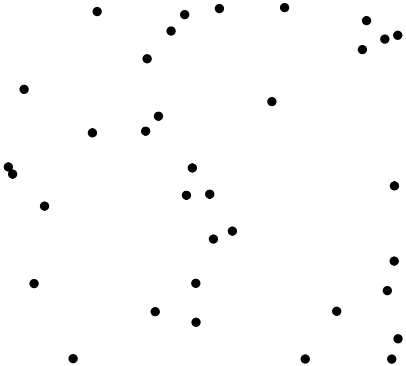
DFG FORSCHERGRUPPE
Geometry & Physics of Spatial Random Systems

Central limit theorems for random tessellations and random graphs

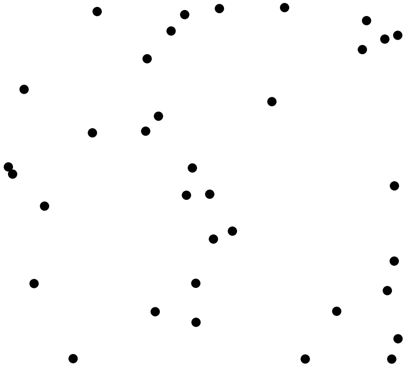
Matthias Schulte

Poisson process in $[0,1]^d$



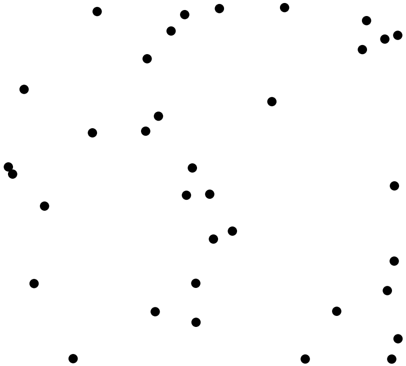
- $(X_i)_{1 \leq i \leq M}$ with independent $X_1, X_2, \ldots \sim \text{Uniform}([0,1]^d)$ and $M \sim \text{Poisson}(t)$, $t \geq 0$, i.e. $\mathbb{P}(M = k) = \frac{t^k}{k!} e^{-t}$, $k \in \mathbb{N} \cup \{0\}$.

Poisson process in $[0,1]^d$



- $(X_i)_{1 \leq i \leq M}$ with independent $X_1, X_2, \dots \sim \text{Uniform}([0,1]^d)$ and $M \sim \text{Poisson}(t)$, $t \geq 0$, i.e. $\mathbb{P}(M = k) = \frac{t^k}{k!} e^{-t}$, $k \in \mathbb{N} \cup \{0\}$.
- Define $\eta = \sum_{i=1}^{M} \delta_{X_i}$, where δ_x is the Dirac measure at $x \in \mathbb{R}^d$, i.e., $\eta(A)$ is the number of points of $(X_i)_{1 \leq i \leq M}$ in $A \in \mathcal{B}(\mathbb{R}^d)$.

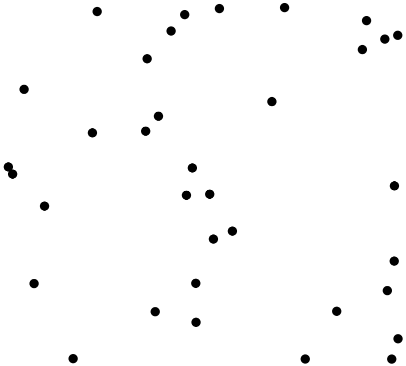
Poisson process in $[0, 1]^d$



Observe that

- $\eta(A_1), \dots, \eta(A_n)$ independent for disjoint $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$

Poisson process in $[0,1]^d$



Observe that

- $\eta(A_1), \dots, \eta(A_n)$ independent for disjoint $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$
- $\eta(A) \sim \text{Poisson}(t \, \text{Vol}(A \cap [0,1]^d))$, $A \in \mathcal{B}(\mathbb{R}^d)$

Poisson process

Karlsruhe Institute of Technology

Definition:

A random counting measure η on a measurable space $(\mathbb{X}, \mathcal{X})$ is a Poisson process with σ-finite intensity measure λ if

- $\eta(A_1), \dots, \eta(A_n)$ are independent for all disjoint sets $A_1, \dots, A_n \in \mathcal{X}$, $n \in \mathbb{N}$,
- $\eta(A)$ is Poisson distributed with parameter $\lambda(A)$ for all $A \in \mathcal{X}$.

Poisson process

Karlsruhe Institute of Technology

Definition:

A random counting measure η on a measurable space $(\mathbb{X}, \mathcal{X})$ is a Poisson process with σ-finite intensity measure λ if

- $\eta(A_1), \ldots, \eta(A_n)$ are independent for all disjoint sets $A_1, \ldots, A_n \in \mathcal{X}$, $n \in \mathbb{N}$,
- $\eta(A)$ is Poisson distributed with parameter $\lambda(A)$ for all $A \in \mathcal{X}$.

In the following we identify η with its support and think of it as a random configuration of points.

Poisson process

Karlsruhe Institute of Technology

Definition:

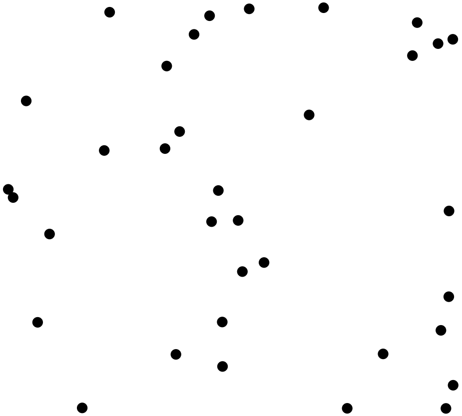
A random counting measure η on a measurable space $(\mathbb{X}, \mathcal{X})$ is a Poisson process with σ-finite intensity measure λ if

- $\eta(A_1), \dots, \eta(A_n)$ are independent for all disjoint sets $A_1, \dots, A_n \in \mathcal{X}$, $n \in \mathbb{N}$,
- $\eta(A)$ is Poisson distributed with parameter $\lambda(A)$ for all $A \in \mathcal{X}$.

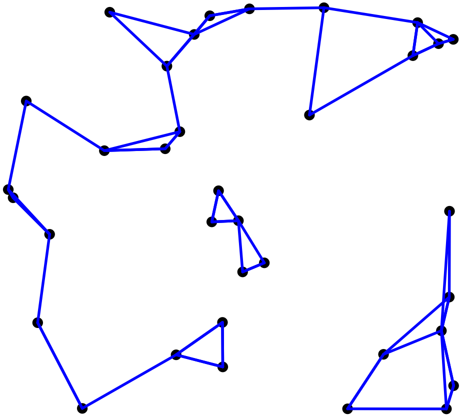
In the following we identify η with its support and think of it as a random configuration of points.

Example:
$\mathbb{X} = \mathbb{R}^d$, $\lambda = t\,\mathrm{Vol}$, $t \geq 0$: stationary Poisson process of intensity t in $\mathbb{R}^d$

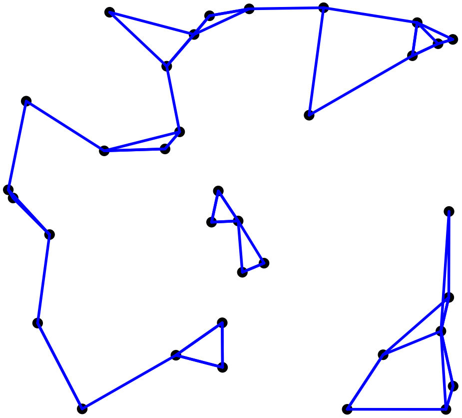
k-Nearest Neighbour Graph



k-Nearest Neighbour Graph

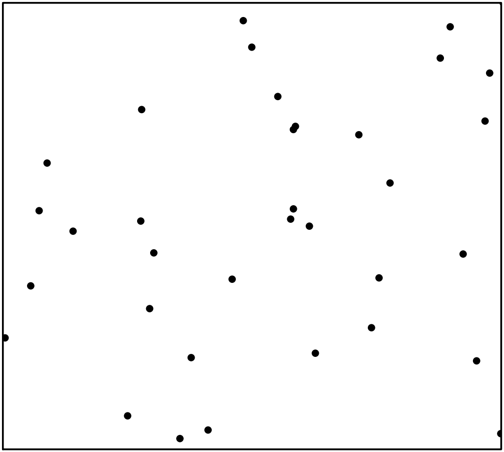


k-Nearest Neighbour Graph

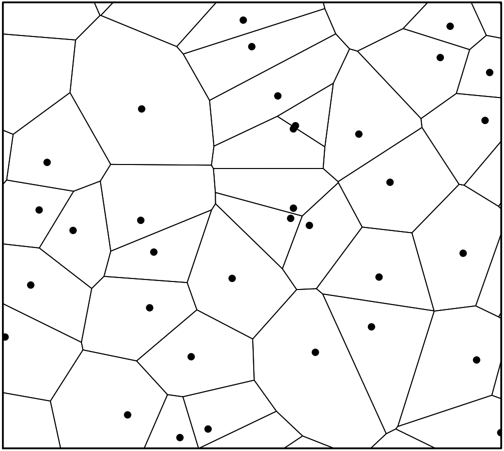


What is the edge length of the *k*-nearest neighbour graph?

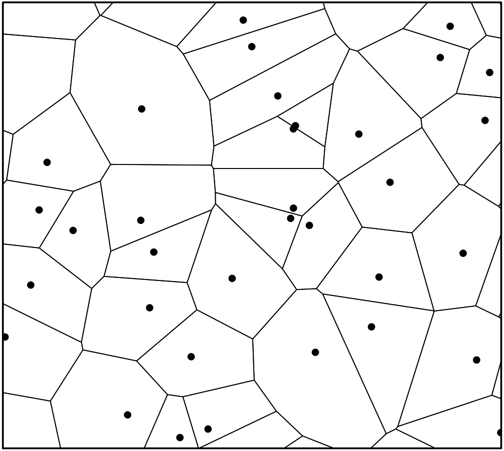
Poisson-Voronoi tessellation



Poisson-Voronoi tessellation



Poisson-Voronoi tessellation



What is the edge length of the Poisson-Voronoi tessellation within the observation window?

Classical central limit Theorem

Theorem:

Let $(Y_i)_{i\in\mathbb{N}}$ be i.i.d. random variables with $\mathbb{E}Y_1^2 < \infty$, let $S_n = \sum_{i=1}^n Y_i$, $n \in \mathbb{N}$, and let N be a standard Gaussian random variable, i.e.,

$$\mathbb{P}(N \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)\, du, \quad x \in \mathbb{R}.$$

Then

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \to N \quad \text{in distribution as} \quad n \to \infty,$$

that is,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \leq x\right) = \mathbb{P}(N \leq x), \quad x \in \mathbb{R}.$$

Classical central limit Theorem

Theorem:

Let $(Y_i)_{i\in\mathbb{N}}$ be i.i.d. random variables with $\mathbb{E}Y_1^2 < \infty$, let $S_n = \sum_{i=1}^n Y_i$, $n \in \mathbb{N}$, and let N be a standard Gaussian random variable, i.e.,

$$\mathbb{P}(N \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)\, du, \quad x \in \mathbb{R}.$$

Then

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \to N \quad \text{in distribution as} \quad n \to \infty,$$

that is,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}} \leq x \right) = \mathbb{P}(N \leq x), \quad x \in \mathbb{R}.$$

Does something similar hold for the edge length of the k-nearest neighbour graph or the Poisson-Voronoi tessellation?

Probability distances

For two random variables X_1 and X_2 we define the Kolmogorov distance

$$d_K(X_1, X_2) := \sup_{x \in \mathbb{R}} |\mathbb{P}(X_1 \leq x) - \mathbb{P}(X_2 \leq x)|$$

and the Wasserstein distance

$$d_W(X_1, X_2) := \sup_{h \in \text{Lip}(1)} |\mathbb{E}h(X_1) - \mathbb{E}h(X_2)|,$$

where Lip(1) is the set of all functions $h : \mathbb{R} \to \mathbb{R}$ with a Lipschitz constant not greater than one.

Convergence in d_K or d_W implies convergence in distribution.

Berry-Esseen-Bound

Karlsruhe Institute of Technology

Theorem: Berry 1941, Esseen 1942

Let $(Y_i)_{i\in\mathbb{N}}$ be i.i.d. random variables with $\mathbb{E}|Y_1|^3 < \infty$, let $S_n = \sum_{i=1}^n Y_i$, $n \in \mathbb{N}$, and let N be a standard Gaussian random variable. Then there is a constant $C > 0$ such that

$$d_K\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}, N\right) \leq \frac{C}{\sqrt{n}} \frac{\mathbb{E}|Y_1 - \mathbb{E}Y_1|^3}{\sqrt{\operatorname{Var} Y_1}^3}, \quad n \in \mathbb{N}.$$

Berry-Esseen-Bound

Theorem: Berry 1941, Esseen 1942

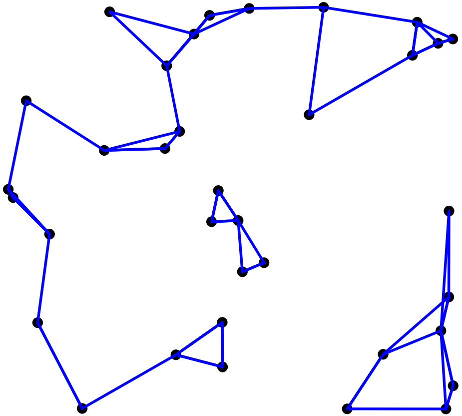
Let $(Y_i)_{i\in\mathbb{N}}$ be i.i.d. random variables with $\mathbb{E}|Y_1|^3 < \infty$, let $S_n = \sum_{i=1}^n Y_i$, $n \in \mathbb{N}$, and let N be a standard Gaussian random variable. Then there is a constant $C > 0$ such that

$$d_K\left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{Var} S_n}}, N\right) \leq \frac{C}{\sqrt{n}} \frac{\mathbb{E}|Y_1 - \mathbb{E}Y_1|^3}{\sqrt{\operatorname{Var} Y_1}^3}, \quad n \in \mathbb{N}.$$

Aim of this talk:

Berry-Esseen bounds for problems from stochastic geometry

k-Nearest Neighbour Graph



k-Nearest Neighbour Graph

η_t homogeneous Poisson process of intensity t in a compact convex set H

$$L_t^{(\alpha)} = \frac{1}{2} \sum_{(x_1,x_2)\in\eta_{t,\neq}^2} \mathbf{1}\{\text{ edge between } x_1 \text{ and } x_2 \text{ in } NNG_k(\eta_t)\} \|x_1 - x_2\|^\alpha$$

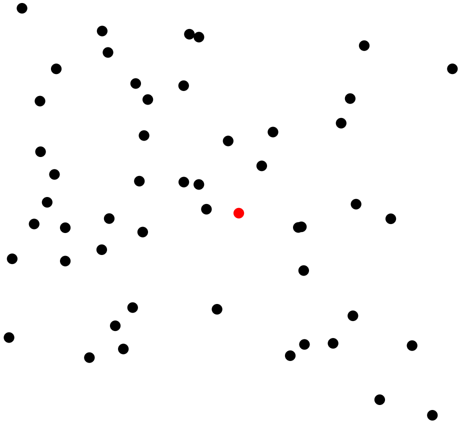
Theorem: Last/Peccati/S. 2014+

Let N be a standard Gaussian random variable. Then there are constants C_α, $\alpha \geq 0$, only depending on k, H and α such that

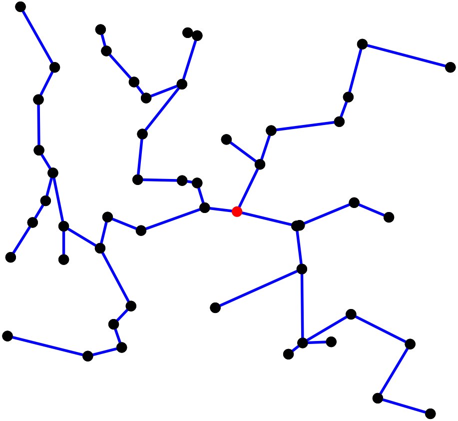
$$d_K\left(\frac{L_t^{(\alpha)} - \mathbb{E}L_t^{(\alpha)}}{\sqrt{\operatorname{Var} L_t^{(\alpha)}}}, N\right) \leq C_\alpha t^{-1/2}, \quad t \geq 1.$$

This improves the rates $(\ln(t))^{1+3/4} t^{-1/4}$ by Avram/Bertsimas (1993) and $(\ln(t))^{3d} t^{-1/2}$ by Penrose/Yukich (2005).

Radial spanning tree



Radial spanning tree



Radial spanning tree

Karlsruhe Institute of Technology

η_t homogeneous Poisson process of intensity t in a compact convex set H with $0 \in H$

$$L_t^{(\alpha)} = \frac{1}{2} \sum_{(x_1,x_2)\in\eta^2_{t,\neq}} \mathbf{1}\{\text{ edge between } x_1 \text{ and } x_2 \text{ in } RST(\eta_t)\} \|x_1 - x_2\|^\alpha$$

Theorem: Schulte/Thäle 2014

Let N be a standard Gaussian random variable. Then there are constants C_α, $\alpha \geq 0$, only depending on H and α such that

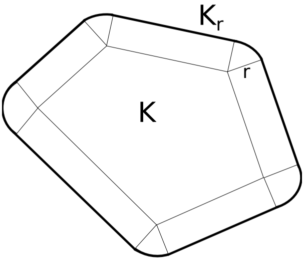
$$d_K\left(\frac{L_t^{(\alpha)} - \mathbb{E}L_t^{(\alpha)}}{\sqrt{\operatorname{Var} L_t^{(\alpha)}}}, N\right) \leq C_\alpha t^{-1/2}, \quad t \geq 1.$$

Intrinsic Volumes

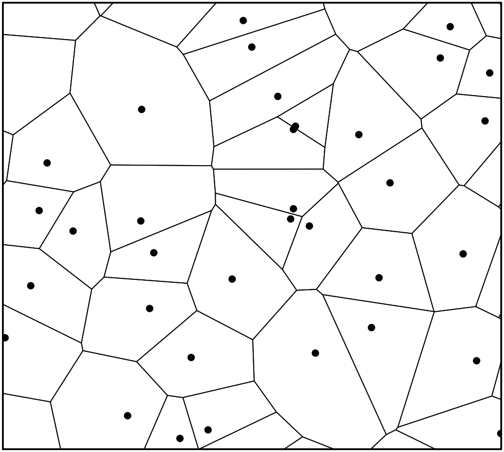
- $\mathcal{K}^d$ compact convex sets in $\mathbb{R}^d$
- The intrinsic volumes $V_i : \mathcal{K}^d \to \mathbb{R}$ are given by the Steiner formula

$$\mathrm{Vol}(K_r) = \mathrm{Vol}(K + rB^d) = \sum_{i=0}^{d} \kappa_{d-i} r^{d-i} V_i(K), \quad K \in \mathcal{K}^d, \quad r \geq 0.$$

- V_0: Euler characteristic, V_{d-1}: half the surface area, V_d: volume



Poisson-Voronoi tessellation



Poisson-Voronoi tessellation

η_t stationary Poisson process of intensity t in $\mathbb{R}^d$, X_t^k k-faces of the induced Voronoi tessellation, H compact convex set with $\mathrm{Vol}(H) > 0$,

$$V_t^{(k,i)} := \sum_{G \in X_t^k} V_i(G \cap H).$$

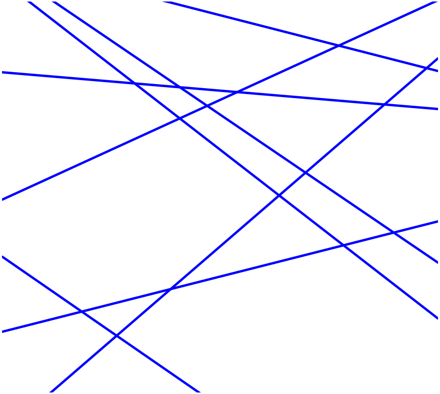
Theorem: Last/Peccati/S. 2014+

Let N be a standard Gaussian random variable. Then there are constants $c_{i,k}$, $k \in \{0, \dots, d\}$, $i \in \{0, \dots, \min\{k, d-1\}\}$, such that

$$d_K\left(\frac{V_t^{(k,i)} - \mathbb{E} V_t^{(k,i)}}{\sqrt{\mathrm{Var}\, V_t^{(k,i)}}}, N\right) \leq c_{k,i} t^{-1/2}, \quad t \geq 1.$$

See also Avram/Bertsimas 1993, Heinrich 1994, Penrose/Yukich 2005.

Poisson hyperplane tessellation



Let η_t be a Poisson hyperplane process with intensity measure $t\Lambda$, $t \geq 1$.
Let Λ be such that the hyperplanes of η_t are in general position a.s.

Poisson hyperplane tessellation

Let X_t^k be the k-faces of the hyperplane tessellation induced by η_t, H compact convex set with $\mathrm{Vol}(H) > 0$,

$$V_t^{(k,i)} := \sum_{G \in X_t^k} V_i(G \cap H).$$

Theorem: S. 2015

Let N be a standard Gaussian random variable. Then there are constants $c_{i,k}$, $k \in \{0, \dots, d-1\}$, $i \in \{0, \dots, k\}$, such that

$$d_K\left(\frac{V_t^{(k,i)} - \mathbb{E} V_t^{(k,i)}}{\sqrt{\mathrm{Var}\, V_t^{(k,i)}}}, N\right) \leq c_{k,i} t^{-1/2}, \quad t \geq 1.$$

Framework

- $(\mathbb{X}, \mathcal{X})$ measurable space with σ-finite measure λ
- $\mathbf{N}$ set of all σ-finite counting measures on $\mathbb{X}$
- η Poisson process with intensity measure λ
- Poisson functional $F = f(\eta)$ with $f : \mathbf{N} \to \mathbb{R}$ measurable
- For $x, x_1, x_2 \in \mathbb{X}$ we define

$$D_x F = f(\eta + \delta_x) - f(\eta)$$
$$D^2_{x_1,x_2} F = f(\eta + \delta_{x_1} + \delta_{x_2}) - f(\eta + \delta_{x_1}) - f(\eta + \delta_{x_2}) + f(\eta).$$

- We write $F \in \operatorname{dom} D$ if $F \in L^2_\eta$ and

$$\mathbb{E} \int_{\mathbb{X}} (D_x F)^2 \, \lambda(\mathrm{d}x) < \infty.$$

Variance inequalities

Theorem: Wu 2000, Last/Penrose 2011

For $F \in L^2_\eta$,

$$\int_{\mathbb{X}} (\mathbb{E} D_x F)^2 \, \lambda(dx) \leq \operatorname{Var} F \leq \mathbb{E} \int_{\mathbb{X}} (D_x F)^2 \, \lambda(dx).$$

The upper bound is called Poincaré inequality.

Second order Poincaré inequality

Karlsruhe Institute of Technology

Theorem: Last/Peccati/S. 2014+

Let $F \in \operatorname{dom} D$ be such that $\mathbb{E}F = 0$ and $\operatorname{Var} F = 1$, and let N be a standard Gaussian random variable. Then,

$$d_W(F, N) \leq \gamma_1 + \gamma_2 + \gamma_3,$$

where

$$\gamma_1 := 2\Big[\int_{\mathbb{X}^3} \big(\mathbb{E}[(D_{x_1}F\, D_{x_2}F)^2]\mathbb{E}[(D^2_{x_1,x_3}F)^2(D^2_{x_2,x_3}F)^2]\big)^{\frac{1}{2}} \lambda^3(\mathrm{d}(x_1, x_2, x_3))\Big]^{\frac{1}{2}},$$

$$\gamma_2 := \Big[\int_{\mathbb{X}^3} \mathbb{E}(D^2_{x_1,x_3}F)^2(D^2_{x_2,x_3}F)^2\, \lambda^3(\mathrm{d}(x_1, x_2, x_3))\Big]^{\frac{1}{2}},$$

$$\gamma_3 := \int_{\mathbb{X}} \mathbb{E}|D_x F|^3\, \lambda(\mathrm{d}x).$$

Second order Poincaré inequality

Karlsruhe Institute of Technology

Theorem: Last/Peccati/S. 2014+

Let $F \in \operatorname{dom} D$ be such that $\mathbb{E}F = 0$ and $\operatorname{Var} F = 1$, and let N be a standard Gaussian random variable. Then,

$$d_K(F, N) \leq \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5 + \gamma_6,$$

where

$$\gamma_4 := \frac{1}{2} \left[\mathbb{E}F^4\right]^{\frac{1}{4}} \int_{\mathbb{X}} \left[\mathbb{E}(D_x F)^4\right]^{\frac{3}{4}} \lambda(\mathrm{d}x),$$

$$\gamma_5 := \left[\int_{\mathbb{X}} \mathbb{E}(D_x F)^4 \, \lambda(\mathrm{d}x) \right]^{\frac{1}{2}},$$

$$\gamma_6 := \left[\int_{\mathbb{X}^2} 6 \left[\mathbb{E}(D_{x_1} F)^4\right]^{\frac{1}{2}} \left[\mathbb{E}(D^2_{x_1,x_2} F)^4\right]^{\frac{1}{2}} + 3\mathbb{E}(D^2_{x_1,x_2} F)^4 \, \lambda^2(\mathrm{d}(x_1, x_2)) \right]^{\frac{1}{2}}.$$

Thank you!