Statistical aspects of determinantal point processes

Jesper Møller,
Department of Mathematical Sciences, Aalborg University.

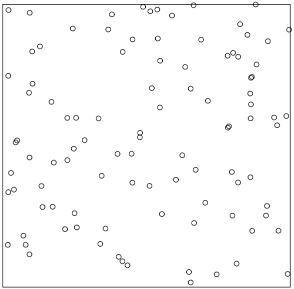
Joint work with
Frédéric Lavancier,
Laboratoire de Mathématiques Jean Leray, Nantes,
and
Ege Rubak,
Department of Mathematical Sciences, Aalborg University.

May 6, 2012

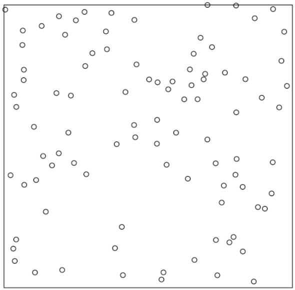
Introduction

- Determinantal point processes (DPP) form a class of repulsive point processes.
- They were introduced in their general form by O. Macchi in 1975 to model fermions (i.e. particules with repulsion) in quantum mechanics.
- Particular cases include the law of the eigenvalues of certain random matrices (Gaussian Unitary Ensemble, Ginibre Ensemble,...)
- Most theoretical studies have been published in the 2000's.
- The statistical aspects have so far been largely unexplored.

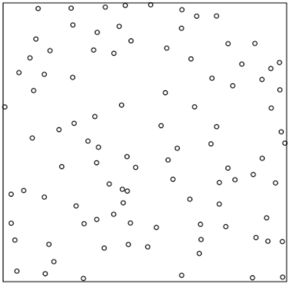
Examples



Poisson



DPP



DPP with stronger repulsion

Statistical motivation

Do DPP's constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

Statistical motivation

Do DPP's constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

$\longrightarrow$ Answer: **YES**.

Statistical motivation

Do DPP's constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

$\longrightarrow$ Answer: **YES**.

In fact:

- DPP's can be easily simulated.
- There are closed form expressions for the moments.
- There is a closed form expression for the density of a DPP on any bounded set.
- Inference is feasible, including likelihood inference.

Statistical motivation

Do DPP's constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

$\longrightarrow$ Answer: **YES**.

In fact:

- DPP's can be easily simulated.
- There are closed form expressions for the moments.
- There is a closed form expression for the density of a DPP on any bounded set.
- Inference is feasible, including likelihood inference.

These properties are unusual for Gibbs point processes which are commonly used to model inhibition (e.g. the Strauss process).

Notation

- We view a spatial point process X on $\mathbb{R}^d$ as a random locally finite subset of $\mathbb{R}^d$.

Notation

- We view a spatial point process X on $\mathbb{R}^d$ as a random locally finite subset of $\mathbb{R}^d$.
- For any borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.

Notation

- We view a spatial point process X on $\mathbb{R}^d$ as a random locally finite subset of $\mathbb{R}^d$.
- For any borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.
- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order product density function* of X.
  Intuitively,
  $$\rho^{(n)}(x_1, \ldots, x_n)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n$$
  is the probability that for each $i = 1, \ldots, n$,
  X has a point in a region around x_i of volume $\mathrm{d}x_i$.

Notation

- We view a spatial point process X on $\mathbb{R}^d$ as a random locally finite subset of $\mathbb{R}^d$.
- For any borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.
- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order product density function* of X.

  Intuitively,

  $$\rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

  is the probability that for each $i = 1, \ldots, n$, X has a point in a region around x_i of volume $\mathrm{d}x_i$.

  In particular $\rho = \rho^{(1)}$ is the *intensity function*.

Definition of a determinantal point process

For any function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, denote $[C](x_1, \dots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

Ex.: $[C](x_1) = C(x_1, x_1) \quad [C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}.$

Definition of a determinantal point process

For any function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, denote $[C](x_1, \dots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

Ex.: $[C](x_1) = C(x_1, x_1) \quad [C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}$.

Definition

X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if its product density functions satisfy

$$\rho^{(n)}(x_1, \dots, x_n) = \det[C](x_1, \dots, x_n), \quad n = 1, 2, \dots$$

Definition of a determinantal point process

For any function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, denote $[C](x_1, \ldots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

Ex.: $[C](x_1) = C(x_1, x_1) \quad [C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}.$

Definition

X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if its product density functions satisfy

$$\rho^{(n)}(x_1, \ldots, x_n) = \det[C](x_1, \ldots, x_n), \quad n = 1, 2, \ldots$$

The Poisson process with intensity $\rho(x)$ is the special case where $C(x, x) = \rho(x)$ and $C(x, y) = 0$ if $x \neq y$.

Definition of a determinantal point process

For any function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, denote $[C](x_1, \ldots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

Ex.: $[C](x_1) = C(x_1, x_1) \quad [C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}$.

Definition

X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if its product density functions satisfy

$$\rho^{(n)}(x_1, \ldots, x_n) = \det[C](x_1, \ldots, x_n), \quad n = 1, 2, \ldots$$

The Poisson process with intensity $\rho(x)$ is the special case where $C(x, x) = \rho(x)$ and $C(x, y) = 0$ if $x \neq y$.

For existence, conditions on the kernel C are mandatory, e.g. C must satisfy: for all $x_1, \ldots, x_n$, $\det[C](x_1, \ldots, x_n) \geq 0$.

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1, \ldots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some } i \neq j,$$

$\Longrightarrow$ **the points of X repel each other.**

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1, \dots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some } i \neq j,$$

$\Longrightarrow$ **the points of X repel each other**.

- The intensity of X is $\rho(x) = C(x, x)$.

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1, \ldots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some } i \neq j,$$

$\Longrightarrow$ **the points of X repel each other.**

- The intensity of X is $\rho(x) = C(x, x)$.
- The pair correlation function is

$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)}$$

- Thus $g \leq 1$ (i.e. repulsiveness).

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1, \ldots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some } i \neq j,$$

$\Longrightarrow$ **the points of X repel each other**.

- The intensity of X is $\rho(x) = C(x, x)$.
- The pair correlation function is

$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)}$$

- Thus $g \leq 1$ (i.e. repulsiveness).
- If $X \sim \mathrm{DPP}(C)$, then $X_B \sim \mathrm{DPP}_B(C_B)$

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1, \ldots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some } i \neq j,$$

$\Longrightarrow$ **the points of X repel each other.**

- The intensity of X is $\rho(x) = C(x, x)$.
- The pair correlation function is

$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)}$$

- Thus $g \leq 1$ (i.e. repulsiveness).
- If $X \sim \mathrm{DPP}(C)$, then $X_B \sim \mathrm{DPP}_B(C_B)$
- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

First properties

- From the definition, if C is continuous,

$$\rho^{(n)}(x_1,\dots,x_n)\approx 0 \quad \text{whenever} \quad x_i\approx x_j \quad \text{for some } i\neq j,$$

$\Longrightarrow$ **the points of X repel each other**.

- The intensity of X is $\rho(x)=C(x,x)$.
- The pair correlation function is

$$g(x,y):=\frac{\rho^{(2)}(x,y)}{\rho(x)\rho(y)}=1-\frac{C(x,y)C(y,x)}{C(x,x)C(y,y)}$$

- Thus $g\leq 1$ (i.e. repulsiveness).
- If $X\sim \mathrm{DPP}(C)$, then $X_B\sim \mathrm{DPP}_B(C_B)$
- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.
- There exists at most one $\mathrm{DPP}(C)$.

Existence

In all that follows we assume

$$\text{(C1)} \qquad C \text{ is a continuous complex covariance function.}$$

Existence

In all that follows we assume

(C1) C is a continuous complex covariance function.

By Mercer's theorem, for any compact set $S \subset \mathbb{R}^d$, C restricted to $S \times S$, denoted C_S, has a spectral representation,

$$C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

where $\lambda_k^S \geq 0$ and $\int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, \mathrm{d}x = \mathbf{1}_{\{k=l\}}$.

Existence

In all that follows we assume

(C1) C is a continuous complex covariance function.

By Mercer's theorem, for any compact set $S \subset \mathbb{R}^d$, C restricted to $S \times S$, denoted C_S, has a spectral representation,

$$C_S(x, y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x, y) \in S \times S,$$

where $\lambda_k^S \geq 0$ and $\int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, \mathrm{d}x = \mathbf{1}_{\{k=l\}}$.

Theorem (Macchi, 1975; Hough et al., 2009; our paper)

Under (C1), existence of DPP(C) is equivalent to that

$$\lambda_k^S \leq 1 \textit{ for all compact } S \subset \mathbb{R}^d \textit{ and all } k.$$

Density on a compact set S

Let $X_S \sim \mathrm{DPP}_S(C_S)$ with $S \subset \mathbb{R}^d$ compact.
Recall that $C_S(x, y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}$.

Density on a compact set S

Let $X_S \sim \mathrm{DPP}_S(C_S)$ with $S \subset \mathbb{R}^d$ compact.
Recall that $C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}$.

Theorem (Macchi, 1975)

Assuming $\lambda_k^S < 1$, for all k, then X_S is absolutely continuous with respect to the homogeneous Poisson process on S with unit intensity, and has density

$$f(\{x_1,\ldots,x_n\}) = \exp(|S| - D) \det[\tilde{C}](x_1,\ldots,x_n),$$

where $D = -\sum_{k=1}^{\infty} \log(1-\lambda_k^S)$ and $\tilde{C} : S \times S \to \mathbb{C}$ is given by

$$\tilde{C}(x,y) = \sum_{k=1}^{\infty} \frac{\lambda_k^S}{1-\lambda_k^S} \phi_k^S(x) \overline{\phi_k^S(y)}$$

Let $X_S \sim \mathrm{DPP}_S(C_S)$ where $S \subset \mathbb{R}^d$ is compact.

We want to simulate X_S.

Recall that $C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}$.

Let $X_S \sim \text{DPP}_S(C_S)$ where $S \subset \mathbb{R}^d$ is compact.

We want to simulate X_S.

Recall that $C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x)\overline{\phi_k^S(y)}$.

Theorem (Hough et al., 2006)

For $k \in \mathbb{N}$, let B_k be independent Bernoulli r.v.'s with means λ_k^S. Define

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Then $DPP_S(C_S) \stackrel{d}{=} DPP_S(K)$.

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Let $M = \max\{k \geq 0; B_k \neq 0\}$.
Note that M is a.s. finite, since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Let $M = \max\{k \geq 0; B_k \neq 0\}$.
Note that M is a.s. finite, since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

1. simulate a realization $M = m$ (by the inversion method);

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Let $M = \max\{k \geq 0; B_k \neq 0\}$.
Note that M is a.s. finite, since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

1. simulate a realization $M = m$ (by the inversion method);
2. generate the Bernoulli variables $B_1, \ldots, B_{m-1}$ (these are independent of $\{M = n\}$);

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Let $M = \max\{k \geq 0; B_k \neq 0\}$.
Note that M is a.s. finite, since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

1. simulate a realization $M = m$ (by the inversion method);
2. generate the Bernoulli variables $B_1, \ldots, B_{m-1}$ (these are independent of $\{M = n\}$);
3. simulate the point process $\mathrm{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

So simulating X_S is equivalent to simulate $\mathrm{DPP}_S(K)$ with

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Let $M = \max\{k \geq 0; B_k \neq 0\}$.
Note that M is a.s. finite, since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

1. simulate a realization $M = m$ (by the inversion method);
2. generate the Bernoulli variables $B_1, \ldots, B_{m-1}$ (these are independent of $\{M = n\}$);
3. simulate the point process $\mathrm{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

In step 3, the kernel K becomes a *projection kernel*, and w.l.o.g.

$$K(x,y) = \sum_{k=1}^{n} \phi_k^S(x)\overline{\phi_k^S(y)}$$

where $n = \#\{1 \leq k \leq M : B_k = 1\}$.

Simulation of determinantal projection processes

Denoting $\boldsymbol{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T$, we have

$$K(x,y) = \sum_{k=1}^{n} \phi_k^S(x)\overline{\phi_k^S(y)} = \boldsymbol{v}(y)^* \boldsymbol{v}(x)$$

Simulation of determinantal projection processes

Denoting $\boldsymbol{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T$, we have

$$K(x,y) = \sum_{k=1}^{n} \phi_k^S(x)\overline{\phi_k^S(y)} = \boldsymbol{v}(y)^* \boldsymbol{v}(x)$$

The point process $\mathrm{DPP}_S(K)$ has a.s. n points $(X_1, \ldots, X_n)$ that can be simulated by the following Gram-Schmidt procedure:

Simulation of determinantal projection processes

Denoting $\boldsymbol{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T$, we have

$$K(x,y) = \sum_{k=1}^{n} \phi_k^S(x)\overline{\phi_k^S(y)} = \boldsymbol{v}(y)^*\boldsymbol{v}(x)$$

The point process $\mathrm{DPP}_S(K)$ has a.s. n points $(X_1, \ldots, X_n)$ that can be simulated by the following Gram-Schmidt procedure:

---

sample X_n from the distribution with density $p_n(x) = \|\boldsymbol{v}(x)\|^2/n$.
set $\boldsymbol{e}_1 = \boldsymbol{v}(X_n)/\|\boldsymbol{v}(X_n)\|$.
for $i = (n-1)$ to 1 **do**

  sample X_i from the distribution (given $X_{i+1}, \ldots, X_n$) :

$$p_i(x) = \frac{1}{i}\left[\|\boldsymbol{v}(x)\|^2 - \sum_{j=1}^{n-i} |\boldsymbol{e}_j^*\boldsymbol{v}(x)|^2\right], \quad x \in S$$

  set $\boldsymbol{w}_i = \boldsymbol{v}(X_i) - \sum_{j=1}^{n-i} \left(\boldsymbol{e}_j^*\boldsymbol{v}(X_i)\right)\boldsymbol{e}_j$, $\boldsymbol{e}_{n-i+1} = \boldsymbol{w}_i/\|\boldsymbol{w}_i\|$

Simulation of determinantal projection processes

Denoting $\boldsymbol{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T$, we have

$$K(x,y) = \sum_{k=1}^{n} \phi_k^S(x)\overline{\phi_k^S(y)} = \boldsymbol{v}(y)^* \boldsymbol{v}(x)$$

The point process $\mathrm{DPP}_S(K)$ has a.s. n points $(X_1, \ldots, X_n)$ that can be simulated by the following Gram-Schmidt procedure:

---

sample X_n from the distribution with density $p_n(x) = \|\boldsymbol{v}(x)\|^2/n$.
set $\boldsymbol{e}_1 = \boldsymbol{v}(X_n)/\|\boldsymbol{v}(X_n)\|$.
for $i = (n-1)$ to 1 **do**

  sample X_i from the distribution (given $X_{i+1}, \ldots, X_n$) :

$$p_i(x) = \frac{1}{i}\left[\|\boldsymbol{v}(x)\|^2 - \sum_{j=1}^{n-i} |\boldsymbol{e}_j^* \boldsymbol{v}(x)|^2\right], \quad x \in S$$

  set $\boldsymbol{w}_i = \boldsymbol{v}(X_i) - \sum_{j=1}^{n-i} \left(\boldsymbol{e}_j^* \boldsymbol{v}(X_i)\right) \boldsymbol{e}_j$, $\boldsymbol{e}_{n-i+1} = \boldsymbol{w}_i/\|\boldsymbol{w}_i\|$

Theorem

$\{X_1, \ldots, X_n\}$ *generated as above has distribution* $DPP_S(K)$.

Illustration of simulation algorithm

Example: Let $S = [-1/2, 1/2]^2$ and

$$\phi_k(x) = \mathrm{e}^{2\pi \mathrm{i} k \cdot x}, \quad k \in \mathbb{Z}^2,\ x \in S,$$

for a set of indices $k_1, \ldots, k_n$ in $\mathbb{Z}^2$.
So the projection kernel writes

$$K(x, y) = \sum_{j=1}^{n} \mathrm{e}^{2\pi \mathrm{i} k_j \cdot (x-y)}$$

and $X_S \sim \mathrm{DPP}_S(K)$ is homogeneous and has a.s. n points.

Illustration of simulation algorithm

Step 1. The first point is sampled uniformly on S

Illustration of simulation algorithm

Step 1. The first point is sampled uniformly on S

Step 2. The next point is sampled w.r.t. the following density :

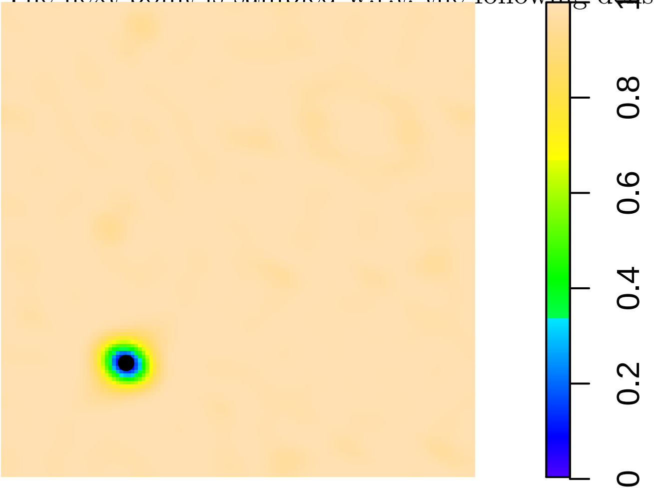


Illustration of simulation algorithm

Step 3. The next point is sampled w.r.t. the following density :

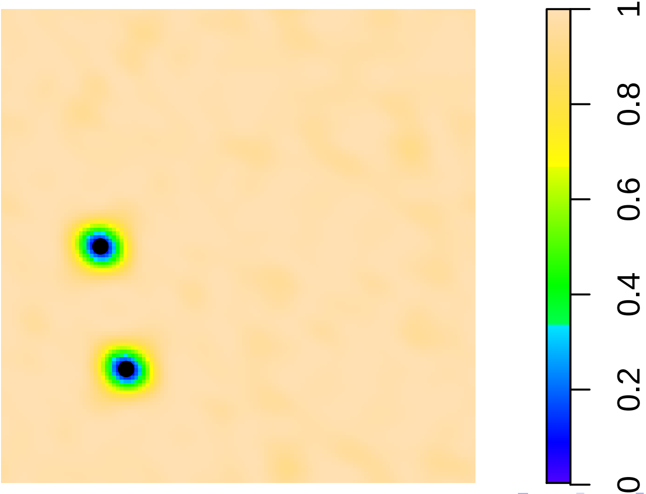


Illustration of simulation algorithm

etc.

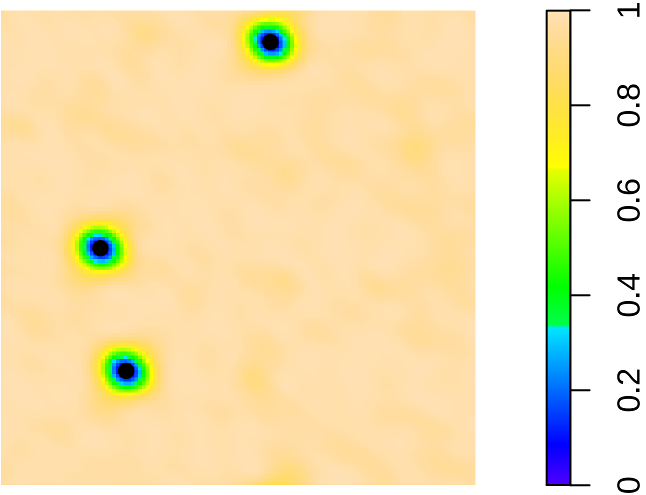


Illustration of simulation algorithm

etc.

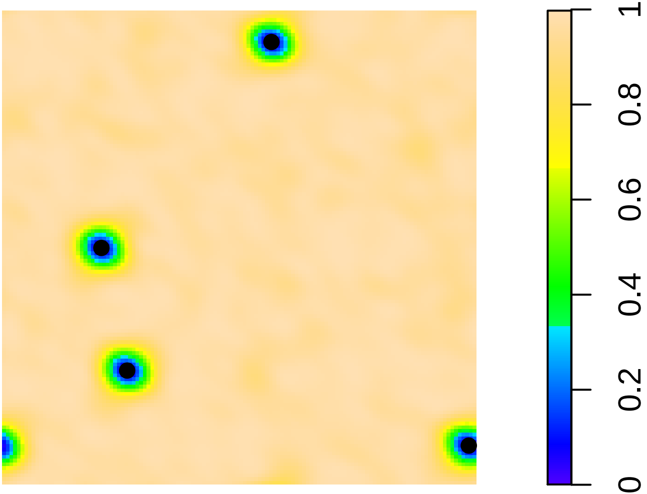


Illustration of simulation algorithm

etc.

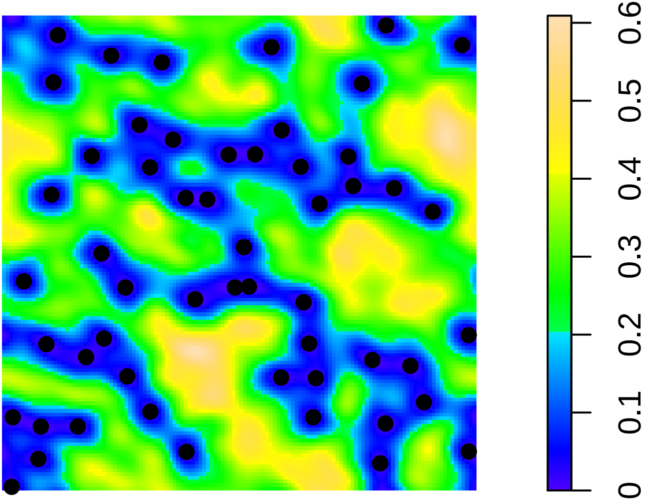
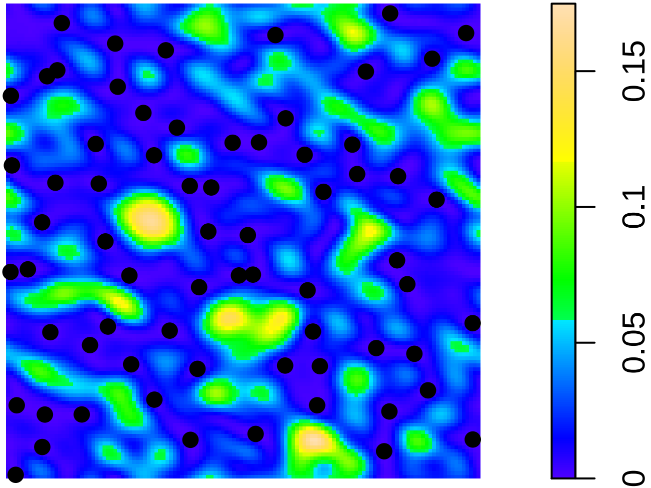


Illustration of simulation algorithm

etc.



Stationary models

We focus on a kernel C of the form

$$C(x,y) = C_0(x-y), \quad x,y \in \mathbb{R}^d.$$

(C1) C_0 is a continuous covariance function
Moreover, if $C_0 \in L^2(\mathbb{R}^d)$ we can define its Fourier transform

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x \cdot t} \, \mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Stationary models

We focus on a kernel C of the form

$$C(x,y) = C_0(x-y), \quad x, y \in \mathbb{R}^d.$$

(C1) C_0 is a continuous covariance function
Moreover, if $C_0 \in L^2(\mathbb{R}^d)$ we can define its Fourier transform

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x \cdot t} \, \mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Theorem

Under (C1), if $C_0 \in L^2(\mathbb{R}^d)$, then existence of DPP($C_0$) is equivalent to

$$\varphi \leq 1.$$

Stationary models

We focus on a kernel C of the form

$$C(x,y) = C_0(x-y), \quad x,y \in \mathbb{R}^d.$$

(C1) C_0 is a continuous covariance function
Moreover, if $C_0 \in L^2(\mathbb{R}^d)$ we can define its Fourier transform

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x\cdot t}\,\mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Theorem

Under (C1), if $C_0 \in L^2(\mathbb{R}^d)$, then existence of DPP$(C_0)$ is equivalent to

$$\varphi \leq 1.$$

To construct parametric families of DPP :
Consider parametric families of C_0 and rescale so that $\varphi \leq 1$.
$\rightarrow$ This will induce a bound on the parameter space.

Several parametric families of covariance function are available, with closed form expressions for their Fourier transform.

- For $d = 2$, the circular covariance function with range α is given by

$$C_0(x) = \rho \frac{2}{\pi} \left(\arccos(\|x\|/\alpha) - \|x\|/\alpha \sqrt{1 - (\|x\|/\alpha)^2} \right) \mathbf{1}_{\|x\|<\alpha}.$$

  DPP(C_0) exists iff $\varphi \leq 1 \Leftrightarrow \rho\alpha^2 \leq 4/\pi$.

  $\Rightarrow$ Tradeoff between the intensity ρ and the range of repulsion α.

Several parametric families of covariance function are available, with closed form expressions for their Fourier transform.

- For $d = 2$, the circular covariance function with range α is given by

$$C_0(x) = \rho \frac{2}{\pi} \left(\arccos(\|x\|/\alpha) - \|x\|/\alpha \sqrt{1 - (\|x\|/\alpha)^2} \right) \mathbf{1}_{\|x\|<\alpha}.$$

  DPP(C_0) exists iff $\varphi \leq 1 \Leftrightarrow \rho\alpha^2 \leq 4/\pi$.

  $\Rightarrow$ Tradeoff between the intensity ρ and the range of repulsion α.

- Whittle-Matérn (includes Exponential and Gaussian) :

$$C_0(x) = \rho \, \frac{2^{1-\nu}}{\Gamma(\nu)} \|x/\alpha\|^{\nu} K_{\nu}(\|x/\alpha\|), \quad x \in \mathbb{R}^d.$$

  DPP(C_0) exists iff $\rho \leq \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)(2\sqrt{\pi}\alpha)^d}$.

Several parametric families of covariance function are available, with closed form expressions for their Fourier transform.

- For $d = 2$, the circular covariance function with range α is given by

  $$C_0(x) = \rho \frac{2}{\pi} \left(\arccos(\|x\|/\alpha) - \|x\|/\alpha \sqrt{1 - (\|x\|/\alpha)^2} \right) \mathbf{1}_{\|x\|<\alpha}.$$

  DPP(C_0) exists iff $\varphi \leq 1 \Leftrightarrow \rho\alpha^2 \leq 4/\pi$.

  $\Rightarrow$ Tradeoff between the intensity ρ and the range of repulsion α.

- Whittle-Matérn (includes Exponential and Gaussian) :

  $$C_0(x) = \rho \frac{2^{1-\nu}}{\Gamma(\nu)} \|x/\alpha\|^\nu K_\nu(\|x/\alpha\|), \quad x \in \mathbb{R}^d.$$

  DPP(C_0) exists iff $\rho \leq \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)(2\sqrt{\pi}\alpha)^d}$.

- Generalized Cauchy

  $$C_0(x) = \frac{\rho}{(1 + \|x/\alpha\|^2)^{\nu+d/2}}, \quad x \in \mathbb{R}^d.$$

  DPP(C_0) exists iff $\rho \leq \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)(\sqrt{\pi}\alpha)^d}$.

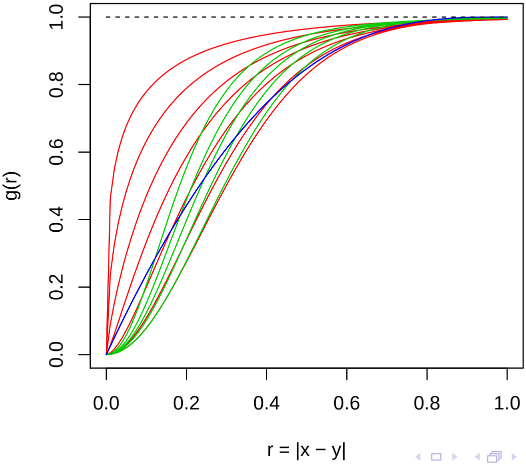
Pair correlation functions of DPP(C_0) for previous models :

In blue : C_0 is the circular covariance function.

In red : C_0 is Whittle-Matérn, for different values of ν

In green : C_0 is generalized Cauchy, for different values of ν

The parameter α is chosen such that the range of corr. ≈ 1.



Spectral approach

- Specify a parametric class of integrable functions $\varphi_\theta : \mathbb{R}^d \to [0,1]$ (spectral densities).
- This is all we need for having a well-defined DDP.
- Is convenient for simulation and for (approximate) density calculations as seen later.

Spectral approach

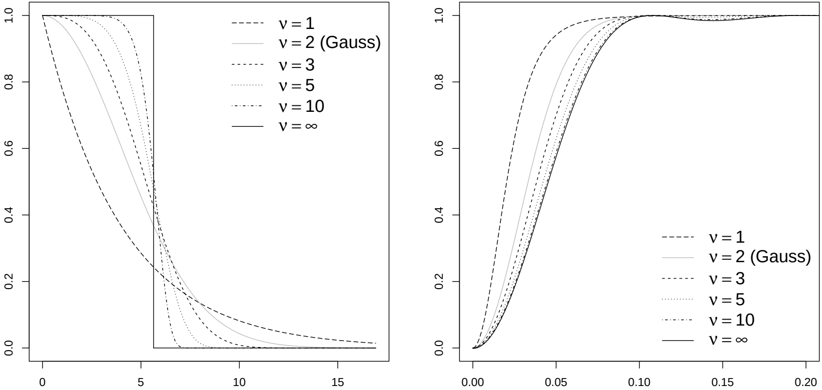
- Specify a parametric class of integrable functions $\varphi_\theta : \mathbb{R}^d \to [0,1]$ (spectral densities).
- This is all we need for having a well-defined DDP.
- Is convenient for simulation and for (approximate) density calculations as seen later.
- Example: *power exponential spectral model*:

$$\varphi_{\rho,\nu,\alpha}(x) = \rho \frac{\Gamma(d/2+1)\nu\alpha^d}{d\pi^{d/2}\Gamma(d/\nu)} \exp\left(-\|\alpha x\|^\nu\right)$$

with

$$\rho > 0, \quad \nu > 0, \quad 0 < \alpha < \alpha_{\max}(\rho,\nu) := \left(\frac{2\pi^{d/2}\Gamma(d/\nu+1)}{\rho\Gamma(d/2)}\right)^{1/d}.$$

Power exponential spectral model: (isotropic) spectral densities and pair correlation functions



Spectral densities

Pair correlation functions

Approximation of stationary models

Consider a stationary kernel C_0 and $X \sim \mathrm{DPP}(C_0)$.

- The simulation and the density of X_S requires the expansion

$$C_S(x,y) = C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

but in general λ_k^S and ϕ_k^S are not expressible on closed form.

Approximation of stationary models

Consider a stationary kernel C_0 and $X \sim \mathrm{DPP}(C_0)$.

- The simulation and the density of X_S requires the expansion

$$C_S(x,y) = C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

but in general λ_k^S and ϕ_k^S are not expressible on closed form.

- Consider the unit box $S = [-\frac{1}{2}, \frac{1}{2}]^d$ and the Fourier expansion

$$C_0(y-x) = \sum_{k \in \mathbb{Z}^d} c_k \mathrm{e}^{2\pi \mathrm{i} k \cdot (y-x)}, \quad y - x \in S.$$

The Fourier coefficients are

$$c_k = \int_S C_0(u)\mathrm{e}^{-2\pi \mathrm{i} k \cdot u} \, \mathrm{d}u \approx \int_{\mathbb{R}^d} C_0(u)\mathrm{e}^{-2\pi \mathrm{i} k \cdot u} \, \mathrm{d}u = \varphi(k)$$

which is a good approximation if $C_0(u) \approx 0$ for $|u| > \frac{1}{2}$.

Approximation of stationary models

Consider a stationary kernel C_0 and $X \sim \mathrm{DPP}(C_0)$.

- The simulation and the density of X_S requires the expansion

$$C_S(x,y) = C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x)\overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

but in general λ_k^S and ϕ_k^S are not expressible on closed form.

- Consider the unit box $S = [-\frac{1}{2}, \frac{1}{2}]^d$ and the Fourier expansion

$$C_0(y-x) = \sum_{k \in \mathbb{Z}^d} c_k \mathrm{e}^{2\pi \mathrm{i} k\cdot(y-x)}, \quad y-x \in S.$$

The Fourier coefficients are

$$c_k = \int_S C_0(u)\mathrm{e}^{-2\pi \mathrm{i} k\cdot u}\,\mathrm{d}u \approx \int_{\mathbb{R}^d} C_0(u)\mathrm{e}^{-2\pi \mathrm{i} k\cdot u}\,\mathrm{d}u = \varphi(k)$$

which is a good approximation if $C_0(u) \approx 0$ for $|u| > \frac{1}{2}$.

- Example: For the circular covariance, this is true whenever $\rho > 5$.

Approximation of stationary models

The approximation of $\mathrm{DPP}(C_0)$ on S is then $\mathrm{DPP}_S(C_{\mathrm{app},0})$ with

$$C_{\mathrm{app},0}(x-y) = \sum_{k \in \mathbb{Z}^d} \varphi(k) \mathrm{e}^{2\pi \mathrm{i}(x-y)\cdot k}, \quad x, y \in S,$$

where φ is the Fourier transform of C_0.

Approximation of stationary models

The approximation of DPP(C_0) on S is then DPP$_S(C_{\text{app},0})$ with

$$C_{\text{app},0}(x-y) = \sum_{k\in\mathbb{Z}^d} \varphi(k) e^{2\pi i (x-y)\cdot k}, \quad x, y \in S,$$

where φ is the Fourier transform of C_0.

This approximation allows us

- to simulate DPP(C_0) on S;
- to compute the (approximated) density of DPP(C_0) on S.

Consider a stationary and isotropic parametric DPP(C), i.e.

$$C(x,y) = C_0(x-y) = \rho R_\alpha(\|x-y\|),$$

with $R_\alpha(0) = 1$.

The first and second moments are easily deduced:

- The intensity is ρ.
- The pair correlation function is

$$g(x,y) = g_0(\|x-y\|) = 1 - R_\alpha^2(\|x-y\|).$$

- Ripley's K-function is easily expressible in terms of R_α: if $d = 2$,

$$K_\alpha(r) := 2\pi \int_0^r t g_0(t)\,\mathrm{d}t = \pi r^2 - 2\pi \int_0^r t|R_\alpha(t)|^2\,\mathrm{d}t.$$

Inference

Parameter estimation can be conducted as follows.

1. Estimate ρ by #{obs. points}/area of obs. window.
2. Estimate α
   - either by **minimum contrast** estimator (MCE):

$$\hat{\alpha} = \text{argmin}_\alpha \int_0^{r_{\max}} \left| \sqrt{\widehat{K}(r)} - \sqrt{K_\alpha(r)} \right|^2 \, \mathrm{d}r$$

   - or by **maximum likelihood** estimator: given $\hat{\rho}$, the likelihood is deduced from the kernel approximation.

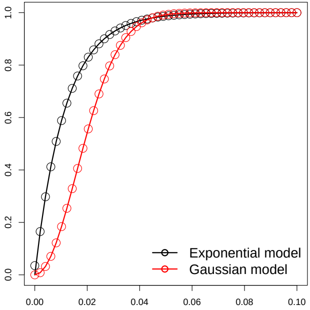
Two model examples

- Exponential model with $\rho = 200$ and $\alpha = 0.014$:

$$C_0(x) = \rho \exp(-\|x\|/\alpha).$$

- Gaussian model with $\rho = 200$ and $\alpha = 0.02$:

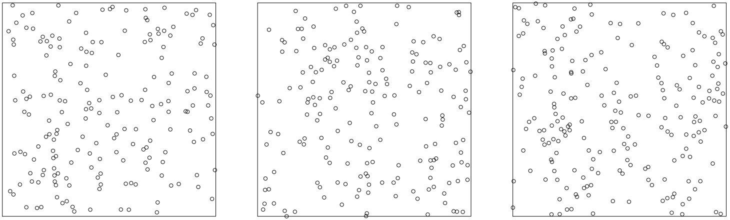
$$C_0(x) = \rho \exp(-\|x/\alpha\|^2).$$



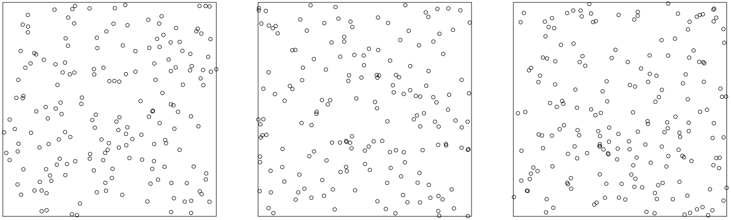
– Solid lines: theoretical pair correlation function

∘ Circles: pair correlation from the approximated kernel

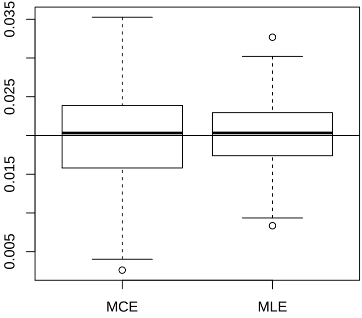
Samples from the Gaussian model on $[0,1]^2$:



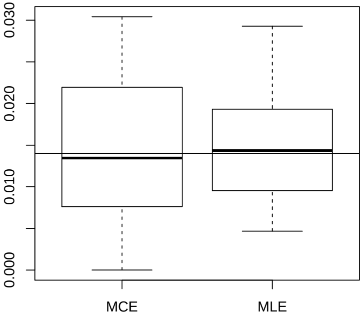
Samples from the exponential model on $[0,1]^2$:



Estimation of α from 200 realisations

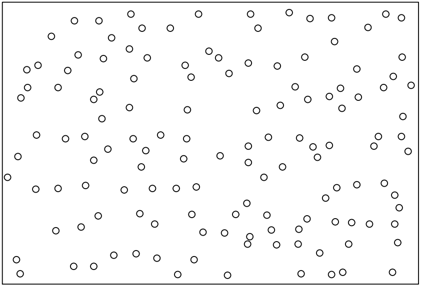


Gaussian model



Exponential model

Example: 134 Norwegian pine trees observed in a 56×38 m region



Møller and Waagpetersen (2004): a five parameter multiscale process is fitted using elaborate MCMC MLE methods.

Here we fit a more parsimonious DPP models.

First,

- Whittle-Matérn model;
- Cauchy model;
- Gaussian model: the best fit, but plots of summary statistics indicate a lack of fit.

First,

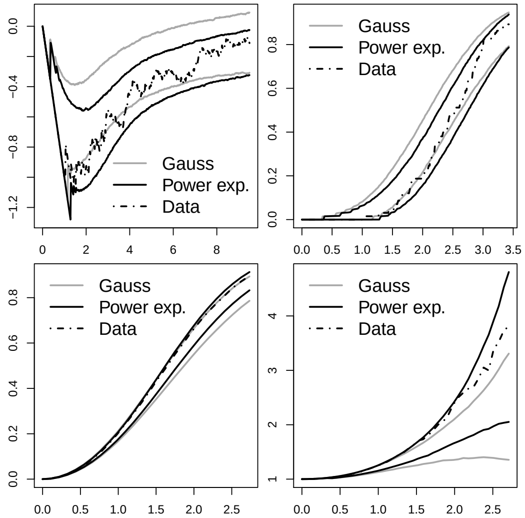
- Whittle-Matérn model;
- Cauchy model;
- Gaussian model: the best fit, but plots of summary statistics indicate a lack of fit.

Second,

- power exponential spectral model: provides a much better fit, with

$$\hat{\nu} = 10, \quad \hat{\alpha} = 6.36 \ \approx \ \alpha_{\max} = 6.77$$

i.e. close to the "most repulsive possible stationary DPP".



Clockwise from top left: $L(r) - r$; $G(r)$; $F(r)$; $J(r)$. Simulated 2.5% and 97.5% envelopes are based on 4000 realizations of the fitted Gaussian model resp. power exponential spectral model.

Conclusions

- DPP's provide flexible parametric models of repulsive point processes.

Conclusions

- DPP's provide flexible parametric models of repulsive point processes.

- DPP's possess the following appealing properties:
  - Easily simulated.
  - Closed form expressions for the moments.
  - Closed form expression for the density of a DPP on any bounded set.
  - Inference is feasible, including likelihood inference.

Conclusions

- DPP's provide flexible parametric models of repulsive point processes.

- DPP's possess the following appealing properties:
  - Easily simulated.
  - Closed form expressions for the moments.
  - Closed form expression for the density of a DPP on any bounded set.
  - Inference is feasible, including likelihood inference.

$\Rightarrow$ Promising alternative to repulsive Gibbs point processes.

References

Hough, J. B., M. Krishnapur, Y. Peres, and B. Viràg (2006).
Determinantal processes and independence.
Probability Surveys 3, 206–229.

Macchi, O. (1975).
The coincidence approach to stochastic point processes.
Advances in Applied Probability 7, 83–122.

McCullagh, P. and J. Møller (2006).
The permanental process.
Advances in Applied Probability 38, 873–888.

Scardicchio, A., C. Zachary, and S. Torquato (2009).
Statistical properties of determinantal point processes in high-dimensional Euclidean spaces.
Physical Review E 79(4).

Shirai, T. and Y. Takahashi (2003).
Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes.
Journal of Functional Analysis 2, 414–463.

Soshnikov, A. (2000).
Determinantal random point fields.
Russian Math. Surveys 55, 923–975.