Overcrowding estimates for nodal volume of centered SGPs

Lakshmi Priya

arXiv:2012.10857

Department of Mathematics
Indian Institute of Science, Bangalore

Plan of the talk

- Introduction.
- Overcrowding in dimension one.
- Overcrowding in higher dimensions.

The setting

SGPs and their nodal sets

For $d \geq 1$, a **centered SGP** X on $\mathbb{R}^d$ is a centered Gaussian process whose distribution is translation invariant, i.e., $\forall t_j, s \in \mathbb{R}^d$

$$(X_{t_1+s}, \ldots, X_{t_n+s}) \stackrel{d}{=} (X_{t_1}, \ldots, X_{t_n}) \sim \mathcal{N}(0, \Sigma_{t_1,\ldots,t_n}).$$

Spectral measure μ of X is the unique finite (for us $\mu(\mathbb{R}^d) = 1$) positive symmetric Borel measure on $\mathbb{R}^d$ s.t.

$$\mathbb{E}[X_s X_t] =: k(s-t) = \widehat{\mu}(s-t) = \int_{\mathbb{R}^d} e^{-i\langle s-t, w\rangle} d\mu(w).$$

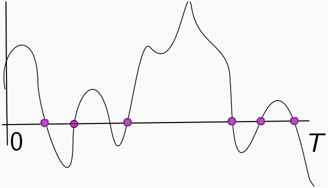
Nodal/Zero set of X, $\mathcal{Z}(X) := X^{-1}\{0\}$.

Bulinskaya's lemma $\Rightarrow$ a.s. X has **no singular zeros**. Hence,

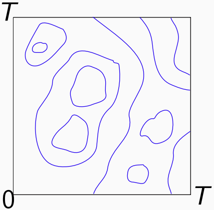
$d = 1 \Rightarrow$ every zero has multiplicity 1,

$d = 2$, Implicit function thm. $\Rightarrow \mathcal{Z}(X) = \bigsqcup$ smooth curves, if X is C^∞.

Zero count & Nodal length



$\mathcal{N}_T := \#\{[0,T] \cap \mathcal{Z}(X)\}$



$\mathcal{L}_T :=$ length$\{[0,T]^2 \cap \mathcal{Z}(X)\}$

If $m \in \mathbb{N}$, $\mathcal{N}_m = \mathcal{N}_{[0,1]} + \mathcal{N}_{[1,2]} + \ldots + \mathcal{N}_{[m-1,m]}$, sum of m ident. dist. r.v.

Hence $\mathbb{E}[\mathcal{N}_m] = \alpha m$, where $\alpha = \mathbb{E}[\mathcal{N}_1]$.

Stationarity $\Rightarrow \mathbb{E}[\mathcal{N}_T] = \alpha T$ and $\mathbb{E}[\mathcal{L}_T] = \beta T^2$.

Overcrowding Question: $\mathbb{P}(\mathcal{N}_T \gg \mathbb{E}[\mathcal{N}_T])$ and $\mathbb{P}(\mathcal{L}_T \gg \mathbb{E}[\mathcal{L}_T])$?

SGP $\overset{?}{\longleftrightarrow}$ Spectral measure $\overset{?}{\longleftrightarrow}$ Zero count

Ex1: Let $\mu = \frac{a}{2}(\delta_{-2} + \delta_2) + \frac{b}{2}(\delta_{-7} + \delta_7) + \frac{c}{2}(\delta_{-30} + \delta_{30})$, w/ $a + b + c = 1$

$$X_t = (\xi_a \cos 2t + \eta_a \sin 2t) + (\xi_b \cos 7t + \eta_b \sin 7t) + (\xi_c \cos 30t + \eta_c \sin 30t),$$

$$\{\xi_s, \eta_s\} \text{ independent and } \xi_s, \eta_s \sim \mathcal{N}(0, s).$$

So X is a random superposition of waves w/ frequencies $\leftrightarrow$ supp(μ) and random amplitude, phase $\leftrightarrow$ mass assigned by μ.

Zero count: In Ex1, behaviour of X depends on the values of a, b, c.

- If $c \gg a, b$, w.h.p. $X_t \approx A\cos(30t + \phi)$, hence many oscillations and many zeros.
- If $a \gg b, c$, w.h.p. $X_t \approx A\cos(2t + \phi)$, hence fewer oscillations and fewer zeros.

Conclusion: Heavy tail of $\mu \to$ many oscillations $\to$ possibily many zeros.

Tail of $\mu \overset{?}{\longleftrightarrow}$ zero count of X (for a general μ)

- For any μ, topological $\mathrm{supp}(X) = \mathcal{F}L^2_{\mathrm{symm}}(\mu)$, where $L^2_{\mathrm{symm}}(\mu) = \{f : \mathbb{R} \to \mathbb{C} \mid f \in L^2(\mu),\ f(-t) = \overline{f(t)}\}$.

- **Ex2:** Let $\mu \sim \mathrm{unif}[-14, 14]$.

  Naive exp.: Every $f \in \mathcal{F}L^2_{\mathrm{symm}}(\mu)$ doesn't oscillate as much as $\sin 14t$.

  Reality: $\mathcal{F}L^2_{\mathrm{symm}}(\mu) = C(\mathbb{R})$, but probabilistically ok!

  $$\text{Borel–TIS/Dudley} \Rightarrow \text{Tail of } \|X^{(n-1)}\|_{[0,1]} \ll \text{Tail of } \mathcal{N}(14^n, (14^n)^2),$$

  $$\boxed{\text{w.h.p. } \|X^{(n)}\|_{[0,1]} \lesssim n14^n \longleftrightarrow \|(\sin 14t)^{(n)}\|_{[0,1]} = 14^n}$$

- For a more general μ, Tail of $\|X^{(n-1)}\|_{[0,1]} \ll$Tail of $\mathcal{N}(C_n, C_n^2)$, where $C_m := \int_{\mathbb{R}} |x|^m d\mu(x)$.

- Heavy tails of $\mu \Rightarrow$ possibly many zeros holds here also.

Statistics of $\mathcal{N}_T$: Known results

	Work of	Conditions on μ or k	Result
Moment conditions	Kac–Rice	$C_2 < \infty$	$\mathbb{E}[\mathcal{N}_T] = (\sqrt{C_2}/\pi)T$
	Beljaev	$C_m < \infty$	$\mathbb{E}[\mathcal{N}_T^m] < \infty$
	Nualart Wschebor	$C_p < \infty$, for some $p > 2m$	somewhat explicit bound for $\mathbb{E}[\mathcal{N}_T^m]$
	Our results	$C_m < \infty$, $\forall m$ & $(A1)$	Overcrowding estimates
Mixing+ moment conditions	Cuzik	$k, k'' \in L^2(\mathbb{R})$ Geman condition	$\mathrm{Var}\mathcal{N}_T \asymp T$ CLT
	BDFZ	$\mathrm{Supp}(\mu)$ compact $\int_{\mathbb{R}} \lvert k(x)\rvert dx < \infty$	$\mathbb{P}(\lvert\mathcal{N}_T - \mathbb{E}[\mathcal{N}_T]\rvert > \eta T) \lesssim e^{-c_\eta T}$

BDFZ: Basu, Dembo, Feldheim, Zeitouni.

Assumptions on the spectral measure

Mixing/decay of k: Events in well separated intervals are *almost independent* and hence $\mathcal{N}_T = \mathcal{N}_{[0,1]} + \mathcal{N}_{[1,2]} + \ldots + \mathcal{N}_{[T-1,T]}$ is approximately a sum of identically distributed M-dependent r. v.

Decay of $k \longleftrightarrow$ Smoothness of the density of μ (restrictive condition!).

Example: Result of BDFZ does **not** apply to $\mu \sim \text{unif}[-1,1]$.

Finiteness of moments of μ:

Lighter tails of $\mu \Rightarrow$ fewer oscillations $\Rightarrow$ fewer zeros.

(A1): $d\mu(x) = f(x)dx + d\mu_s(x)$, where $f \not\equiv 0$.

Overcrowding estimates in dimension one

Theorem 1: If X is a centered SGP on $\mathbb{R}$ with spectral measure μ. Then,

μ has finite moments + μ satisfies $(A1)$ $\implies$ Overcrowding estimates in terms of moments C_n.

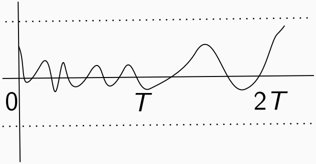
Growth of C_n	**Example of** μ	**Constraints on** T and n	$\log \mathbb{P}(\mathcal{N}_T \geq n)$
$C_n \leq q^n$, for $q > 0$	Any μ with $\text{supp}(\mu) \subseteq [-q, q]$	$n \geq CT$, for $C \gg 1$	$\asymp -n^2 \log(\frac{n}{T})$
$C_n \leq n^{\alpha n}$, for $\alpha \in (0,1)$	$\mu \sim \mathcal{N}(0,1)$, with $\alpha = 1/2$	$n \geq T^{1/\kappa}$, for $\kappa \in (0, 1-\alpha)$	$\asymp -n^2 \log n$
	$\mu \sim \mathcal{N}(0,1)$	$n \geq T^{2+\epsilon}$	

Overcrowding of zeros

A deterministic idea to understand overcrowding...

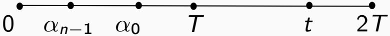
Lemma ($\sim$ *Azais*, *Nualart*, *Wschebor*)

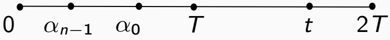
Let $f : [0, 2T] \to \mathbb{R}$ *be a smooth function. If* $\|f^{(n)}\|_{[0,2T]} \leq M$ *and* $\mathcal{N}_T(f) \geq n$*, then* $\|f\|_{[T,2T]} \leq M(2T)^n/n!$.



$$\left.\begin{array}{l} |f^{(n)}| \text{ not too large } + \\ f \text{ has many zeros} \end{array}\right\} \Rightarrow |f| \text{ small}$$

Proof. Since b/w every two zeros of a smooth function, there is a zero of its derivative, there exist $0 \leq \alpha_{n-1} \leq \ldots \leq \alpha_0 \leq T$ s.t. $f^{(k)}(\alpha_k) = 0$.





$f^{(k)}(\alpha_k) = 0$

We have $\forall t \in [\alpha_{n-1}, 2T]$,

$$f^{(n-1)}(t) - \cancel{f^{(n-1)}(\alpha_{n-1})} = \int_{\alpha_{n-1}}^{t} f^{(n)}(s)ds,$$

$$|f^{(n-1)}(t)| \leq Mt.$$

For $t \in [\alpha_{n-2}, 2T]$, we similarly have

$$f^{(n-2)}(t) - \cancel{f^{(n-2)}(\alpha_{n-2})} = \int_{\alpha_{n-2}}^{t} f^{(n-1)}(s)ds,$$

$$|f^{(n-2)}(t)| \leq \int_{\alpha_{n-2}}^{t} Ms\ ds,$$

$$\leq Mt^2/2!.$$

Continuing in a similar manner gives the conclusion. $\square$

… and a corollary to study overcrowding of random zeros

Corollary

For $n \in \mathbb{N}$, $M, T > 0$ and any a.s. smooth random function F, we have

$$\mathbb{P}(\mathcal{N}_T > n) \;\leq\; \underbrace{\mathbb{P}(\|F\|_{[T,2T]} \leq M\tfrac{(2T)^n}{n!})}_{\boxed{\text{small ball prob.}}} \;+\; \underbrace{\mathbb{P}(\|F^{(n)}\|_{[0,2T]} \geq M)}_{\boxed{\text{Borel–TIS/Dudley's bound}}},$$

We will choose M as follows:

- M small enough so that the first term indeed corresponds to a *small ball event*.
- M large enough so that event in the second term is unlikely.

Finiteness of moments of $\mu \Rightarrow$ tail bounds for $\|X^{(n)}\|$

$X^{(n)}$ is also a smooth centered SGP with spectral density $x^{2n} d\mu(x)$.

Borel–TIS inequality, Dudley integral give tail bounds for $\|X^{(n)}\|_{[0,2T]}$.

The pseudo-metric d_n on $\mathbb{R}$, induced by $X^{(n)}$ satisfies

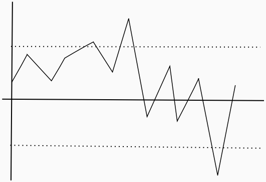
$$d_n(t,s) \le \sqrt{C_{2n+2}}\, |t-s|.$$

With this, Dudley's integral and the concentration result, for $2T \le n$

$$\mathbb{P}(\|X^{(n)}\|_{[0,2T]} \ge n\sqrt{C_{2n}}x) \le e^{-x^2}.$$

$(A1) \Rightarrow$ a small ball probability (SBP) estimate

Most of the known SBP estimates are for non-smooth processes and rightly so; the *sharp turns* make it difficult for the process to be confined to a small ball.



Under assumption $(A1)$, we get SBP estimates.

$$(A1) : d\mu(x) = f(x)dx + d\mu_s(x), \text{ where } f \not\equiv 0.$$

Interior of $\text{supp}(f) \neq \phi \Rightarrow$ topological $\text{supp.}(X) = \mathcal{F}L^2_{\text{symm}}(\mu) = C(\mathbb{R})$. Hence X comes very close to behaving like a non-smooth process.

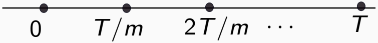
SBP estimate

Lemma (∼ Krishnapur–Maddaly)

Let $T > 0, m \in \mathbb{N}$. Assume μ satisfies $(A1)$, then there is $b \in (0,1)$, $C > 1$ such that if $T \le bm$, we have

$$\mathbb{P}(\|X\|_{[0,T]} \le \eta) \le (Cm/T)^{m^2}\, \eta^m.$$

Proof idea. Consider the lattice $\{jT/m : j \in \mathbb{Z}\}$, let $Y_j = X_{jT/m}$, then Y is a centered SGP on $\mathbb{Z}$ w/ spectral measure ν on $[-\pi, \pi]$ satisfying $(A1)$.



Let $t_j := jT/m$, then

$$\begin{aligned}
\mathbb{P}(\|X\|_{[0,T]} \le \eta) &\le \mathbb{P}(|X_{t_j}| \le \eta, \forall j \in [m]), \\
&= \int_{[-\eta,\eta]^m} \frac{1}{(\sqrt{2\pi})^m |\Sigma|^{1/2}} e^{-\frac{\langle \Sigma^{-1} x, x\rangle}{2}}\, dx, \\
&\le \frac{(2\eta)^m}{(\sqrt{2\pi})^m |\Sigma|^{1/2}} \le \frac{(2\eta)^m}{(\sqrt{2\pi})^m |\lambda|^{m/2}},
\end{aligned}$$

SBP estimate contd.

w/ $\Sigma = \mathrm{Cov}(X_{t_1}, \ldots, X_{t_m})$ and λ the least eigenvalue of Σ. Hence it suffices to get a lower bound for λ. For $v \in \mathbb{R}^m$

$$\langle \Sigma v, v\rangle = \sum \mathbb{E}[X_{t_\ell} X_{t_n}] v_\ell v_n = \sum \int_{-\pi}^{\pi} e^{-i\langle \ell - n, x\rangle} v_\ell v_n \, d\nu(x),$$

$$= \int_{-\pi}^{\pi} |\sum v_n e^{-inx}|^2 \, d\nu(x),$$

since $d\nu = g dx + d\nu_s$, $g \geq \delta \mathbb{1}_{\mathcal{J}}$, for some $\delta > 0$ and $\mathcal{J} \subseteq [-\pi, \pi]$,

$$\gtrsim \int_{\mathcal{J}} |\sum v_n e^{-inx}|^2 \, dx,$$

$$\underset{\text{Turan's lemma}}{\gtrsim} \int_{-\pi}^{\pi} |\sum v_n e^{-inx}|^2 \, dx = 2\pi \|v\|^2.$$

This gives a lower bound for λ since $\lambda = \inf_{v \neq 0} \langle \Sigma v, v\rangle / \langle v, v\rangle$. $\square$

Overcrowding estimates

$$\begin{cases} \textbf{Recall}\text{: } \mathbb{P}(\mathcal{N}_T > n) \leq \mathbb{P}(\|X\|_{[0,T]} \leq M\frac{(2T)^n}{n!}) + \mathbb{P}(\|X^{(n)}\|_{[0,2T]} \geq M), \\ \mathbb{P}(\|X^{(n)}\|_{[0,2T]} \geq n\sqrt{C_{2n}}x) \leq e^{-x^2} \text{ and } \mathbb{P}(\|X\|_{[0,T]} \leq \eta) \leq \left(\frac{Cm}{T}\right)^{m^2} \eta^m. \end{cases}$$

Choosing $M = n^3\sqrt{C_{2n}}$, term $\leq e^{-n^4}$ and (forgetting insignificant terms)

$$\text{term} \leq \left(\frac{m}{T}\right)^{m^2} \left(\sqrt{C_{2n}}\left(\frac{T}{n}\right)^n\right)^m.$$

Ex: If μ is s.t. $\operatorname{supp}(\mu) \subseteq [-q, q]$ for some $q > 1$, then $C_n \leq q^n$. For $m = n/2$,

$$\begin{aligned} \text{term} &\leq \left(\frac{n}{T}\right)^{n^2/4} \left(\frac{qT}{n}\right)^{n^2/2} = \left(\frac{q^2T}{n}\right)^{n^2/4}, \\ &\leq \exp\left(-\frac{n^2}{4}\log\left(\frac{n}{q^2T}\right)\right), \end{aligned}$$

which is useful when $n \geq CT$, for some large enough C.

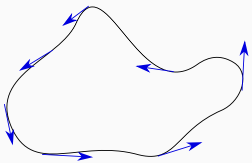
Overcrowding of the nodal set in dimension 2

How to measure nodal length of a random function?

Traditional way to measure length of a curve:

- Parametrize it, $\gamma : [0,1] \to \mathbb{R}^2$.
- Evaluate $\int_0^1 \|\dot{\gamma}(t)\| \, dt$.

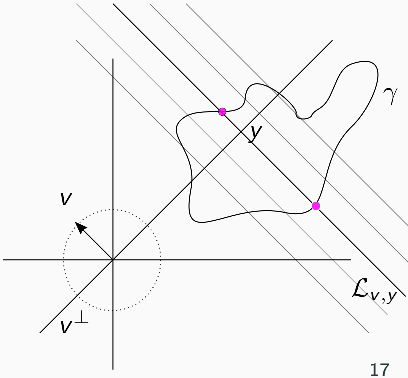
Not useful for random functions.



Crofton's formula for length:

$$\text{length}(\gamma) = c \int_{\mathbb{S}^1} \int_{v^\perp \simeq \mathbb{R}} \#\{\gamma \cap \mathcal{L}_{v,y}\} \, dy dv.$$

We understand (to some extent) zero count in one dimension, hence this might help. But there are infinitely many lines on which we need to know the # of intersections to get len(γ).



Discretizing Crofton's formula: Step 1

We *discretize* Crofton's formula.

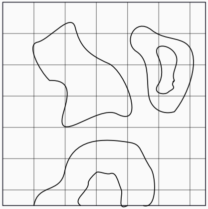
$$\text{length}(\gamma) = c \int_{\mathbb{S}^1} \int_{v^\perp \simeq \mathbb{R}} \#\{\gamma \cap \mathcal{L}_{v,y}\}\, dy dv.$$

There are infintely many directions $v \in \mathbb{S}^1$. We can use area/coarea formula to reduce to considering just two perpendicular directions

$$\text{length}(\gamma) \leq C \left(\int_{\mathbb{R}} \#\{\gamma \cap \mathcal{L}_{e_1,t}\}\, dt + \int_{\mathbb{R}} \#\{\gamma \cap \mathcal{L}_{e_2,t}\}\, dt \right),$$

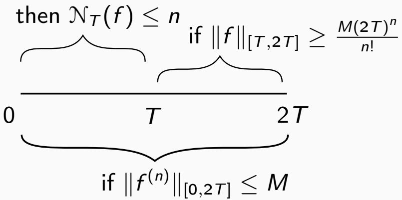
hence $\mathcal{L}_T(f) \leq C \left(\int_0^T \#\{\mathcal{Z}(f) \cap \mathcal{L}_{e_1,t}\}\, dt + \int_0^T \#\{\mathcal{Z}(f) \cap \mathcal{L}_{e_2,t}\}\, dt \right).$

Even now, there are infinitely many lines on which we need to know the zero count.



Discretizing Crofton's formula: Step 2

Recall the idea for zero count:

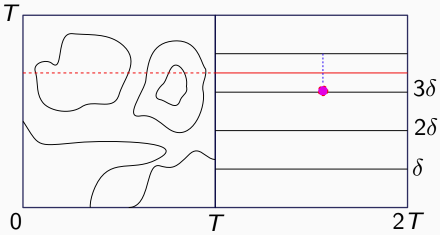


Let $f : [0, n]^2 \to \mathbb{R}$ and $\delta = (2T)^n/n!$.
Every horizontal line in $[0, T]^2$ has $\leq n$ zeros if

a. $\|\partial_1^n f\|_{[0,n]^2} \leq M/2$,

b. $\|f(\cdot, t)\|_{[T,2T]} \geq M\delta,\ \forall t \in [0, T]$.

Cond. $\overset{\text{replace}}{\longleftrightarrow}$ $c.$ and $d.$



c. $\|\partial_2 f\|_{[0,n]^2} \leq M/2.$ and d. $\|f(\cdot, t)\|_{[T,2T]} \geq M\delta,\ \forall t \in \{\delta, 2\delta, \ldots, T/\delta\}$.

$c.$ and $d.$ $\Rightarrow$ $\|f(\cdot, t)\|_{[T,2T]} \geq M\delta/2,\ \forall t \in [0, T]$.

From last slide,

$$a.\ \|\partial_1^n f\|_{[0,n]^2} \leq M/2,\ \ c.\ \|\partial_2 f\|_{[0,n]^2} \leq M/2,$$

$$d.\ \|f(\cdot,t)\|_{[T,2T]} \geq M\delta,\ \forall t \in \{\delta, 2\delta, \dots, T/\delta\}.$$

Hence, a., c., d, and analogous condts for $e_2 \ \Rightarrow \mathcal{L}_T(f) \lesssim nT$.

thus, $\mathbb{P}(\mathcal{L}_T \gtrsim nT) \leq \mathbb{P}(\neg a.) + \mathbb{P}(\neg c.) + \mathbb{P}(\neg d.) +$

analogous terms from e_2.

Terms $\longleftrightarrow$ Borel–TIS/Dudley,

term $\longleftrightarrow$ SBP estimates on finitely many lines.

Overcrowding estimates for nodal length

Theorem 2: Let X be a centered SGP on $\mathbb{R}^2$ with spectral measure μ.

μ has finite moments + μe_1 and μe_2 satisfy $(A1)$	$\Longrightarrow$	Overcrowding estimates for $\mathcal{L}_T$ in terms of moments.

Consequence: Similar to one dimesional overcrowding estimates, we get estimates in this case too.

Ex: For μ which is compactly supported, $\exists C \gg 1$ s.t. if $\ell \geq CT^2$

$$\mathbb{P}(\mathcal{L}_T \geq \ell) \lesssim \exp\left(-\frac{\ell^2}{T^2}\log\frac{\ell}{T^2}\right).$$

Questions galore...

...and here are some of them:

- Compared to one dimension, much less is known about nodal volume in higher dimensions. We show that the nodal volume in $[0, T]^d$ has light tails, can we establish *exponential concentration for nodal volume in higher dim*?

- We show $\mathcal{N}_T$ has light tails even without any regularity of k. *How essential is the regularity of covariance k to establish exponential conct./CLT in one dimension?*

Thank You!