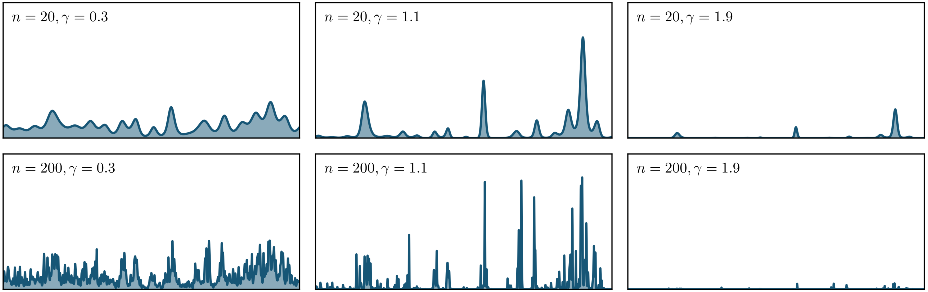
# Invariance principle for the Gaussian multiplicative chaos

## *via a high-dimensional CLT with low-rank increments*

Mriganka Basu Roy Chowdhury (UC Berkeley)

*jointly with Shirshendu Ganguly*

February 19, 2025



# Introduction

- Starting with an i.i.d. sequence of random variables $a_1, a_2, \ldots, a'_1, a'_2, \ldots,$ sampled from some common law $\mathcal{L}$ (centered, variance 1, with exponential moments), let us form the *random Fourier series*

$$S_n(t) = \sum_{k=1}^{n} k^{-\theta} \Big\{ a_k \cos(2\pi k t) + a'_k \sin(2\pi k t) \Big\} \quad t \in [0,1].$$

- How does this look like in the $n \to \infty$ limit, for some exponent $\theta \in (0,1)$?

- Starting with an i.i.d. sequence of random variables $a_1, a_2, \ldots, a'_1, a'_2, \ldots$, sampled from some common law $\mathcal{L}$ (centered, variance 1, with exponential moments), let us form the *random Fourier series*

$$S_n(t) = \sum_{k=1}^{n} k^{-\theta} \Big\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \Big\} \quad t \in [0,1].$$

- How does this look like in the $n \to \infty$ limit, for some exponent $\theta \in (0,1)$?
- To get started, consider the variance of $S_n(t)$:

$$\mathrm{Var}(S_n(t)) \quad = \quad \sum_{k=1}^{n} k^{-2\theta} \Big\{ \cos^2(2\pi kt) + \sin^2(2\pi kt) \Big\} = \sum_{k=1}^{n} k^{-2\theta}.$$

- Starting with an i.i.d. sequence of random variables $a_1, a_2, \ldots, a'_1, a'_2, \ldots$, sampled from some common law $\mathcal{L}$ (centered, variance 1, with exponential moments), let us form the *random Fourier series*

$$S_n(t) = \sum_{k=1}^{n} k^{-\theta} \Big\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \Big\} \quad t \in [0,1].$$

- How does this look like in the $n \to \infty$ limit, for some exponent $\theta \in (0,1)$?
- To get started, consider the variance of $S_n(t)$:

$$\mathrm{Var}(S_n(t)) \;=\; \sum_{k=1}^{n} k^{-2\theta} \Big\{ \cos^2(2\pi kt) + \sin^2(2\pi kt) \Big\} = \sum_{k=1}^{n} k^{-2\theta}.$$

- Thus, one expects a qualitative change in the behavior of $S_n(t)$ at $\theta = 1/2$, when $\mathrm{Var}(S_n(t)) = \Theta(\log n)$. This is the regime we will focus on.

- A calculation shows that $S_n$ is log-correlated in the following sense. For $t, s \in [0, 1]$,

$$\mathrm{Cov}(S_n(t), S_n(s)) = \min\left\{\log\frac{1}{|t-s|}, \mathrm{Var}(S_n(0))\right\} + \text{something uniformly bounded},$$

- To unpack this a bit, recall that $\mathrm{Var}(S_n(0)) = O(\log n)$. So informally,

$$\mathrm{Cov}(S_n(t), S_n(s)) = \begin{cases} \log n, & \text{if } |t-s| \ll n^{-1}, \\ \log\frac{1}{|t-s|}, & \text{if } |t-s| \gg n^{-1}, \end{cases}$$

up to some bounded error. *Observe that the singularity occurs only when $s$ and $t$ are arbitrarily close.*

- A calculation shows that $S_n$ is log-correlated in the following sense. For $t, s \in [0,1]$,

$$\mathrm{Cov}(S_n(t), S_n(s)) = \min\left\{\log\frac{1}{|t-s|}, \mathrm{Var}(S_n(0))\right\} + \text{something uniformly bounded},$$

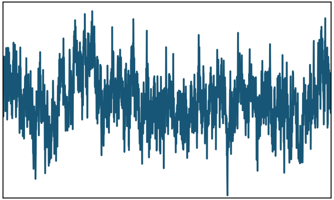
- To unpack this a bit, recall that $\mathrm{Var}(S_n(0)) = O(\log n)$. So informally,

$$\mathrm{Cov}(S_n(t), S_n(s)) = \begin{cases} \log n, & \text{if } |t-s| \ll n^{-1}, \\ \log\frac{1}{|t-s|}, & \text{if } |t-s| \gg n^{-1}, \end{cases}$$

up to some bounded error. *Observe that the singularity occurs only when $s$ and $t$ are arbitrarily close.*

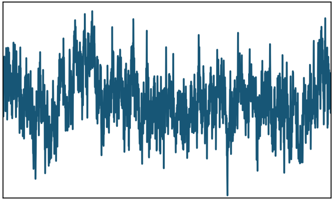
- Due to diverging one-point variance, referred to in physics as an *ultraviolet divergence*, this series does not converge as a function. However, the limit can be viewed as a *random distribution*, in $H^{-\varepsilon}$ for any $\varepsilon > 0$.

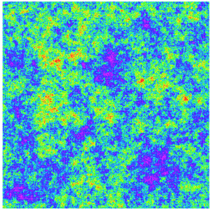
The random Fourier series model (left) is an example of a *log-correlated Gaussian field*, instances of which appear widely in mathematics and statistical physics.



The random Fourier series model (left) is an example of a *log-correlated Gaussian field*, instances of which appear widely in mathematics and statistical physics.

The canonical example is the **Gaussian free field (GFF)** (right, in heightmap, due to Marek Biskup) in 2D (the log-correlation follows from the logarithmic behavior of the Green's function in 2D).





- Log-correlated fields (LCFs) come up in diverse contexts, including the study of fluid turbulence in physics, log-characteristic polynomials of random unitary matrices, and the Riemann zeta function.
- LCFs also have a rich universality theory, especially in the Gaussian case, where the universality of the maxima and related processes have been studied extensively recently.

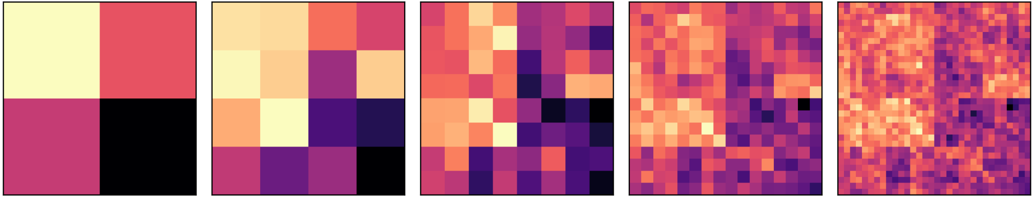
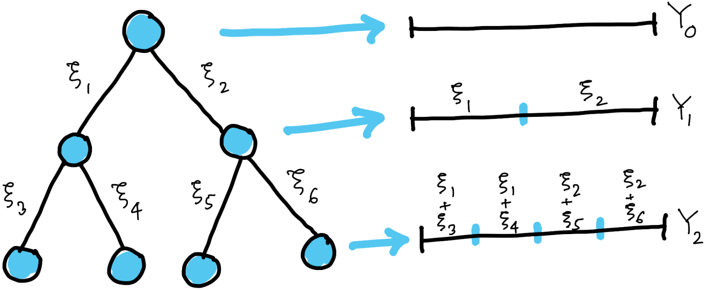
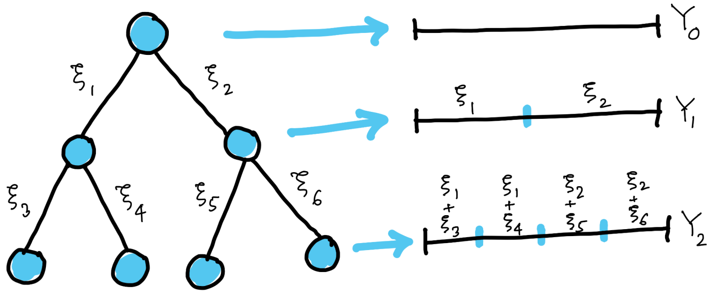


Figure: Structure of a generic LCF (lighter colors are higher values). There is a unit randomness at each "scale", inducing an additional unit correlation between two points separated by a distance smaller than the scale. Since only log-many scales are bigger than the distance between two points, the covariance is logarithmic.

- The informal structure described above inspires the definition of a “model” log-correlated process, called the *branching random walk*.

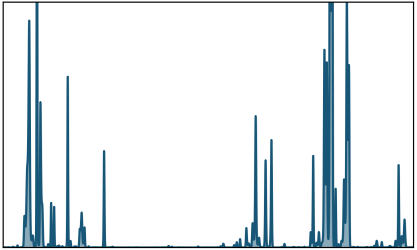


- The informal structure described above inspires the definition of a "model" log-correlated process, called the *branching random walk*.




- Observe that if $s, t \in (0,1)$ are such that $|s-t| \in [2^{-k-1}, 2^{-k})$, then $\mathrm{Cov}(Y_n(s), Y_n(t)) = k\sigma^2$ for all sufficiently large $n$, indicating a log-correlated structure.
- To match with our definition of log-correlation, which required $\log(1/|x-y|)$ covariance, we will then need $\log(1/2^{-k})$ to match $k\sigma^2$, which will be the case when $\sigma^2 = \log 2$.

Inspired by Mandelbrot's theories on intermittency in fluid turbulence, Kahane (1985) constructed the *Gaussian multiplicative chaos*, as a way to define the exponential of log-correlated *Gaussian processes* as a random measure. These measures are supposed to exhibit **intermittency**, i.e., stretches of calm punctuated by bursts of fractal behavior.



- Let us briefly outline the construction for $S_n(t)$, assuming for the moment that the noise variables are standard Gaussian.

- Let us briefly outline the construction for $S_n(t)$, assuming for the moment that the noise variables are standard Gaussian.
- For any *intermittency exponent* $\gamma \geq 0$, define the density (not probability density!)

$$\rho_n(t) = \rho_{\gamma,n}(t) = \frac{e^{\gamma S_n(t)}}{\mathbb{E}e^{\gamma S_n(t)}} = \exp\left\{\gamma S_n(t) - \frac{\gamma^2}{2} \cdot \mathrm{Var}(S_n(0))\right\}.$$

The additional factor of $\exp\left(-\frac{\gamma^2}{2} \cdot \mathrm{Var}(S_n(0))\right)$ is a *renormalization* factor, which ensures that $\mathbb{E}\rho_n(t) = 1$.

- Thus, $\mathbb{E}\int \rho_n(t) = 1$. Thinking of $\rho_n$ as a "probability density on average", we also define the random measure

$$\mu_n = \mu_{\gamma,n}(A) = \int_A \rho_n(t) \, \mathrm{d}t.$$

- Thus, $\mathbb{E}\int \rho_n(t) = 1$. Thinking of $\rho_n$ as a "probability density on average", we also define the random measure

$$\mu_n = \mu_{\gamma,n}(A) = \int_A \rho_n(t)\, \mathrm{d}t.$$

- It turns out that $\mu_n(A)$ is a martingale for each $A$, with a nontrivial limit $\mu(A)$ as long as $\gamma < \sqrt{2}$ (for $\gamma \geq \sqrt{2}$, the limit is zero this is because the expectation comes from scales larger than those in a typical sample). This random measure is the *Gaussian multiplicative chaos* (GMC) for the "process" $S(t)$.

- Thus, $\mathbb{E}\int \rho_n(t) = 1$. Thinking of $\rho_n$ as a "probability density on average", we also define the random measure

$$\mu_n = \mu_{\gamma,n}(A) = \int_A \rho_n(t)\,\mathrm{d}t.$$

- It turns out that $\mu_n(A)$ is a martingale for each $A$, with a nontrivial limit $\mu(A)$ as long as $\gamma < \sqrt{2}$ (for $\gamma \geq \sqrt{2}$, the limit is zero this is because the expectation comes from scales larger than those in a typical sample). This random measure is the *Gaussian multiplicative chaos* (GMC) for the "process" $S(t)$.
- This general procedure for exponentiating log-correlated Gaussian processes has also been used as a key tool in quantum field theory, for instance in the construction of the Liouville quantum gravity (LQG) measure, as the exponential of the 2D GFF.

$$S_n(t) = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \left[ g_k \cos(2\pi k t) + g_k' \sin(2\pi k t) \right]$$

$$S_n(t) = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \left[ g_k \cos(2\pi k t) + g_k' \sin(2\pi k t) \right] \xrightarrow[\text{as dist.}]{\lim} S(t)$$

distribution log-corr.

$$S_n(t) = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \left[ g_k \cos(2\pi k t) + g_k' \sin(2\pi k t) \right] \xrightarrow[\text{as dist.}]{\lim} S(t)$$

distribution
log-corr.

exp

$$\mu_{\gamma,n}(dt) = \rho_n(t)\,dt = \frac{e^{\gamma S_n(t)}}{\mathbb{E}\, e^{\gamma S_n(t)}}\,dt$$

smooth measure

$$S_n(t) = \sum_{k=1}^{n} \frac{1}{\sqrt{k}} \left[ g_k \cos(2\pi k t) + g_k' \sin(2\pi k t) \right] \xrightarrow[\text{as dist.}]{\lim} S(t)$$

$S(t)$: distribution, log-corr.

$S_n(t)$ $\xrightarrow{\exp}$ $\mu_{\gamma,n}(dt)$; $S(t)$ $\xrightarrow{\text{GMC}}$ $\mu_\gamma(dt)$

$$\mu_{\gamma,n}(dt) = \rho_n(t)\,dt = \frac{e^{\gamma S_n(t)}}{\mathbb{E}\, e^{\gamma S_n(t)}}\,dt \xrightarrow[\text{as martingales}]{\lim} \mu_\gamma(dt)$$

$\mu_{\gamma,n}(dt)$: smooth measure

$\mu_\gamma(dt)$: random, fractal measure

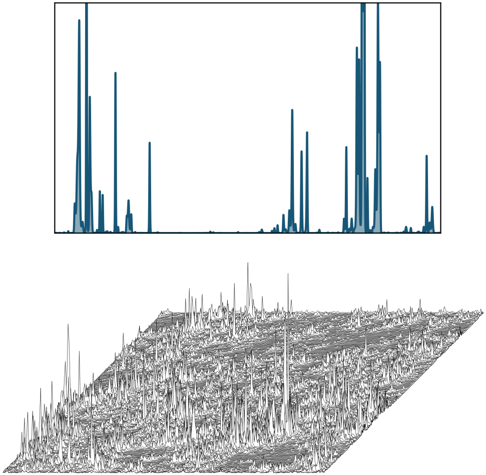


Figure: GMC for the random Fourier series (top), and the GMC for the 2D Gaussian free field, also known as the Liouville Quantum Gravity measure (bottom). LQG picture due to Marek Biskup.

- A couple of natural questions arise:
  - Is there a similar construction for non-Gaussian noise?

- A couple of natural questions arise:
  - Is there a similar construction for non-Gaussian noise? And if so,
  - is there an invariance principle connecting the *non-Gaussian* multiplicative chaos to the GMC?

- A couple of natural questions arise:
  - Is there a similar construction for non-Gaussian noise? And if so,
  - is there an invariance principle connecting the *non-Gaussian* multiplicative chaos to the GMC?
- In 2016, Junnila showed that the basic objects involved are well-defined. That is,
  - the corresponding Fourier series is log-correlated, and
  - under some suitable conditions, the GMC construction described earlier goes through, yielding a multiplicative chaos measure.

- A couple of natural questions arise:
  - Is there a similar construction for non-Gaussian noise? And if so,
  - is there an invariance principle connecting the *non-Gaussian* multiplicative chaos to the GMC?
- In 2016, Junnila showed that the basic objects involved are well-defined. That is,
  - the corresponding Fourier series is log-correlated, and
  - under some suitable conditions, the GMC construction described earlier goes through, yielding a multiplicative chaos measure.

These developments then open up the possibility of studying the behavior of these newly constructed non-Gaussian measures.

# Main result

## Theorem (B.R.C.-Ganguly, 2025)

*Let $\mu_a$ be the multiplicative chaos corresponding to the random Fourier series with noise variables $\{a_i, a_i'\}$ sampled from $\mathcal{L}$ and $\mu_g$ be the analogous GMC constructed out of standard Gaussians $\{g_i, g_i'\}$.*
*Then for* **any** $\gamma \in (0, \sqrt{2})$ *one may couple the $a$ and $g$ variables such that*

$$\mu_g \ll \mu_a \ll \mu_g,$$

*i.e., they are mutually absolutely continuous a.s.*

## Theorem (B.R.C.-Ganguly, 2025)

*Let* $\mu_a$ *be the multiplicative chaos corresponding to the random Fourier series with noise variables* $\{a_i, a_i'\}$ *sampled from* $\mathcal{L}$ *and* $\mu_g$ *be the analogous GMC constructed out of standard Gaussians* $\{g_i, g_i'\}$*.*
*Then for* **any** $\gamma \in (0, \sqrt{2})$ *one may couple the a and g variables such that*

$$\mu_g \ll \mu_a \ll \mu_g,$$

*i.e., they are mutually absolutely continuous a.s.*

- This resolves a recent conjecture due to Kim and Kriechbaum, who only proved this for $\gamma > 1$.

## Theorem (B.R.C.-Ganguly, 2025)

*Let $\mu_a$ be the multiplicative chaos corresponding to the random Fourier series with noise variables $\{a_i, a_i'\}$ sampled from $\mathcal{L}$ and $\mu_g$ be the analogous GMC constructed out of standard Gaussians $\{g_i, g_i'\}$.*
*Then for* **any** $\gamma \in (0, \sqrt{2})$ *one may couple the a and g variables such that*

$$\mu_g \ll \mu_a \ll \mu_g,$$

*i.e., they are mutually absolutely continuous a.s.*

- This resolves a recent conjecture due to Kim and Kriechbaum, who only proved this for $\gamma > 1$.
- Note that the two chaos measures *cannot be coupled to be equal in general* (and thus absolute continuity is a natural target), because any coupling error between $a_1$ and $g_1$ (and $a_1'$ and $g_1'$) will cause a relative shift between the two log-correlated fields. *This error cannot be "corrected" by subsequent terms.*

As indicated in the title for this talk, the proof strategy crucially involves a high-dimensional CLT. We will spend the next few slides to see how the problem can be reduced to such a question.

Define the partial sums

$$S_{n,a}(t) = \sum_{k=1}^{n} k^{-\frac{1}{2}} \left\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \right\},$$
$$S_{n,g}(t) = \sum_{k=1}^{n} k^{-\frac{1}{2}} \left\{ g_k \cos(2\pi kt) + g'_k \sin(2\pi kt) \right\},$$

Define the partial sums

$$S_{n,a}(t) = \sum_{k=1}^{n} k^{-\frac{1}{2}} \Big\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \Big\},$$
$$S_{n,g}(t) = \sum_{k=1}^{n} k^{-\frac{1}{2}} \Big\{ g_k \cos(2\pi kt) + g'_k \sin(2\pi kt) \Big\},$$

so that the corresponding multiplicative chaos measures are

$$\mu_a = \lim_{n\to\infty} \mu_{n,a}, \quad \mu_g = \lim_{n\to\infty} \mu_{n,g},$$

where the finite $n$ measures are defined as

$$\mu_{n,a}(A) = \int_A \rho_{n,a}(t)\ \mathrm{d}t = \int_A e^{\gamma S_{n,a}(t)} \Big/ \ \mathbb{E} e^{\gamma S_{n,a}(t)}\ \mathrm{d}t,$$
$$\mu_{n,g}(A) = \int_A \rho_{n,g}(t)\ \mathrm{d}t = \int_A e^{\gamma S_{n,g}(t)} \Big/ \ \mathbb{E} e^{\gamma S_{n,g}(t)}\ \mathrm{d}t.$$

A candidate for the derivative $\frac{\mathrm{d}\mu_a}{\mathrm{d}\mu_g}$ is the limit of the finite $n$-derivatives

A candidate for the derivative $\frac{d\mu_a}{d\mu_g}$ is the limit of the finite $n$-derivatives

$$R_n(t) = \frac{d\mu_{n,a}}{d\mu_{n,g}}(t) = \frac{e^{\gamma S_{n,a}(t)} \Big/ \mathbb{E} e^{\gamma S_{n,a}(t)}}{e^{\gamma S_{n,g}(t)} \Big/ \mathbb{E} e^{\gamma S_{n,g}(t)}}.$$

A candidate for the derivative $\frac{\mathrm{d}\mu_a}{\mathrm{d}\mu_g}$ is the limit of the finite $n$-derivatives

$$R_n(t) = \frac{\mathrm{d}\mu_{n,a}}{\mathrm{d}\mu_{n,g}}(t) = \frac{e^{\gamma S_{n,a}(t)} \Big/ \mathbb{E}e^{\gamma S_{n,a}(t)}}{e^{\gamma S_{n,g}(t)} \Big/ \mathbb{E}e^{\gamma S_{n,g}(t)}}.$$

It is not hard to show that $\mathbb{E}e^{\gamma S_{n,a}(t)} \approx \mathbb{E}e^{\gamma S_{n,g}(t)}$ (since this is a one-dimensional CLT problem involving only the marginal at time $t$), so

$$R_n(t) \approx \exp\left(\gamma\Big\{S_{n,a}(t) - S_{n,g}(t)\Big\}\right).$$

A candidate for the derivative $\frac{d\mu_a}{d\mu_g}$ is the limit of the finite $n$-derivatives

$$R_n(t) = \frac{d\mu_{n,a}}{d\mu_{n,g}}(t)$$
$$= \frac{e^{\gamma S_{n,a}(t)} \Big/ \mathbb{E}e^{\gamma S_{n,a}(t)}}{e^{\gamma S_{n,g}(t)} \Big/ \mathbb{E}e^{\gamma S_{n,g}(t)}}.$$

It is not hard to show that $\mathbb{E}e^{\gamma S_{n,a}(t)} \approx \mathbb{E}e^{\gamma S_{n,g}(t)}$ (since this is a one-dimensional CLT problem involving only the marginal at time $t$), so

$$R_n(t) \approx \exp\left(\gamma\left\{S_{n,a}(t) - S_{n,g}(t)\right\}\right).$$

Thus, the goal is:

**Couple the $a$ and $g$ variables such that the difference $S_{n,a} - S_{n,g}$ is uniformly controlled.**

Recall that $S_{n,a}(t)$ accumulates $O(1)$ variance per dyadic scale, i.e.,

$$S_{n,a}(t) = \sum_{\ell=1}^{\log_2 n} X_{\ell,a}(t), \quad \text{(sum over dyadic scales)}$$

Recall that $S_{n,a}(t)$ accumulates $O(1)$ variance per dyadic scale, i.e.,

$$S_{n,a}(t) = \sum_{\ell=1}^{\log_2 n} X_{\ell,a}(t), \quad \text{(sum over dyadic scales)}$$

where $X_{\ell,a}(t)$ is the $\ell$-th scale

$$X_{\ell,a}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} \Big\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \Big\},$$

with variance $\sum_{k=2^{\ell-1}}^{2^\ell-1} k^{-1} \approx \log 2$.

Recall that $S_{n,a}(t)$ accumulates $O(1)$ variance per dyadic scale, i.e.,

$$S_{n,a}(t) = \sum_{\ell=1}^{\log_2 n} X_{\ell,a}(t), \quad \text{(sum over dyadic scales)}$$

where $X_{\ell,a}(t)$ is the $\ell$-th scale

$$X_{\ell,a}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} \Big\{ a_k \cos(2\pi kt) + a'_k \sin(2\pi kt) \Big\},$$

with variance $\sum_{k=2^{\ell-1}}^{2^\ell-1} k^{-1} \approx \log 2$.

**We will attempt to couple the $a$ and $g$ variables, scale-by-scale, such that the error per scale, $|X_{\ell,a} - X_{\ell,g}|$ is controlled,** *in a summable fashion*. **For instance, a uniform bound of $\ell^{-2}$ will suffice.** **This is plausible because $X_{\ell,a}$ is more homogenous with increasing $\ell$.**

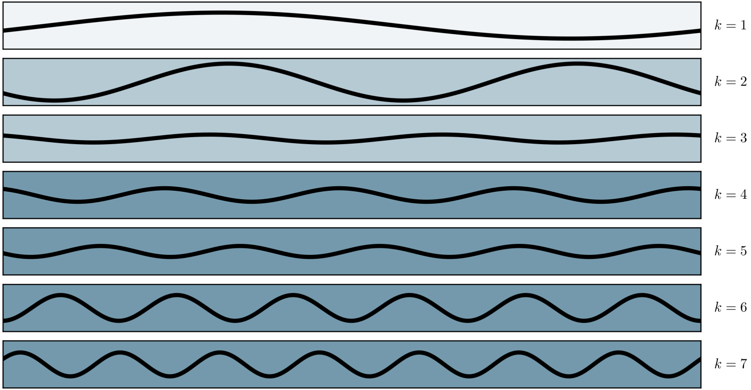



Figure: Sinusoids $a_k \cos(2\pi kt) + a'_k \sin(2\pi kt)$, colored by the dyadic scale in which $k$ belongs. These are scaled and added up to form $S_{7,a}(t)$.

# Reducing to a finite-dimensional problem

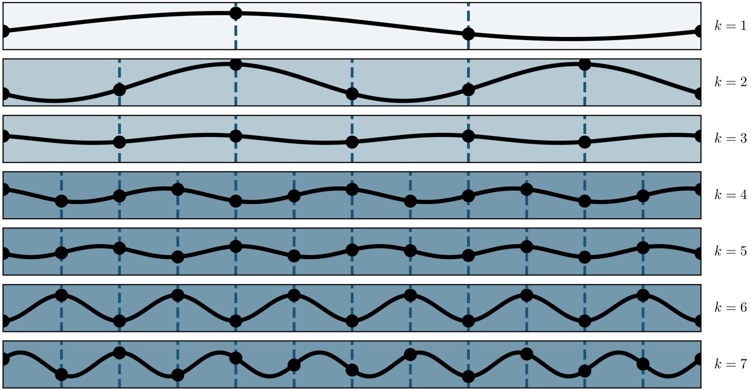
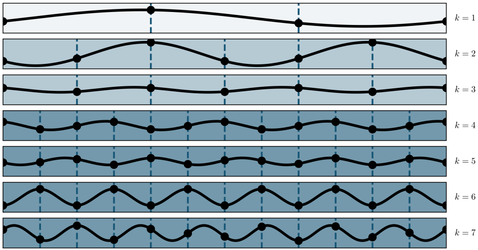



Figure: We discretize $[0, 1]$ (approximately) dyadically in space, with resolution depending on the scale. This produces a *hierarchical model*. It will suffice to couple only at the mesh points.

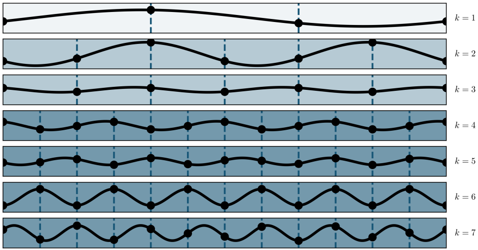
The mesh resolution is chosen approximately dyadically, so that the oscillation of the sinusoids are approximately constant over each interval.



The mesh resolution is chosen approximately dyadically, so that the oscillation of the sinusoids are approximately constant over each interval.

**Thus, it is enough to match the values of $S_{n,a}(t)$ and $S_{n,g}(t)$ at the endpoints of each interval, as marked by the dots.**

We will now focus on coupling in *one scale*, say scale $\ell$.



Focusing just on the cosine terms (the sine part is analogous), that is, we will couple $C_{\ell,a}(t)$ and $C_{\ell,g}(t)$ where

$$C_{\ell,a}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} a_k \cos(2\pi k t), \quad C_{\ell,g}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} g_k \cos(2\pi k t).$$

Focusing just on the cosine terms (the sine part is analogous), that is, we will couple $C_{\ell,a}(t)$ and $C_{\ell,g}(t)$ where

$$C_{\ell,a}(t) = \sum_{k=2^{\ell-1}}^{2^\ell-1} k^{-1/2} a_k \cos(2\pi kt), \quad C_{\ell,g}(t) = \sum_{k=2^{\ell-1}}^{2^\ell-1} k^{-1/2} g_k \cos(2\pi kt).$$

In view of the discretization, let us define vectors $w_1, \ldots, w_N$ (where $N = 2^{\ell-1}$ is the size of the scale) with values given by the cosines at the mesh points. Then, it suffices to couple

$$A_\ell = \sum_{k=2^{\ell-1}}^{2^\ell-1} \frac{1}{\sqrt{k}} w_k a_k, \quad G_\ell = \sum_{k=2^{\ell-1}}^{2^\ell-1} \frac{1}{\sqrt{k}} w_k g_k,$$

such that the error (as a function of $\ell$) is summable, w.h.p.

The coefficient $\frac{1}{\sqrt{k}}$ can be converted into a global $\frac{1}{\sqrt{2^{\ell-1}}}$ coefficient, with no change in behavior, since $k$ varies by at most a constant factor in each scale.

Focusing just on the cosine terms (the sine part is analogous), that is, we will couple $C_{\ell,a}(t)$ and $C_{\ell,g}(t)$ where

$$C_{\ell,a}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} a_k \cos(2\pi kt), \quad C_{\ell,g}(t) = \sum_{k=2^{\ell-1}}^{2^\ell - 1} k^{-1/2} g_k \cos(2\pi kt).$$

In view of the discretization, let us define vectors $w_1, \ldots, w_N$ (where $N = 2^{\ell-1}$ is the size of the scale) with values given by the cosines at the mesh points. Then, it suffices to couple

$$A_\ell = \sum_{k=2^{\ell-1}}^{2^\ell - 1} \frac{1}{\sqrt{k}} w_k a_k, \quad G_\ell = \sum_{k=2^{\ell-1}}^{2^\ell - 1} \frac{1}{\sqrt{k}} w_k g_k,$$

such that the error (as a function of $\ell$) is summable, w.h.p.

The coefficient $\frac{1}{\sqrt{k}}$ can be converted into a global $\frac{1}{\sqrt{2^{\ell-1}}}$ coefficient, with no change in behavior, since $k$ varies by at most a constant factor in each scale.

**This is a high-dimensional CLT problem, where the increments are low-rank.**

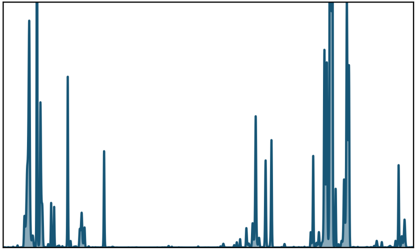
However, the vectors $w_k$ are of dimension $d$ given by the mesh size, which is also $\approx 2^\ell$, same order as that of the number of vectors $N$.

In such a situation, a CLT usually fails. Informally, a CLT requires sufficient "homogenization" in every direction, which cannot happen in general **unless** $d \ll N$. We will go over an example demonstrating this later.

However, the vectors $w_k$ are of dimension $d$ given by the mesh size, which is also $\approx 2^\ell$, same order as that of the number of vectors $N$.

In such a situation, a CLT usually fails. Informally, a CLT requires sufficient "homogenization" in every direction, which cannot happen in general **unless** $d \ll N$. We will go over an example demonstrating this later.

A hint to overcome this is to use the *intermittency of the multiplicative chaos*. As noted earlier in the case of the branching random walk, these measures are supported on $\gamma$-thick points.



**Thick points are polynomially rarer, with their count scaling like** $2^{\ell(1-\gamma^2/2)}$**.**

**The message:** It suffices to couple the processes only near thick points of either field, since the two measures are zero elsewhere.

This last observation enables a reduction to the following generic CLT question:

*Given vectors $v_1, \ldots, v_n \in \mathbb{R}^d$ (restrictions of $w_k$ to a subset of size $d$), and noise variables $\{a_k\}_{k=1}^n$ and $\{g_k\}_{k=1}^n$, can one construct a coupling such that*

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty$$

*is small?* *We need to couple in $L^\infty$ because we wish to control the Radon-Nikodym derivative uniformly across all thick points.*

**Note that in our case, $d$ is the number of thick points, i.e., $d \approx 2^{\ell(1-\gamma^2/2)}$ for level $\ell$, while $n \approx 2^\ell$. Then $d$ is polynomially smaller than $n$, but to cover $\gamma$ close to $0$, the ratio $d/n$ can be $n^\varepsilon$ for arbitrarily small $\varepsilon > 0$.**

# A new high-dimensional CLT

## Theorem ($L^\infty$ coupling CLT, exponential tails – B.R.C.-Ganguly, 2025)

*Let* $v_1, \ldots, v_n$ *be* $n$ *vectors in* $\mathbb{R}^d$ *with* $\|v_k\|_2^2 \leq d$ *for all* $k$ *(in our application, we actually have* $\|v_k\|_\infty \leq 1$ *since* sin *and* cos *are bounded).*

### Theorem ($L^\infty$ coupling CLT, exponential tails – B.R.C.-Ganguly, 2025)

*Let $v_1, \ldots, v_n$ be $n$ vectors in $\mathbb{R}^d$ with $\|v_k\|_2^2 \leq d$ for all $k$ (in our application, we actually have $\|v_k\|_\infty \leq 1$ since* sin *and* cos *are bounded). Suppose $a_1, \ldots, a_n$ are i.i.d. random variables sampled from a centered law $\mathcal{L}$ with unit variance and exponential tails, and let $g_1, \ldots, g_n$ be i.i.d. standard Gaussians.*

## Theorem ($L^\infty$ coupling CLT, exponential tails – B.R.C.-Ganguly, 2025)

*Let* $v_1, \ldots, v_n$ *be* $n$ *vectors in* $\mathbb{R}^d$ *with* $\|v_k\|_2^2 \le d$ *for all* $k$ *(in our application, we actually have* $\|v_k\|_\infty \le 1$ *since* sin *and* cos *are bounded). Suppose* $a_1, \ldots, a_n$ *are i.i.d. random variables sampled from a centered law* $\mathcal{L}$ *with unit variance and exponential tails, and let* $g_1, \ldots, g_n$ *be i.i.d. standard Gaussians. Then there is a coupling of* $\{a_k\}$ *and* $\{g_k\}$ *such that*

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4},$$

*with probability* $\ge 1 - O(n^{-100})$.

## Theorem ($L^\infty$ coupling CLT, exponential tails – B.R.C.-Ganguly, 2025)

*Let* $v_1, \ldots, v_n$ *be* $n$ *vectors in* $\mathbb{R}^d$ *with* $\|v_k\|_2^2 \le d$ *for all* $k$ *(in our application, we actually have* $\|v_k\|_\infty \le 1$ *since* sin *and* cos *are bounded). Suppose* $a_1, \ldots, a_n$ *are i.i.d. random variables sampled from a centered law* $\mathcal{L}$ *with unit variance and exponential tails, and let* $g_1, \ldots, g_n$ *be i.i.d. standard Gaussians. Then there is a coupling of* $\{a_k\}$ *and* $\{g_k\}$ *such that*

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4},$$

*with probability* $\ge 1 - O(n^{-100})$. *Here* $U$ *is defined as*

$$U = \frac{1}{n} \sum_{k=1}^{n} v_k v_k^T .$$

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T .$$

- **Exponential tails.** The result indeed works for any *stretched exponential tail*, with no change to the probability, except the poly log factor. The optimal tail assumption is left as an interesting open question.

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

- **Exponential tails.** The result indeed works for any *stretched exponential tail*, with no change to the probability, except the poly log factor. The optimal tail assumption is left as an interesting open question.
- **Dimension bound $d \le n$.** A simple example can be furnished with $v_i = \sqrt{d} e_{i \bmod d}$. The coordinates are independent. To lower bound the $L^\infty$ norm, observe that the $L^2$ lower bound simply is $\sqrt{d}$ times the $L^2$ error in one-coordinate which is $\sqrt{d/n}$. Fact: (a) The Wasserstein-2 distance factorizes as

$$W_2(\mu \otimes \mu', \nu \otimes \nu') = \sqrt{W_2^2(\mu, \nu) + W_2^2(\mu', \nu')}$$

and (b) a sum of $n$ i.i.d. subexponential variables can be coupled to $\mathcal{N}(0, n)$ with $O(1)$ error, w.h.p. Thus, the overall $L^2$ lower bound is $d/\sqrt{n}$, and consequently the $L^\infty$ error is at least $\sqrt{d/n}$.

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

- $1/4$ is likely not the optimal exponent, $1/2$ is more likely.

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

- $1/4$ is likely not the optimal exponent, $1/2$ is more likely.
- $U$ contains information about the "isotropy" of the vectors. Smaller norms mean that the vectors are spread out.

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

Note that if $\|U\|$ is sub-polynomial, the error is small as soon as $d/n$ is polynomially small, which allows us to cover $\gamma$ close to 0. Recall that in our application, $d \approx 2^{(1-\gamma^2/2)\ell}$ and $n \approx 2^\ell$. We crucially rely on the Fourier structure of $U$ to obtain a polynomial in $\ell$ bound on the spectral norm.

$$\left\| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} v_k g_k \right\|_\infty \lesssim \log^4 n \cdot \left( \frac{d}{n} \cdot (\|U\| + 1) \right)^{1/4}, \quad U = \sum_{k=1}^{n} \frac{1}{n} v_k v_k^T.$$

Note that if $\|U\|$ is sub-polynomial, the error is small as soon as $d/n$ is polynomially small, which allows us to cover $\gamma$ close to 0. Recall that in our application, $d \approx 2^{(1-\gamma^2/2)\ell}$ and $n \approx 2^\ell$. We crucially rely on the Fourier structure of $U$ to obtain a polynomial in $\ell$ bound on the spectral norm.

Further, note that the trivial bound on $\|U\|$ is $\mathrm{tr}(U) = \frac{1}{n} \sum_{i=1}^{n} \|v_i\|^2 \leq d$. Using this bound in our result, we recover the threshold $d^2/n$, which is small, in our application, only when $\gamma > 1$, recovering the result of Kim and Kriechbaum.

Coupling CLTs have been studied extensively in the past, although results in $L^\infty$ are rare. The only one we are aware of is the following result due to Yurinskii, which was used in Kim and Kriechbaum's work, as well as extensively in the broader literature.

Coupling CLTs have been studied extensively in the past, although results in $L^\infty$ are rare. The only one we are aware of is the following result due to Yurinskii, which was used in Kim and Kriechbaum's work, as well as extensively in the broader literature.

### Theorem (Yurinskii coupling, 1978)

*In the same setting as earlier,*

$$\mathbb{P}\left(\left\|\frac{1}{\sqrt{n}}\sum_{k=1}^{n} v_k a_k - \frac{1}{\sqrt{n}}\sum_{k=1}^{n} v_k g_k\right\|_\infty > \delta\right) \lesssim \min_{t\geq 0}\left(\frac{t^2}{(\delta\sqrt{n})^3}\cdot\sum_{k=1}^{n}\|v_k\|_2^2\,\|v_k\|_\infty + f(t)\right),$$

*where* $f(t) = \mathbb{P}\left(\max_{j=1,\ldots,d}|Z_j| > t\right)$ *with* $Z_1,\ldots,Z_d$ *i.i.d. standard Gaussians.*

Specializing to our case with $\|v_k\|_\infty \leq 1$ and $\delta = o(1)$, the effective bound is $\frac{d}{\sqrt{n}}$ (upto polylog factors), which is small only when $d^2 \ll n$.

On the other hand, there has been significant recent progress on $L^2$-coupling CLTs, better known as Wasserstein-2 CLTs, under i.i.d. assumptions:

On the other hand, there has been significant recent progress on $L^2$-coupling CLTs, better known as Wasserstein-2 CLTs, under i.i.d. assumptions:

### Theorem (Zhai '16, Eldan-Mikulincer-Zhai '18)

*Let $X_1, \dots, X_n$ be i.i.d. random variables in $\mathbb{R}^d$ with $\mathbb{E}X_1 = 0, \|X_1\|_2 \leq \beta$, and $\mathrm{Cov}(X_1) = \Sigma$. Then there is a coupling of $X_1, \dots, X_n$ and $Z \sim \mathcal{N}(0, \Sigma)$ such that*

$$\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i - Z \right\|_2 \lesssim \sqrt{\frac{d}{n} \cdot \beta^2 \log n}.$$

On the other hand, there has been significant recent progress on $L^2$-coupling CLTs, better known as Wasserstein-2 CLTs, under i.i.d. assumptions:

## Theorem (Zhai '16, Eldan-Mikulincer-Zhai '18)

*Let $X_1, \ldots, X_n$ be i.i.d. random variables in $\mathbb{R}^d$ with $\mathbb{E}X_1 = 0, \|X_1\|_2 \le \beta$, and $\mathrm{Cov}(X_1) = \Sigma$. Then there is a coupling of $X_1, \ldots, X_n$ and $Z \sim \mathcal{N}(0, \Sigma)$ such that*

$$\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i - Z \right\|_2 \lesssim \sqrt{\frac{d}{n} \cdot \beta^2 \log n}.$$

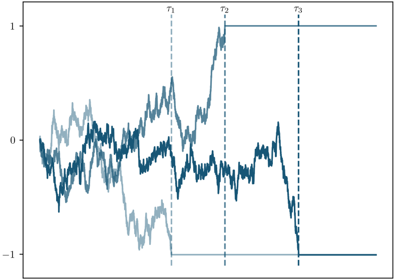
**This result is optimal, up to logarithmic factors.** Similar to an example earlier, a $d$-dimensional random walk with steps $\sqrt{d} e_i$ will suffice.

Zhai's 2016 result uses the Talagrand transportation inequality, combined with a Lindeberg strategy, meanwhile Eldan, Mikulincer, Zhai's 2018 result uses Wasserstein-2 bounds for *martingale embeddings*. Both of these strategies fail in our case, due to the lack of any $L^2$-technology.

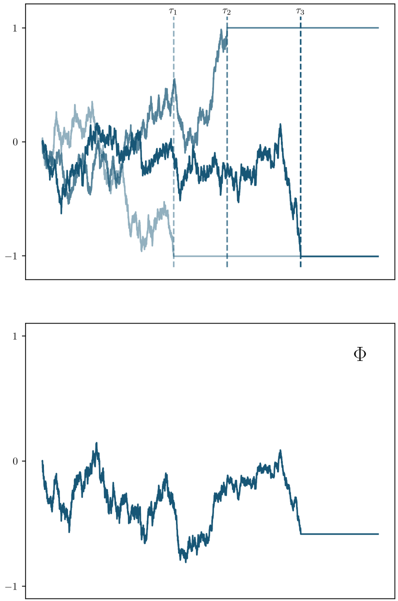
Zhai's 2016 result uses the Talagrand transportation inequality, combined with a Lindeberg strategy, meanwhile Eldan, Mikulincer, Zhai's 2018 result uses Wasserstein-2 bounds for *martingale embeddings*. Both of these strategies fail in our case, due to the lack of any $L^2$-technology.

Also, in general, a Lindeberg-type strategy that does not account for the cancellations induced by different vectors is not expected to succeed. In fact, Yurinskii's 1978 result uses a similar strategy, and as we saw, fails to capture the $d/n$ threshold.

Instead, inspired by Strassen's argument for a one-dimensional CLT, we employ a path-wise approach, by first Skorokhod-embedding the random variables. That is, we consider independent Brownian motions $B^1, \ldots, B^n$, and construct stopping times $\tau_1, \ldots, \tau_n$ such that $B^i_{\tau_i} \sim \mathcal{L}$, the common law of the $a_k$s.



Instead, inspired by Strassen's argument for a one-dimensional CLT, we employ a path-wise approach, by first Skorokhod-embedding the random variables. That is, we consider independent Brownian motions $B^1, \ldots, B^n$, and construct stopping times $\tau_1, \ldots, \tau_n$ such that $B^i_{\tau_i} \sim \mathcal{L}$, the common law of the $a_k$s.




Considering the stopped processes $B^i_{t\wedge\tau_i}$ as shown, we can then represent $n^{-1/2}\sum_{k=1}^n v_k a_k$ as $\Phi_\infty$, where $\Phi_t$ is the process

$$\Phi_t = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} v_k B^k_{t\wedge\tau_k}.$$

In fact, $\Phi$ can be written differentially as

$$d\Phi_t = \sqrt{\Gamma_t}\, dW_t, \quad \Phi_0 = 0,$$

for another Brownian motion $W$ (using a straightforward change of variables), where

$$\Gamma_t = \frac{1}{n} \sum_{k=1}^{n} v_k v_k^T \mathbf{1}_{\tau_k > t}.$$

In fact, $\Phi$ can be written differentially as

$$\mathrm{d}\Phi_t = \sqrt{\Gamma_t}\,\mathrm{d}W_t, \quad \Phi_0 = 0,$$

for another Brownian motion $W$ (using a straightforward change of variables), where

$$\Gamma_t = \frac{1}{n}\sum_{k=1}^{n} v_k v_k^T \mathbf{1}_{\tau_k > t}.$$

It is natural to also consider the “averaged process” $\Psi_t$ given by

$$\mathrm{d}\Psi_t = \sqrt{\mathbb{E}\Gamma_t}\,\mathrm{d}W_t, \quad \Psi_0 = 0.$$

Due to the deterministic covariance structure, $\Psi_\infty$ is a Gaussian vector, and, as easily checked via a quadratic variation computation, $\Psi_\infty$ has the same covariance as $\Phi_\infty$. Thus, we have a coupling between $n^{-1/2}\sum_{k=1}^{n} v_k a_k = \Phi_\infty$ and $n^{-1/2}\sum_{k=1}^{n} v_k g_k = \Psi_\infty$.

The coupling error $\|\Phi_\infty - \Psi_\infty\|_\infty$ can be analyzed by studying the gap process $\Phi_t - \Psi_t$ which admits the differential representation $\left(\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right)\, dW_t$. This reduces to understanding the error

$$\left\|\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right\|, \quad \text{where } \Gamma_t = \frac{1}{n}\sum_{k=1}^{n} v_k v_k^T \mathbf{1}_{\tau_k > t}.$$

The coupling error $\|\Phi_\infty - \Psi_\infty\|_\infty$ can be analyzed by studying the gap process $\Phi_t - \Psi_t$ which admits the differential representation $\left(\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right) \, \mathrm{d}W_t$. This reduces to understanding the error

$$\left\|\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right\|, \quad \text{where } \Gamma_t = \frac{1}{n}\sum_{k=1}^{n} v_k v_k^T \mathbf{1}_{\tau_k > t}.$$

The quantity $\|\Gamma_t - \mathbb{E}\Gamma_t\|$ can be controlled via a standard matrix Bernstein inequality, but the appearance of the matrix square-root poses a challenge.

The coupling error $\|\Phi_\infty - \Psi_\infty\|_\infty$ can be analyzed by studying the gap process $\Phi_t - \Psi_t$ which admits the differential representation $\left(\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right)\, \mathrm{d}W_t$. This reduces to understanding the error

$$\left\|\sqrt{\Gamma_t} - \sqrt{\mathbb{E}\Gamma_t}\right\|, \quad \text{where } \Gamma_t = \frac{1}{n}\sum_{k=1}^{n} v_k v_k^T \mathbf{1}_{\tau_k > t}.$$

The quantity $\|\Gamma_t - \mathbb{E}\Gamma_t\|$ can be controlled via a standard matrix Bernstein inequality, but the appearance of the matrix square-root poses a challenge. Thus, it remains to prove a *perturbation bound for the matrix square-root*, i.e., a bound on quantities of the form:

$$\left\|\sqrt{U+E} - \sqrt{U}\right\|$$

where $U$ is a fixed positive definite matrix, and $E$ is a mean zero, random perturbation.

# Matrix square-roots

While matrix perturbations have been studied extensively in the literature, to the best of our knowledge, every bound on $\left\|\sqrt{U+E}-\sqrt{U}\right\|$ depends on the *smallest eigenvalue of* $U$.

While matrix perturbations have been studied extensively in the literature, to the best of our knowledge, every bound on $\left\|\sqrt{U+E}-\sqrt{U}\right\|$ depends on the *smallest eigenvalue of* $U$. For instance, Schmitt (1992) proves

$$\left\|\sqrt{U+E}-\sqrt{U}\right\| \lesssim \frac{\|E\|}{\sqrt{\lambda_{\min}(U)}},$$

analogous to the Taylor expansion $\left|\sqrt{u+e}-\sqrt{u}\right| \approx \frac{e}{\sqrt{u}}$. But this bound is effective only when $\lambda_{\min}(U)$ is large.

While matrix perturbations have been studied extensively in the literature, to the best of our knowledge, every bound on $\left\|\sqrt{U+E}-\sqrt{U}\right\|$ depends on the *smallest eigenvalue of* $U$. For instance, Schmitt (1992) proves

$$\left\|\sqrt{U+E}-\sqrt{U}\right\| \lesssim \frac{\|E\|}{\sqrt{\lambda_{\min}(U)}},$$

analogous to the Taylor expansion $\left|\sqrt{u+e}-\sqrt{u}\right| \approx \frac{e}{\sqrt{u}}$. But this bound is effective only when $\lambda_{\min}(U)$ is large.

However, in our application, it seems quite difficult to prove, and likely false, that $\lambda_{\min}(U)$ is large. On the other hand, for positive real numbers

$$\left|\sqrt{u+e}-\sqrt{u}\right| \leq \sqrt{e}, \quad \text{always.}$$

While matrix perturbations have been studied extensively in the literature, to the best of our knowledge, every bound on $\left\|\sqrt{U+E}-\sqrt{U}\right\|$ depends on the *smallest eigenvalue of* $U$. For instance, Schmitt (1992) proves

$$\left\|\sqrt{U+E}-\sqrt{U}\right\| \lesssim \frac{\|E\|}{\sqrt{\lambda_{\min}(U)}},$$

analogous to the Taylor expansion $\left|\sqrt{u+e}-\sqrt{u}\right| \approx \frac{e}{\sqrt{u}}$. But this bound is effective only when $\lambda_{\min}(U)$ is large.

However, in our application, it seems quite difficult to prove, and likely false, that $\lambda_{\min}(U)$ is large. On the other hand, for positive real numbers

$$\left|\sqrt{u+e}-\sqrt{u}\right| \leq \sqrt{e}, \quad \text{always.}$$

Therefore, one may speculate an unconditional bound involving $\sqrt{\|E\|}$.

### Lemma (Matrix square root perturbation – B.R.C.-Ganguly, 2025)

*Let $U, E$ be symmetric matrices such that $U$ and $U + E$ are both positive semi-definite. Then*

$$\left\|\sqrt{U+E} - \sqrt{U}\right\| \leq 3\sqrt{\|E\|}.$$

This lemma is key to our proof strategy, but we also expect it to be of independent interest.

Thank you!