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.

June 27, 2012
<table>
<thead>
<tr>
<th></th>
<th>Introduction</th>
<th>Definition, existence and basic properties</th>
<th>Simulation</th>
<th>Parametric models</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Introduction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><strong>Definition, existence and basic properties</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><strong>Simulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><strong>Parametric models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td><strong>Inference</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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. particles 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.
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. particles 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.
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. particles 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.
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. particles 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.
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. particles 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
Do DPP’s constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

→ 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).
Statistical motivation

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

→ 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).
Do DPP’s constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

→ 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).
Do DPP’s constitute a *tractable* and *flexible* class of models for *repulsive* point processes?

→ 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).
<table>
<thead>
<tr>
<th>Introduction</th>
<th>Definition</th>
<th>Simulation</th>
<th>Parametric models</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>2 Definition, existence and basic properties</td>
<td>3 Simulation</td>
<td>4 Parametric models</td>
<td>5 Inference</td>
</tr>
</tbody>
</table>
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 \( X \) (the density for the \( n \)'th order factorial moment measure).

Intuitively,
\[
\rho^{(n)}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n
\]
is the probability that for each \( i = 1, \ldots, n \), \( X \) has a point in a region around \( x_i \) of volume \( dx_i \).

In particular \( \rho = \rho^{(1)} \) is the intensity function.
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 $X$ (the density for the $n$’th order factorial moment measure).

Intuitively,
\[
\rho^{(n)}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n
\]
is the probability that for each $i = 1, \ldots, n$, $X$ has a point in a region around $x_i$ of volume $dx_i$.

In particular $\rho = \rho^{(1)}$ is the intensity function.
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 $X$ (the density for the $n$'th order factorial moment measure).

Intuitively,

$$\rho^{(n)}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n$$

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

In particular $\rho = \rho^{(1)}$ is the intensity function.
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 $X$ (the density for the $n$’th order factorial moment measure).

Intuitively, 
$$\rho^{(n)}(x_1, \ldots, x_n) \, dx_1 \cdots dx_n$$

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

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 $[C](x_1, \ldots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

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

Definition

$X$ is a determinantal point process with kernel $C$, denoted $X \sim \text{DPP}(C)$, 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 $x_1, \ldots, x_n$, $\det[C](x_1, \ldots, x_n) \geq 0$. 

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

**Definition**

$X$ is a *determinantal point process* with kernel $C$, denoted $X \sim \text{DPP}(C)$, 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 $x_1, \ldots, x_n$, $\det[C](x_1, \ldots, x_n) \geq 0$. 
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.: $[C](x_1) = C(x_1, x_1)$  
$[C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}$.

Definition

$X$ is a **determinantal point process** with kernel $C$, denoted $X \sim \text{DPP}(C)$, 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 $x_1, \ldots, x_n$, $\det[C](x_1, \ldots, x_n) \geq 0$. 

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

Definition

$X$ is a **determinantal point process** with kernel $C$, denoted $X \sim \text{DPP}(C)$, 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,

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

$\implies$ **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 $C$ is Hermitean.
- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$
- Any smooth transformation or independent thinning of a
  DPP is still a DPP with explicit given kernel.
- There exists at most one DPP$(C)$.
First properties

- From the definition, if $C$ is continuous,

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

$\implies$ 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 $C$ is Hermitean.

- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$

- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

- There exists at most one DPP$(C)$.
First properties

- From the definition, if $C$ is continuous,
  \[ \rho^{(n)}(x_1, \ldots, x_n) \approx 0 \text{ whenever } x_i \approx x_j \text{ for some } i \neq j, \]
  \[ \implies \text{the points of } X \text{ repel each other.} \]

- 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 $C$ is Hermitean.

- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$

- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

- There exists at most one DPP $(C)$. 
First properties

- From the definition, if $C$ is continuous,

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

$\implies$ 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 $C$ is Hermitean.

- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$

- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

- There exists at most one DPP$(C)$.
First properties

- From the definition, if $C$ is continuous,

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

$\implies$ 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 $C$ is Hermitean.

- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$

- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

- There exists at most one DPP$(C)$. 

First properties

- From the definition, if $C$ is continuous,
  
  \[ \rho^{(n)}(x_1, \ldots, x_n) \approx 0 \quad \text{whenever} \quad x_i \approx x_j \quad \text{for some} \quad i \neq j, \]

  \[ \implies \text{the points of } X \text{ repel each other.} \]

- 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 $C$ is Hermitean.

- If $X \sim \text{DPP}(C)$, then $X_B \sim \text{DPP}_B(C_B)$

- Any smooth transformation or independent thinning of a DPP is still a DPP with explicit given kernel.

- There exists at most one DPP($C$).
Henceforth 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 \( \int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, dx = 1_{\{k=l\}} \).

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

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

\[ \lambda_k^S \leq 1 \text{ for all compact } S \subset \mathbb{R}^d \text{ and all } k. \]
Henceforth assume

\[(C1) \quad C \text{ is a continuous complex covariance function.}\]

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 \(\int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, dx = 1_{\{k=l\}}\).

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

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

\[\lambda_k^S \leq 1 \text{ for all compact } S \subset \mathbb{R}^d \text{ and all } k.\]**
Henceforth 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 \( \int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, dx = 1_{\{k=l\}}. \)

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

*Under (C1), existence of DPP(\( C \)) is equivalent to that \( \lambda_k^S \leq 1 \) for all compact \( S \subset \mathbb{R}^d \) and all \( k \).*
Density on a compact set $S$

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

**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

$$\tilde{C}(x, y) = \sum_{k=1}^{\infty} \frac{\lambda_k^S}{1 - \lambda_k^S} \phi_k^S(x)\phi_k^S(y)$$
Density on a compact set $S$

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

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

**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

$$\tilde{C}'(x, y) = \sum_{k=1}^{\infty} \frac{\lambda_k^S}{1 - \lambda_k^S} \phi_k^S(x) \phi_k^S(y)$$
<table>
<thead>
<tr>
<th></th>
<th>Introduction</th>
<th>Definition, existence and basic properties</th>
<th>Simulation</th>
<th>Parametric models</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>Definition, existence and basic properties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Parametric models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Inference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 $\lambda_k^S$. 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 $\text{DPP}_S(C_S) \overset{d}{=} \text{DPP}_S(K)$. 
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 \( C_S(x, y) = \sum_{k=1}^{\infty} \lambda^S_k \phi^S_k(x) \phi^S_k(y) \).

**Theorem (Hough et al., 2006)**

*For \( k \in \mathbb{N} \), let \( B_k \) be independent Bernoulli r.v.’s with means \( \lambda^S_k \). Define

\[
K(x, y) = \sum_{k=1}^{\infty} B_k \phi^S_k(x) \phi^S_k(y), \quad (x, y) \in S \times S.
\]

Then \( \text{DPP}_S(C_S) \overset{d}{=} \text{DPP}_S(K) \).
So simulating $X_S$ is equivalent to simulate $\text{DPP}_S(K)$ with

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

Note that $M := \sup\{k \geq 0 : B_k \neq 0\}$ is a.s. finite, since

$$\mathbb{E} \sum B_k = \sum \lambda_k^S = \int_S C(x, x) \, dx < \infty.$$

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 $\{M = m\}$).
3. Simulate the point process $\text{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

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

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

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

$\bar{\phi}_k^S$
So simulating \( X_S \) is equivalent to simulate \( \text{DPP}_S(K) \) with

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

Note that \( M := \sup\{k \geq 0 : B_k \neq 0\} \) is a.s. finite, since

\[
E \sum B_k = \sum \lambda_k^S = \int S C(x, x) \, dx < \infty.
\]

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 \( \{M = m\} \)).
3. Simulate the point process \( \text{DPP}_S(K) \) given \( B_1, \ldots, B_M \) and \( M = m \).

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

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

where \( n = \#\{1 \leq k \leq M : B_k = 1\} \).
So simulating $X_S$ is equivalent to simulate $\text{DPP}_S(K)$ with

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

Note that $M := \sup\{k \geq 0 : B_k \neq 0\}$ is a.s. finite, since

$$\mathbb{E} \sum B_k = \sum \lambda_k^S = \int_S C(x, x) \, dx < \infty.$$ 

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 $\{M = m\}$).
3. Simulate the point process $\text{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

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

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

where $n = \#\{1 \leq k \leq M : B_k = 1\}$. 
So simulating $X_S$ is equivalent to simulate $\text{DPP}_S(K)$ with

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

Note that $M := \sup\{k \geq 0 : B_k \neq 0\}$ is a.s. finite, since
\[ E \sum B_k = \sum \lambda_k^S = \int_S C(x, x) \, dx < \infty. \]

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 $\{M = m\}$).
3. Simulate the point process $\text{DPP}_S(K)$ 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 $n = \#\{1 \leq k \leq M : B_k = 1\}$. 
So simulating $X_S$ is equivalent to simulate $\text{DPP}_S(K)$ with

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

Note that $M := \sup\{k \geq 0 : B_k \neq 0\}$ is a.s. finite, since 

$$\mathbb{E} \sum B_k = \sum \lambda_k^S = \int_S C(x, x) \, dx < \infty.$$ 

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 $\{M = m\}$).
3. Simulate the point process $\text{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

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

$$K(x, y) = \sum_{k=1}^{n} \phi^S_k(x) \bar{\phi}^S_k(y)$$

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

14 / 39
So simulating $X_S$ is equivalent to simulate $\text{DPP}_S(K)$ with

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

Note that $M := \sup\{k \geq 0 : B_k \neq 0\}$ is a.s. finite, since

$$E \sum B_k = \sum \lambda_k^S = \int_S C(x, x) \, dx < \infty.$$ 

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 \{\emph{M} = m\}).
3. Simulate the point process $\text{DPP}_S(K)$ given $B_1, \ldots, B_M$ and $M = m$.

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

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

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

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

$$K(x, y) = \sum_{k=1}^{n} \phi_k^S(x) \phi_k^S(y) = \mathbf{v}(y)^* \mathbf{v}(x), \quad \mathbf{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T.$$ 

**Theorem**

The set $\{X_1, \ldots, X_n\}$ generated as above has distribution $\text{DPP}_S(K)$. 

sample $X_n$ from the distribution with density $p_n(x) = \|\mathbf{v}(x)\|^2/n$; set $e_1 = \mathbf{v}(X_n)/\|\mathbf{v}(X_n)\|$;

for $i = (n - 1)$ to 1 do

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

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

set $w_i = \mathbf{v}(X_i) - \sum_{j=1}^{n-i} (\mathbf{e}_j^* \mathbf{v}(X_i)) \mathbf{e}_j, \quad e_{n-i+1} = w_i/\|w_i\|$
Simulation of determinantal projection processes

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

$$K(x, y) = \sum_{k=1}^{n} \phi_k^S(x) \overline{\phi_k^S(y)} = \mathbf{v}(y)^* \mathbf{v}(x), \quad \mathbf{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T.$$

**sample** $X_n$ from the distribution with density $p_n(x) = \|\mathbf{v}(x)\|^2 / n$;

**set** $e_1 = \mathbf{v}(X_n) / \|\mathbf{v}(X_n)\|$;

**for** $i = (n - 1)$ to 1 **do**

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

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

**set** $w_i = \mathbf{v}(X_i) - \sum_{j=1}^{n-i} (e_j^* \mathbf{v}(X_i)) e_j$, $e_{n-i+1} = w_i / \|w_i\|$

**Theorem**

$\{X_1, \ldots, X_n\}$ generated as above has distribution $\text{DPP}_S(K)$. 
Simulation of determinantal projection processes

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

\[
K(x, y) = \sum_{k=1}^{n} \phi_k^S(x) \overline{\phi_k^S(y)} = \mathbf{v}(y)^* \mathbf{v}(x), \quad \mathbf{v}(x) = (\phi_1^S(x), \ldots, \phi_n^S(x))^T.
\]

**sample** \( X_n \) from the distribution with density \( p_n(x) = \parallel \mathbf{v}(x) \parallel^2 / n \);
**set** \( e_1 = \mathbf{v}(X_n) / \parallel \mathbf{v}(X_n) \parallel \);
**for** \( i = (n - 1) \) **to** 1 **do**

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

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

**set** \( w_i = \mathbf{v}(X_i) - \sum_{j=1}^{n-i} (e_j^* \mathbf{v}(X_i)) e_j, \quad e_{n-i+1} = w_i / \parallel w_i \parallel \)

**Theorem**

\( \{X_1, \ldots, X_n\} \) generated as above has distribution \( \text{DPP}_S(K) \).
Illustration of simulation algorithm

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

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

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

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

and $X_S \sim \text{DPP}_S(K)$ 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:
Step 3. The next point is sampled w.r.t. the following density:
Illustration of simulation algorithm

eetc.
Illustration of simulation algorithm

etc.
Illustration of simulation algorithm

etc.
Illustration of simulation algorithm

etc.
1 Introduction

2 Definition, existence and basic properties

3 Simulation

4 Parametric models

5 Inference
Stationary models

Consider a stationary kernel

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

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

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

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 $\varphi \leq 1$. → This will induce a bound on the parameter space.
Stationary models

Consider a stationary kernel

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

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

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

**Theorem**

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

\[ \varphi \leq 1. \]

To construct parametric families of DPP:
Consider parametric families of \( C_0 \) and rescale so that \( \varphi \leq 1. \)
→ *This will induce a bound on the parameter space.*
Stationary models

Consider a stationary kernel

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

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

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

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 \( \varphi \leq 1. \)
\[ \rightarrow \text{This will induce a bound on the parameter space.} \]
Consider a stationary kernel
\[ C(x, y) = C_0(x - y), \quad x, y \in \mathbb{R}^d. \]

Recall (C1): \( C_0 \) is a continuous covariance function. Moreover, if \( C_0 \in L^2(\mathbb{R}^d) \) we can define its Fourier transform
\[
\varphi(x) = \int C_0(t)e^{-2\pi i x \cdot t} \, dt, \quad x \in \mathbb{R}^d.
\]

**Theorem**

*Under (C1), if \( C_0 \in L^2(\mathbb{R}^d) \), then existence of DPP(\( C_0 \)) is equivalent to\n\[
\varphi \leq 1.
\]

To construct parametric families of DPP:
Consider parametric families of \( C_0 \) and rescale so that \( \varphi \leq 1. \)
\[ \rightarrow \text{This will induce a bound on the parameter space.} \]
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 $\alpha$ is given by
  
  \[
  C_0(x) = \rho \frac{2}{\pi} \left( \arccos(\|x\|/\alpha) - \|x\|/\alpha \sqrt{1 - (\|x\|/\alpha)^2} \right) 1_{\|x\| < \alpha}.
  \]

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

  $\Rightarrow$ Tradeoff between the intensity $\rho$ and the range of repulsion $\alpha$.

- 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:
  
  \[
  C_0(x) = \rho \left( 1 + \|x/\alpha\|^2 \right)^{\nu+d/2}, \quad x \in \mathbb{R}^d.
  \]

  DPP($C_0$) exists iff $\rho \leq \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)(\sqrt{\pi} \alpha)^d}$. 
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 $\alpha$ is given by
  \[ C_0(x) = \rho \frac{2}{\pi} \left( \arccos(\|x\|/\alpha) - \|x\|/\alpha \sqrt{1 - (\|x\|/\alpha)^2} \right) 1_{\|x\|<\alpha}. \]

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

$\Rightarrow$ Tradeoff between the intensity $\rho$ and the range of repulsion $\alpha$.

- 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 $\rho \leq \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)(\sqrt{\pi\alpha})^d}$.
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 $\alpha$ is given by
  \[
  C_0(x) = \rho \frac{2}{\pi} \left( \arccos \left( \frac{\|x\|}{\alpha} \right) - \frac{\|x\|}{\alpha} \sqrt{1 - \left( \frac{\|x\|}{\alpha} \right)^2} \right) 1_{\|x\| < \alpha}.
  \]
  DPP$(C_0)$ exists iff $\varphi \leq 1 \iff \rho \alpha^2 \leq 4/\pi$.
  $\Rightarrow$ Tradeoff between the intensity $\rho$ and the range of repulsion $\alpha$.

- 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:
  \[
  C_0(x) = \rho \frac{1}{\left(1 + \|x/\alpha\|^2\right)^{\nu + d/2}}, \quad x \in \mathbb{R}^d.
  \]
  DPP$(C_0)$ exists iff $\rho \leq \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)(\sqrt{\pi} \alpha)^d}$.
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 $\alpha$ is given by
  \[ C_0(x) = \rho \frac{2}{\pi} \left( \arccos\left(\frac{\|x\|}{\alpha}\right) - \frac{\|x\|}{\alpha} \sqrt{1 - \left(\frac{\|x\|}{\alpha}\right)^2} \right) \mathbf{1}_{\|x\| < \alpha}. \]

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

  $\Rightarrow$ Tradeoff between the intensity $\rho$ and the range of repulsion $\alpha$.

- Whittle-Matérn (includes Exponential and Gaussian):
  \[ C_0(x) = \rho \frac{2^{1-\nu}}{\Gamma(\nu)} \|x/\alpha\|^\nu K_\nu\left(\|x/\alpha\|\right), \quad x \in \mathbb{R}^d. \]

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

- Generalized Cauchy:
  \[ C_0(x) = \frac{\rho}{\left(1 + \|x/\alpha\|^2\right)^{\nu + d/2}}, \quad x \in \mathbb{R}^d. \]

  $\text{DPP}(C_0)$ 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 when the scaling parameter $\alpha$ is chosen such that the range of corr. $\approx 1$:

**In blue** : $C_0$ is the **circular** covariance function.

**In red** : $C_0$ is **Whittle-Matérn**, for different values of $\nu$.

**In green** : $C_0$ is generalized **Cauchy**, for different values of $\nu$. 

---

![Graph showing pair correlation functions for different covariance functions]
Quantifying repulsiveness of stationary point processes

One proposal: For stationary and isotropic point processes, $X_1$ exhibits stronger repulsiveness than $X_2$ if $\rho_1 = \rho_2$ and $g_1(\|\cdot\|) \leq g_2(\|\cdot\|) \leq 1$ (repulsive case).

Remark: it becomes difficult to use this when comparing e.g. a Whittle-Matérn model with a generalized Cauchy model.

Another proposal: for any stationary point process defined on $\mathbb{R}^d$, with intensity $\rho > 0$, and pair correlation function $g(x, y) = g(x - y) \leq 1$,

$$\mu = \rho \int [1 - g(x)] \, dx$$

is a rough measure for repulsiveness.

Interpretation: $\mu$ is the limit as $r \to \infty$ of the difference between the expected number of events within distance $r$ from $o$ under respectively $P$ and $P^o$.
Quantifying repulsiveness of stationary point processes

One proposal: For stationary and isotropic point processes, $X_1$ exhibits stronger repulsiveness than $X_2$ if $\rho_1 = \rho_2$ and $g_1(\|\cdot\|) \leq g_2(\|\cdot\|) \leq 1$ (repulsive case).

Remark: it becomes difficult to use this when comparing e.g. a Whittle-Matérn model with a generalized Cauchy model.

Another proposal: for any stationary point process defined on $\mathbb{R}^d$, with intensity $\rho > 0$, and pair correlation function $g(x, y) = g(x - y) \leq 1$,

$$\mu = \rho \int [1 - g(x)] \, dx$$

is a rough measure for repulsiveness.

Interpretation: $\mu$ is the limit as $r \to \infty$ of the difference between the expected number of events within distance $r$ from $o$ under respectively $P$ and $P^!_o$. 
Quantifying repulsiveness of stationary point processes

One proposal: For stationary and isotropic point processes, $X_1$ exhibits stronger repulsiveness than $X_2$ if $\rho_1 = \rho_2$ and $g_1(\|\cdot\|) \leq g_2(\|\cdot\|) \leq 1$ (repulsive case).

Remark: it becomes difficult to use this when comparing e.g. a Whittle-Matérn model with a generalized Cauchy model.

Another proposal: for any stationary point process defined on $\mathbb{R}^d$, with intensity $\rho > 0$, and pair correlation function $g(x, y) = g(x - y) \leq 1$,

$$\mu = \rho \int [1 - g(x)] \, dx$$

is a rough measure for repulsiveness.

Interpretation: $\mu$ is the limit as $r \to \infty$ of the difference between the expected number of events within distance $r$ from $o$ under respectively $P$ and $P_o^!$. 
Quantifying repulsiveness of stationary point processes

One proposal: For stationary and isotropic point processes, \( X_1 \) exhibits stronger repulsiveness than \( X_2 \) if \( \rho_1 = \rho_2 \) and \( g_1(\| \cdot \|) = g_2(\| \cdot \|) \leq 1 \) (repulsive case).

Remark: it becomes difficult to use this when comparing e.g. a Whittle-Matérn model with a generalized Cauchy model.

Another proposal: for any stationary point process defined on \( \mathbb{R}^d \), with intensity \( \rho > 0 \), and pair correlation function \( g(x, y) = g(x - y) \leq 1 \),

\[
\mu = \rho \int [1 - g(x)] \, dx
\]

is a rough measure for repulsiveness.

Interpretation: \( \mu \) is the limit as \( r \to \infty \) of the difference between the expected number of events within distance \( r \) from \( o \) under respectively \( P \) and \( P_o^1 \).
**Definition:** For stationary point processes $X_1$ and $X_2$ with pair correlation functions $\leq 1$,

$X_1$ is *overall more repulsive than* $X_2$ if $\rho_1 = \rho_2$ and $\mu_1 \geq \mu_2$.

- In the isotropic case, this definition is in agreement with our former concept:

  $$g_1 \leq g_2 \Rightarrow \mu_1 \geq \mu_2 \text{ if } \rho_1 = \rho_2.$$

- $0 \leq \mu \leq 1$.

- For a stationary Poisson process, $\mu = 0$. 
Definition of the “overall degree of repulsiveness” and some properties

**Definition:** For stationary point processes $X_1$ and $X_2$ with pair correlation functions $\leq 1$,

\[ X_1 \text{ is overall more repulsive than } X_2 \text{ if } \rho_1 = \rho_2 \text{ and } \mu_1 \geq \mu_2. \]

- In the isotropic case, this definition is in agreement with our former concept:
  \[ g_1 \leq g_2 \Rightarrow \mu_1 \geq \mu_2 \text{ if } \rho_1 = \rho_2. \]

- $0 \leq \mu \leq 1$.

- For a stationary Poisson process, $\mu = 0$. 
Definition of the “overall degree of repulsiveness” and some properties

**Definition:** For stationary point processes \( X_1 \) and \( X_2 \) with pair correlation functions \( \leq 1 \),

\[ X_1 \text{ is overall more repulsive than } X_2 \text{ if } \rho_1 = \rho_2 \text{ and } \mu_1 \geq \mu_2. \]

- In the isotropic case, this definition is in agreement with our former concept:

\[ g_1 \leq g_2 \Rightarrow \mu_1 \geq \mu_2 \text{ if } \rho_1 = \rho_2. \]

- \( 0 \leq \mu \leq 1. \)

- For a stationary Poisson process, \( \mu = 0. \)
The case of stationary DPP’s

For a stationary DPP,

\[ \mu = \rho \int [1 - g(x)] \, dx = \frac{1}{\rho} \int |C_0(x)|^2 \, dx = \frac{1}{\rho} \int |\varphi(x)|^2 \, dx. \]

So \( \mu \) is maximal iff \( \varphi \) is an indicator function with support on a Borel subset of \( \mathbb{R}^d \) of volume \( \rho \).

Obvious choice:

\[ \varphi(x) = \begin{cases} 
1 & \text{if } \|x\| \leq r \\
0 & \text{otherwise}
\end{cases} \]

where \( r^d = \rho d \Gamma(d/2)/(2\pi^{d/2}) \), since \( \rho = \int \varphi(x) \, dx \).

\( d = 1: \ C_0 \propto \ \text{sinc function}; \)
\( d = 2: \ C_0 \propto \ \text{’jinc-like’ function.} \)
The case of stationary DPP’s

For a stationary DPP,

$$\mu = \rho \int [1 - g(x)] \, dx = \frac{1}{\rho} \int |C_0(x)|^2 \, dx = \frac{1}{\rho} \int |\varphi(x)|^2 \, dx.$$

So $\mu$ is maximal iff $\varphi$ is an indicator function with support on a Borel subset of $\mathbb{R}^d$ of volume $\rho$.

Obvious choice:

$$\varphi(x) = \begin{cases} 1 & \text{if } \|x\| \leq r \\ 0 & \text{otherwise} \end{cases}$$

where $r^d = \rho d \Gamma(d/2)/(2\pi^{d/2})$, since $\rho = \int \varphi(x) \, dx$.

($d = 1$: $C_0 \propto$ sinc function;
$d = 2$: $C_0 \propto$ ’jinc-like’ function.)
The case of stationary DPP’s

- For a stationary DPP,

\[ \mu = \rho \int [1 - g(x)] \, dx = \frac{1}{\rho} \int |C_0(x)|^2 \, dx = \frac{1}{\rho} \int |\phi(x)|^2 \, dx. \]

- So \( \mu \) is maximal iff \( \phi \) is an indicator function with support on a Borel subset of \( \mathbb{R}^d \) of volume \( \rho \).

- Obvious choice:

\[ \phi(x) = \begin{cases} 
1 & \text{if } \|x\| \leq r \\
0 & \text{otherwise}
\end{cases} \]

where \( r^d = \rho d \Gamma(d/2)/(2\pi^{d/2}) \), since \( \rho = \int \phi(x) \, dx \).

\( d = 1: \quad C_0 \propto \text{sinc function}; \)
\( d = 2: \quad C_0 \propto \text{’jinc-like’ function}. \)
Modelling approach based on spectral densities

- Specify a parametric class of integrable functions $\varphi_\theta : \mathbb{R}^d \rightarrow [0, 1]$ (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.
- But it may be difficult to determine $C_{0,\theta} = \mathcal{F}^{-1} \varphi_\theta$ and hence closed form expressions for $g$ and $K$ may not be available.
- 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

$$\rho > 0, \quad \nu > 0, \quad 0 < \alpha \leq \alpha_{\text{max}}(\rho, \nu) := \left(\frac{2\pi^{d/2}\Gamma(d/\nu + 1)}{\rho\Gamma(d/2)}\right)^{1/d}.$$
Modelling approach based on spectral densities

- Specify a parametric class of integrable functions $\varphi_{\theta} : \mathbb{R}^d \rightarrow [0, 1]$ (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.
- But it may be difficult to determine $C_{0,\theta} = \mathcal{F}^{-1}\varphi_{\theta}$ and hence closed form expressions for $g$ and $K$ may not be available.

- 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

$$\rho > 0, \quad \nu > 0, \quad 0 < \alpha \leq \alpha_{\text{max}}(\rho, \nu) := \left(\frac{2\pi^{d/2}\Gamma(d/\nu + 1)}{\rho\Gamma(d/2)}\right)^{1/d}.$$
Modelling approach based on spectral densities

- Specify a parametric class of integrable functions \( \varphi_\theta : \mathbb{R}^d \rightarrow [0, 1] \) (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.
- But it may be difficult to determine \( C_{0,\theta} = \mathcal{F}^{-1}\varphi_\theta \) and hence closed form expressions for \( g \) and \( K \) may not be available.
- 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

\[
\rho > 0, \quad \nu > 0, \quad 0 < \alpha \leq \alpha_{\text{max}}(\rho, \nu) := \left(\frac{2\pi^{d/2}\Gamma(d/\nu + 1)}{\rho \Gamma(d/2)}\right)^{1/d}.
\]
Power exponential spectral model: (isotropic) spectral densities and pair correlation functions ($d = 2$)
Approximation of stationary DPP’s models

Consider a stationary kernel $C_0$ and $X \sim \text{DPP}(C_0)$.
- The simulation and the density of $X_S$ requires the expansion

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

but in general $\lambda_k^S$ and $\phi_k^S$ are not expressible on closed form.
- Consider the unit box $S = [-\frac{1}{2}, \frac{1}{2}]^d$ and the Fourier expansion

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

The Fourier coefficients are

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

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 DPP’s models

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

• The simulation and the density of $X_S$ requires the expansion

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

but in general $\lambda_k^S$ and $\phi_k^S$ are not expressible on closed form.

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

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

The Fourier coefficients are

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

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 DPP’s models

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

- 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 $\lambda_k^S$ and $\phi_k^S$ are not expressible on closed form.

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

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

The Fourier coefficients are

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

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 $\rho > 5$. 


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 $\varphi$ 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 $S$. 
Approximation of stationary DPP 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 $\varphi$ 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 $S$. 
<table>
<thead>
<tr>
<th></th>
<th>Introduction</th>
<th>Definition</th>
<th>Simulation</th>
<th>Parametric models</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Definition, existence and basic properties</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Parametric models</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Inference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Consider a stationary and isotropic parametric DPP($C$), e.g.

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

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

The first and second moments are easily deduced:

- The intensity is $\rho$.

- 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 $R_\alpha$:
  if e.g. $d = 2$,

\[ K_\alpha(r) := 2\pi \int_0^r t g_0(t) \, dt = \pi r^2 - 2\pi \int_0^r t |R_\alpha(t)|^2 \, dt. \]
Consider a stationary and isotropic parametric DPP($C$), e.g.

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

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

The first and second moments are easily deduced:

- The intensity is $\rho$.
- 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 $R_\alpha$: if e.g. $d = 2$,
  $$K_\alpha(r) := 2\pi \int_0^r t g_0(t) \, dt = \pi r^2 - 2\pi \int_0^r t |R_\alpha(t)|^2 \, dt.$$
Consider a stationary and isotropic parametric DPP($C$), e.g. 
\[ C(x, y) = C_0(x - y) = \rho R_\alpha(\|x - y\|), \]
with $R_\alpha(0) = 1$.

The first and second moments are easily deduced:

- The intensity is $\rho$.
- 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 $R_\alpha$: if e.g. $d = 2$, 
  \[ K_\alpha(r) := 2\pi \int_0^r tg_0(t) \, dt = \pi r^2 - 2\pi \int_0^r t|R_\alpha(t)|^2 \, dt. \]
Parameter estimation

When e.g.

\[ C(x, y) = \rho R_\alpha(\|x - y\|), \]

1. estimate \( \rho \) by \( \#\{\text{obs. points}\}/[\text{area of obs. window}] \);

2. estimate \( \alpha \)

   - either by \textbf{minimum contrast} estimator (MCE):
     \[ \hat{\alpha} = \arg\min_\alpha \int_0^{r_{\max}} \left| \sqrt{\hat{K}(r)} - \sqrt{K_\alpha(r)} \right|^2 \, dr \]

   - or by \textbf{maximum likelihood} estimator: given \( \hat{\rho} \), the likelihood is deduced from the kernel approximation.
Parameter estimation

When e.g.

\[ C(x, y) = \rho R_\alpha(\|x - y\|), \]

1. estimate \( \rho \) by \( \frac{\#\{\text{obs. points}\}}{\text{area of obs. window}} \);
2. estimate \( \alpha \)
   - either by \textbf{minimum contrast} estimator (MCE):
     \[ \hat{\alpha} = \arg\min_{\alpha} \int_0^{r_{\text{max}}} \left| \sqrt{\hat{K}(r)} - \sqrt{K_\alpha(r)} \right|^2 \, dr \]
   - or by \textbf{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 $\alpha$ from 200 realisations

Gaussian model

Exponential model
Example: 134 Norwegian pine trees observed in a $56 \times 38$ 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, for $C_0$,

- either Whittle-Matérn model; or generalized Cauchy model; or Gaussian model (the limit of both).

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

Second, for $\varphi$,

- power exponential spectral model.

- A much better fit, with

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

i.e. close to the “most repulsive possible stationary DPP”.
First, for $C_0$,

- either Whittle-Matérn model; or generalized Cauchy model; or Gaussian model (the limit of both).

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

Second, for $\varphi$,

- power exponential spectral model.

- 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.

- 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.

⇒ Promising alternative to repulsive Gibbs 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.

⇒ Promising alternative to repulsive Gibbs 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.

⇒ Promising alternative to repulsive Gibbs point processes.
References


