High-frequency asymptotics for isotropic fields on the sphere

Giovanni Peccati (Luxembourg University)

We shall focus on **probabilistic approximations**, that is, mathematical statements allowing one to estimate the distance between the laws of two random elements.

We shall focus on **probabilistic approximations**, that is, mathematical statements allowing one to estimate the distance between the laws of two random elements.

Probabilistic approximations are naturally associated with **limit theorems**, like the **central limit theorem** or the **circular and semicircular laws** in random matrix theory, and are one of the leading themes of the whole theory of probability.

Introduction

The aim of this lecture is to explore some recent developments in this area, mainly focussing on **Gaussian approximations.**

Introduction

The aim of this lecture is to explore some recent developments in this area, mainly focussing on **Gaussian approximations.**

Strong motivations come from the asymptotic analysis of random fields defined on **homogenous spaces**, in connection with the statistical analysis of cosmological data.

The aim of this lecture is to explore some recent developments in this area, mainly focussing on **Gaussian approximations.**

Strong motivations come from the asymptotic analysis of random fields defined on **homogenous spaces**, in connection with the statistical analysis of cosmological data.

Behind the scenes: tools from classical **Gaussian analysis**, **group representations**, **information theory**, and many other probabilistic techniques. Joint works with several coauthors (D. Marinucci, I. Nourdin, D. Nualart, Y. Swan ...).

Introductory example

The following example is classically motivated by cosmological data analysis.

- We consider a (centered) **Gaussian field** on $\mathbb{S}^2$:

$$T = \{T(x) : x \in \mathbb{S}^2\}$$

Introductory example

The following example is classically motivated by cosmological data analysis.

- We consider a (centered) **Gaussian field** on $\mathbb{S}^2$:

$$T = \{T(x) : x \in \mathbb{S}^2\}$$

- We assume that T is **isotropic**, that is:

$$T(x) \overset{\text{LAW}}{=} T(gx),$$

for every rotation $g \in SO(3)$.

Introductory example

The following example is classically motivated by cosmological data analysis.

- We consider a (centered) **Gaussian field** on $\mathbb{S}^2$:

$$T = \{T(x) : x \in \mathbb{S}^2\}$$

- We assume that T is **isotropic**, that is:

$$T(x) \overset{\text{LAW}}{=} T(gx),$$

for every rotation $g \in SO(3)$.

- The collection of **spherical harmonics** is written:

$$\{Y_{lm} : l \geq 0, \; m = -l, ..., l\}.$$

Introductory example

Fact 1

The field T can be decomposed as:

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} \times Y_{lm}(x) \quad ; \quad a_{lm} := \int_{\mathbb{S}^2} T(z) \overline{Y_{lm}(z)} dz.$$

Introductory example

Fact 1

The field T can be decomposed as:

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} \times Y_{lm}(x) \quad ; \quad a_{lm} := \int_{\mathbb{S}^2} T(z) \overline{Y_{lm}(z)} dz.$$

Remark

A physical interpretation: any device measuring T with an angular resolution of $180^\circ / L_{MAX}$ can roughly observe the truncated harmonic development of T, namely $\sum_{l=0}^{L_{MAX}} \sum_{m=-l}^{l} a_{lm} Y_{lm}$.

Introductory example

Fact 2

The quantity $C_l = E\left|a_{lm}\right|^2$ *depends uniquely on* l *(and not on* m*).*

Introductory example

Fact 2

The quantity $C_l = E\,|a_{lm}|^2$ *depends uniquely on* l *(and not on* m*).*

Fact 3

Under Gaussianity+Isotropy, the law of T *is completely determined by the* ***power spectrum*** $\{C_l : l \geq 0\}$*. (See Baldi and Marinucci (2010) for a sort of 'reverse' statement)*

Now consider a well-behaved function F (interesting case: F is non linear, for instance F is a Hermite polynomial), and define the **Gaussian subordinated** (aka 'Gaussian-related') field:

$$F[T](x) := F(T(x)).$$

Introductory example

Now consider a well-behaved function F (interesting case: F is non linear, for instance F is a Hermite polynomial), and define the **Gaussian subordinated** (aka ‘Gaussian-related’) field:

$$F[T](x) := F(T(x)).$$

For instance, noting $c = E[T(0)^2]$,

$$F[T(x)] = T(x) + \alpha_1(T(x)^2 - c) + \alpha_2(T(x)^3 - 3cT(x)).$$

Introductory example

The process $F[T]$ is, in general not Gaussian, and the associated harmonic coefficients become genuine **'non-dogs'** (they lose Gaussianity *and* independence).

Introductory example

The process $F[T]$ is, in general not Gaussian, and the associated harmonic coefficients become genuine **'non-dogs'** (they lose Gaussianity *and* independence).

More precisely, one has that $F[T]$ is always isotropic and admits the harmonic expansion

$$F[T](x) = \sum_{l \geq 0} \left\{ \sum_{m=-l}^{l} a_{lm}(F) Y_{lm}(x) \right\} := \sum_{l \geq 0} F[T]_l(x),$$

with

$$a_{lm}(F) = \int_{\mathbb{S}^2} F(T(x)) \overline{Y_{lm}(x)} dx.$$

The field $F[T]_l$ is the lth **frequency component** of $F[T]$.

Introductory example

Problem (Marinucci and P., 2011)

Find conditions on F and on the power spectrum $\{C_l : l \geq 0\}$ in order to have that the (normalized) frequency components $F[T]_l$ converge to a Gaussian limit when $l \to \infty$ (at least, componentwise). Quantify the speed of convergence in some appropriate distance.

The phenomenon we are looking for is called a **high-resolution** (or **high-frequency**) Central Limit Theorem: connected e.g. to hypothesis testing and max-likelihood estimation.

Remark

In Marinucci and Peccati (2010), it is proved that, in many models, high-frequency asymptotic Gaussianity implies **ergodicity**, *that is,*

$$E\left[\left(\frac{\hat{C}_l}{C(F)_l}-1\right)^2\right]\longrightarrow 0,$$

where $C(F)$ is the power spectrum of $F[T]$, and

$$\hat{C}_l=\frac{1}{2l+1}\sum_{m=-l}^{l}|a_{lm}(F)|^2.$$

Features

- The problem involves non-linear functionals of an infinite-dimensional random field (in this case, Gaussian).

Features

- The problem involves non-linear functionals of an infinite-dimensional random field (in this case, Gaussian).
- There is no clear “time arrow”.

Features

- The problem involves non-linear functionals of an infinite-dimensional random field (in this case, Gaussian).
- There is no clear “time arrow”.
- The problem cannot be reduced to the classical central limit theorem (sum of independent random variables).

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Discovered in 1965 by Penzias and Wilson, the CMB is a microwave radiation bathing the earth from every direction, at a temperature of around 2.73 K. Its tiny variations are called **anisotropies.**

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Discovered in 1965 by Penzias and Wilson, the CMB is a microwave radiation bathing the earth from every direction, at a temperature of around 2.73 K. Its tiny variations are called **anisotropies.**
- It is a relic radiation, providing maps of the primordial Universe **300.000 years after the Big Bang**, before the formation of any of the current structures.

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Discovered in 1965 by Penzias and Wilson, the CMB is a microwave radiation bathing the earth from every direction, at a temperature of around 2.73 K. Its tiny variations are called **anisotropies.**
- It is a relic radiation, providing maps of the primordial Universe **300.000 years after the Big Bang**, before the formation of any of the current structures.
- If one associates with each $x \in \mathbb{S}^2$ a direction of measurement of the CMB, one obtains a spherical field:

$$T^*(x), \quad x \in \mathbb{S}^2.$$

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Some cosmological models imply that the standardized field

$$T^{**}(x) = \frac{T^{*}(x) - \overline{T^{*}}}{\overline{T^{*}}} \quad (\overline{T^{*}} \text{ is the mean over } \mathbb{S}^2)$$

is indeed the realization of a **Gaussian or close to Gaussian isotropic random field** (the isotropy comes from the cosmological principle).

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Some cosmological models imply that the standardized field

$$T^{**}(x) = \frac{T^*(x) - \overline{T^*}}{\overline{T^*}} \quad (\overline{T^*} \text{ is the mean over } \mathbb{S}^2)$$

is indeed the realization of a **Gaussian or close to Gaussian isotropic random field** (the isotropy comes from the cosmological principle).
- There is just **a single observation** of the CMB (there is just one Universe)

Motivations from Cosmology (Cosmic Microwave Background Radiation)

- Some cosmological models imply that the standardized field

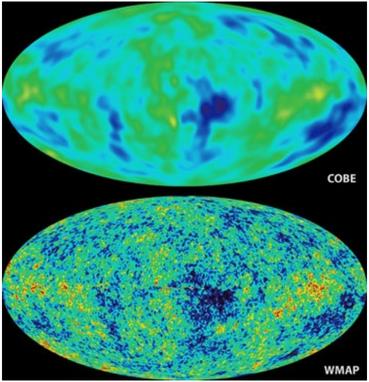
$$T^{**}(x) = \frac{T^*(x) - \overline{T^*}}{\overline{T^*}} \quad (\overline{T^*} \text{ is the mean over } \mathbb{S}^2)$$

  is indeed the realization of a **Gaussian or close to Gaussian isotropic random field** (the isotropy comes from the cosmological principle).
- There is just **a single observation** of the CMB (there is just one Universe)
- Each satellite experiment can observe $T^{**}(x)$ **up to a certain angular resolution** $180^\circ / L_{\mathrm{MAX}}$, so that each observed CMB map has the form

$$T^{**}(x) \simeq \sum_{l=1}^{L_{\mathrm{MAX}}} T_l^{**}(x) .$$

Motivations from Cosmology (Cosmic Microwave Background Radiation)

COBE (1993; $L_{MAX} \sim 20$) and WMAP (2003; $L_{MAX} \sim 600/800$)



Formalisation of the problem

Let $X = \{X_i : i \in I\}$ be a (normalised) Gaussian family. And let

$$F = (F_1, ..., F_d)$$

be a vector of random variables such that each F_k belongs to the linear span of polynomial random variables of the type $P(X_{i_1}, ..., X_{i_m})$, where $d^0 P \leq M$.

Formalisation of the problem

Let $X = \{X_i : i \in I\}$ be a (normalised) Gaussian family. And let

$$F = (F_1, ..., F_d)$$

be a vector of random variables such that each F_k belongs to the linear span of polynomial random variables of the type $P(X_{i_1}, ..., X_{i_m})$, where $d^0 P \leq M$. Let

$$Z = (Z_1, ..., Z_d)$$

be a Gaussian vector with the same covariance structure as F.

Formalisation of the problem

Let $X = \{X_i : i \in I\}$ be a (normalised) Gaussian family. And let

$$F = (F_1, ..., F_d)$$

be a vector of random variables such that each F_k belongs to the linear span of polynomial random variables of the type $P(X_{i_1}, ..., X_{i_m})$, where $d^0 P \leq M$. Let

$$Z = (Z_1, ..., Z_d)$$

be a Gaussian vector with the same covariance structure as F.

Problem. Find (sharp) estimates on the following crucial quantities: (1) The **total variation distance** between F and Z, (2) The **1-Wasserstein distance** between F and Z, and (3) the **relative entropy** of F.

Recall (assuming F and Z have densities, say f_F and f_Z):

- The **total variation distance** is

$$\begin{aligned} d_{TV}(F,Z) &= \sup |P[F \in A] - P[Z \in A]| \\ &= \frac{1}{2}\int_{\mathbb{R}^d} |f_F(x) - f_Z(x)|\, dx. \end{aligned}$$

Recall (assuming F and Z have densities, say f_F and f_Z):

- The **total variation distance** is

$$\begin{aligned} d_{TV}(F,Z) &= \sup |P[F \in A] - P[Z \in A]| \\ &= \frac{1}{2}\int_{\mathbb{R}^d} |f_F(x) - f_Z(x)|\, dx. \end{aligned}$$

- The **1-Wasserstein distance** is

$$d_W(F,Z) = \sup_{\|h'\|_\infty \le 1} |E[h(F)] - E[h(Z)]|$$

Recall (assuming F and Z have densities, say f_F and f_Z):

- The **total variation distance** is

$$\begin{aligned} d_{TV}(F,Z) &= \sup |P[F \in A] - P[Z \in A]| \\ &= \frac{1}{2}\int_{\mathbb{R}^d} |f_F(x) - f_Z(x)|\, dx. \end{aligned}$$

- The **1-Wasserstein distance** is

$$d_W(F,Z) = \sup_{\|h'\|_\infty \le 1} |E[h(F)] - E[h(Z)]|$$

- The **relative entropy** of F is

$$D(F\|Z) = \int_{\mathbb{R}^d} f_F(x) \log \frac{f_F(x)}{f_Z(x)} dx \ge 2 d_{TV}(F,Z)^2,$$

where the last relation is known as **Pinsker inequality.**

Related research

- Remarkable upper bounds for Gaussian approximations (among hundreds!): **Lindberg** (1923), **Berry** (1941), **Esseen** (1942), **Trotter** (1950), **Stein** (1971, 1984), **Barron** (1984), **Bolthausen** (1984), **Arstein, Ball, Barthe, Naor** (2005).

Related research

- Remarkable upper bounds for Gaussian approximations (among hundreds!): **Lindberg** (1923), **Berry** (1941), **Esseen** (1942), **Trotter** (1950), **Stein** (1971, 1984), **Barron** (1984), **Bolthausen** (1984), **Arstein, Ball, Barthe, Naor** (2005).
- Apart from **limit theorems on homogeneous spaces**, applications e.g. to: **fractional processes**, **Gaussian polymers, density estimates**, **random matrices**, **stochastic geometry** (random geometric graphs, tesselations).

Related research

- This general framework also includes non-linear functionals of **needlet coefficients** (Baldi, Kerchyakarian, Marinucci, Picard (2009ab), Lan and Marinucci (2009, 2010), ...), as well as estimates for **empirical polyspectra** of spherical Gaussian fields, like e.g. (up to constants).

Related research

- This general framework also includes non-linear functionals of **needlet coefficients** (Baldi, Kerchyakarian, Marinucci, Picard (2009ab), Lan and Marinucci (2009, 2010), ...), as well as estimates for **empirical polyspectra** of spherical Gaussian fields, like e.g. (up to constants).
- Our techniques are based on suitable generalisations of the **Poincaré inequality** and of the **de Bruijn's identity** of information theory.

Related research

- This general framework also includes non-linear functionals of **needlet coefficients** (Baldi, Kerchyakarian, Marinucci, Picard (2009ab), Lan and Marinucci (2009, 2010), ...), as well as estimates for **empirical polyspectra** of spherical Gaussian fields, like e.g. (up to constants).
- Our techniques are based on suitable generalisations of the **Poincaré inequality** and of the **de Bruijn's identity** of information theory.
- Basic message: Gaussianity emerges from the **reduction of kurtosis**!

Dimension 1 and the Poincaré inequality

- Let us focus on the following (much simpler) setting: let $X = (X_1, ..., X_k)$ be a Gaussian vector with independent $\mathcal{N}(0,1)$ components, and consider a smooth mapping $f : \mathbb{R}^k \to \mathbb{R}$. We want to measure the distance between the law of $f(X)$ and that of a Gaussian random variable.

- Let us focus on the following (much simpler) setting: let $X = (X_1, ..., X_k)$ be a Gaussian vector with independent $\mathcal{N}(0, 1)$ components, and consider a smooth mapping $f : \mathbb{R}^k \to \mathbb{R}$. We want to measure the distance between the law of $f(X)$ and that of a Gaussian random variable.
- The famous **Poincaré inequality** states that

$$\mathbf{Var}(f(X)) = Ef(X)^2 - E^2 f(X) \leq E\|\nabla f(X)\|^2_{\mathbb{R}^k},$$

providing a rough measure of the **concentration** of $f(X)$ around its mean.

Dimension 1 and the Poincaré inequality

- Let us focus on the following (much simpler) setting: let $X = (X_1, ..., X_k)$ be a Gaussian vector with independent $\mathcal{N}(0,1)$ components, and consider a smooth mapping $f : \mathbb{R}^k \to \mathbb{R}$. We want to measure the distance between the law of $f(X)$ and that of a Gaussian random variable.
- The famous **Poincaré inequality** states that

$$\mathbf{Var}(f(X)) = Ef(X)^2 - E^2 f(X) \leq E\|\nabla f(X)\|^2_{\mathbb{R}^k},$$

  providing a rough measure of the **concentration** of $f(X)$ around its mean.
- Question: can we measure the distance between the law of $f(X)$ and the law of a Gaussian random variable by controlling the magnitude of quantities related to

$$\mathbf{Hess}\, f(X) = \left(\frac{\partial^2}{\partial x_i \partial x_j} f(X) \right)_{i,j=1,...,d} ?$$

Second order Poincaré

The use of a powerful probabilistic technique, known as the **Stein's method**, allows one to deduce the following bound:

Proposition

Let $Z \sim \mathcal{N}(0,1)$ be centered and have unit variance. Then:

$$d_{TV}(f(X), Z) \leq \sqrt{10} \times E[\|\mathbf{Hess} f(X)\|_{op}^4]^{1/4} \times E[\|\nabla f(X)\|_{\mathbb{R}^k}^4]^{1/4}.$$

The use of a powerful probabilistic technique, known as the **Stein's method**, allows one to deduce the following bound:

Proposition

Let $Z \sim \mathcal{N}(0,1)$ *be centered and have unit variance. Then:*

$$d_{TV}(f(X), Z) \leq \sqrt{10} \times E[\|\mathbf{Hess} f(X)\|_{op}^4]^{1/4} \times E[\|\nabla f(X)\|_{\mathbb{R}^k}^4]^{1/4}.$$

This was first proved by Chatterjee (2007), and then generalised to an infinite-dimensional setting by Nourdin, Peccati and Reinert (2010).

Fourh Moment Theorem

These bounds take a very neat form when one deals with polynomial random variables. Denote by $\mathcal{H}_q$ the span (called the qth **Wiener chaos**) of random variables of the type $H_q(X_i)$, where H_q is the qth Hermite polynomial.

Fourh Moment Theorem

These bounds take a very neat form when one deals with polynomial random variables. Denote by $\mathcal{H}_q$ the span (called the qth **Wiener chaos**) of random variables of the type $H_q(X_i)$, where H_q is the qth Hermite polynomial.

Theorem (Nourdin & Peccati, 2009)

For $q \geq 2$, let $F \in \mathcal{H}_q$ have unit variance, and $Z \sim \mathcal{N}(0,1)$. Then

$$d_{TV}(F,Z) \leq \frac{2}{\sqrt{3}}\sqrt{E[F^4] - E[Z^4]} = \frac{2}{\sqrt{3}}\sqrt{E[F^4] - 3}$$

Fourh Moment Theorem

These bounds take a very neat form when one deals with polynomial random variables. Denote by $\mathcal{H}_q$ the span (called the qth **Wiener chaos**) of random variables of the type $H_q(X_i)$, where H_q is the qth Hermite polynomial.

Theorem (Nourdin & Peccati, 2009)

For $q \geq 2$, let $F \in \mathcal{H}_q$ have unit variance, and $Z \sim \mathcal{N}(0,1)$. Then

$$d_{TV}(F,Z) \leq \frac{2}{\sqrt{3}}\sqrt{E[F^4] - E[Z^4]} = \frac{2}{\sqrt{3}}\sqrt{E[F^4] - 3}$$

This statement contains findings by **Nualart and Peccati** (2005) and **Nualart and Ortiz-Latorre** (2008).

Fourh Moment Theorem

These bounds take a very neat form when one deals with polynomial random variables. Denote by $\mathcal{H}_q$ the span (called the qth **Wiener chaos**) of random variables of the type $H_q(X_i)$, where H_q is the qth Hermite polynomial.

Theorem (Nourdin & Peccati, 2009)

For $q \geq 2$, let $F \in \mathcal{H}_q$ have unit variance, and $Z \sim \mathcal{N}(0,1)$. Then

$$d_{TV}(F,Z) \leq \frac{2}{\sqrt{3}}\sqrt{E[F^4]-E[Z^4]} = \frac{2}{\sqrt{3}}\sqrt{E[F^4]-3}$$

This statement contains findings by **Nualart and Peccati** (2005) and **Nualart and Ortiz-Latorre** (2008).

Computing the moments of chaotic random variables is done by resorting to the enumeration of **Feynman-type diagrams**. Some surprising implications of this fact emerge when studying high-frequency limit theorems.

Wasserstein bounds in any dimension

Let $F = (F_1, ..., F_d)$ be a normalised vector of uncorrelated random variables, such that F_l belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. The following estimate follows from Nourdin and Rosinski (2012), and is based on estimates by Nourdin, Peccati and Réveillac (2010).

Wasserstein bounds in any dimension

Let $F = (F_1, ..., F_d)$ be a normalised vector of uncorrelated random variables, such that F_l belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. The following estimate follows from Nourdin and Rosinski (2012), and is based on estimates by Nourdin, Peccati and Réveillac (2010).

Theorem

Let $Z = (Z_1, ..., Z_d)$ be a vector of i.i.d. $\mathcal{N}(0, 1)$ random variables. Then,

$$d_W(F, Z) \leq d^{1/2}\sqrt{E[\|F\|^4] - E[\|Z\|^4]},$$

where $\|\cdot\|$ stands for the Euclidean norm.

Let $F = (F_1, ..., F_d)$ be a normalised vector of uncorrelated random variables, such that F_l belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. The following estimate follows from Nourdin and Rosinski (2012), and is based on estimates by Nourdin, Peccati and Réveillac (2010).

Theorem

Let $Z = (Z_1, ..., Z_d)$ be a vector of i.i.d. $\mathcal{N}(0, 1)$ random variables. Then,

$$d_W(F, Z) \leq d^{1/2}\sqrt{E[\|F\|^4] - E[\|Z\|^4]},$$

where $\|\cdot\|$ stands for the Euclidean norm.

Fact: most techniques 'break down' when one wants to estimate the multidimensional total variation distance.

Entropic estimates

Let $F_n = (F_{1,n}, ..., F_{d,n})$ be a sequence of normalised vectors, such that $F_{l,n}$ belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. Let Z_n be a sequence of random vectors with the same covariance as F_n, and write

$$\Delta_n := E[\|F_n\|^4] - E[\|Z_n\|^4]$$

Entropic estimates

Let $F_n = (F_{1,n}, ..., F_{d,n})$ be a sequence of normalised vectors, such that $F_{l,n}$ belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. Let Z_n be a sequence of random vectors with the same covariance as F_n, and write

$$\Delta_n := E[\|F_n\|^4] - E[\|Z_n\|^4]$$

Theorem (Nourdin, Peccati & Swan, 2013)

Assume that F_n has a density for n large enough and also that $\Delta_n \to 0$.

Entropic estimates

Let $F_n = (F_{1,n}, ..., F_{d,n})$ be a sequence of normalised vectors, such that $F_{l,n}$ belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. Let Z_n be a sequence of random vectors with the same covariance as F_n, and write

$$\Delta_n := E[\|F_n\|^4] - E[\|Z_n\|^4]$$

Theorem (Nourdin, Peccati & Swan, 2013)

Assume that F_n has a density for n large enough and also that $\Delta_n \to 0$. Then,

$$D(F_n \| Z_n) = O(1)\,\Delta_n |\log \Delta_n|, \quad n \to \infty.$$

Entropic estimates

Let $F_n = (F_{1,n}, ..., F_{d,n})$ be a sequence of normalised vectors, such that $F_{l,n}$ belongs to the span of some $\{H_{q_l}(X_i)\}$ for every $l = 1, ..., d$. Let Z_n be a sequence of random vectors with the same covariance as F_n, and write

$$\Delta_n := E[\|F_n\|^4] - E[\|Z_n\|^4]$$

Theorem (Nourdin, Peccati & Swan, 2013)

Assume that F_n has a density for n large enough and also that $\Delta_n \to 0$. Then,

$$D(F_n\|Z_n) = O(1)\,\Delta_n|\log \Delta_n|, \quad n \to \infty.$$

Based on a suitable reformulation of **de Bruijn's formula** of information theory, providing a representation of the derivative of the entropy along the Ornstein-Uhlenbeck semigroup, in terms of **Fisher information**.

Back to the initial example

Let $T = \{T(x) : x \in \mathbb{S}^2\}$ be an isotropic Gaussian field on the sphere (centered, with unit variance), with power spectrum $\{C_l : l \geq 0\}$.

Back to the initial example

Let $T = \{T(x) : x \in \mathbb{S}^2\}$ be an isotropic Gaussian field on the sphere (centered, with unit variance), with power spectrum $\{C_l : l \geq 0\}$.

For $q \geq 2$, let H_q be the qth Hermite polynomial, and define the subordinated field

$$H_q[T](x) = H_q(T(x)) = \sum_{l \geq 0} \sum_{m=-l}^{l} a_{lm}(H_q) Y_{lm}(x) = \sum_{l \geq 0} H_q[T]_l(x).$$

Back to the initial example

Let $T = \{T(x) : x \in \mathbb{S}^2\}$ be an isotropic Gaussian field on the sphere (centered, with unit variance), with power spectrum $\{C_l : l \geq 0\}$.
For $q \geq 2$, let H_q be the qth Hermite polynomial, and define the subordinated field

$$H_q[T](x) = H_q(T(x)) = \sum_{l \geq 0} \sum_{m=-l}^{l} a_{lm}(H_q) Y_{lm}(x) = \sum_{l \geq 0} H_q[T]_l(x).$$

Fix $x \in \mathbb{S}^2$. Let $v_l^2 = \mathbf{Var}(H_q[T]_l(x))$, and $Z \sim \mathcal{N}(0, 1)$.

Back to the initial example

Let $T = \{T(x) : x \in \mathbb{S}^2\}$ be an isotropic Gaussian field on the sphere (centered, with unit variance), with power spectrum $\{C_l : l \geq 0\}$.
For $q \geq 2$, let H_q be the qth Hermite polynomial, and define the subordinated field

$$H_q[T](x) = H_q(T(x)) = \sum_{l \geq 0} \sum_{m=-l}^{l} a_{lm}(H_q) Y_{lm}(x) = \sum_{l \geq 0} H_q[T]_l(x).$$

Fix $x \in \mathbb{S}^2$. Let $v_l^2 = \mathbf{Var}(H_q[T]_l(x))$, and $Z \sim \mathcal{N}(0,1)$. We want to estimate the RHS of the inequality

$$d_{TV}(\widetilde{H_q}[T]_l(x), Z) \leq \frac{2}{\sqrt{3}} \sqrt{\mathbb{E}[\widetilde{H_q}[T](x)^4] - 3}.$$

An application of the diagrams/moments connection shows that the quantity $\mathbb{E}[\widetilde{H_q}[T]_l(x)^4]-3$ is indeed a convolution of **generalized Gaunt integrals** of the type

$$\int_{\mathbb{S}^2} Y_{l_1 m_1}(z)\cdots Y_{l_q m_q}(z)\,\overline{Y_{lm}(z)}dz.$$

These integrals can be computed in terms of convolutions of the elements of unitary **Clebsch-Gordan matrices** $\mathbf{C}_{l_1 l_2}$, defined by the relation,

$$\mathbf{D}_{l_1}\otimes\mathbf{D}_{l_2}=\mathbf{C}_{l_1 l_2}\left[\bigoplus_{l=|l_1-l_2|}^{l_1+l_2}\mathbf{D}_l\right]\mathbf{C}^*_{l_1 l_2},$$

where $\mathbf{D}_l$, $l\geq 0$, indicates the lth **Wigner matrix**.

Algebraic/exponential dualities

Some checkable conditions:

Theorem (Marinucci & Peccati, 2008)

One observes the basic duality:

- *If* $C_l \sim l^{\alpha} \exp(-\beta l)$*, then a high-frequency CLT takes place (with a speed of convergence of the order* $l^{-1/2}$*).*
- *If* $C_l \sim l^{-\gamma}$*, then a high-frequency CLT does not hold.*

Algebraic/exponential dualities

Some checkable conditions:

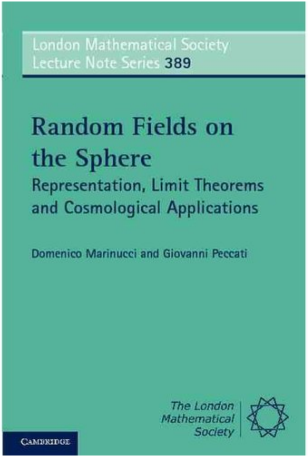
Theorem (Marinucci & Peccati, 2008)

One observes the basic duality:

- *If* $C_l \sim l^{\alpha} \exp(-\beta l)$*, then a high-frequency CLT takes place (with a speed of convergence of the order* $l^{-1/2}$*).*
- *If* $C_l \sim l^{-\gamma}$*, then a high-frequency CLT does not hold.*

Also: the same method yields estimates on the Gaussian approximation of empirical bispectra, with a speed of convergence of the order of $\ell^{-1/2}$.

Advertising, part I



Advertising, part II