J. Appl. Prob. **56,** 139–152 (2019)
*doi:*10.1017/jpr.2019.10

ON NEGATIVE ASSOCIATION OF SOME FINITE POINT PROCESSES ON GENERAL STATE SPACES

GÜNTER LAST,* *Karlsruhe Institute of Technology*
RYSZARD SZEKLI, ** *University of Wroclaw*

Abstract

We study negative association for mixed sampled point processes and show that negative association holds for such processes if a random number of their points fulfils the ultra log-concave (ULC) property. We connect the negative association property of point processes with directionally convex dependence ordering, and show some consequences of this property for mixed sampled and determinantal point processes. Some applications illustrate the general theory.

Keywords: Finite point process; mixed sampled point process; negative association; strong Rayleigh measure; ULC

2010 Mathematics Subject Classification: Primary 60K25
Secondary 60J25

1. Introduction

The questions studied in this paper are motivated by several negative dependence properties which are present in combinatorial probability, stochastic processes, statistical mechanics, reliability, and statistics. We focus our study on the theory of point processes because it is a natural tool in many of these fields. For each of these fields, it seems desirable to get a better understanding of what it means for a collection of random variables to be *repelling* or mutually negatively dependent. It is known that it is not possible to copy the theory of positively dependent random variables.

Negative association was introduced by Joag-Dev and Proschan [11]. Negative association has a distinct advantage over the other types of negative dependence, namely, nondecreasing functions of disjoint sets of negatively associated random variables are also negatively associated. This closure property does not hold for other types of negative dependence.

Pemantle [25] in his negative dependence study confined himself to binary-valued random variables. The list of examples that motivated him to develop techniques for proving that measures have negative dependence properties, such as negative association, include uniform random spanning trees, simple exclusion processes, random cluster models, and the occupation status of competing urns.

In Borcea *et al.* [6] several conjectures related to negative dependence made by Liggett [20], Pemantle [25], and Wagner [30] were solved; also Lyons' main results [21] on negative

Received 28 March 2018; revision received 4 December 2018.
* Postal address: Department of Mathematics, Karlsruhe Institute of Technology, Englerstr. 2, D-76131, Karlsruhe, Germany.
** Postal address: University of Wrocław, Mathematical Institute, pl. Grunwaldzki 2/4, 50-384, Wrocław, Poland. Email address: ryszard.szekli@uwr.edu.pl

association for determinantal probability measures induced by positive contractions were extended. The authors used several new classes of negatively dependent measures for zero–one valued vectors related to the theory of polynomials and to determinantal measures (for example, strongly Rayleigh measures related to the notion of proper position for multivariate stable polynomials).

For point processes, a negative association result is known in a fairly general setting for determinantal point processes on locally compact complete separable metric spaces generated by locally trace-class positive contractions on natural L^2-space (see, e.g. [22, Theorem 3.7]). A broad list of interesting examples of determinantal point processes can be found in [28]. Negative dependence for finite point processes via determinantal and/or strongly Rayleigh measures has interesting applications in various applied fields, such as machine learning, computer vision, computational biology, natural language processing, combinatorial bandit learning, neural network compression, and matrix approximations (see, e.g. [2], [13], [17], [18], and the references therein).

Another approach to the study of dependence has been used in finance models. Positive and negative dependences for a random vector may be seen as some stochastic ordering relations of this vector with some vector with independent coordinates. Such stochastic orderings are called *dependence orderings* (see [12] or [23]). Related results in the theory of point processes and stochastic geometry, where directionally convex ordering is used to express more clustering in point patterns, have been obtained in [4] and [5].

Apart from the negative association property of determinantal point processes, not much is known about the negative association property of other point processes. We show the negative association property for mixed sampled point processes under an ultra log-concave (ULC) assumption on the distribution of the number of points in these point processes. In order to obtain the negative association property in this general class of point processes, we use some results from the theory of strongly Rayleigh measures on the unit cube (see Theorems 3.2 and 3.3). Consequences of the negative association property of point processes in the theory of dependence orderings of point processes are described in a separate section (see Proposition 4.1). We stress that in order to obtain comparisons in terms of dependence orderings, it is enough to use a weaker property than negative association, which we denote by wNA.

2. Negative association and related definitions

We recall the definition and basic properties of negative association.

Definition 2.1. A random vector $\boldsymbol{X} = (X_1, \dots, X_n)$ is negatively associated (NA) if, for every subset $A \subseteq \{1, \dots, n\}$,

$$\operatorname{cov}\left((X_i, i \in A), g(X_j, j \in A^c)\right) \leq 0,$$

whenever f and g are real nondecreasing Borel functions for which the covariance exists.

We also use negative association to refer to the set of random variables $\{X_1, \dots, X_n\}$, or to the underlying distribution of $\boldsymbol{X}$.

Negative association possesses the following properties (see [11]).

(i) A pair (X, Y) of random variables is NA if and only if

$$\mathbb{P}(X \leq x, Y \leq y) \leq \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y),$$

i.e. (X, Y) is negatively quadrant dependent (NQD).

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

(ii) For disjoint subsets $A_1, \dots, A_m$ of $\{1, \dots, n\}$, and nondecreasing positive Borel functions $f_1, \dots, f_m$, $\boldsymbol{X}$ is NA implies that

$$\mathbb{E} \prod_{i=1}^{m} f_i(\boldsymbol{X}_{A_i}) \le \prod_{i=1}^{m} \mathbb{E} f_i(\boldsymbol{X}_{A_i}),$$

where $\boldsymbol{X}_{A_i} = (X_j, j \in A_i)$.

(iii) Any (at least two-element) subset of NA random variables is NA.

(iv) If $\boldsymbol{X}$ has independent components then it is NA.

(v) Increasing (nondecreasing) real functions defined on disjoint subsets of a set of NA random variables are NA.

(vi) If $\boldsymbol{X}$ is NA and $\boldsymbol{Y}$ is NA, and $\boldsymbol{X}$ is independent of $\boldsymbol{Y}$, then $(\boldsymbol{X}, \boldsymbol{Y})$ is NA.

We shall utilize a slightly broader class than NA in our formulations on dependence orderings. We define this new class of distributions as an analogue of the weak association in sequence (WAS) class introduced in [27]. We say that a random vector $\boldsymbol{X}$ (or its distribution) is weakly negatively associated (wNA) if

$$\operatorname{cov}(\mathbf{1}_{\{X_i > t\}}, f(X_{i+1}, \dots, X_n)) \le 0 \tag{2.1}$$

for all real nondecreasing functions f and $t \in \mathbb{R}$, $i = 1, \dots, n-1$.

This condition is equivalent to $[(X_{i+1}, \dots, X_n) \mid X_i > t)] <_{\text{st}} (X_{i+1}, \dots, X_n)$ for all $t \in \mathbb{R}$ and $i = 1, \dots, n-1$, where '$<_{\text{st}}$' denotes the usual strong stochastic ordering on $\mathbb{R}^n$. For the definitions of stochastic orderings; see [29, Chapter 2]. A number of positively or negatively dependent systems of random variables is considered in [7] and [23].

3. Negative association for mixed sampled point processes

We first introduce some basic point process notation; see, e.g. [15]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathbb{X}$ be a complete separable metric space equipped with the Borel σ-field $\mathcal{X}$. Denote by $\boldsymbol{N}$ the space of all measures μ on $(\mathbb{X}, \mathcal{X})$ such that $\mu(B) \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ for all bounded $B \in \mathcal{X}$. An example is the Dirac measure δ_x for a point $x \in \mathbb{X}$, given by $\delta_x(B) := \mathbf{1}_B(x)$. A more general example is a finite sum of Dirac measures. A *point process* η is a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\boldsymbol{N}, \mathcal{N})$, where $\mathcal{N}$ is the smallest σ-field on $\boldsymbol{N}$ such that $\mu \mapsto \mu(B)$ is measurable for all $B \in \mathcal{X}$.

We define the negative association property of point processes as follows.

Definition 3.1. A point process η is NA if, for each collection of disjoint bounded sets $B_1, \dots, B_n \in \mathcal{X}$, the vector $(\eta(B_1), \dots, \eta(B_n))$ is NA as defined for random vectors.

For a Borel set $A \subseteq \mathbb{X}$, let $\mathcal{N}_A$ denote the σ-field on $\boldsymbol{N}$ generated by the functions $\mu \mapsto \mu(B)$ for Borel $B \subseteq A$. The natural (inclusion) partial order on $\boldsymbol{N}$ allows us to define $f \colon \boldsymbol{N} \to \mathbb{R}$, which is increasing. We say that a point process η has *negative associations* if $\mathbb{E}(f(\eta)g(\eta)) \le \mathbb{E}(f(\eta))\mathbb{E}(g(\eta))$ for every pair f, g of real bounded increasing functions that are measurable with respect to complementary subsets A, A^c of $\mathbb{X}$, meaning that a function is measurable with respect to A if it is measurable with respect to $\mathcal{N}_A$. Clearly, if η has *negative associations* then η is NA. In the case of a locally compact space $\mathbb{X}$ the converse was shown in [22, Lemma 3.6, Theorem 3.9]; see also [26, Theorem A.1] for the case $\mathbb{X} = \mathbb{R}^d$. A different proof which works for general random measures on general Polish spaces is given in [16].

Let us recall Theorem 3.7 of [22]. Let λ be a Radon measure on a locally compact complete separable metric space $\mathbb{X}$. Let K be a locally trace-class positive contraction on $L_2(\mathbb{X}, \lambda)$. By η_K we denote the determinantal point process generated by K; for details, see [22, Section 3.2].

Theorem 3.1. *The determinantal point process η_K defined above has negative associations.*

Apart from determinantal point processes not much is known about the negative association property of point processes. Therefore we concentrate our efforts on characterizing the negative association property for an elementary but very useful class of finite point processes with an independent and identically distributed (i.i.d.) location of points. More precisely, our main focus in this paper is on the class of so-called *mixed sampled point processes* on $(\mathbb{X}, \mathcal{X})$, defined by

$$\eta = \sum_{i=1}^{\tau} \delta_{X_i}, \tag{3.1}$$

where $(X)_{i\geq 1}$ is i.i.d. with distribution F, and τ is an $\mathbb{N}_0$-valued random variable, independent of $(X_i)_{i\geq 1}$. For such a process, given any finite partition $A_1, \ldots, A_k \in \mathcal{X}$ of $\mathbb{X}$, conditionally on τ, the joint distribution of the number of points is given by

$$\mathbb{P}(\eta(A_1) = n_1, \ldots, \eta(A_k) = n_k \mid \tau = N) = \binom{N}{n_1 \cdots n_k} F(A_1)^{n_1} \cdots F(A_k)^{n_k},$$

and, unconditionally,

$$\mathbb{P}\,(\eta(A_1) = n_1, \ldots, \eta(A_k) = n_k) = \sum_{N=0}^{\infty} \mathbb{P}(\tau = N) \binom{N}{n_1 \cdots n_k} F(A_1)^{n_1} \cdots F(A_k)^{n_k}.$$

The joint probability generating function is therefore given for $z_1, \ldots, z_n \in [0, 1]$ by

$$\mathbb{E}\,(z_1^{\eta(A_1)} \cdots z_k^{\eta(A_k)}) = P_\tau\,(F(A_1)z_1 + \cdots + F(A_k)z_k),$$

where $P_\tau(z) = \mathbb{E}\,(z^\tau),\ z \in [0, 1]$.

First, we consider mixed point processes defined by (3.1) for which random variables τ are of the form

$$\tau = \sum_{i=1}^{n} U_i,$$

where $n \in \mathbb{N}$ and $U_1, \ldots, U_n$ are independent Bernoulli variables with possibly different success probabilities. The class of random variables which are the sums of n independent Bernoulli variables we denote by $\mathbb{Q}_n$. Moreover, we let

$$\mathcal{Q} := \mathrm{cl}\left(\bigcup_{n=1}^{\infty} \mathbb{Q}_n\right) \tag{3.2}$$

denote the class of all random variables with values in $\{0, 1, \ldots\}$ appearing as limits in distribution of variables from $\mathbb{Q}_n,\ n \geq 1$, i.e. the weak closure of $\bigcup_{n=1}^{\infty} \mathbb{Q}_n$. The main results of this paper are contained in Theorems 3.2 and 3.3.

Theorem 3.2. *Suppose that η is a mixed sampled point process on $(\mathbb{X}, \mathcal{X})$, defined by (3.1), for which $\tau \in \mathcal{Q}$. Then η is NA.*

Proof. Let $B_1, \ldots, B_n \in \mathcal{X}$ be a partition of $\mathbb{X}$, and $q_i := F(B_i),\ i = 1, \ldots, m$. Define

$$\boldsymbol{Z}_i := (\mathbf{1}_{\{X_i \in B_1\}}, \ldots, \mathbf{1}_{\{X_i \in B_m\}}) \tag{3.3}$$

to be the vector generated by the ith sample $X_i \in \mathbb{X},\ i \geq 1$. Note that each $\boldsymbol{Z}_i$ has multinomial distribution with success parameters $q_1, \ldots, q_m$, and the number of trials equals 1, and as such is NA. Moreover, the $\boldsymbol{Z}_i,\ i \geq 1$, are independent. Let $\boldsymbol{U} = (U_1, \ldots, U_n)$ be a vector of zero–one valued, independent random variables which is independent of $\boldsymbol{Z}_i,\ i \geq 1$. The vector composed as $(\boldsymbol{U}, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n)$ is NA because of properties (iv) and (vi) of negative association.

Now, using property (v), we find that the vector $(U_1\boldsymbol{Z}_1, \ldots, U_n\boldsymbol{Z}_n)$ is NA as a monotone transformation (multiplication) of disjoint coordinates of $(\boldsymbol{U}, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n)$. Again, using property (v), this time for $(U_1\boldsymbol{Z}_1, \ldots, U_n\boldsymbol{Z}_n)$, and using appropriate addition, we deduce that the vector $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ has the same distribution as $\sum_{i=1}^{\tau} \boldsymbol{Z}_i$, where $\tau := \sum_{i=1}^n U_i$. Defining η by (3.1) we hence see that $(\eta(B_1), \ldots, \eta(B_m))$ is NA. This completes the proof for $\tau \in \mathcal{Q}_n$ for arbitrary $n \in \mathbb{N}$.

For $\tau \in \mathcal{Q}$, there exists a sequence $\tau_k \xrightarrow{\mathrm{D}} \tau,\ k \to \infty$, for $\tau_k \in \bigcup_{n=1}^{\infty} \mathcal{Q}_n$, and

$$\mathbb{E}\bigg(f\bigg(\sum_{i=1}^{\tau_k} \boldsymbol{Z}_i\bigg)g\bigg(\sum_{i=1}^{\tau_k} \boldsymbol{Z}_i\bigg)\bigg) \leq \mathbb{E}\bigg(f\bigg(\sum_{i=1}^{\tau_k} \boldsymbol{Z}_i\bigg)\bigg)\mathbb{E}\bigg(g\bigg(\sum_{i=1}^{\tau_k} \boldsymbol{Z}_i\bigg)\bigg)$$

for f, g supported by disjoint coordinates, which are nondecreasing and bounded. Letting $k \to \infty$ gives

$$\mathbb{E}\bigg(f\bigg(\sum_{i=1}^{\tau} \boldsymbol{Z}_i\bigg)g\bigg(\sum_{i=1}^{\tau} \boldsymbol{Z}_i\bigg)\bigg) \leq \mathbb{E}\bigg(f\bigg(\sum_{i=1}^{\tau} \boldsymbol{Z}_i\bigg)\bigg)\mathbb{E}\bigg(g\bigg(\sum_{i=1}^{\tau} \boldsymbol{Z}_i\bigg)\bigg).$$

Since each nondecreasing function can be monotonically approximated by nondecreasing and bounded functions, we obtain the negative association property of η. □

The class $\mathcal{Q}$ can be completely characterized; see, e.g. [1].

Lemma 3.1. *We have*

$$\tau \in \mathcal{Q} \quad \text{if and only if} \quad \tau \stackrel{\mathrm{D}}{=} \tau_1 + \tau_2,$$

where τ_1 and τ_2 are independent, τ_1 has a Poisson distribution, and $\tau_2 \stackrel{\mathrm{D}}{=} \sum_{i=1}^{\infty} U_i$ for independent zero–one valued variables U_i with $\mathbb{P}(U_i = 1) \geq 0,\ i \geq 1,$ and such that $\sum_{i=1}^{\infty} \mathbb{P}(U_i = 1) < \infty$.

It is interesting to note that hypergeometric random variables belong to the class $\mathcal{Q}$; see, e.g. [10].

We say that a real sequence $(a_i)_{i=0}^n$ has no internal zeros if the indices of its non-ero terms form a discrete interval. Following [25] we shall use the following class of sequences and distributions.

Definition 3.2. A finite real sequence $(a_i)_{i=0}^n$ of nonnegative real numbers with $a_i \neq 0$ for $1 \leq i \leq n-1$ (no internal zeros) is *ultra log-concave* (ULC(n)) if

$$\bigg(\frac{a_i}{\binom{n}{i}}\bigg)^2 \geq \frac{a_{i-1}}{\binom{n}{i-1}} \frac{a_{i+1}}{\binom{n}{i+1}}, \qquad i = 1, \ldots, n-1.$$

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

Define

$$\mathcal{S}_n := \{\tau : (\mathbb{P}(\tau = i))_{i=0}^n \text{ is ULC}(n)\}$$

to be the class of random variables whose probability functions have the above property. It is known that if a nonnegative sequence $(a_i)_{i=0}^n$ is ULC(n) and a nonnegative sequence $(b_i)_{i=0}^m$ is ULC(m), then the convolution of these sequences is ULC($m+n$) (see [19, Theorem 2]). Let

$$\mathcal{S} := \text{cl}\left(\bigcup_{n=1}^{\infty} \mathcal{S}_n\right). \tag{3.4}$$

Sums of independent variables from the class $\mathcal{S}$ are in $\mathcal{S}$. We shall see below that $\mathcal{Q} \subseteq \mathcal{S}$. Utilizing the class $\mathcal{S}$, Theorem 3.2 can be generalized with the use of elementary symmetric functions.

Theorem 3.3. *Suppose that η is a mixed sampled point process on $(\mathbb{X}, \mathcal{X})$, defined by (3.1), for which $\tau \in \mathcal{S}$. Then η is NA.*

Proof. Assume first that $\tau \in \mathcal{S}_n$. Let $B_1, \dots, B_n \in \mathcal{X}$ be a partition of $\mathbb{X}$, and let $q_i := F(B_i)$, $i = 1, \dots, m$. Let $\boldsymbol{Z}_i$, $i \geq 1$ be defined by (3.3). Note that each $\boldsymbol{Z}_i$ has multinomial distribution with success parameters $q_1, \dots, q_m$, and the number of trials equals 1, and as such is NA. Moreover, $\boldsymbol{Z}_i$, $i \geq 1$, are independent. For fixed $n \in \mathbb{N}$, we now define a vector $\boldsymbol{U} = (U_1, \dots, U_n)$ with $\{0, 1\}$-valued coordinates, independent of $\boldsymbol{Z}_i$, $i \geq 1$. It is enough to define the distribution of $\boldsymbol{U}$. For the generating function of τ, $P_\tau(z) = \mathbb{E}(z^\tau)$, we define the distribution of $\boldsymbol{U}$ by providing its multidimensional generating function. It is obtained by substituting in $P_\tau(z)$ for each $k = 0, \dots, n$,

$$z^k := \binom{n}{k}^{-1} e_k(z_1, \dots, z_n),$$

where $e_k(z_1, \dots, z_n)$ is the kth elementary symmetric polynomial. Note that, by the definition of the elementary symmetric polynomials, for each k, the function $\binom{n}{k}^{-1} e_k(z_1, \dots, z_n)$ of variables $z_1, \dots, z_n$ is the multivariate generating function of a vector of n $\{0, 1\}$-valued variables which contains exactly k values 1 with the same probability $\binom{n}{k}^{-1}$ defined for all possible selections of the k coordinates at which the values 1 occur. The distribution of $\boldsymbol{U}$ defined in such a way is the mixture with the coefficients $a_k := \mathbb{P}(\tau = k)$ of the distributions corresponding to $\binom{n}{k}^{-1} e_k(z_1, \dots, z_n)$, $k = 0, \dots, n$. Since each function e_k is symmetric in variables $z_1, \dots, z_n$, the same is true for the generating function of $\boldsymbol{U} = (U_1, \dots, U_n)$; therefore, $(U_1, \dots, U_n)$ are exchangeable. Moreover, $\sum_{i=1}^n U_i \stackrel{\text{D}}{=} \tau$, since, by setting $z_1 = \cdots = z_n := z$, we obtain $P_\tau(z)$. The sequence $(\mathbb{P}(\tau = i))_{i=0}^n$ is called the *rank sequence* of the vector $\boldsymbol{U} = (U_1, \dots, U_n)$ (see [6, Definition 2.8]).

From our assumption, the rank sequence for $\boldsymbol{U} = (U_1, \dots, U_n)$ is ULC(n) and from Theorem 2.7 of [25], we deduce that $\boldsymbol{U}$ is NA. Now the vector composed as $(\boldsymbol{U}, \boldsymbol{Z}_1, \dots, \boldsymbol{Z}_n)$ is NA because of property (vi). Using property (v), we find that the vector $(U_1\boldsymbol{Z}_1, \dots, U_n\boldsymbol{Z}_n)$ is NA as a monotone transformation (multiplication) of disjoint coordinates of $(\boldsymbol{U}, \boldsymbol{Z}_1, \dots, \boldsymbol{Z}_n)$. Again, using property (v), this time for $(U_1\boldsymbol{Z}_1, \dots, U_n\boldsymbol{Z}_n)$, and using addition, we deduce that the vector $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ has the same distribution as $\sum_{i=1}^\tau \boldsymbol{Z}_i$, which in turn has the same distribution as $(\eta(B_1), \dots, \eta(B_m))$. This completes the proof for $\tau \in \mathcal{S}_n$ for arbitrary $n \in \mathbb{N}$. For $\tau \in \mathcal{S}$, we apply a limiting argument analogous to that used in the proof of Theorem 3.2. □

The following lemma may be regarded as known since it is an immediate consequence of the classical Newton inequalities (see, e.g. [24] for a new look at Newton's inequalities). We formulate it in the setting of the classes of random variables introduced in this paper.

Lemma 3.2. *For the classes of random variables defined in (3.2) and (3.4), we have*

$$\mathcal{Q} \subset \mathcal{S}.$$

Proof. Suppose that $\tau \in \mathcal{Q}_n$. Then, for its generating function,

$$P_\tau(z) = (1 - p_1 + p_1 z) \cdots (1 - p_n + p_n z) = p_1 \cdots p_n \Big(\frac{1-p_1}{p_1} + z \Big) \cdots \Big(\frac{1-p_n}{p_n} + z \Big).$$

For $a_k := (1 - p_k)/p_k$, we have

$$P_\tau(z) = p_1 \cdots p_n (a_1 + z) \cdots (a_n + z) = p_1 \cdots p_n [x^n + c_1 x^{n-1} + \cdots + c_n],$$

where

$$c_n = a_1 + \cdots + a_n, \qquad c_2 = a_1 a_2 + \cdots + a_{n-1} a_n, \qquad \ldots, \qquad c_n = a_1 \cdots a_n,$$

i.e. the coefficients c_k, $k = 0, \ldots, n$, are given by the corresponding elementary symmetric polynomials in the variables a_i, $i = 1, \ldots, n$. It is known from the classical Newton inequalities that the sequence $(c_i)_{i=0}^n$ is ULC(n), and, since $\mathbb{P}(\tau = k) = p_1 \cdots p_n c_{n-k}$, we conclude that the sequence $(\mathbb{P}(\tau = k))_{k=0}^n$ is also ULC(n), and, therefore, $\tau \in \mathcal{S}_n$, which immediately implies the inclusion $\mathcal{Q} \subset \mathcal{S}$. □

We note that the arguments utilized in Theorem 3.3 can be used for random vectors with arbitrary positive values.

Proposition 3.1. *Assume that $\mathbf{Z}_i = (Z_i^1, \ldots, Z_i^m)$, $i \geq 1$, is a sequence of i.i.d. random vectors with components in $\mathbb{R}_+$ such that, for each $i \geq 1$,*

$$\sum_{j=1}^m \mathbf{1}_{\{Z_i^j > 0\}} \leq 1,$$

that is, at most one of the components can be positive. Then, for $\tau \in \mathcal{S}$, which is independent of $\mathbf{Z}_i$, $i \geq 1$, the vector $\boldsymbol{W} := \sum_{i=1}^\tau \mathbf{Z}_i$ is NA.

Proof. We use basically the same argument as used in the proof of Theorem 3.3. Let $\boldsymbol{U} = (U_1, \ldots, U_n)$, independent of $\mathbf{Z}_i$, $i \geq 1$, be the vector of $\{0, 1\}$-valued random variables obtained by its generating function as follows. In $P_\tau(z)$, substitute $z^k := \binom{n}{k}^{-1} e_k(z_1, \ldots, z_n)$, $k = 1, \ldots, n$, where $e_k(z_1, \ldots, z_n)$ are the elementary symmetric polynomials. This substitution defines a generating function of variables $z_1, \ldots, z_n$. It is then immediate that $\sum_{i=1}^n U_i \stackrel{\mathrm{D}}{=} \tau$, that is, the sequence $(\mathbb{P}(\tau = i))_{i=0}^n$ is the rank sequence for (symmetric) $\boldsymbol{U} = (U_1, \ldots, U_n)$. From our assumption, the rank sequence for $\boldsymbol{U} = (U_1, \ldots, U_n)$ is ULC(n) and, from Theorem 2.7 of [25], we deduce that $\boldsymbol{U}$ is NA. Now the vector composed as $(\boldsymbol{U}, \mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ is NA because of the following lemma and property (vi).

Lemma 3.3. *Assume that $\mathbf{Z} = (Z^1, \ldots, Z^m)$ is a random vector with components in $\mathbb{R}_+$. Assume that $\sum_{j=1}^m \mathbf{1}_{\{Z^j > 0\}} \leq 1$, that is, at most one of the components can be positive. Then $\mathbf{Z}$ is NA.*

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

Proof of Lemma 3.3. In order to show that $\operatorname{cov}(f(Z^1, \ldots, Z^k), g(Z^{k+1}, \ldots, Z^m)) \geq 0$ for nondecreasing f and g, it suffices to assume that $f(0) = g(0) = 0$. Otherwise, we can consider $f - f(0)$ and $g - g(0)$. Because only one of the coordinates can be nonzero, we obtain $\mathbb{E}(f(Z^1, \ldots, Z^k)g(Z^{k+1}, \ldots, Z^m)) = 0$, while the product of the expectations is nonnegative since $f \geq 0$ and $g \geq 0$. □

Now, using property (v), we deduce that the vector $(U_1\boldsymbol{Z}_1, \ldots, U_n\boldsymbol{Z}_n)$ is NA because it is a monotone transformation (multiplication) of disjoint independent coordinates of $(\boldsymbol{U}, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n)$. Again, using property (v), this time for $(U_1\boldsymbol{Z}_1, \ldots, U_n\boldsymbol{Z}_n)$, and using addition, we deduce that the vector $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\boldsymbol{Z}_i$ has the same distribution as $\sum_{i=1}^\tau \boldsymbol{Z}_i$. □

The above proposition can be used to study random measures other than point processes. We shall pursue this topic elsewhere.

4. Dependence orderings for point processes

An extensive study of dependence orderings for multivariate point processes on $\mathbb{R}$ is contained in [14]. Related results in the theory of point processes and stochastic geometry, where the directionally convex ordering is used to express more clustering in point patterns, are obtained by Błaszczyszyn and Yogeshwaran [5]; see also the references therein. We shall use the negative association property of point processes to obtain comparisons of dependence in point processes. More precisely, we recall some basic facts on dependence orderings of vectors and their relation to negative association which can be directly utilized for point processes.

4.1. Dependence orderings and negative correlations for vectors

For a function $f\colon \mathbb{R}^n \to \mathbb{R}$, define the difference operator Δ_i^ϵ, $\epsilon > 0$, $1 \leq i \leq n$, by

$$\Delta_i^\epsilon f(\boldsymbol{x}) = f(\boldsymbol{x} + \epsilon \boldsymbol{e}_i) - f(\boldsymbol{x}),$$

where $\boldsymbol{e}_i$ is the ith unit vector. Then f is called *supermodular* if, for all $1 \leq i < j \leq n$ and $\epsilon, \delta > 0$,

$$\Delta_j^\delta \Delta_i^\epsilon f(\boldsymbol{x}) \geq 0$$

for all $\boldsymbol{x} \in \mathbb{R}$, and *directionally convex* if this inequality holds for all $1 \leq i \leq j \leq n$. Let $\mathcal{F}^{\mathrm{sm}}$ and $\mathcal{F}^{\mathrm{dcx}}$ denote the classes of supermodular and directionally convex functions. Then, of course, $\mathcal{F}^{\mathrm{dcx}} \subseteq \mathcal{F}^{\mathrm{sm}}$. Typical examples from the $\mathcal{F}^{\mathrm{dcx}}$ class of functions are $f(\boldsymbol{x}) = \psi(\sum_{i=1}^n x_i)$ for ψ convex, or $f(\boldsymbol{x}) = \max_{1 \leq i \leq n} x_i$, but there are many other useful functions in this class; see, for example, [3].

The corresponding stochastic orderings are defined by $\boldsymbol{X} <_{\mathrm{sm}} \boldsymbol{Y}$ if $\mathbb{E}f(\boldsymbol{X}) \leq \mathbb{E}f(\boldsymbol{Y})$ for all $f \in \mathcal{F}^{\mathrm{sm}}$, and analogously for $\boldsymbol{X} <_{\mathrm{dcx}} \boldsymbol{Y}$. For differentiable functions f, we obtain $f \in \mathcal{F}^{\mathrm{sm}}$ if and only if $(\partial^2 f / \partial x_i x_j) \geq 0$ for $i < j$, and $f \in \mathcal{F}^{\mathrm{dcx}}$ if and only if this inequality holds for $i \leq j$ (see [23, Theorems 3.9.3 and 3.12.2]). While comparison of $\boldsymbol{X}$ and $\boldsymbol{Y}$ with respect to '$<_{\mathrm{sm}}$' implies (and is restricted to the case of) identical marginals $X_i \overset{\mathrm{D}}{=} Y_i$, the comparison with respect to the smaller class $\mathcal{F}^{\mathrm{dcx}}$ implies convexly increasing marginals $X_i <_{\mathrm{cx}} Y_i$ (which means by definition that $\mathbb{E}\psi(X_i) \leq \mathbb{E}\psi(Y_i)$ for all $\psi\colon \mathbb{R} \to \mathbb{R}$ convex). Both of these orderings belong to the class of so-called *dependence orderings* (see, e.g. [12]) which is defined by a list of suitable properties, among them the property that $\operatorname{cov}(X_i, X_j) \leq \operatorname{cov}(Y_i, Y_j)$.

In [27] another dependence ordering $\boldsymbol{X} <_{\text{wcs}} \boldsymbol{Y}$ (weakly conditional increasing in sequence order) was introduced by the condition

$$\operatorname{cov}(\mathbf{1}_{\{X_i>t\}}, f(X_{i+1}, \ldots, X_n)) \leq \operatorname{cov}(\mathbf{1}_{\{Y_i>t\}}, f(Y_{i+1}, \ldots, Y_n))$$

for all monotonically nondecreasing f, and all $t \in \mathbb{R},\ 1 \leq i \leq n-1$.

All dependence orderings can be used to define some classes of distributions with negative (or positive covariances) when applied to vectors with independent components. More precisely, let $\boldsymbol{X}^*$ denote a vector with independent components, and such that $X_i^* \stackrel{\text{D}}{=} X_i$. Using this approach, definition (2.1) of wNA is equivalent to the relation $\boldsymbol{X} <_{\text{wcs}} \boldsymbol{X}^*$. It is clear that wNA is further equivalent to

$$[(X_{i+1}, \ldots, X_n) \mid X_i > t)] <_{\text{st}} (X_{i+1}, \ldots, X_n)$$

for all $i = 1, \ldots, n-1,\ t > 0$, where $[(X_{i+1}, \ldots, X_n) \mid X_i > t)]$ denotes a random vector which has the distribution of $(X_{i+1}, \ldots, X_n)$ conditioned on the event $\{X_i > t\}$, and '$<_{\text{st}}$' is the usual (strong) stochastic order. Directly from the definition of the negative association property we see that the weakly negative association property is weaker than the negative association property. For $\boldsymbol{X}$ being wNA, Theorem 4.1 implies that, for example (see also [8] for the negative association case), $\sum_{i=1}^n X_i <_{\text{cx}} \sum_{i=1}^n X_i^*$, and $\max_{1\leq k\leq n} \sum_{i=1}^k X_i <_{\text{icx}} \max_{1\leq k\leq n} \sum_{i=1}^k X_i^*$, where '$<_{\text{icx}}$' is defined similarly to '$<_{\text{cx}}$' but with the use of nondecreasing convex functions. Taking other supermodular functions, it is possible to get maximal inequalities for wNA vectors as in [8].

The following theorem from [27] connects the above-defined orderings.

Theorem 4.1. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be n-dimensional random vectors.*

(1) *If $X_i \stackrel{\text{D}}{=} Y_i,\ 1 \leq i \leq n$, then $\boldsymbol{X} <_{\text{wcs}} \boldsymbol{Y}$ implies that $\boldsymbol{X} <_{\text{sm}} \boldsymbol{Y}$.*

(2) *If $X_i <_{\text{cx}} Y_i,\ 1 \leq i \leq n$, then $\boldsymbol{X} <_{\text{wcs}} \boldsymbol{Y}$ implies that $\boldsymbol{X} <_{\text{dcx}} \boldsymbol{Y}$.*

For the '$<_{\text{dcx}}$' ordering, we get the following corollary from Theorem 4.1. This corollary will be used later for point processes with the negative association property.

Corollary 4.1. *Suppose that $\boldsymbol{X}$ is wNA and $\boldsymbol{Y}^*$ has independent coordinates with $X_i <_{\text{cx}} Y_i^*$. Then $\boldsymbol{X} <_{\text{dcx}} \boldsymbol{Y}^*$.*

4.2. Negative association and dependence orderings for point processes

Using Theorem 4.1, we are able to compare the covariance structure of some point processes. To be more precise, we need a couple of definitions.

Definition 4.1. Two point processes η_1 and η_2 on $\mathbb{X}$ are ordered in the directionally convex order (dcx) (weakly conditional increasing sequence order (wcs))

$$\eta_1 <_{\text{dcx}} \eta_2 \ (\eta_1 <_{\text{wcs}} \eta_2)$$
$$\text{if and only if} \quad (\eta_1(B_1), \ldots, \eta_1(B_n)) <_{\text{dcx}} (<_{\text{wcs}}) (\eta_2(B_1), \ldots, \eta_2(B_n)),$$

as defined for random vectors, for all bounded Borel sets $B_1, \ldots, B_n,\ n \geq 1$.

Similarly to Definition 3.1 we define a weaker version of negative association for point processes.

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

Definition 4.2. A point process η is wNA if, for each collection of disjoint bounded sets $B_1, \dots, B_n \in \mathcal{X}$, the vector $(\eta(B_1), \dots, \eta(B_n))$ is wNA, as defined for random vectors.

We now propose a direct consequence of the above definitions and Theorem 4.1. It will be used in the next section for determinantal and mixed sampled point processes.

Proposition 4.1. *Suppose that η_1 is a point process on $(\mathbb{X}, \mathcal{X})$ with a locally finite intensity measure which is wNA. Let η_2 be a Poisson process with intensity measure $\mathbb{E}(\eta_1)$. Assume that $\eta_1(B) <_{\mathrm{cx}} \eta_2(B)$ for all bounded Borel sets B. Then $\eta_1 <_{\mathrm{dcx}} \eta_2$.*

For illustration, we now show that the ordering '$<_{\mathrm{dcx}}$' can be used to obtain comparisons of moment measures and void probabilities for point processes on $\mathbb{R}^d$. These results can be modified with similar arguments to hold on general state spaces.

The following consequence of the '$<_{\mathrm{dcx}}$' ordering for point processes is known from [3] and [5]. Given a point process η on $\mathbb{X}$ and $n \in \mathbb{N}$, we let $\mathbb{E}(\eta^n)$ denote the nth moment measure of η, that is, the measure $B \mapsto \mathbb{E}(\eta^n(B))$ on $(\mathbb{R}^d)^n$ equipped with the Borel σ-field.

Lemma 4.1. *Let η_1 and η_2 be two point processes on $\mathbb{R}^d$. If $\eta_1 <_{\mathrm{dcx}} \eta_2$ then the following statements holds.*

1. *(Moment measures) If the measure $\mathbb{E}\eta_2^n$ is σ-finite then $\mathbb{E}(\eta_1^n(B)) \le \mathbb{E}(\eta_2^n(B))$ for all bounded Borel sets $B \subset (\mathbb{R}^d)^n$.*
2. *(Void probabilities) $\mathbb{P}(\eta_1(B) = 0) \le \mathbb{P}(\eta_2(B) = 0)$ for all bounded Borel sets B.*

Using the '$<_{\mathrm{wcs}}$' criterion for '$<_{\mathrm{dcx}}$' from Theorem 4.1, we obtain the following result.

Corollary 4.2. *Let η_1 and η_2 be two point processes on $\mathbb{R}^d$. If, for all bounded Borel sets B, $\eta_1(B) <_{\mathrm{cx}} \eta_2(B)$ and $\eta_1 <_{\mathrm{wcs}} \eta_2$, then the comparisons of moment measures and void probabilities from the above lemma hold.*

An interesting case for such comparisons is when η_2 is a Poisson point process.

Proposition 4.2. *Suppose that η is a point process on $\mathbb{R}^d$ which is wNA and has a locally finite intensity measure. Then the following statements hold.*

1. *(Moment measures) $\mathbb{E}(\eta(B_1) \cdots \eta(B_n)) \le \mathbb{E}(\eta(B_1)) \cdots \mathbb{E}(\eta(B_n))$ for all disjoint, bounded Borel sets $B_1, \dots, B_n$, which, for simple point processes η, implies that*
$$\mathbb{E}\Big(\exp\Big(\int_{\mathbb{R}^d} h(\boldsymbol{x})\eta(\mathrm{d}\boldsymbol{x})\Big)\Big) \le \exp\Big(\int_{\mathbb{R}^d} (\mathrm{e}^{h(\boldsymbol{x})} - 1)\mathbb{E}\eta(\mathrm{d}\boldsymbol{x})\Big)$$
for all measurable $h \ge 0$.
2. *(Void probabilities) $\mathbb{P}(\eta(B) = 0) \le \exp(-\mathbb{E}\eta(B))$ for all bounded Borel sets B, which, for simple point processes η, implies that*
$$\mathbb{E}\Big(\exp\Big(-\int_{\mathbb{R}^d} h(\boldsymbol{x})\eta(\mathrm{d}\boldsymbol{x})\Big)\Big) \le \exp\Big(\int_{\mathbb{R}^d} (\mathrm{e}^{-h(\boldsymbol{x})} - 1)\mathbb{E}\eta(\mathrm{d}\boldsymbol{x})\Big)$$
for all measurable $h \ge 0$.

Proof. Let $B_1, \dots, B_n$ be disjoint Borel sets. The vector $\boldsymbol{X} := (\eta(B_1), \dots, \eta(B_n))$ is wNA, by our assumptions. Let $\boldsymbol{X}^*$ denote a corresponding independent version. By the definition of wNA, we have $\boldsymbol{X} <_{\mathrm{wcs}} \boldsymbol{X}^*$, so that Theorem 4.1 shows that $X <_{\mathrm{sm}} \boldsymbol{X}^*$. From the definition of '$<_{\mathrm{sm}}$', $\mathbb{E}(\eta(B_1) \cdots \eta(B_n))) \le \mathbb{E}(\eta(B_1)) \cdots \mathbb{E}(\eta(B_n))$. Now, from Proposition 1 of [5], this implies the second assertion of the first part.

To prove the second part of the proposition, let B and B' be disjoint bounded Borel sets. By assumption, $(\eta(B), \eta(B'))$ is wNA. Directly from the definition of wNA we conclude that $\mathbb{P}(\eta(B)=0, \eta(B')=0) \leq \mathbb{P}(\eta(B)=0)\mathbb{P}(\eta(B')=0)$. Now, from Proposition 3.1 of [4], $\mathbb{P}(\eta(B)=0) \leq \exp(-\mathbb{E}\eta(B))$ for all bounded Borel sets B. Moreover, from Proposition 2 of [5], this inequality is equivalent to the second inequality asserted in the second part. □

4.3. '$<_{\mathrm{dcx}}$' comparisons for determinantal and mixed sampled point processes

In the following corollary we shall assume that $\mathbb{X}$ is locally compact and that λ is a Radon measure on $\mathbb{X}$. Let K be a locally trace-class positive contraction on $L_2(\mathbb{X}, \lambda)$, and let η_K be the determinantal point process generated by K. From Proposition 4.1 we obtain the following corollary.

Corollary 4.3. *Suppose that η_K is the determinantal point process described above. Then*

$$\eta_K <_{dcx} \eta_2,$$

where η_2 denotes a Poisson point process with intensity measure $\mathbb{E}\eta_K$.

Proof. Fix a bounded Borel set B. From Theorem 3.1 we know that η_K is NA, so in order to get the conclusion of this corollary it is enough (see Proposition 4.1) to show that $\eta_K(B) <_{\mathrm{cx}} \eta_2(B)$. From [9, Proposition 9] we know that $\eta_K(B)$ is distributed as a sum of independent Bernoulli random variables. From the definition of the log-concave ordering, '$<_{\mathrm{lc}}$' in [31], it follows that $\eta_K(B) <_{\mathrm{lc}} \eta_2(B)$, which in turn implies that $\eta_K(B) <_{\mathrm{cx}} \eta_2(B)$; see Theorem 1 of [31]. □

The above corollary for the case of jointly observable sets and $\mathbb{X}=\mathbb{R}^d$ was proved in [4, Proposition 5.3] using a different argument.

For mixed sampled point processes on general spaces, we obtain the following comparison result.

Proposition 4.3. *Suppose that η_1 is a mixed sampled point process on $(\mathbb{X}, \mathcal{X})$ defined by (3.1) for which $\tau \in \mathcal{S}$. Then*

$$\eta_1 <_{\mathrm{dcx}} \eta_2,$$

where η_2 denotes a Poisson point process on $\mathbb{X}$, with the intensity measure $\mathbb{E}\eta_1$.

Proof. From Theorem 3.3 we know that η_1 is NA, and from Proposition 4.1 we shall get the conclusion of the present proposition if we show that, for such processes, $\eta_1(B) <_{\mathrm{cx}} \eta_2(B)$. From the definition of mixed sampled point processes we know that $\eta_1(B)$ is distributed as a random sum $\sum_{i=1}^{\tau} U_i$, where $(U_i, i \geq 1)$ is an i.i.d. sequence of Bernoulli, i.e. $\{0, 1\}$-valued variables with success probability $F(B)$. Since $\tau \in \mathcal{S}_n$, from the definition of the log-concave ordering '$<_{\mathrm{lc}}$' in [31], it follows that $\tau <_{\mathrm{lc}} \kappa$, where κ has a Poisson distribution with mean $\mathbb{E}(\tau)$. Therefore, $\tau <_{cx} \kappa$; see Theorem 1 of [31]. It follows that $\sum_{i=1}^{\tau} U_i <_{\mathrm{cx}} \sum_{i=1}^{\kappa} U_i$, where κ is now assumed to be independent of $(U_i, i \geq 1)$ (see, e.g. [14, Corollary 4.5]). From this we get $\eta_1(B) <_{\mathrm{cx}} \eta_2(B)$, since $\mathbb{E}\eta_1(B) = \mathbb{E}(\tau)\mathbb{E}(U_i)$ and $\sum_{i=1}^{\kappa} U_i$ has Poisson distribution. For arbitrary $\tau \in \mathcal{S}$, we apply weak approximation by random variables from $\mathcal{S}_n$, $n \geq 1$. □

From Proposition 4.2, we obtain the following corollary.

Corollary 4.4. *Suppose that η is a mixed sampled point process on $\mathbb{R}^d$, defined by* (3.1), for which $\tau \in \mathcal{S}$. Then the following statements hold

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

1. (Moment measures) $\mathbb{E}(\eta(B_1)\cdots\eta(B_n))) \le \mathbb{E}(\eta(B_1))\cdots\mathbb{E}(\eta(B_n))$, for all disjoint, bounded Borel sets $B_1, \dots, B_n,\ n \ge 1$.
2. (Void probabilities) $\mathbb{P}(\eta(B)=0) \le \exp(-\mathbb{E}\eta(B))$ for all bounded Borel sets $B,\ n \ge 1$.

We illustrate this with an example of a direct approach to the comparison of void probabilities.

Example 4.1. (*Comparison of binomial mixed sampled p.p. with Poisson p.p. on* $\mathbb{R}^d$.) Let η be a mixed sampled point process on $\mathbb{R}^d$ with τ being a binomial distributed random variable. We shall compare void probabilities for this process with void probabilities of a Poisson point process with the same intensity measure. In general, for a simple point process η on $\mathbb{R}^d$, to test whether $\mathbb{P}(\eta(B)=0) \le \exp(-\mathbb{E}(\eta(B)))$, it is enough to check (see [5, Proposition 3.1]) that

$$\mathbb{P}(\eta(B)=0,\, \eta(B')=0) \le \mathbb{P}(\eta(B)=0))\mathbb{P}(\eta(B')=0)$$

for disjoint B, B'. For arbitrary, measurable disjoint sets B, B', we have

$$\begin{aligned}\mathbb{P}(\eta(B)=\eta(B')=0) &= \mathbb{P}(\tau=0) + \sum_{n=1}^{\infty}\mathbb{P}(\tau=n)\mathbb{P}(X_1 \notin B\cup B', \dots, X_n \notin B\cup B') \\ &= \mathbb{P}(\tau=0) + \sum_{n=1}^{\infty}\mathbb{P}(\tau=n)(1-F(B\cup B'))^n \\ &= P_\tau(1-F(B\cup B')) \\ &= P_\tau(1-(F(B)+F(B')).\end{aligned}$$

Since τ has a binomial distribution with, say, parameters n (number of trials) and $p \in (0, 1)$ (success probability), then $P_\tau(1-s) = (p(1-s)+(1-p))^n$. It is easy to see by differentiation that $\phi(s) := -\log P_\tau(1-s)$ is then an increasing and convex function such that $\phi(0)=0$. It is known that such a function is superadditive; therefore, $\phi(s+t) \ge \phi(s)+\phi(t)$, and then $P_\tau(1-(F(B)+F(B')) \le P_\tau(1-F(B))P_\tau(1-F(B'))$. In this case, for disjoint B, B' we obtain

$$\mathbb{P}(\eta(B)=\eta(B')=0) \le \mathbb{P}(\eta(B)=0)P(\eta(B')=0).$$

Therefore, for this process, we obtain $\mathbb{P}(\eta_1(B)=0) \le \exp(-\mathbb{E}(\eta(B)))$.

5. Some applications

Let η be a point process on a complete, separable metric space $\mathbb{X}$. Using the Chebyshev inequality, we have

$$\mathbb{P}(|\eta(B)-\mathbb{E}(\eta(B))\ |\ \ge \epsilon) \le \frac{\operatorname{var}(\eta(B))}{\epsilon^2}$$

for all bounded Borel sets B and $\epsilon > 0$.

Similarly, using the Chernoff bound,

$$\mathbb{P}(\eta(B)-\mathbb{E}(\eta(B)) \ge \epsilon) \le \mathrm{e}^{-t(\mathbb{E}(\eta(B))+\epsilon)}\mathbb{E}(\mathrm{e}^{t\eta(B)})$$

for any $t, a > 0$, and the upper bounds can be replaced by the values taken from a process larger in '$<_{\mathrm{dcx}}$' than η, directly using the definition of '$<_{\mathrm{dcx}}$' recalled in Section 4.1. If η is determinantal or a NA mixed sampled point process, the corresponding larger process is a

https://doi.org/10.1017/jpr.2019.10 Published online by Cambridge University Press

Poisson process which can be used to obtain upper bounds and concentration inequalities using Proposition 4.3.

Similarly, for all bounded Borel sets B and $\epsilon, t > 0$, we have

$$\mathbb{P}(\mathbb{E}(\eta(B)) - \eta(B) \geq \epsilon) \leq \mathrm{e}^{t(\mathbb{E}(\eta(B))-\epsilon)}\mathbb{E}(\mathrm{e}^{-t\eta(B)}).$$

Using Corollary 2 of [8], we can obtain Kolmogorov-type inequalities from the negative association property of η.

Corollary 5.1. *Suppose that η is a mixed sampled point process on $\mathbb{X}$, for which $\tau \in \mathcal{S}$. Then, for any increasing sequence b_k, $k \geq 1$, of positive numbers, any collection of disjoint bounded Borel sets $B_1, \dots, B_n \in \mathcal{X}$, and $\epsilon > 0$,*

$$\mathbb{P}\bigg(\max_{k\leq n}\bigg|\frac{1}{b_k}\sum_{i=1}^{k}(\eta(B_i) - \mathbb{E}(\eta(B_i))\bigg| \geq \epsilon\bigg) \leq 8\epsilon^{-2}\sum_{i=1}^{n}\frac{\mathrm{var}\,(\eta(B_i))}{b_i^2},$$

$$\mathbb{P}\bigg(\max_{m\leq k\leq n}\bigg|\frac{1}{b_k}\sum_{i=1}^{k}(\eta(B_i) - \mathbb{E}(\eta(B_i)))\bigg| \geq \epsilon\bigg) \leq 32\epsilon^{-2}\bigg(\sum_{i=m+1}^{n}\frac{\mathrm{var}\,(\eta(B_i))}{b_i^2} + \sum_{i=1}^{m}\frac{\mathrm{var}\,(\eta(B_i))}{b_m^2}\bigg) \quad \textit{for any integer } m < n.$$

Acknowledgements

This work was in part supported by the German Research Foundation (DFG) under grant LA965/9-2, awarded as part of the DFG-Forschergruppe FOR 1548 ‘Geometry and Physics of Spatial Random Systems’, and by the National Science Centre, Poland, under grant NCN 2015/19/B/ST1/01152.

References

[1] ALEMAN, A., BELIAEV, D. AND HEDENMALM, H. (2004). Real zero polynomials and Pólya–Schur type theorems. *J. Anal. Math.* **94**, 49–60.

[2] ANARI, N., GHARAN, S. O. AND REZAEI, A. (2016). Monte Carlo Markov chain algorithms for sampling strongly Rayleigh distributions and determinantal point processes. *J. Mach. Learn. Res.* **49**, 103–115.

[3] BŁASZCZYSZYN, B. AND YOGESHWARAN, D. (2009). Directionally convex ordering of random measures, shot noise fields, and some applications to wireless communications. *Adv. Appl. Prob.* **41**, 623–646.

[4] BŁASZCZYSZYN, B. AND YOGESHWARAN, D. (2014). On comparison of clustering properties of point processes. *Adv. Appl. Prob.* **46**, 1–20.

[5] BŁASZCZYSZYN, B. AND YOGESHWARAN, D. (2015). Clustering comparison of point processes with applications to random geometric models. In *Stochastic Geometry, Spatial Statistics and Random Fields* (Lecture Notes Math. **2120**), Springer, pp. 31–71.

[6] BORCEA, J. BRÄNDÉN, P. AND LIGGETT, T. (2009). Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* **22**, 521–567.

[7] BULINSKI, A. AND SHASHKIN, A. (2007). *Limit Theorems for Associated Random Fields and Related Systems*. World Scientific.

[8] CHRISTOFIDES, T. C. AND VAGGELATOU, E. (2004). A connection between supermodular ordering and positive/negative association. *J. Multivariate Anal.* **88**, 138–151.

[9] HOUGH, J. B., KRISHNAPUR, M., PERES, Y. AND VIRÁG, B. (2006). Determinantal processes and independence. *Probab. Surveys* **3**, 206–229.

[10] HUI, S. AND PARK, C. (2014). The representation of hypergeometric random variables using independent Bernoulli random variables. *Commun. Statist. Theory Methods* **43**, 4103–4108.

[11] JOAG-DEV, K. AND PROSCHAN, F. (1983). Negative association of random variables with applications. *Ann. Statist.* **11**, 286–295.

[12] JOE, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. CRC Press.
[13] KULESZA, A. AND TASKAR, B. (2012). Determinantal point processes for machine learning. *Found. Trends Mach. Learn.* **5**, 123–286.
[14] KULIK, R. AND SZEKLI, R. (2005). Dependence orderings for some functionals of multivariate point processes. *J. Multivariate Anal.* **92**, 145–173.
[15] LAST, G. AND PENROSE, M. (2017). *Lectures on the Poisson Process*. Cambridge University Press.
[16] LAST, G., SZEKLI, R. AND YOGESHWARAN, D. (2018). Some remarks on associated random fields, random measures and point processes. *ALEA*, to appear. arXiv: 1903.06004.
[17] LI, C., JEGELKA, S. AND SRA, S. (2015). Efficient sampling for k-determinantal point processes. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS)* (Proc. Mach. Learn. Res. **51**), pp. 1328–1337.
[18] LI, C., SRA, S. AND JEGELKA, S. (2016). Fast mixing Markov chains for strongly Rayleigh measures, DPPs, and constrained sampling. In *Advances in Neural Information Processing Systems 29*, eds D. D. Lee *et al.*, Curran Associates, pp. 4188–4196.
[19] LIGGETT, T. (1997). Ultra log-concave sequences and negative dependence. *J. Combinatorial Theory A* **79**, 315–325.
[20] LIGGETT, T. (2002). Negative correlations and particle systems. *Markov Process. Relat. Fields* **8**, 547–564.
[21] LYONS, R. (2003). Determinantal probability measures. *Publ. Math. Inst. Hautes Études Sci.* **98**, 167–212.
[22] LYONS, R. (2014). Determinantal probability: basic properties and conjectures. In *Proceedings of the International Congress of Mathematicians 2014*, Vol. IV, pp. 137–161.
[23] MÜLLER, A. AND STOYAN, D. (2002). *Comparison Methods for Stochastic Models and Risks.* John Wiley, Chichester.
[24] NICULESCU, C. (2000). A new look at Newton's inequalities. *J. Inequal. Pure Appl. Math.* **1,** 17, 14pp.
[25] PEMANTLE, R. (2000). Towards a theory of negative dependence. *J. Math. Phys.* **41**, 1371–1390.
[26] POINAS, A., DELYON, B. AND LAVANCIER, F. (2019). Mixing properties and central limit theorem for associated point processes. To appear in *Bernoulli*.
[27] RÜSCHENDORF, L. (2004). Comparison of multivariate risks and positive dependence. *J. Appl. Prob.* **41**, 391–406.
[28] SOSHNIKOV, A. (2000). Determinantal random point fields. *Russian Math. Surveys* **55**, 923.
[29] SZEKLI, R. (1995). *Stochastic Ordering and Dependence in Applied Probability* (Lecture Notes Statist. **97**). Springer.
[30] WAGNER, D. G. (2008). Negatively correlated random variables and Mason's conjecture for independent sets in matroids. *Ann. Combinatorics* **12**, 211–239.
[31] WHITT, W. (1985). Uniform conditional variability ordering of probability distributions. *J. Appl. Prob.* **22**, 619–633.