corollary 1.6, we deduce the existence of $c_1, c_2 > 0$ such that

$$\mathbb{P}[N(B)/n^{d/2} \leq (c_N - \varepsilon)\mathrm{Vol}(B)] \leq c_1(e^{-c_2 r^{-d}} + \alpha_{r,n}),$$

where $\alpha_{r,n}$ denotes the supremum of α-mixing coefficients $\alpha_{\mathrm{top}}(B_1, B_2)$ among all pairs of disjoint stratified sets B_1 and B_2 contained in B and separated by a distance at least r. By Theorem 2.15 (see also the proof of Corollary 2.17), there exist $k, c_3, c_4 > 0$ such that

$$\alpha_{r,n} \leq c_3 n^k e^{-c_4 r^2 n}.$$

Setting $r = n^{-1/(2+d)}$ yields the desired bound.

In the general case, we simply choose a finite number of disjoint stratified sets B_i that are each contained within an open hemisphere. Since by super-additivity $N(B)/n^{d/2} \leq c_N\mathrm{Vol}(B) - \varepsilon$ implies that $N(B_i)/n^{d/2} \leq c_N\mathrm{Vol}(B_i) - \varepsilon/k$ for some B_i, the argument goes through in this case as well. $\qquad\square$

6.3. **Decorrelation for topological counts.** Corollary 2.20 is a direct consequence of the following general result, applied to the random variables $X = N(B_1)$ and $Y = N(B_2)$:

Proposition 6.2 (See [24, Theorem 17.2.2]). *Let X and Y be random variables and define the α-mixing coefficient associated to their σ-algebras*

$$\alpha(X, Y) = \sup_{A \in \sigma(X), B \in \sigma(Y)} |\mathbb{P}[A \cap B] - \mathbb{P}[A]\mathbb{P}[B]|.$$

Suppose further that

$$\mathbb{E}[X^{2+\delta}] < c \quad and \quad \mathbb{E}[Y^{2+\delta}] < c$$

for positive constants $\delta, c > 0$. Then

$$|\mathrm{Cov}(X, Y)| \leq 8c^{2/(2+\delta)}\alpha(X, Y)^{2/(2+\delta)}.$$

Appendix A. Gaussian computations

In this section we gather results about finite-dimensional Gaussian vectors. If X is a Gaussian vector in finite-dimensional vector space equipped with a scalar product, let $\mathrm{DC}(X)$ be the determinant of its covariance.

Lemma A.1. *Let X, Y be jointly Gaussian vectors such that Y is non-degenerate. Then $\mathrm{DC}(X|Y)$ does not depend on Y, and*

$$\mathrm{DC}(X, Y) = \mathrm{DC}(X|Y)\mathrm{DC}(Y).$$

Proof. This is an easy consequence of the Gaussian regression formula ([3, Proposition 1.2]). $\quad\square$

Lemma A.2. *Given two independent Gaussian vectors X and Y of the same dimension,*

$$\mathrm{DC}(X + Y) \geq \mathrm{DC}(X) + \mathrm{DC}(Y).$$

Proof. In terms of covariance matrices, this amounts to saying that given A, B two symmetric non-negative matrices of the same size,

$$\det(A + B) \geq \det(A) + \det(B),$$

which follows from the Minkowski inequality. $\qquad\square$

Lemma A.3. *Let X, Y, Z be jointly Gaussian vectors such that Y and Z have the same dimension and Z is independent of (X, Y). Then*

$$\mathrm{DC}(X, Y + Z) \geq \mathrm{DC}(X, Y) + \mathrm{DC}(X)\mathrm{DC}(Z).$$

Proof. Let us assume that X is non-degenerate; the general case follows by continuity. By Lemma A.1

$$\mathrm{DC}(X, Y + Z) = \mathrm{DC}(X)\mathrm{DC}(Y + Z | X).$$

Applying Lemma A.2,

$$\mathrm{DC}(Y + Z | X) \geq \mathrm{DC}(Y | X) + \mathrm{DC}(Z | X) = \mathrm{DC}(Y | X) + \mathrm{DC}(Z),$$

with the final equality since Z is independent of X. Hence

$$\mathrm{DC}(X, Y + Z) \geq \mathrm{DC}(X)\mathrm{DC}(Y | X) + \mathrm{DC}(X)\mathrm{DC}(Z) = \mathrm{DC}(X, Y) + \mathrm{DC}(X)\mathrm{DC}(Z),$$

where the equality holds by Lemma A.1. $\qquad\square$

Lemma A.4. *Let X_1 and X_2 be respectively $d_1 \times d_1$ and $d_2 \times d_2$ random matrices, and let $Y = (Y_1, Y_2) \in \mathbb{R}^2$ and $Z \in \mathbb{R}^{d_1 + d_2}$ be random vectors. Suppose that (Y, Z) is a non-degenerate Gaussian vector and, conditionally on $Z = 0$, X_1 and X_2 have entries that are jointly Gaussian with Y. Let $\varphi_{Y,Z}$ denote the density of (Y, Z). Then there exists a constant $c > 0$, depending only on d_1 and d_2, such that*

(A.1)
$$\varphi_{Y,Z}(0)\, \mathbb{E}[|\det(X_1)\det(X_2)| \,|\, Y = 0, Z = 0]$$

is bounded above by the maximum, over $i \in \{1, 2\}$, of

$$\frac{c\, \mathbb{E}\big[\|X_i\|_{op}^2 \,|\, Z = 0\big]^{d_i}}{\sqrt{\mathrm{DC}(Y, Z)}} \max\left\{1, \Big(\frac{\max_k \mathrm{Var}[Y_k]\, \sqrt{\mathrm{DC}(Z)}}{\sqrt{\mathrm{DC}(Y, Z)}}\Big)^{2d_i}\right\},$$

where $\|\cdot\|_{op}$ denotes the (L^2-)operator norm.

Proof. Let c denote a positive constant, depending only on d_1 and d_2, that may change from line to line. In the proof we use repeated the fact that conditioning on part of a Gaussian vector reduces the variance of all coordinates. By the Cauchy-Schwarz inequality and an elementary bound on the determinant, (A.1) is bounded above by

$$c\, \varphi_{Y,Z}(0) \max_{i, j_1, j_2} \mathbb{E}\big[(X_i)_{j_1, j_2}^{2d_i} \,|\, Y = 0, Z = 0\big].$$

Since a normally distributed random variable $Z \sim \mathcal{N}(\mu, \sigma^2)$ satisfies

$$\mathbb{E}[Z^{2d_i}] \leq c \max\{\mu^{2d_i}, \sigma^{2d_i}\},$$

and since the variance of a random variable is less than its second moment,

$$\mathbb{E}[(X_i)_{j_1, j_2}^{2d_i} \,|\, Y = 0, Z = 0] \leq c \max\left\{\mathbb{E}\big[(X_i)_{j_1, j_2}^2 | Z = 0\big]^{d_i}, \mathbb{E}\big[(X_i)_{j_1, j_2} \,|\, Y = 0, Z = 0\big]^{2d_i}\right\}.$$

Let $\Sigma_{Y|Z}$ and Σ_Z denote the covariance matrices of $Y|Z$ and Z respectively. By conditioning on $Z = 0$ and applying Lemma A.1, we have that

$$\varphi_{Y,Z}(0) = c\frac{e^{-\frac{1}{2}\mathbb{E}[Y|Z=0]^T\Sigma_{Y|Z}^{-1}\mathbb{E}[Y|Z=0]}}{\sqrt{\mathrm{DC}(Y|Z)}}\frac{e^{-\frac{1}{2}\mathbb{E}[Z]^T\Sigma_Z^{-1}\mathbb{E}[Z]}}{\sqrt{\mathrm{DC}(Z)}}$$

$$\leq c\frac{e^{-\frac{1}{2}\mathbb{E}[Y|Z=0]^T\Sigma_{Y|Z}^{-1}\mathbb{E}[Y|Z=0]}}{\sqrt{\mathrm{DC}(Y,Z)}} \leq \frac{c}{\sqrt{\mathrm{DC}(Y,Z)}}.$$

Since, moreover,

$$\max_{j_1,j_2}\mathbb{E}\big[(X_i)_{j_1,j_2}^2 \mid Z = 0\big] \leq \mathbb{E}[\|X_i\|_{\mathrm{op}}^2 \mid Z = 0],$$

it suffices to show that

$$\sup_{\ell\in\mathbb{R}^2}\Big\{\mathbb{E}\big[(X_i)_{j_1,j_2} \mid Y - E[Y|Z=0] = \ell, Z = 0\big]^{2d_i}e^{-\frac{1}{2}\ell^T\Sigma_{Y|Z}^{-1}\ell}\Big\}$$

is bounded above by

$$c\,\mathbb{E}\big[(X_i)_{j_1,j_2}^2 \mid Z = 0\big]^{d_i}\max\Big\{1, \Big(\frac{\max_k\mathrm{Var}\big[Y_k\big]\sqrt{\mathrm{DC}(Z)}}{\sqrt{\mathrm{DC}(Y,Z)}}\Big)^{2d_i}\Big\}.$$

To show this, decompose $\Sigma_{Y|Z}^{-1} = U^T\Lambda^{-1}U$, where $U = (u_{k_1,k_2})$ is a 2×2 orthogonal matrix and $\Lambda = \mathrm{Diag}(\lambda_k)$ is the 2×2 diagonal matrix of (positive) eigenvalues of $\Sigma_{Y|Z}$. Abbreviating $W = (w_k) := U\mathrm{Cov}[(X_i)_{j_1,j_2}Y_k|Z=0]$ and replacing ℓ by $U\ell$, by the Gaussian regression formula ([3, Proposition 1.2]) we have that

$$\sup_{\ell}\Big\{\mathbb{E}\big[(X_i)_{j_1,j_2} \mid Y - \mathbb{E}[Y|Z=0] = \ell, Z = 0\big]^{2d_i}e^{-\frac{1}{2}\ell^T\Sigma_{Y|Z}^{-1}\ell}\Big\}$$

$$= \sup_{\ell}\Big\{\big(\mathbb{E}[(X_i)_{j_1,j_2} \mid Z = 0] + W^T\Lambda^{-1}\ell\big)^{2d_i}e^{-\frac{1}{2}\ell^T\Lambda^{-1}\ell}\Big\}$$

$$\leq c\max\Big\{\sup_{\ell}\Big\{\mathbb{E}[(X_i)_{j_1,j_2} \mid Z = 0]^{2d_i}e^{-\frac{1}{2}\ell^T\Lambda^{-1}\ell}\Big\}, \sup_{\ell}\Big\{\big(W^T\Lambda^{-1}\ell\big)^{2d_i}e^{-\frac{1}{2}\ell^T\Lambda^{-1}\ell}\Big\}\Big\}$$

$$\leq c\max\Big\{\mathbb{E}[(X_i)_{j_1,j_2}^2 \mid Z = 0]^{d_i}, \sup_{\ell}\Big\{\big(W^T\Lambda^{-1}\ell\big)^{2d_i}e^{-\frac{1}{2}\ell^T\Lambda^{-1}\ell}\Big\}\Big\}.$$

Differentiating in ℓ, the maxima of the expression on the right is attained at

$$\ell = (\ell_1, \ell_1) = \begin{cases} \frac{\pm\sqrt{2d_i}}{\sqrt{w_1^2\lambda_1^{-1}+w_2^2\lambda_2^{-1}}}(w_1, w_2), & (w_1, w_2) \neq (0,0), \\ (0,0), & (w_1, w_2) = (0,0), \end{cases}$$

and yields a maximum value of

$$(2d_i/e)^{d_i}(w_1^2\lambda_1^{-1} + w_2^2\lambda_2^{-1})^{d_i} \leq c\Big(\max_k w_k^2\max_k\lambda_k^{-1}\Big)^{d_i}.$$

Since the eigenvalues of a positive-definite real symmetric matrix are bounded by a constant times the maximum diagonal entry,

$$\max_k\lambda_k^{-1} = \frac{\max_k\lambda_k}{\det(\Lambda)} \leq c\frac{\max_k\mathrm{Var}\big[Y_k \mid Z = 0\big]}{\det(\Sigma_{Y|Z})} \leq c\frac{\max_k\mathrm{Var}\big[Y_k\big]}{\det(\Sigma_{Y|Z})} = c\frac{\max_k\mathrm{Var}\big[Y_k\big]\mathrm{DC}(Z)}{\mathrm{DC}(Y,Z)},$$

where in the last step we used Lemma A.1. Moreover, since U has entries bounded above in absolute value by one (being orthogonal), and by the Cauchy-Schwarz inequality,

$$\max_k|w_k| \leq c\max_k|\mathrm{Cov}[(X_i)_{j_1,j_2}Y_k \mid Z = 0]| \leq c\,\mathbb{E}[(X_i)_{j_1,j_2}^2 \mid Z = 0]^{1/2}\max_k\mathrm{Var}\big[Y_k\big]^{1/2}.$$

Combining we have the result. $\qquad\square$

Lemma A.5. *Let (Y_1, Y_2) denote a $(d_1 + d_2)$-dimensional non-degenerate Gaussian vector. For each $t \in [0, 1]$, define $Y^t = (Y_1, tY_2 + \sqrt{1 - t^2}\tilde{Y}_2)$ where $\tilde{Y}_2$ is a copy of Y_2 independent of (Y_1, Y_2). Then*

$$DC(Y^1) \le DC(Y^t) \le DC(Y^0).$$

Proof. Observe that $DC(Y^t)$ has the block form

$$\begin{bmatrix} A & tB \\ tB^T & C \end{bmatrix},$$

where A and C are (strictly) positive-definite. Since A is positive-definite and $BC^{-1}B^T$ is symmetric and positive-definite, there exists a P such that

$$A = P^T P \quad \text{and} \quad BC^{-1}B^T = P^T DP,$$

where $D = \text{Diag}((d_i)_i)$ is a positive diagonal matrix. Hence

$$DC(Y^t) = \det(C) \det(A - t^2 BC^{-1}B^T) = \det(C) \det(P)^2 \, \Pi_i(1 - t^2 d_i)$$

which, since $d_i > 0$, is decreasing in $t \in [0, 1]$. $\square$

APPENDIX B. PROOF OF PITERBARG'S FORMULA

In the proof of Piterbarg's formula, we will use the classical fact that the density function $\varphi(x; \Sigma)$ of a (possibly non-centred) Gaussian vector with covariance Σ satisfies

(B.1) $\dfrac{1}{2}\dfrac{\partial^2}{\partial x_i^2}\varphi(x; \Sigma) = \dfrac{\partial}{\partial \Sigma_{ii}}\varphi(x; \Sigma)$ and $\dfrac{\partial^2}{\partial x_i \partial x_j}\varphi(x; \Sigma) = \dfrac{\partial}{\partial \Sigma_{ij}}\varphi(x; \Sigma), \; i \ne j.$

Proof of Lemma 2.22. Let $(f_i)_{i \ge 1}$ and $(g_i)_{i \ge 1}$ be sequences of smooth compactly supported functions on $\mathbb{R}^m$ that converge to $\mathbb{1}_A$ and $\mathbb{1}_B$ in the sense of tempered distributions. Following the proof of [39, Theorem 1.4], by writing the derivative with respect to t in terms of derivatives with respect to the elements of the covariant matrix, and then by using the identity (B.1) and integrating by parts, we obtain

$$\frac{d}{dt}\mathbb{E}\left[f_i(X_t)g_i(Y_t)\right] = \sum_{k=1}^m \int_{\mathbb{R}^{2m}} \partial_{x_k} f_i(x) \partial_{y_k} g_i(y) \, \gamma_t(x, y) \mathrm{d}x \, \mathrm{d}y$$

$$= \sum_{k=1}^m \int_{\mathbb{R}^{2m}} f_i(x) g_i(y) \, \partial_{x_k} \partial_{y_k} \gamma_t(x, y) \mathrm{d}x \, \mathrm{d}y;$$

(all the other terms disappear since the only covariances that depend on t are $\text{Cov}(X_{t,k}, Y_{t,k}) = t$). Passing to the limit as $i \to \infty$ gives that

$$\frac{d}{dt}\mathbb{P}\left[Z_t \in A \times B\right] = \frac{d}{dt}\mathbb{E}\left[\mathbb{1}_A(X_t)\mathbb{1}_B(Y_t)\right] = \sum_{k=1}^m \int_{\mathbb{R}^{2m}} \mathbb{1}_A(x)\mathbb{1}_B(y)\partial_{x_k}\partial_{y_k}\gamma_t(x, y)\mathrm{d}x \, \mathrm{d}y.$$

By Gauss's theorem, applied both in the x_k and in the y_k variables we have

$$\int_{\mathbb{R}^{2m}} \mathbb{1}_A(x)\mathbb{1}_B(y)\partial_{x_k}\partial_{y_k}\gamma_t(x, y)\mathrm{d}x \, \mathrm{d}y = \int_{\partial A \times \partial B} \nu_A(x)_k \nu_B(y)_k \gamma_t(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

where $\nu_A(x)_k$ is the k-th component of $\nu_A(x)$, and $\mathrm{d}x$ on the right-hand side of the equation is the volume element on ∂_A (and similarly for B and y). Since $\sum_{k=1}^m (\nu_A(x))_k (\nu_B(y))_k = \langle \nu_A(x), \nu_B(y) \rangle$,

$$\frac{d}{dt}\mathbb{P}\left[Z_t \in A \times B\right] = \int_{\partial A \times \partial B} \langle \nu_A(x), \nu_B(y) \rangle \, \gamma_t(x, y)\mathrm{d}x \, \mathrm{d}y,$$

which proves the first part of the statement.

To prove the last part of the lemma let us consider X to be a translation of the standard Gaussian vector by μ. Let Y be an independent copy of X. We can define $X_t = X$ and $Y_t = t(X - \mu) + \sqrt{1 - t^2}(Y - \mu) + \mu$. It is easy to see that these vectors satisfy the assumptions in the first part of the lemma. Note that in this case $Z_0 = (X, Y)$ and $Z_1 = (X, X)$. Integrating with respect to t from 0 to 1 we have

$$\int_0^1 \int_{\partial A \times \partial B} \langle \nu_A(x), \nu_B(y) \rangle \gamma_t(x, y) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t = \mathbb{P}[(X, X) \in A \times B] - \mathbb{P}[(X, Y) \in A \times B].$$

Since $\mathbb{P}[(X, X) \in A \times B] = \mathbb{P}[X \in A \cap B]$ and $\mathbb{P}[(X, Y) \in A \times B] = \mathbb{P}[X \in A]\mathbb{P}[X \in B]$, this proves the second part of the statement. $\qquad\square$

Appendix C. On stratified critical points

Here we prove two elementary lemmas about stratified critical points. Recall that M is a smooth manifold and $(B, \mathcal{F})$ is a stratified set of M.

Lemma C.1. *Let $(u_k, x_k)_{k \in \mathbb{N}}$ be a sequence in $C^1(M) \times B$ converging to a limit $(u, x) \in C^1(M) \times B$. Assume that, for each $k \in \mathbb{N}$, x_k is a stratified critical point of u_k. Then x is a stratified critical point of u.*

Lemma C.1 implies that the discriminant $\mathfrak{D}_B$ is C^1-closed. Moreover, taking $u_k = u$ for all k, it implies that the set of stratified critical points of u in B is compact.

Proof of Lemma C.1. Without loss of generality, we may assume that there exist $F, F' \in \mathcal{F}$ such that $x_k \in F$ for each $k \in \mathbb{N}$ and $x \in F'$. If $F' = F$ then the sequence $(u_k|_F)_k$ converges to $u|_F$ in C^1 so $d_x u|_F = 0$ and x is a stratified critical point of u. Otherwise, $F' < F$ and $d_x u$ vanishes on $T_x F|_{F'}$ which contains $T_x F'$, so x is a critical point of $u|_{F'}$. $\qquad\square$

Lemma C.2. *Let $u \in C^2(M)$ and let $x \in B$ be a non-degenerate stratified critical point of u. Then x is isolated in the set of stratified critical points in B.*

Lemma C.2 shows that Definition 3.1 is the natural definition of non-degenerate critical points in the setting of stratified sets.

Proof of Lemma C.2. Let x be a stratified critical point of u belonging to $F' \in \mathcal{F}$. Assume that there exists a sequence $(x_k)_{k \in \mathbb{N}}$ of stratified critical points of u distinct from x converging to x; let us show that x is degenerate. Without loss of generality, we may assume that there exists $F \in \mathcal{F}$ such that, for each $k \in \mathbb{N}$, $x_k \in F$. If $F' = F$ then x is a degenerate critical point of $u|_{F'}$. Otherwise, $F' < F$ and $d_x u$ vanishes on $T_x F|_{F'}$, in which case x is a degenerate stratified critical point of u. $\qquad\square$

References

[1] K.S. Alexander. Boundedness of level lines for two-dimensional random fields. *Ann. Probab.*, 24(4):1653–1674, 1996.

[2] F. Aurzada and T. Simon. Persistence probabilities and exponents. In L.N. Andersen et. al., editor, *Lévy Matters V*, pages 183–224. Springer, 2015.

[3] J.M. Azaïs and M. Wschebor. *Level Sets and Extrema of Random Processes and Fields*. John Wiley & Sons, Inc., Hoboken, NJ, 2009.

[4] R. Basu, A. Dembo, N. Feldheim, and O. Zeitouni. Exponential concentration for zeroes of stationary Gaussian processes. *?Int. Math. Res. Not.*, to appear.

[5] V. Beffara and D. Gayet. Percolation of random nodal lines. *Publ. Math. IHES*, 126:131–176, 2017.

[6] V. Beffara and D. Gayet. Percolation without FKG. *arXiv preprint, arXiv:1710.10644*, 2017.

[7] D. Beliaev, M. McAuley, and S. Muirhead. On the number of excursion sets of planar Gaussian fields. *Probab. Theory Related Fields*, to appear.

[8] D. Beliaev and S. Muirhead. Discretisation schemes for level sets of planar Gaussian fields. *Commun. Math. Phys.*, 359:869–913, 2018.

[9] D. Beliaev, S. Muirhead, and I. Wigman. Russo-Seymour-Welsh estimates for the Kostlan ensemble of random polynomials. *arXiv preprint, arXiv:1709.08961*, 2017.

[10] D. Beliaev and I. Wigman. Volume distribution of nodal domains of random band-limited functions. *Probab. Theory Related Fields*, 172(1–2):453–492, 2018.

[11] E. Bogomolny and C. Schmit. Random wavefunctions and percolation. *J. Phys. A: Math. Theor.*, 40:14033–14043, 2007.

[12] B. Bollobás and O. Riordan. *Percolation.* Cambridge: Cambridge University Press, 2006.

[13] R.C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144, 2005.

[14] Y. Canzani and P. Sarnak. On the topology of the zero sets of monochromatic random waves. *arXiv preprint, arXiv:1412.4437*, 2014.

[15] S. Chatterjee. Chaos, concentration, and multiple valleys. *arXiv preprint, arXiv:0810.4221*, 2008.

[16] A. Dembo and S. Mukherjee. No zero-crossings for random polynomials and the heat equation. *Ann. Probab.*, 43(1):85–118, 2015.

[17] P. Doukhan. *Mixing: Properties and Examples.* Springer-Verlag New York, 1994.

[18] R. D. Edwards and R. C. Kirby. Deformations of spaces of imbeddings. *Ann. Math. (2)*, 93:63–88, 1971.

[19] M. Ekström. A general central limit theorem for strong mixing sequences. *Stat. Probab. Lett.*, 94(3):236–238, 2014.

[20] N. Feldheim, O. Feldheim, and S. Nitzan. Persistence of Gaussian stationary processes: a spectral perspective. *arXiv preprint, arXiv:1709.00204*, 2017.

[21] D. Gayet and J.Y. Welschinger. Lower estimates for the expected Betti numbers of random real hypersurfaces. *J. London Math. Soc. (2)*, 90(1):105–120, 2014.

[22] M. Golubitsky and V. Guillemin. *Stable mappings and their singularities. 2nd corr. printing.*, volume 14. Springer, New York, NY, 1980.

[23] M. Goresky and R. MacPherson. *Stratified Morse Theory*, volume 14. Berlin etc.: Springer-Verlag, 1988.

[24] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[25] I.A. Ibragimov and Y.V. Linnik. *Independent and stationary sequences of random variables.* Wolters-Noordhoff, 1971.

[26] M. Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 18:29–35, 1943.

[27] H. Kesten. Scaling relations for 2D-percolation. *Commun. Math. Phys.*, 109(1):109–156, 1987.

[28] A. N. Kolmogorov and Ju. A. Rozanov. On a strong mixing condition for stationary Gaussian processes. *Teor. Verojatnost. i Primenen.*, 5:222–227, 1960.

[29] E. Kostlan. On the distribution of roots of random polynomials. In *From Topology to Computation: Proceedings of the Smalefest*, pages 419–431. Springer, 1993.

[30] E. Kostlan. On the expected number of real roots of a system of random polynomial equations. In *Foundations of computational mathematics: Proceedings of the Smalefest 2000*, pages 149–188. World Sci. Publishing, River Edge, NJ, 2002.

[31] P. Kurlberg and I. Wigman. Variation of the Nazarov-Sodin constant for random plane waves and arithmetic random waves. *Adv. Math.*, 330:516–552, 2018.

[32] Z. Lin and C. Lu. *Limit Theory for Mixing Dependent Random Variables.* Springer Netherlands, 1996.

[33] R.M. Loynes. Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.*, 36(3):993–999, 1965.

[34] J.N. Mather. Stratifications and mappings. Dynamical Syst., Proc. Sympos. Univ. Bahia, Salvador 1971, 195-232 (1973)., 1973.

[35] S.A. Molchanov and A.K. Stepanov. Percolation in random fields. II. *Theor. Math. Phys.*, 55(3):592–599, 1983.

[36] S. Muirhead and H. Vanneuville. The sharp phase transition for level set percolation of smooth planar Gaussian fields. *Ann. Inst. Henri Poincaré Probab. Stat.*, to appear.

[37] F. Nazarov and M. Sodin. Fluctuations in random complex zeroes: asymptotic normality revisited. *Int. Math. Res. Not.*, 2011(24):720–5759, 2011.

[38] 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. Geo.*, 12(3):205–278, 2016.

[39] L.I. Nicolaescu. *An invitation to Morse theory. 2nd ed.* Berlin: Springer, 2nd ed. edition, 2011.

[40] V.I. Piterbarg. *Asymptotic Methods in the Theory of Gaussian Processes and Fields.* Providence, RI: AMS, 1996.

[41] L.D. Pitt. Positively correlated normal variables are associated. *Ann. Probab.*, 10(2):496–499, 1982.

[42] S.O. Rice. Mathematical analysis of random noise. *Bell Sys. Tech. Jour.*, 24:46–156, 1945.

[43] A. Rivera. Hole probability for nodal sets of the cut-off Gaussian free field. *Adv. Math.*, 319:1–39, 2017.

[44] A. Rivera and H. Vanneuville. Quasi-independence for nodal lines. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(3):1679–1711, 2019.

[45] A. Rivera and H. Vanneuville. The critical threshold for Bargmann-Fock percolation. *Ann. Henri Lebesgue*, to appear.

[46] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. U.S.A.*, 42(1):43–47, 1956.

[47] P. Sarnak and I. Wigman. Topologies of nodal sets of random band limited functions. In D. Jiang, F. Shahidi, and D. Soudry, editors, *Advances in the Theory of Automorphic Forms and Their L-functions*, pages 351–365. American Mathematical Society, 2016.

[48] H. Vanneuville. Reading group on random nodal lines at ETHZ. `https://metaphor.ethz.ch/x/2019/hs/401-4600-69L/Reading_group_nodal_1_2.pdf`, 2019.

[49] A. Weinrib. Long-range correlated percolation. *Phys. Rev. B*, 29(1):387, 1984.