A $\beta$-mixing inequality for point processes induced by their intensity functions

Arnaud Poinas

IRMAR, Campus de Beaulieu, Bat. 22/23, 263 avenue du Général Leclerc, 35042 Rennes, France

Abstract

We prove a general inequality on $\beta$-mixing coefficients of point processes depending uniquely on their $n$-th order intensity functions. We apply this inequality in the case of determinantal point processes and show that its rate of convergence is optimal.

Keywords: lower sum transform, determinantal point processes

1. Introduction

In asymptotic inference for dependent random variables, it is necessary to quantify the dependence between $\sigma$-algebras. The first measures of dependence that have been introduced are the alpha-mixing coefficients [25] and the beta-mixing coefficients [26]. They have been used to establish moment inequalities, exponential inequalities and central limit theorems for stochastic processes (see [7, 12, 24] for more details about mixing) with wide applications in statistics, see for instance [6, 10]. In this paper, we focus on spatial point processes. As detailed below, for these models, alpha-mixing has been widely studied and exploited in the literature, but not beta-mixing in spite of its stronger properties. In a lesser extent, some alternative measures of dependence have also been used for spatial point processes, namely Brillinger mixing [4, 14] (which only applies to stationary point processes) and association [18, 22].

The main models used in spatial point processes are Gibbs point processes, Cox processes and determinantal point processes, see [21] for a recent review. An $\alpha$-mixing inequality is established for Gibbs point processes in the Dobrushin uniqueness region in [13]. It has been used to show asymptotic normality of maximum likelihood and pseudo-likelihood estimates [16]. Similarly, some inhomogeneous Cox processes like the Neyman-Scott process have also been showed to satisfy $\alpha$-mixing inequalities in [28]. These inequalities are at the core of asymptotic inference results in [8, 23, 28]. Finally, an $\alpha$-mixing inequality has also been showed for determinantal point processes in [22] and used to get the asymptotic normality of a wide class of estimators of these models.

On the other hand, $\beta$-mixing is a stronger property than $\alpha$-mixing. It implies stronger covariance inequalities [24, 11] as well as a coupling theorem known as Berbee’s Lemma [3] used in various limit theorems (for example in [2, 27]). Nevertheless, it rarely appears in the literature in comparison to $\alpha$-mixing. This is especially true for point processes where there has been no $\beta$-mixing property established for any specific class of point process models to the author’s knowledge even if there exists some general results for $\beta$-mixing point processes [14]. Our goal is to establish a general inequality for the $\beta$-mixing coefficients of a point process in terms of its intensity functions.

We begin in Section 2 by recalling the basic definitions and properties of the $\alpha$-mixing and $\beta$-mixing coefficients and we introduce the lower sum transform which is the main technical tool that we use throughout the paper. Then, a general inequality for the $\beta$-mixing coefficients of a point process that depends only on its $n$-th order intensity functions is proved in Section 3. As an example, we deduce a $\beta$-mixing inequality in the special case of determinantal point processes (DPPs) in Section 4 and we show that its rate of convergence is optimal for a wide class of DPPs.
2. Preliminaries

2.1. Intensities of point processes

In this paper, we consider simple point processes on \((\mathbb{R}^d, \mu)\) equipped with the euclidean norm \(\| \cdot \|\) where \(d\) is a fixed integer and \(\mu\) is a locally finite measure (more information on spatial point processes can be found in [12, 24]). We denote by \(\Omega\) (resp. \(\Omega_f\)) the set of locally finite (resp. finite) point configurations in \(\mathbb{R}^d\). For all functions \(f : \Omega_f \rightarrow \mathbb{R}, n \in \mathbb{N}\) and \(x = (x_1, \cdots, x_n) \in (\mathbb{R}^d)^n\), we write \(f(x)\) for \(f(x_1, \cdots, x_n)\) by an abuse of notation. Finally, we write \(|A|\) for the cardinal of a finite set \(A\) and \(\|f\|_\infty\) for the uniform norm of a function \(f\).

We begin by recalling that intensity functions are defined the following way (see [20]).

**Definition 2.1.** Let \(X\) be a simple point process on \(\mathbb{R}^d\) and \(n \geq 1\) be an integer. If there exists a non negative function \(\rho_n : (\mathbb{R}^d)^n \rightarrow \mathbb{R}\) such that

\[
\mathbb{E} \left[ \sum_{x_1, \cdots, x_n \in X} f(x_1, \cdots, x_n) \right] = \int_{(\mathbb{R}^d)^n} f(x)\rho_n(x)\,d\mu^n(x).
\]

for all locally integrable functions \(f : (\mathbb{R}^d)^n \rightarrow \mathbb{R}\) then \(\rho_n\) is called the \(n\)th order intensity function of \(X\).

In the rest of the paper, all point processes will be considered to admit bounded \(n\)-th order intensity function for all \(n \geq 1\).

2.2. Mixing

In this part, we recall generalities about mixing in the general case of two sub \(\sigma\)-algebras and in the more specific case of point processes.

2.2.1. The general case

Consider a probability space \((\mathcal{X}, \mathcal{F}, \mathbb{P})\) and \(\mathcal{A}, \mathcal{B}\) two sub \(\sigma\)-algebras of \(\mathcal{F}\). Let \(\mathbb{P}_\mathcal{A}\) and \(\mathbb{P}_\mathcal{B}\) be the respective restrictions of \(\mathbb{P}\) to \(\mathcal{A}\) and \(\mathcal{B}\) and define the probability \(\mathbb{P}_\mathcal{A} \otimes \mathbb{P}_\mathcal{B}\) on the product \(\sigma\)-algebra by \(\mathbb{P}_\mathcal{A} \otimes \mathbb{P}_\mathcal{B}(A \times B) = \mathbb{P}(A \cap B)\) for all \(A \in \mathcal{A}\) and \(B \in \mathcal{B}\). The \(\alpha\)-mixing and \(\beta\)-mixing coefficients (also called strong-mixing and absolute regularity coefficients) are defined as the following measures of dependence between \(\mathcal{A}\) and \(\mathcal{B}\) [12, 24]:

\[
\alpha(\mathcal{A}, \mathcal{B}) := \sup \{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A}, B \in \mathcal{B} \}, \tag{2}
\]

\[
\beta(\mathcal{A}, \mathcal{B}) := |\mathbb{P}_\mathcal{A} \otimes \mathbb{P}_\mathcal{B} - \mathbb{P}_\mathcal{A} \otimes \mathbb{P}_\mathcal{B}|_{TV}, \tag{3}
\]

where \(| \cdot |_{TV}\) is the total variation distance of a signed norm. If \(\mathcal{A}\) and \(\mathcal{B}\) are generated by random variables \(X\) and \(Y\) then \(\beta(\mathcal{A}, \mathcal{B})\) is the total variation distance between the law of \((X, Y)\) and the law of \((X, Y')\) for some independent copy \(Y'\) of \(Y\). Therefore, the definition of the \(\beta\)-mixing coefficients is also equivalent to

\[
\beta(\mathcal{A}, \mathcal{B}) := \frac{1}{2} \sup_{\|f\|_{\infty} \leq 1} \mathbb{E}|f(X, Y) - f(X, Y')|. \tag{3'}
\]

Finally, this definition is also equivalent in this case to the more commonly used one:

\[
\beta(\mathcal{A}, \mathcal{B}) := \frac{1}{2} \sup \left\{ \sum_{i,j} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)| \right\}, \tag{3''}
\]

where the supremum is taken over all finite partition \((A_i)_{i \in I}\) of \(\mathcal{A}\) and all finite partition \((B_j)_{j \in J}\) of \(\mathcal{B}\). In particular, \(\alpha\)-mixing and \(\beta\)-mixing coefficients satisfy the relation \(2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B})\).
2.2.2. The case of point processes

For a given point process $X$ and a bounded set $A \subset \mathbb{R}^d$, we denote by $\mu(A) := \int_A \mathrm{d}\mu(x)$ the volume of $A$ and $\mathcal{E}(A)$ the $\sigma$-algebra generated by $X \cap A$. Finally, for all $A, B \subset \mathbb{R}$, we write $\text{dist}(A, B)$ for the infimum of $\|y - x\|$ where $(x, y) \in A \times B$. The $\beta$-mixing coefficients of the point process $X$ are then defined by

$$\beta_{p,q}(r) := \sup\{\beta(\mathcal{E}(A), \mathcal{E}(B)) : \mu(A) \leq p, \mu(B) \leq q, \text{dist}(A, B) > r\},$$

and we say that the point process $X$ is beta-mixing if $\beta_{p,q}(r)$ vanishes when $r \to +\infty$ for all $p, q > 0$.

Our goal is to prove that under appropriate assumptions over the intensity functions $\rho_n$ of $X$ we have a $\beta$-mixing property.

2.3. Lower sum transform

The main tool we use throughout this paper is the so-called lower sum operator (see [1]). Notice that as shown in Example 4.19 in [1], this operator admits the following inverse transform.

**Proposition 2.3** ([1, Theorem 4.18]). The operator (4) admits an inverse transform $\hat{f}$, called the lower difference of $f$, defined by

$$\hat{f} : \Omega \rightarrow -\sum_{Z \subset X} f(Z).$$

As shown in Example 4.19 in [1], this operator admits the following inverse transform.

These definitions extend to functions over $\Omega_f^2$ by defining

$$\hat{f} : (X_1, X_2) \mapsto -\sum_{Z_1 \subset X_1, Z_2 \subset X_2} f(Z_1, Z_2)$$

and

$$\hat{f} : (X_1, X_2) \mapsto -\sum_{Z_1 \subset X_1, Z_2 \subset X_2} (-1)^{|X_1 \setminus Z_1| + |X_2 \setminus Z_2|} f(Z_1, Z_2).$$

These operators allow us to give an explicit expression for the expectation of a functional of a point process with respect to its intensity functions.

**Proposition 2.4.** If $X$ is an almost surely finite point process such that $\mathbb{E}[|X|] < +\infty$, then

$$\mathbb{E}[f(X)] = \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{(\mathbb{R}^d)^n} \hat{f}(x)\rho_n(x)\mathrm{d}\mu^n(x)$$

for all bounded functions $f : \Omega_f \rightarrow \mathbb{R}$. Moreover, if $X'$ is a point process independent from $X$ satisfying the same assumptions than $X$ and with $n$-th order intensity functions $\rho_n'$, then

$$\mathbb{E}[f(X, X')] = \sum_{m,n=0}^{+\infty} \frac{1}{m!n!} \int_{(\mathbb{R}^d)^{m+n}} \hat{f}(x, y)\rho_m(x)\rho_n'(y)\mathrm{d}\mu^m(x)\mathrm{d}\mu^n(x)$$

for all bounded functions $f : \Omega_f^2 \rightarrow \mathbb{R}$. 

3
Proof. Using the bound $|\hat{f}(x)| \leq \|f\|_\infty 2^{|x|}$ we get

$$\sum_{n \geq 0} \mathbb{E} \left[ \sum_{Z \subset X, |Z| = n} \hat{f}(Z) \right] \leq \sum_{n \geq 0} \mathbb{E} \left[ 2^{|X|} \left( \frac{|X|}{n} \right) \right] |f|_\infty = \|f\|_\infty \mathbb{E} \left[ 4^{|X|} \right] < +\infty. \tag{8}$$

Since we can write

$$f(X) = \hat{f}(X) = \sum_{Z \subset X} \hat{f}(Z) = \sum_{n \geq 0} \sum_{Z \subset X, |Z| = n} \hat{f}(Z) \text{ a.s.,}$$

then

$$\mathbb{E}[f(X)] = \sum_{n \geq 0} \mathbb{E} \left[ \sum_{Z \subset X, |Z| = n} \hat{f}(Z) \right] = \sum_{n \geq 0} \frac{1}{n!} \int_{(\mathbb{R}^d)^n} \hat{f}(x) \rho_n(x) d\mu^n(x).$$

Similarly, for all functions $f : \Omega \to \mathbb{R}$ we have

$$\mathbb{E}[f(X, X') | X'] = \mathbb{E} [ \mathbb{E} [f(X, X') | X'] ] = \mathbb{E} \left[ \sum_{m=0}^{+\infty} \frac{1}{m!} \int_{(\mathbb{R}^d)^m} \left( \sum_{z \subset X} (-1)^{m-|z|} f(z, X') \right) \rho_m(x) d\mu^m(x) \right],$$

where all inversion of expectation with sum and integrals can be justified in a similar fashion than (8). \hfill \Box

3. $\beta$-mixing of point processes with known intensity functions

Our main result is the following inequality showing that if all $\rho_m(x) \rho_n(y) - \rho_{m+n}(x, y)$ vanish fast enough when $\|y - x\| \to +\infty$ for all $m, n \in \mathbb{N}$, then the underlying point process is $\beta$-mixing.

**Theorem 3.1.** Let $X$ be a simple point process on $(\mathbb{R}^d, \mu)$ such that $\mathbb{E}[4^{|X\cap A|}] < +\infty$ for all bounded subsets $A \subset \mathbb{R}^d$. Then, for all $p, q, r \in \mathbb{R}_+$,

$$\beta_{p,q}(r) \leq \sup_{\mu(A) < p, \mu(B) < q} \left( \sum_{m,n=0}^{+\infty} \frac{2^{n+m}}{m! n!} \int_{A^m \times B^n} |\rho_m(x) \rho_n(y) - \rho_{m+n}(x, y)| d\mu^m(x) d\mu^n(y) \right). \tag{9}$$

Before giving the proof of Theorem 3.1 we need the following lemmas showing the behaviour of $f(X \cap A, X \cap B)$ and $f(X \cap A, X' \cap B)$ under the lower difference operator. For all bounded subsets $A \subset \mathbb{R}^d$ and functions $f : \Omega \to \mathbb{R}$ we denote by $f_A$ the function $f : X \to f(X \cap A)$.

**Lemma 3.2.** Let $A \subset \mathbb{R}^d$ and $f : \Omega \to \mathbb{R}$. Then,

$$\tilde{f}_A(X) = f(X) 1_{X \subset A}.$$

**Proof.** If $X \subset A$ then the result is trivial. Otherwise, there exists $x \in X \setminus A$ and we can write

$$\tilde{f}_A(X) = \sum_{Z \subset X, Z \ni x} (-1)^{|X\setminus Z|} f(Z \cap A) + \sum_{Z \subset X, Z \notin x} (-1)^{|X\setminus Z|} f(Z \cap A)$$

$$= \sum_{Z \subset X, Z \ni x} (-1)^{|X\setminus Z|} f(Z \cap A) + \sum_{Z \subset X, Z \notin x} (-1)^{|X\setminus Z|} f(Z \cap A)$$

$$= 0. \hfill \Box$$
It is easy to extend this result to multivariate functions: The lower difference of \( (X_1, X_2) \to f(X_1 \cap A_1, X_2 \cap A_2) \) is \( \tilde{f}(X_1, X_2) \mathbb{1}_{\{X_1 \in A_1\}} \mathbb{1}_{\{X_2 \in A_2\}} \).

**Lemma 3.3.** For all \( f : \Omega_f^d \to \mathbb{R} \) and \( A, B \) disjoint subsets of \( \mathbb{R}^d \), let us define the function \( g : X \to f(X \cap A, X \cap B) \). The lower difference of \( g \) satisfies
\[
\tilde{g}(X) = \tilde{f}(X \cap A, X \cap B) \mathbb{1}_{\{X \in A \cup B\}}.
\]

**Proof.** Using Lemma 3.2, we get that \( \tilde{g}(X) = 0 \) whenever \( X \notin (A \cup B) \). Otherwise, since \( A \) and \( B \) are disjoint sets,
\[
\tilde{g}(X) = \sum_{Z \subseteq X} (-1)^{|X \setminus Z|} f(Z \cap A, Z \cap B) = \sum_{U \subseteq X \cap A} \sum_{V \subseteq X \cap B} (-1)^{|(X \cap A) \setminus U| + |(X \cap B) \setminus V|} f(U, V)
\]
which, by definition, is equal to \( \tilde{f}(X \cap A, X \cap B) \).

We now have the necessary tools required for the proof of Theorem 3.1.

**Proof of Theorem 3.1.** Let \( p, q > 0 \) and \( A, B \) be two disjoint subsets of \( \mathbb{R}^d \) such that \( \mu(A) \leq p \) and \( \mu(B) \leq q \). Consider a function \( f : \Omega_f^d \to \mathbb{R} \) such that \( \|f\|_\infty = 1 \). By Definition (3) of the \( \beta \)-mixing coefficient, we need to bound the expression
\[
|\mathbb{E}[f(X \cap A, X \cap B)] - \mathbb{E}[f(X \cap A, X' \cap B)]|
\]
where \( X' \) is an independent copy of \( X \). Since \( X \cap A, X' \cap B \) and \( X \cap B \) are finite a.s., we can apply (7) which, combined with Lemma 3.2, gives us
\[
\mathbb{E}[f(X \cap A, X' \cap B)] = \sum_{m,n=0}^{+\infty} \frac{1}{m!n!} \int_{A^m \times B^n} \tilde{f}(x, y)\rho_m(x)\rho_n(y)d\mu^m(x)d\mu^n(y).
\]
On the other hand, by combining (9) with Lemma 3.3, we get
\[
\mathbb{E}[f(X \cap A, X \cap B)] = \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{(A \cup B)^n} \tilde{f}(x \cap A, x \cap B)\rho_n(x)d\mu^n(x).
\]
Since \( A \) and \( B \) are disjoint sets and by symmetry of \( \tilde{f}(x \cap A, x \cap B)\rho_n(x) \), we can simplify the above expression into
\[
\mathbb{E}[f(X \cap A, X \cap B)] = \sum_{n=0}^{+\infty} \frac{1}{n!} \left( \frac{n}{m!} \right) \int_{A^m \times B^{n-m}} \tilde{f}(x, y)\rho_n(x, y)d\mu^m(x)d\mu^{n-m}(y)
\]
\[
= \sum_{m,n=0}^{+\infty} \frac{1}{m!n!} \int_{A^m \times B^n} \tilde{f}(x, y)\rho_{m+n}(x, y)d\mu^m(x)d\mu^n(y).
\]
Combining these two results together yields that \( |\mathbb{E}[f(X \cap A, X \cap B)] - \mathbb{E}[f(X \cap A, X' \cap B)]| \) is equal to
\[
\left| \sum_{m,n=0}^{+\infty} \frac{1}{m!n!} \int_{A^m \times B^n} \tilde{f}(x, y)(\rho_m(x)\rho_n(y) - \rho_{m+n}(x, y))d\mu^m(x)d\mu^n(y) \right|
\]
which is bounded by
\[
\sum_{m,n=0}^{+\infty} \frac{2^{n+m}}{m!n!} \int_{A^m \times B^n} |\rho_m(x)\rho_n(y) - \rho_{m+n}(x, y)|d\mu^m(x)d\mu^n(y)
\]
where we used the bound \( |\tilde{f}(x, y)| \leq 2^{|x|+|y|} \).
4. Application to determinantal point processes

We can directly apply Theorem 3.1 to determinantal point processes. First introduced in [19] to model fermion systems, DPPs are a broad class of repulsive point processes. We recall that a DPP $X$ with kernel $K : (\mathbb{R}^d)^2 \to \mathbb{R}$ is defined by its intensity functions

$$
\rho_n(x_1, \ldots, x_n) = \det(K[x]) \quad \forall x \in (\mathbb{R}^d)^n, \ \forall n \in \mathbb{N}
$$

where we denote by $K[x]$ the matrix $(K(x_i, x_j))_{1 \leq i,j \leq n}$. For existence, $K$ is assumed to be a locally square integrable hermitian measurable function such that its associated integral operator $K$ is locally of trace class with eigenvalues in $[0,1]$ (see [15] for more details about DPPs).

The application of Theorem 3.1 to DPPs gives us the following $\beta$-mixing condition:

**Theorem 4.1.** Let $X$ be a DPP with kernel $K$ and define

$$
\omega(r) := \sup_{\|y-x\| \geq r} |K(x, y)|.
$$

If $K$ is bounded and $\omega(r) \to 0$ then $X$ is $\beta$-mixing. In particular,

$$
\beta_{p,q}(r) \leq 4pq(1 + 2p\|K\|_{\infty})(1 + 2q\|K\|_{\infty})e^{2\|K\|_{\infty}(p+q)}\omega(r)^2.
$$

Furthermore, if there exists $N \in \mathbb{N}$ such that $\text{rank}(K) \leq N$ then

$$
\beta_{p,q}(r) \leq 4pqN^29^N\omega(r)^2.
$$

**Proof.** Since $\mathbb{E}[\chi_{[X \cap A]}] < +\infty$ for all bounded sets $A$ (see [22, Lemma B.5]) then the $\beta$-mixing coefficients of $X$ satisfy [22] by Theorem 3.1. Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$, we need to control $|\det(K[x])\det(K[y]) - \det(K[x, y])|$ where $\|x - y\| \geq r$. By [22, Lemma B.4], we get the bound

$$
|\det(K[x])\det(K[y]) - \det(K[x, y])| \leq nm\|K\|_{\infty}^{n+m-2}\sum_{i=1}^n \sum_{j=1}^m K(x_i, y_j)^2.
$$

Injecting this bound into (11) gives us

$$
\beta_{p,q}(r) \leq \sum_{n,m=0}^{+\infty} \frac{n^2m^{2n+m}p^{n-1}q^{m-1}\|K\|_{\infty}^{n+m-2}}{n!m!} \sup_{\text{dist}(A,B) \geq r} \int_{A \times B} |K(x, y)|^2 d\mu(x)d\mu(y) \tag{11}
$$

$$
\leq \sum_{n,m=0}^{+\infty} \frac{n^2m^{2n+m}p^np^m\|K\|_{\infty}^{n+m-2}}{n!m!} \omega(r)^2
$$

$$
= 4pq(1 + 2p\|K\|_{\infty})(1 + 2q\|K\|_{\infty})e^{2(p+q)\|K\|_{\infty}}\omega(r)^2.
$$

In particular, if $\omega(r)$ vanishes when $r \to +\infty$ then $X$ is $\beta$-mixing.

In the case where $\text{rank}(K) \leq N$ then $|X| \leq N$ almost surely (see [15, Theorem 4.5.3]) thus $\rho_n(x) = 0$ for $\mu$-almost all $x \in (\mathbb{R}^d)^n$ and $n > N$. Moreover, using the inequality

$$
|\det(K[x])\det(K[y]) - \det(K[x, y])| \leq \sum_{i=1}^n \Delta_i(K[x]) \sum_{j=1}^m \Delta_j(K[y]) \sum_{i=1}^n \sum_{j=1}^m K(x_i, y_j)^2
$$

we get

$$
\beta_{p,q}(r) \leq \sum_{n,m=0}^{+\infty} \frac{n^2m^{2n+m}p^{n-1}q^{m-1}\|K\|_{\infty}^{n+m-2}}{n!m!} \omega(r)^2
$$

$$
= \sum_{n,m=0}^{+\infty} \frac{n^2m^{2n+m}p^np^m\|K\|_{\infty}^{n+m-2}}{n!m!} \omega(r)^2
$$

$$
= 4pq(1 + 2p\|K\|_{\infty})(1 + 2q\|K\|_{\infty})e^{2(p+q)\|K\|_{\infty}}\omega(r)^2.
$$
that appears at the end of the proof of [22, Lemma B.5], where $\Delta_i(K[x])$ is the $i$-th principal minor of $K[x]$, we get for all $A, B$ such that $\mu(A) \leq p, \mu(B) \leq q$ and $\text{dist}(A, B) \geq r$,
\[
\sum_{m,n=0}^{\infty} \frac{2^{n+m}}{m!n!} \int_{A^n \times B^n} |\rho_m(x)\rho_n(y) - \rho_{m+n}(x,y)| d\mu^m(x)d\mu^n(y) \
\leq \omega(r)^2 \sum_{m,n=0}^{\infty} \frac{2^{n+m}nnm}{m!n!} \int_{A^n \times B^n} \Delta_i(K[x]) \sum_{j=1}^{m} \Delta_j(K[y]) d\mu^m(x)d\mu^n(y) \
\leq pq\omega(r)^2 \sum_{m,n=1}^{\infty} \frac{2^{n+m}nnm}{(m-1)!(n-1)!} \int_{A^{m-1} \times B^{n-1}} \rho_{m-1}(x)\rho_{n-1}(y) d\mu^m(x)d\mu^n(y) \
\leq 4pqN^2\omega(r)^2 \sum_{m,n=0}^{\infty} \frac{2^{n+m}}{m!n!} \int_{A^n \times B^n} \rho_m(x)\rho_n(y) d\mu^m(x)d\mu^n(y). \tag{12}
\]

On the other hand, notice that if we apply \(8\) to the function $f(X) = 3^{\lfloor X \cap A \rfloor}$ we get that
\[
\mathbb{E}[3^{\lfloor X \cap A \rfloor}] = \sum_{n=0}^{\infty} \frac{2^n}{n!} \int_{A^n} \rho_n(x) d\mu(x) \tag{13}
\]
since the lower difference transform of $f$ is equal to
\[
\tilde{f}(x) = \sum_{Z \subseteq X} (-1)^{|X| - |Z|} 3^{|Z|} = \sum_{k=0}^{|X|} \binom{|X|}{k} (-1)^{|X|} 3^k = 2^{|X|}.
\]

Finally, plugging \(13\) back into \(12\) gives us
\[
\beta_{p,q}(r) \leq \sup_{|A| < p, |B| < q, \text{dist}(A,B) > r} 4pqN^2\omega(r)^2 \mathbb{E}[3^{\lfloor X \cap A \rfloor}] \mathbb{E}[3^{\lfloor X \cap B \rfloor}] \leq 4pqN^29^N\omega(r)^2.
\]

In conclusion, the $\beta$-mixing coefficients of DPPs decay at the same rate than $|K(x,y)|^2$ does when $x$ and $y$ deviates from each other. For example, kernels of the Ginibre ensemble or the Gaussian unitary ensemble have an exponential decay (see [15]). Moreover, among translation-invariant kernels used in spatial statistics (see [3, 17]), all kernels of the Laguerre-Gaussian family also have an exponential decay while kernels of the Whittle-Matérn and Cauchy family satisfy $\omega(r) = o(r^{-\delta})$ and kernels of the Bessel family satisfy $\omega(r) = O(r^{-(d+1)/2})$.

It is also worth noticing that Theorem \(4.1\) is optimal in the sense that for a wide class of DPPs, the $\beta$-mixing coefficients $\beta_{p,q}(r)$ do not decay faster than $\sup_{|A| < p, |B| < q, \text{dist}(A,B) > r} \int_{A \times B} |K(x,y)|^2 dxdy$ when $r$ goes to infinity, as stated in the following proposition.

**Proposition 4.2.** Let $X$ be a DPP with a non-negative bounded kernel $K$ such that the eigenvalues of $K$ are all in $[0, M]$ where $M < 1$. Then, for all $p, q, r > 0$,
\[
2(1-M)^{\frac{p+q+1}{2}} \sup_{|A| < p, |B| < q, \text{dist}(A,B) > r} \int_{A \times B} |K(x,y)|^2 dxdy \leq \beta_{p,q}(r)
\]
\[
\leq 4(1 + 2|K|_\infty^2) (1 + 2q|K|_\infty^2)^{2(p+q)} \int_{A \times B} |K(x,y)|^2 dxdy.
\]

**Proof.** The first inequality is a consequence of the fact that $\beta_{p,q}(r) \geq 2\alpha_{p,q}(r)$ and [22, Proposition 4.3]. The second inequality is equivalent to \(11\) once the sum has been developed.
References

[1] M. Aigner, Combinatorial Theory, Springer-Verlag, 1979.

[2] M. Banna, F. Merlevède, P. Youssef, Bernstein-type inequality for a class of dependent random matrices, Random Matrices: Theory Appl. 5, 28pp.

[3] H. Berbee, Random walks with stationary increments and renewal theory, Mathematisch Centrum, 1979.

[4] C. Biscio, F. Lavancier, Brillinger mixing of determinantal point processes and statistical applications, Electron. J. Statist. 10 (2016) 582–607.

[5] C. Biscio, F. Lavancier, Quantifying repulsiveness of determinantal point processes, Bernoulli 22 (2016) 2001–2028.

[6] D. Bosq, Nonparametric Statistics for Stochastic Processes: Estimation and Prediction, Springer Science & Business Media, 2012.

[7] R. Bradley, Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions, Probability Surveys 2 (2005) 107–144.

[8] J. Coeurjolly, J. Møller, Variational approach for spatial point process intensity estimation, Bernoulli 20 (2014) 1097–1125.

[9] D. J. Daley, D. Vere-Jones, An introduction to the theory of point processes. Vol. I, Springer-Verlag, ISBN 0-387-95541-0, 2003.

[10] J. Dedecker, Weak Dependence: With Examples and Applications, Springer Science & Business Media, 2007.

[11] B. Delyon, Limit theorem for mixing processes, Tech. Rep., IRISA, Rennes1, 1990.

[12] P. Doukhan, Mixing: Properties and Examples, Springer-Verlag, 1994.

[13] H. Föllmer, A Covariance Estimate for Gibbs Measures, Journal of Functional Analysis 46 (1982) 387–395.

[14] L. Heinrich, Z. Pawlas, Absolute regularity and Brillinger-mixing of stationary point processes, Lithuanian Math. J. 53 (2013) 293–310.

[15] J. Hough, M. Krishnapur, Y. Peres, B. Virag, Zeros of Gaussian Analytic Functions and Determinantal Point Processes, American Mathematical Society, 2009.

[16] J. Jensen, Asymptotic Normality of Estimates in Spatial Point Processes, Scandinavian Journal of Statistics 20 (1993) 97–109.

[17] F. Lavancier, J. Møller, E. Rubak, Determinantal point process models and statistical inference, Journal of Royal Statistical Society: Series B (Statistical Methodology) 77 (2015) 853–877.

[18] R. Lyons, Determinantal Probability: Basic Properties and Conjectures, Proceedings of the International Congress of Mathematicians, Seoul, Korea IV (2014) 137–161.

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

[20] J. Møller, R. Waagepetersen, Statistical Inference and Simulation for Spatial Point Processes, Chapman and Hall/CRC, 2004.
[21] J. Møller, R. Waagepetersen, Some Recent Developments in Statistics for Spatial Point Patterns, Annual Review of Statistics and Its Application 4 (1) (2017) 317–342.

[22] A. Poinas, B. Delyon, F. Lavancier, Mixing properties and central limit theorem for associated point processes, to appear in Bernoulli, 2017.

[23] M. Prokešová, E. Jensen, Asymptotic Palm likelihood theory for stationary point processes, Ann Inst Stat Math 65 (2013) 387–412.

[24] E. Rio, Asymptotic Theory of Weakly Dependent Random Processes, Springer Berlin Heidelberg, 2017.

[25] M. Rosenblatt, A central limit theorem and a strong mixing condition., Proc. Nat. Acad. Sci. U.S.A. 42 (1956) 43–47.

[26] Y. Rozanov, V. Volkonskii, Some limit theorems for random functions I, Theory Probab. Appl. 4 (1959) 178–197.

[27] G. Viennet, Inequalities for absolutely regular sequences: application to density estimation, Probab. Theory Relat. Fields 107 (1997) 467–492.

[28] R. Waagepetersen, Y. Guan, Two-step estimation for inhomogeneous spatial point processes, Journal of the Royal Statistical Society 71 (2009) 685–702.