On negative association of some finite point processes on general state spaces

Günter Last
Karlsruhe Institute of Technology

Ryszard Szekli *
University of Wrocław

March 15, 2018

Abstract: We study negative association (NA) for mixed sampled point processes and show that NA holds for such processes if the random number of points of them fulfills ULC property. We connect NA property of point processes with dcx dependence ordering and show some consequences of it for mixed sampled and determinantal point processes. Some applications illustrate general theory.

Keywords: Finite point processes; association; negative association; strong Rayleigh measure; ULC;

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 point processes theory which 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.

Early history of various concepts of multivariate negative dependence are based on topics considered in Block and Savits [8], Block, Savits and Shaked [6], Ebrahimi and Ghosh [17] and Karlin and Rinott [24]. One of the fundamental results discussed in [8] is that if a distribution satisfies an intuitive structural condition called *Condition N* (recalled in the next section), then it satisfies all of the other conditions introduced there. *Condition N* is satisfied by the multinomial, hypergeometric and Dirichlet distribution as well as several others. It also implies *negative association* introduced by Joag-Dev and Proschan [22]. A slightly stronger version of *Condition N* implies a condition based on stochastic ordering (NDS) due to Block, Savits and Shaked [7]. Negative association has one distinct advantage over the other types of negative dependence. Non-decreasing functions of disjoint sets of negatively associated random variables are also negatively associated. This closure property does not hold for the other types of negative dependence

*Work supported by National Science Centre, Poland grant 2015/19/B/ST1/01152, address: University of Wrocław, Mathematical Institute, pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.

studied in the above mentioned papers. *Condition N* appears in a natural way in the context of queueing networks implying negative association of the population vector in Gordon-Newell networks and negative association of the sojourn times vector in cyclic networks, see [12], and [13].

Pemantle in [38] in his negative dependence study confined himself to binary-valued random variables, in the hope that eliminating the metric and order properties of the real numbers in favour of the two point set, will better reveal what is essential to the questions about negative dependence. The list of examples that motivated him to develop techniques for proving that measures have negative dependence properties such as negative association include the uniform random spanning tree, where the vector of indicator functions of the events that the edges of a graph belong to randomly chosen spanning tree is a random vector which is negatively associated, which was proved by Feder and Mihail [16]. Similar properties hold for weighted spanning trees. Further items on this list are simple exclusion processes, random cluster models and the occupation status of competing urns. Dubhashi and Ranjan [14] consider the competing urns example in detail and show negative association of the numbers of balls in each bin and some consequences such as Chernoff bounds for various models (for extensions see [20]). From this fact it follows negative association of the indicators of exceeding any prescribed thresholds in bins. Occupation numbers of urns under various probability schemes have appeared in many places. Instead of multinomial probabilities, one can postulate indistinguishability of urns or balls and arrive at Bose-Einstein or other statistics. Negative association arise in the multinomial models, where Mallows [35] was one of the first ones to observe negative dependence.

In Borcea et al. [9] several conjectures related to negative dependence made by Liggett [31], Pemantle [38], and Wagner [45], respectively, were solved and also Lyons' main results [33] on negative 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). The problem of describing natural negative dependence properties that are preserved by symmetric exclusion evolutions has attracted some attention in the theory of interacting particle systems and Markov processes. In [9] the authors provide an answer to the aforementioned problem and show that if the initial distribution of a symmetric exclusion process is strongly Rayleigh, then so is the distribution at any time; therefore, the latter distribution is strongly negatively associated. In particular, this solved an open problem of Pemantle and Liggett stating that the distribution of a symmetric exclusion process at any time, with non-random/deterministic initial configuration is negatively associated, and shows that the same is actually true whenever the initial distribution is strongly Rayleigh. In a later paper [32], Liggett has applied these results to prove convergence to the normal and Poisson laws for various functionals of the symmetric exclusion process. Another possible scenario when utilising negative association is for example : first, show that a model is negatively associated; second, use that negatively associated measures have sub - Gaussian tails; finally use that negative association is known to imply the Chernoff - Hoeffding tail bounds. That kind of approach via strong Rayleigh measures has been realised by Pemantle and Peres [39], and a way of finding negative association via strong

Rayleigh property by Peres et al [40].

For point processes a negative association result is known in a quite general setting for so called determinantal point processes on locally compact Polish spaces generated by locally trace class positive contractions on natural L^2 space, see e.g. [34], Theorem 3.7. A broad list of interesting examples of determinantal point processes can be found in [43]. Negative dependence for finite point processes via determinantal and/or strongly Rayleigh measures have 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 for example [2], [25], [28], [29], and references therein.

Another approach to study dependence has been used in finance models. Positive and negative dependence may be seen as some stochastic ordering relation to independence. Such stochastic orderings are called *dependence orderings* (see [23] or [36]). Typical orderings used are supermodular ordering and directionally convex ordering. Relations of these orderings to association and negative association with some applications to concentration inequalities and to the theory of copula functions are given by Christofides and Vaggelatou [11], and Ruechendorf [42], respectively. 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 Blaszczyszyn and Yogeshwaran [5].

Positive and negative association may be used to obtain information on the distribution of functionals such as the sum of coordinates. Newman [37] shows that under either a positive or a negative dependence assumption the joint characteristic function of the variables $\mathbf{X}$ is well approximated by the product of individual characteristic functions. This allows him to obtain central limit theorems for stationary sequences of associated variables. In the positive association case one needs to assume summable covariances, whereas in the negative case one gets this for free. The list of references on central limit theorems for positively/negatively associated variables is very long, a recent reference dealing with point processes is for example the paper by Poinas et al [41].

The central concept of negative dependence in the present paper is negative association (NA) which we show for some point processes. The theory and application of NA are not simply the duals of the theory and application of positive association, but differ in important respects. Negative association has one distinct advantage over the other known types of negative dependence. Non-decreasing functions of disjoint sets of NA random variables are also NA. Apart from NA property of determinantal point processes not much is known about NA property of other point processes. Therefore it might be interesting to characterise NA property of a different, elementary but very useful class of finite point processes with iid locations of points, so called mixed sampled point processes. In order to obtain NA 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 3.3 and 3.6). Consequences of NA property of point processes to ordering of dependence in point processes are described in Proposition 4.7 and Corollary 4.8.

2 Negative association and related definitions

For the distribution of a real random variable X we say that its density (or probability function) is PF_2 if it is log-concave (discrete log-concave) on its support, see eg. Block et al [6] for a detailed description.

The random vector $\mathbf{X} = (X_1, \dots, X_n)$ satisfies *Condition N* if there exist $n+1$ independent real random variables $S_0, S_1, \dots, S_n$, each having a PF_2 density (or probability function) and a real number s such that

$$\mathbf{X} \stackrel{d}{=} [(S_1, \dots, S_n) | S_0 + S_1 + \cdots + S_n = s]$$

where $\stackrel{d}{=}$ denotes equality in distribution, and $[(S_1, \dots, S_n)|S_0 + S_1 + \cdots + S_n = s]$ denotes a random variable having the distribution of $(S_1, \dots, S_n)$ conditioned on the event $S_0 + S_1 + \cdots + S_n = s$.

The multinomial distribution is conditional distribution of sums of independent Poisson random variables given that their sum is fixed, and the Dirichlet is the conditional distribution of independent gammas given that its sum is fixed.

Condition N is related to many other definitions of negative dependence, and for example, as summarised in [8], it is stronger than being completely RR_2 in pairs, than having RR_2 in pairs measure, than being $S-MRR_2$, and finally stronger than NA. We recall the definition of NA.

Definition 2.1. $\mathbf{X}$ *is negatively associated (NA) if, for every subset* $A \subseteq \{1, \dots, n\}$

$$\mathbb{C}\mathrm{ov}(f(X_i, i \in A), g(X_j, j \in A^c)) \leq 0,$$

whenever f, g *are non-decreasing.*

NA may also refer to the set of random variables $\{X_1, \dots, X_n\}$, or to the underlying distribution of $\mathbf{X}$.

Negative association possesses the following properties (see Joag-Dev and Proschan [22])

(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).

(ii) For disjoint subsets $A_1, \dots, A_m$ of $\{1, \dots, n\}$, and non-decreasing positive functions $f_1, \dots, f_m$, $\mathbf{X}$ is NA implies

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

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

(iii) If $\mathbf{X}$ is NA then it is NOD.

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

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

(vi) Increasing (non-decreasing) functions defined on disjoint subsets of a set of NA random variables are NA.

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

In some applications negative association appears, when the random variables are subjected to conditioning.

Theorem 2.2. *Let $X_1, \ldots, X_n$ be independent, and suppose that*

$$\mathbb{E}(f(\mathbf{X}_A) \mid \sum_{i \in A} X_i = s) \tag{2.1}$$

is increasing in s, for every nondecreasing f, and every $A \subseteq \{1, \ldots, n\}$. Then the distribution of $(\mathbf{X} \mid \sum_{i=1}^{n} X_i = s)$ is NA, for almost all s.

The above theorem takes on added interest, when considered in conjunction with the following theorem from [15]. For a queueing theoretical proof of this theorem see [12].

Theorem 2.3. *Let $X_1, \ldots, X_n$ be mutually independent with PF_2 densities and $S_n = \sum_{i=1}^{n} X_i$. Then*

$$\mathbb{E}(\phi(\mathbf{X}) \mid S_n = s)$$

is increasing in (almost every) s, provided ϕ is non-decreasing.

Corollary 2.4. *If $X_1, \ldots, X_n$ are independent with PF_2 densities then the conditional distribution of $(\mathbf{X} \mid S_n = s)$ is NA, for almost all s.*

Conditioning with respect to sums is not the only way to obtain NA property by conditioning, as it is shown in [19], where conditioning on order statistics were used.

The property of negative association is reasonably useful but hard to verify. Negatively correlated probability measures appear naturally in many different contexts. Here are some examples related to NA property. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be a set of real numbers. A permutation distribution is the joint distribution of the vector $\mathbf{X}$, which takes as values all permutations of $\mathbf{x}$ with equal probabilities $1/n!$. Such a distribution is NA. Negatively correlated normal random variables are NA. See Joag-Dev and Proschan [22]) for these and many other examples. There are some examples directly related to *Condition N* and conditioning. NA property of multinomial distributions can be seen from *Condition N*, since it is the conditional distribution of independent Poisson random variables given their sum. One can see that multivariate hyper-geometric distribution is NA because it is the conditional distribution of independent binomial random variables given their sum, see [6]. The population vector in Gordon-Newell closed queueing networks is NA, see [44], Section 3.8, Theorem E. The sojourn times vector in cyclic queues has got property NA, see [13]. A plethora of negative dependence properties and conjectures for zero-one valued vectors ($CNA, CNA+, JNRD, JNRD+, h-NLC+, S-MRR_2$, ULC, strongly Rayleigh, Rayleigh, PHR) were introduced and studied in [38] and [10] with an application to symmetric exclusion processes. These classes are related to determinantal probability measures described for example in [33] and [34]. We shall not recall all of these definitions, and only recall NA property of determinantal point processes.

We shall utilize a slightly broader class than NA in our formulations on dependence orderings. We say that a random vector $\mathbf{X}$ (or its distribution) is sNA (negatively associated in sequence) if

$$\mathbb{C}\mathrm{ov}(\mathbf{1}_{(X_i>t}, f(X_{i+1}, \ldots, X_n)) \leq 0,$$

for all f non-decreasing and $t \in \mathbb{R}$.

This condition is equivallent to $[(X_{i+1}, \ldots, X_n) \mid X_i > t)] <_{st} (X_{i+1}, \ldots, X_n)$ for all $t \in \mathbb{R}$ and $i = 1, \ldots, n-1$ Here $<_{st}$ denotes the usual strong stochastic ordering on $\mathbb{R}^n$.

3 NA for mixed sampled point processes

We shall adopt our notation from the book by Last and Penrose [27]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathbb{X}, \mathcal{X})$ a measurable state space. Denote by $\mathbf{N}_{<\infty}(\mathbb{X}) = \mathbf{N}_{<\infty}$ the space of all measures μ on $(\mathbb{X}, \mathcal{X})$ such that $\mu(B) \in \mathbf{Z}_+$ for all $B \in \mathcal{X}$. Let $\mathbf{N}(\mathbb{X}) = \mathbf{N}$ be the space of all measures that can be written as a countable sum of measures from $\mathbf{N}_{<\infty}$. An example is the Dirac measure δ_x for a point $x \in \mathbb{X}$, given by $\delta_x(B) := \mathbf{1}_B(x)$.

We define a *point process* η as a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathbf{N}, \mathcal{N})$ ($\mathcal{N}$ is the smallest σ-field on $\mathbf{N}$ such that $\mu \to \mu(B)$ is measurable for all $B \in \mathcal{X}$).

We restrict our attention in this paper to finite point process with points located in a complete separable metric space $\mathbb{X}$.

We define NA property of point processes as follows.

Definition 3.1. *A point process η is negatively associated (NA) or negatively associated in sequence (sNA) if for each collection of disjoint sets $B_1, \ldots, B_n \in \mathcal{X}$ the vector $(\eta(B_1), \ldots, \eta(B_n))$ is NA or sNA, respectively, as defined for random vectors.*

More general definition of negative association for point processes (random measures) is possible, see [34] or [47].

For a Borel set $A \subseteq \mathbb{X}$, let $\mathcal{N}_{|A}$ denote the σ-field on $\mathbf{N}$ generated by the functions $\mu \to \mu(B)$ for Borel $B \subseteq A$. The natural (inclusion) partial order on $\mathbf{N}$ allows us to define $f : \mathbf{N} \to \mathbb{R}$ which is increasing. We say that a point process η has *negative associations* if $\mathbb{E}(f(\eta)g(\eta)) \leq \mathbb{E}(f(\eta))\mathbb{E}(g(\eta))$ for every pair f, g of 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 $\mathcal{N}_{|A}$. It is clear that with such a general definition we have that if η has *negative associations* then η is NA. Let us recall Theorem 3.7 from [34]. In this theorem η_K denotes the determinantal point process generated by K, for details see [34], section 3.2.

Theorem 3.2. *Let λ be a Radon measure on a locally compact Polish space $\mathbb{X}$. Let K be a locally trace-class positive contraction on $L_2(\mathbb{X}, \lambda)$. Then η_K has negative associations.*

Apart from determinantal point processes not much is known about NA property of point processes. Therefore it might be interesting to characterise NA property of an elementary but very useful class of finite point processes with iid locations of points, more precisely our main focus in this paper is on the class of so called *mixed sampled point processes* on $\mathbb{X}$, defined by

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

where $(X)_{i\geq 1}$ is iid with distribution F, and $\tau \in \mathbb{N} \cup \{0\}$ is finite and independent of $(X_i)_{i\geq 1}$.

For this process, given any finite partition $A_1, \ldots, A_k$ of $\mathbb{X}$, conditionally on τ, the joint distribution of the number of points is given by

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

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 by

$$\mathbb{E}(z_1^{\eta(A_1)} \cdots z_k^{\eta(A_k)}) = P_\tau(F(A_1)z_1 + \cdots + F(A_k)z_k), \tag{3.2}$$

where $P_\tau(z) = \mathbb{E}(z^\tau)$.

We shall use the following distribution classes for $U_1, \ldots, U_i$ independent Bernoulli variables with success probabilities $p_1, \ldots, p_n \in [0,1]$, respectively. The class of random variables which are the sums of n Bernoulli variables we denote by

$$\mathcal{Q}_n := \{\tau : \tau = \sum_{i=1}^{n} U_i\}.$$

The class of all distributions with supports contained in $\{0, 1, \ldots\}$ appearing as weak limits of distributions from $\mathbb{Q}_n$, $n \geq 1$, i.e. the weak closure of $\cup_{n=1}^{\infty} \mathbb{Q}_n$ we denote by

$$\mathcal{Q} := \mathrm{cl}(\cup_{n=1}^{\infty} \mathbb{Q}_n).$$

The main results of this paper are contained in Theorems 3.3 and 3.6.

Theorem 3.3. *Suppose that η is a mixed sampled point process on $\mathbb{X}$, 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 by

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

the zero-one valued vector generated by the i-th sample X_i, $i \geq 1$. Note that each $\mathbf{Z}_i$ has multinomial distribution with success parameters $q_1, \ldots, q_m$ and the number of trials equal 1, and as such is NA. Moreover $\mathbf{Z}_1, \ldots$ are independent. For fixed $n \in \mathbb{N}$ and $p_1, \ldots, p_n \in [0,1]$, let $\mathbf{U} = (U_1, \ldots, U_n)$ be a vector of zero-one valued, independent random variables with $p_i = \mathbb{P}(U_i = 1)$, which is independent of $\mathbf{Z}_1, \ldots$. The vector composed as $(\mathbf{U}, \mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ is NA because of properties (v) and (vii) of NA.

Now using property (vi) we get that the vector $(U_1\mathbf{Z}_1, \ldots, U_n\mathbf{Z}_n)$ is NA as a monotone transformation (multiplication) of disjoint coordinates of $(\mathbf{U}, \mathbf{Z}_1, \ldots, \mathbf{Z}_n)$. Again using property (vi), this time to $(U_1\mathbf{Z}_1, \ldots, U_n\mathbf{Z}_n)$ and using appropriately addition we get that the vector $\sum_{i=1}^n U_i\mathbf{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\mathbf{Z}_i$ has got the same distribution as $\sum_{i=1}^\tau \mathbf{Z}_i$ if $\tau \in \mathcal{Q}_n$, which in turn is the same as for $(\eta(B_1), \ldots, \eta(B_m))$. This finishes the proof for $\tau \in \mathcal{Q}_n$ for arbitrary $n \in \mathbb{N}$.

For $\tau \in \mathcal{Q}$ there exist a sequence $\tau_k \xrightarrow{d} \tau$, $k \to \infty$, for $\tau_k \in \cup_{n=1}^\infty \mathcal{Q}_n$, and

$$\mathbb{E}[f(\sum_{i=1}^{\tau_k} \mathbf{Z}_i)g(\sum_{i=1}^{\tau_k} \mathbf{Z}_i)] \leq \mathbb{E}[f(\sum_{i=1}^{\tau_k} \mathbf{Z}_i)]\mathbb{E}[g(\sum_{i=1}^{\tau_k} \mathbf{Z}_i)],$$

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

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

Since each non-decreasing function can be monotonically approximated by non-decreasing and bounded functions, we get NA property of η. □

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

Lemma 3.4.

$$\tau \in \mathcal{Q} \quad \textit{iff} \quad \tau =^d \tau_1 + \tau_2,$$

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

It is interesting to note that hypergeometric random variables belong to the class $\mathcal{Q}$, see e.g. [21]

We say that a real sequence $(a_i)_{i=0}^n$ has no internal zeros if the indices of its non-zero terms form a discrete interval. Following Pemantle [38] we shall use the following class of sequences and distributions.

Definition 3.5. *W say that a finite real sequence $(a_i)_{i=0}^n$ of non-negative real numbers with no internal zeros is ultra log-concave* (ULC(n)) *if*

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

The class of random variables for which their probability functions have the above property we denote by

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

It is known that if a non-nengative 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 [30], Theorem 2. Let

$$\mathcal{S} := \mathrm{cl}(\cup_{n=1}^\infty \mathcal{S}_n).$$

Sums of independent variables from the class $\mathcal{S}$ are in $\mathcal{S}$. Using $\mathcal{S}$, the Theorem 3.3 can be generalised.

Theorem 3.6. *Suppose that η is a mixed sampled point process on $\mathbb{X}$, 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 $q_i := F(B_i)$, $i = 1, \dots, m$. Define by

$$\mathbf{Z}_i := (\mathbf{1}_{\{X_i \in B_1\}}, \dots, \mathbf{1}_{\{X_i \in B_m\}},$$

the zero-one valued vector generated by the i-th sample X_i, $i \geq 1$. Note that each $\mathbf{Z}_i$ has multinomial distribution with success parameters $q_1, \dots, q_m$ and the number of trials equal 1, and as such is NA. Moreover $\mathbf{Z}_1, \dots$ are independent.

For fixed $n \in \mathbb{N}$, let $\mathbf{U} = (U_1, \dots, U_n)$ be a vector, independent of $\mathbf{Z}_1, \dots,$ of zero-one valued random variables obtained in the following way. For the generating function of τ, $P_\tau(z) = \mathbb{E}(z^\tau)$, we define $\mathbf{U}$ by providing its multidimensional generating function, which 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 $e_k(z_1, \dots, z_n)$ are the elementary symmetric polynomials. Note that immediately from 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 zero-one valued variables which takes on exactly k values one with equal probability $\binom{n}{k}^{-1}$, for all possible selections of k coordinates where the values one are obtained. The distribution of $\mathbf{U}$ 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 $\mathbf{U} = (U_1, \dots, U_n)$, therefore $(U_1, \dots, U_n)$ are exchangeable. Moreover $\sum_{i=1}^n U_i \stackrel{d}{=} \tau$, since by setting $z_1 = \dots = z_n := z$ we obtain $P_\tau(z)$. In other words the sequence $(a_i := \mathbb{P}(\tau = i))_{i=0}^n$ is the rank sequence for the vector $\mathbf{U} = (U_1, \dots, U_n)$.

From our assumption the rank sequence for $\mathbf{U} = (U_1, \dots, U_n)$ is ULC(n) and from Theorem 2.7 in [38] we obtain that $\mathbf{U}$ is NA. Now the vector composed as $(\mathbf{U}, \mathbf{Z}_1, \dots, \mathbf{Z}_n)$ is NA because of property (vii) of NA. Using property (vi) we get that the vector $(U_1\mathbf{Z}_1, \dots, U_n\mathbf{Z}_n)$ is NA as a monotone transformation (multiplication) of disjoint coordinates of $(\mathbf{U}, \mathbf{Z}_1, \dots, \mathbf{Z}_n)$. Again using property (vi), this time to $(U_1\mathbf{Z}_1, \dots, U_n\mathbf{Z}_n)$ and using appropriately addition we get that the vector $\sum_{i=1}^n U_i\mathbf{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\mathbf{Z}_i$ has got the same distribution as $\sum_{i=1}^\tau \mathbf{Z}_i$, which in turn has the same distribution as $(\eta(B_1), \dots, \eta(B_m))$. This finishes the proof for $\tau \in \mathcal{S}_n$, for arbitrary $n \in \mathbb{N}$. For $\tau \in \mathcal{S}$, we apply an analogous argument as in Theorem 3.3. □

The following lemma may be regarded as known since it is an immediate consequence of the Newton inequalities. For convenience we put its formulation in the setting of the introduced in this paper classes of random variables.

Lemma 3.7. *For all $n > 1$*

$$\mathcal{Q}_n \subseteq \mathcal{S}_n.$$

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

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

For $a_k := \frac{1-p_k}{p_k}$ we have

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

where $c_n = a_1 + \cdots + a_n$, $c_2 = a_1a_2 + \cdots + a_{n-1}a_n$, $\cdots$, $c_n = a_1 \cdots a_n$, i.e. c_k, $k = 0, \ldots, n$, coefficients are given by the corresponding elementary symmetric polynomials in variables $a_i, i = 1, \ldots, n$. It is known from Newton's inequalities that for $k = 1, \ldots, n-1$

$$\left(\frac{c_k}{\binom{n}{k}}\right)^2 \geq \frac{c_{k-1}}{\binom{n}{k-1}} \frac{c_{k+1}}{\binom{n}{k+1}},$$

and since $\mathbb{P}(\tau = k) = p_1 \cdots p_n c_{n-k}$, we get that the sequence $(\mathbb{P}(\tau = k))_{k=0}^n$ is ULC(n), and therefore $\tau \in \mathcal{S}_n$. $\square$

It is interesting to note that arguments utilised in Theorem 3.6 can be used for random vectors with arbitrary positive values.

Proposition 3.8. *Assume that $(\mathbf{Z}_i = (Z_i^1, \ldots, Z_i^m), i \geq 1)$ is a sequence of independent, identically distributed 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 if $\tau \in \mathcal{S}$ is independent of $(\mathbf{Z}_i = (Z_i^1, \ldots, Z_i^m), i \geq 1)$ then $\mathbf{W} := \sum_{i=1}^\tau \mathbf{Z}_i$ is NA,*

Proof. We use basically the same argument as in Theorem 3.6. Let $\mathbf{U} = (U_1, \ldots, U_n)$ be a vector, independent of $\mathbf{Z}_1, \ldots$, of zero-one valued random variables obtained from $P_\tau = \mathbb{E}(z^\tau)$ by substituting in $P_\tau(z)$, $z^k := \binom{n}{k}^{-1} e_k(z_1, \ldots, z_n)$, $k = 1, \ldots, n$, where $e_k(z_1, \ldots, z_n)$ are the elementary symmetric polynomials, and taking $\mathbf{U} = (U_1, \ldots, U_n)$ as random variables with the generating function obtained in this way, as a function of variables $z_1, \ldots, z_n$. It is then immediate that $\sum_{i=1}^n U_i =^d \tau$, that is the sequence $(a_i := \mathbb{P}(\tau = i))_{i=0}^n$ is the rank sequence for (symmetric) $\mathbf{U} = (U_1, \ldots, U_n)$. From our assumption we have that the rank sequence for $\mathbf{U} = (U_1, \ldots, U_n)$ is ULC(n) and from Theorem 2.7 in [38] we obtain that $\mathbf{U}$ is NA. Now the vector composed as $(\mathbf{U}, \mathbf{Z}_1, \ldots, \mathbf{Z}_n)$ is NA because of the following lemma and property (vii) of NA.

Lemma 3.9. *Assume that $\mathbf{Z} = (Z^1, \ldots, Z^m)$ is a random vector with components in $\mathbb{R}_+$. Assume that $\sum_{j=1}^m \mathbf{1}\{Z^i > 0\} \leq 1$, that is, at most one of the components can be positive. Then $\mathbf{Z}$ is NA.*

Proof of lemma 3.9. In order to show that $\mathbb{C}\mathrm{ov}(f(Z^1, \ldots, Z^k), g(Z^{k+1}, \ldots, Z^m)) \geq 0$ for non-decreasing f and g, it suffices to assume that $f(0) = g(0) = 0$. Otherwise one can consider $f - f(0)$ and $g - g(0)$. Because only of the coordinates can be non-zero, we get $E[f(Z^1, \ldots, Z^k)g(Z^{k+1}, \ldots, Z^m))] = 0$, while the product of the expectations is non-negative since, $f \geq 0$ and $g \geq 0$. $\square$

Now, using property (vi), we get that the vector $(U_1\mathbf{Z}_1, \ldots, U_n\mathbf{Z}_n)$ is NA as a monotone transformation (multiplication) of disjoint coordinates of $(\mathbf{U}, \mathbf{Z}_1, \ldots, \mathbf{Z}_n)$. Again using property (vi), this time to $(U_1\mathbf{Z}_1, \ldots, U_n\mathbf{Z}_n)$ and using appropriately addition we get that the vector $\sum_{i=1}^n U_i\mathbf{Z}_i$ is NA. It is clear that $\sum_{i=1}^n U_i\mathbf{Z}_i$ has got the same distribution as $\sum_{i=1}^\tau \mathbf{Z}_i$. $\square$

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

4 Dependence orderings for point processes

An extensive study of dependence orderings for multivariate point processes on $\mathbb{R}$ is contained in [26]. 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 Blaszczyszyn and Yogeshwaran [5], see also references therein. We shall use NA property of point processes we study to obtain comparisons related to dependency properties. First we recall some basic facts on dependence orderings of vectors and their relation to NA which can be directly utilized for point processes.

4.1 Dependence orderings and negative correlations for vectors

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

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

where $\mathbf{e}_i$ is the i-th unit vector. Then f is called *super-modular* if for all $1 \le i < j \le n$ and $\epsilon, \delta > 0$

$$\Delta_j^\delta \Delta_i^\epsilon f(\mathbf{x}) \ge 0.$$

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

The corresponding stochastic orderings one defines by $\mathbf{X} <_{sm} \mathbf{Y}$ if $\mathbb{E} f(\mathbf{X}) \le \mathbb{E} f(\mathbf{Y})$ for all $f \in \mathcal{F}^{sm}$, and analogously for $\mathbf{X} <_{dcx} \mathbf{Y}$. For differentiable functions f one obtains $f \in \mathcal{F}^{sm}$ iff $\frac{\partial^2}{\partial x_i x_j} f \ge 0$, for $i < j$, and $f \in \mathcal{F}^{dcx}$ iff this inequality holds for $i \le j$. While comparison of $\mathbf{X}$ and $\mathbf{Y}$ with respect to $<_{sm}$ implies (and is restricted to the case of) identical marginals $X_i =^d Y_i$, the comparison with respect to the smaller class $\mathcal{F}^{dcx}$ implies convexly increasing marginals $X_i <_{cx} Y_i$ (which means by definition that $\mathbb{E}\psi(X_i) \le \mathbb{E}\psi(Y_i)$ for all $\psi : \mathbb{R} \to \mathbb{R}$ convex). Both of these orderings belong to the class of so called *dependency orderings*, see e.g. [23], which is defined by a list of suitable properties, among them the property that $\mathbb{C}\mathrm{ov}(X_i, X_j) \le \mathbb{C}\mathrm{ov}(Y_i, Y_j)$ for such orderings.

In [42] another stochastic ordering related to dependence comparisons was introduced by the condition $\mathbf{X} <_{wcs} \mathbf{Y}$ (*weakly conditional increasing in sequence* order) iff

$$\mathbb{C}\mathrm{ov}(\mathbf{1}_{(X_i > t}, f(X_{i+1}, \ldots, X_n)) \le \mathbb{C}\mathrm{ov}(\mathbf{1}_{(Y_i > t}, f(Y_{i+1}, \ldots, Y_n)),$$

for all f monotonically non-decreasing, and all $t \in \mathbb{R}$, $1 \le i \le n-1$. The following theorem from [42] connects the above defined orderings.

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

a) If $X_i \stackrel{d}{=} Y_i$, $1 \le i \le n$ then $\mathbf{X} <_{wcs} \mathbf{Y}$ implies that $\mathbf{X} <_{sm} \mathbf{Y}$.

b) If $X_i <_{cx} Y_i$, $1 \le i \le n$ then $\mathbf{X} <_{wcs} \mathbf{Y}$ implies that $\mathbf{X} <_{dcx} \mathbf{Y}$.

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, denote by $\mathbf{X}^*$ a vector with independent components, and $X_i^* \stackrel{d}{=} X_i$. Then $\mathbf{X}$ is called *weakly associated in sequence* if $\mathbf{X}^* <_{wcs} \mathbf{X}$. We shall say that $\mathbf{X}$ is *negatively associated in sequence* (sNA) if $\mathbf{X} <_{wcs} \mathbf{X}^*$. We remark that sNA as defined above is equivalent to

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

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 got the distribution of $(X_{i+1}, \ldots, X_n)$ conditioned on the event $\{X_i > t\}$, and $<_{st}$ is the usual (strong) stochastic order. It is clear directly from the definition that NA property implies sNA property. From Theorem 4.1 it follows immediately for $\mathbf{X}$ being sNA that, for example, (see also [11] for the case NA) $\sum_{i=1}^n X_i <_{cx} \sum_{i=1}^n X_i^*$, and $\max_{1\le k\le n} \sum_{i=1}^k X_i <_{icx} \max_{1\le k\le n} \sum_{i=1}^k X_i^*$, where $<_{icx}$ is defined similarly as $<_{cx}$ but with the use of non-decreasing convex functions. Taking other super-modular functions it is possible to get maximal inequalities for sNA vectors as in [11] . For $<_{dcx}$ ordering we get the following corollary from Theorem 4.1 which will be used for point processes with NA property.

Corollary 4.2. *Suppose $\mathbf{X}$ is sNA and $\mathbf{Y}^*$ has independent coordinates with $X_i <_{cx} Y_i^*$ then $\mathbf{X} <_{dcx} \mathbf{Y}^*$.*

4.2 NA and dependence orderings for point processes

We shall connect the approach by dependence orderings with negative association and by using Corollary 4.2 we shall be able to compare dependence for point processes which posses NA property with some other processes. This will imply in particular comparison results for void probabilities and/or moment measures.

Definition 4.3. *Two point processes η_1, η_2 on a complete, separable metric space $\mathbb{X}$ are ordered in directionaly convex order (weakly conditional increasing sequence order)*

$$\eta_1 <_{dcx} \eta_2 \ (\eta_1 <_{wcs} \eta_2) \ \text{ iff } \ (\eta_1(B_1), \ldots, \eta_1\ (B_n)) <_{dcx} (<_{wcs})\ (\eta_2(B_1), \ldots, \eta_2(B_n)),$$

as defined for random vectors, for all disjoint, bounded $B_1, \ldots, B_n \in \mathcal{X}$, $n \ge 1$.

The following consequence of $<_{dcx}$ ordering for point processes is known from [5], Proposition 6.

Lemma 4.4. *Let η_1, η_2 be two point process on $\mathbb{R}^d$. If $\eta_1 <_{dcx} \eta_2$ then*

(a) (moment measures) $\mathbb{E}(\eta_1(B_1)\cdots\eta_1(B_n))) \le \mathbb{E}(\eta_2(B_1)\cdots\eta_2(B_n)))$,

(b) (void probabilities) $\mathbb{P}(\eta_1(B) = 0) \le \mathbb{P}(\eta_2(B) = 0)$, for all bounded Borel sets.

Using $<_{wcs}$ criterion for $< dcx$ from Theorem 4.1, b) we get

Corollary 4.5. *Let η_1, η_2 be two point process on $\mathbb{R}^d$. If $\eta_1(B) <_{cx} \eta_2(B)$ and $\eta_1 <_{wcs} \eta_2$ then the moment measures and void probabilities comparisons from the above lemma hold.*

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

Proposition 4.6. *Suppose η is a simple point process on $\mathbb{R}^d$ which is* sNA *then*

(a) (moment measures) $\mathbb{E}(\eta(B_1)\cdots\eta(B_n))) \le \mathbb{E}(\eta(B_1))\cdots\mathbb{E}(\eta(B_n))$, *which is equivalent to*

$$\mathbb{E}(\exp(\int_{\mathbb{R}^d} h(\mathbf{x})\eta(d\mathbf{x}) \le \exp(\int_{\mathbb{R}^d}(e^{h(\mathbf{x})}-1)\mathbb{E}\eta(d\mathbf{x})),$$

for all $h \ge 0$.

(b) (void probabilities) $\mathbb{P}(\eta(B)=0) \le \exp(-\mathbb{E}\eta(B))$, *for all bounded Borel sets, which is equivalent to*

$$\mathbb{E}(\exp(-\int_{\mathbb{R}^d} h(\mathbf{x})\eta(d\mathbf{x}) \le \exp(\int_{\mathbb{R}^d}(e^{-h(\mathbf{x})}-1)\mathbb{E}\eta(d\mathbf{x})),$$

for all $h \ge 0$.

Using Corollary 4.2 the above result can be extended.

Proposition 4.7. *Suppose η_1 is a simple point process on $\mathbb{X}$ which is* sNA *and $\eta_1(B) <_{cx} Y$, where Y is a random variable with Poisson distribution $Po(\mathbb{E}\eta_1(B))$ for all $B \in \mathcal{X}$ then $\eta_1 <_{dcx} \eta_2$ where η_2 is a Poisson point process with the intensity measure $\mathbb{E}\eta_1$.*

For η_K which is the determinantal point process generated by K (for details see [34], section 3.2) we obtain from Proposition 4.7

Corollary 4.8. *Let λ be a Radon measure on a locally compact Polish space $\mathbb{X}$. Let K be a locally trace-class positive contraction on $L_2(\mathbb{X}, \lambda)$. Then $\eta_K <_{dcx} \eta_2$ where η_2 denotes a Poisson point process with the intensity measure $\mathbb{E}\eta_K$.*

Proof. Fix $B \in \mathcal{X}$. Using Theorem 3.2, we shall show that $\eta(B) <_{cx} Y$, where Y is a random variable with $Po(\mathbb{E}\eta(B))$ distribution. From Hough et al. [18], Proposition 9, we know that $\eta(B)$ is distributed as a sum of independent Bernoulli random variables and therefore it is contained in the class $\mathcal{Q}$. It is immediate from the definition of the log-concave ordering $<_{lc}$ in [46], that $\eta(B) <_{lc} Y$, where Y is described above, which in turn implies that $\eta(B) <_{cx} Y$, see Theorem 1 in [46]. □

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

4.3 Comparisons for mixed sampled point processes

Before we formulate a more general result we start with a particular example.

Example 4.9 (binomial random sum)**.** Comparison of void probabilities $\mathbb{P}(\eta(B)=0)$ is equivalent to the comparison of one dimensional Laplace transforms of $\eta(B)$ for positive arguments. When comparing with a Poisson point process it equivalent to comparison of Laplace functionals as in (b). For simple point processes η on $\mathbb{R}^d$ in order to get (b) it is enough to check (see Proposition 3.1, [5]) whether

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

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

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

If τ has the binomial distribution with 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 differentiaition 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 (see e.g. Bruckner and Ostrow (1963)), therefore $\phi(s + t) \geq \phi(s) + \phi(t)$, and then $P_\tau(1 - (F(B) + F(B')) \leq P_\tau(1 - F(B))P_\tau(1 - F(B'))$. In this case, for disjoint B, B' we obtain

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

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

Proposition 4.10. *Suppose that η is a mixed sampled point process on $\mathbb{X}$, defined by (3.1), for which $\tau \in \mathcal{S}$, then*

(a) (moment measures) $\mathbb{E}(\eta(B_1) \cdots \eta(B_n))) \leq \mathbb{E}(\eta(B_1)) \cdots \mathbb{E}(\eta(B_n))$,

(b) (void probabilities) $\mathbb{P}(\eta(B) = 0) \leq \exp(-\mathbb{E}\eta(B))$, *for all bounded Borel sets,*

Proposition 4.11. *Suppose that η_1 is a mixed sampled point process on $\mathbb{X}$, defined by (3.1), for which $\tau \in \mathcal{S}$, then*

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

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

Proof. From Theorem 3.6, and Proposition 4.7 we shall get the conclusion of the present proposition if we show that for such processes $\eta_1(B) <_{cx} Y$, where Y denotes a random variable with Poisson distribution $Po(\mathbb{E}\eta(B))$. We know that

$$\mathbb{E}(z^{\eta_1(B)}) = P_\tau(F(B)z + F(B^c)), \tag{4.1}$$

where $P_\tau(z) = \mathbb{E}(z^\tau)$. It means that $\eta_1(B)$ is distributed as a random sum $\sum_{i=1}^{\tau} U_i$, where $(U_i, i \geq 1)$ is an iid sequence of Bernoulli, i.e. zero-one valued variables with the success probability $F(B)$. For $\tau \in \mathcal{S}_n$ it is immediate from the definition of the log-concave ordering $<_{lc}$ in [46], that $\tau <_{lc} Y$, which implies that $\tau <_{cx} Y$, see Theorem 1 in [46]. It follows that $\sum_{i=1}^{\tau} U_i <_{cx} \sum_{i=1}^{Y} U_i$, where Y is independent of $(U_i, i \geq 1)$ (see e.g. [26], Corollary 4.5) which implies $\eta_1(B) <_{cx} Y$. For arbitrary $\tau \in \mathcal{S}$ we apply weak approximation by τ's in $\mathcal{S}_n, n \geq 1$. □

5 Some applications

It is clear that $<_{dcx}$ ordering for point processes implies the inequality for the pair correlation functions and Ripleys K functions.

Moreover using Chebyshev's inequality we have

$$\mathbb{P}(|\eta(B) - \mathbb{E}(\eta(B))| \geq \epsilon) \leq \mathbb{V}\mathrm{ar}(\eta(B)/(\epsilon^2),$$

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

Similarly, using Chernoffs bound,

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

for any $t, a > 0$, and the upper bounds can be replaced by the values taken from the dominating in $<_{dcx}$ process. If η is NA then one can use for determinantal or mixed sampled point processes the corresponding Poisson process as an upper bound which gives concentration inequalities.

Similarly, for

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

Using Corollary 2 from [11] we can get, from negative association of η, Kolmogorov type inequalities.

Corollary 5.1. *Suppose that η is a mixed sampled point process on $\mathbb{X}$ for which $\mathbb{P}^\tau \in \mathbb{S}$, then for any increasing sequence (b_k) of positive numbers, any collection of disjoint sets $B_1, \ldots, B_n \in \mathcal{X}$, and $\epsilon > 0$*

(a)

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

(b) for any integer $m < n$

$$\mathbb{P}(\max_{m\leq k\leq n} |\frac{1}{b_k}\sum_{i=1}^{k}(\eta(B_i) - \mathbb{E}(\eta(B_i))| \geq \epsilon) \leq 32\epsilon^{-2}(\sum_{i=m+1}^{n}\frac{\mathbb{V}\mathrm{ar}(\eta(B_i))}{b_i^2} + \sum_{i=1}^{m}\frac{\mathbb{V}\mathrm{ar}(\eta(B_i))}{b_m^2}).$$

References

[1] Alexandru Aleman, Dmitry Beliaev, and Haakan Hedenmalm. Real zero polynomials and pólya-schur type theorems. *Journal d'analyse Mathématique*, 94(1):49–60, 2004.

[2] Nima Anari, Shayan Oveis Gharan, and Alireza Rezaei. Monte carlo markow chain algorithms for sampling strongly rayleigh distributions and determinantal point processes. In *Conference on Learning Theory*, pages 103–115, 2016.

[3] Bartłomiej Błaszczyszyn and D Yogeshwaran. Directionally convex ordering of random measures, shot noise fields, and some applications to wireless communications. *Advances in Applied Probability*, 41(3):623–646, 2009.

[4] Bartłomiej Błaszczyszyn and Dhandapani Yogeshwaran. On comparison of clustering properties of point processes. *Advances in Applied Probability*, 46(1):1–20, 2014.

[5] Bartłomiej Błaszczyszyn and Dhandapani Yogeshwaran. Clustering comparison of point processes, with applications to random geometric models. In *Stochastic Geometry, Spatial Statistics and Random Fields*, pages 31–71. Springer, 2015.

[6] Henry W Block, Thomas H Savits, and Moshe Shaked. Some concepts of negative dependence. *The Annals of Probability*, 10(3):765–772, 1982.

[7] Henry W Block, Thomas H Savits, and Moshe Shaked. A concept of negative dependence using stochastic ordering. *Statistics & Probability Letters*, 3(2):81–86, 1985.

[8] Savits T. Block H. and Shaked M. Negative dependence. *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, 2:206–215, 1982.

[9] Julius Borcea, P Brändén, and T Liggett. Negative dependence and the geometry of polynomials. *J. Am. Math. Soc.*, pages 1–47, 2009.

[10] Julius Borcea, Petter Brändén, and Thomas Liggett. Negative dependence and the geometry of polynomials. *Journal of the American Mathematical Society*, 22(2):521–567, 2009.

[11] Tasos C. Christofides and Eutichia Vaggelatou. A connection between supermodular ordering and positive/negative association. *J. Multivar. Anal.*, 88(1):138–151, 2004.

[12] Hans Daduna and Ryszard Szekli. A queueing theoretical proof of increasing property of pólya frequency functions. *Statistics and Probability Letters*, 26(3):233–242, 1996.

[13] Hans Daduna and Ryszard Szekli. On the correlation structure of closed queueing networks. *Stochastic Models*, 20(1):1–29, 2004.

[14] Devdatt Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *Random Structures & Algorithms*, 13(5):99–124, 1998.

[15] Bradley Efron. Increasing properties of polya frequency function. *The Annals of Mathematical Statistics*, 36(1):272–279, 1965.

[16] Tomás Feder and Milena Mihail. Balanced matroids. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 26–38. ACM, 1992.

[17] Malay Ghosh. Multivariate negative dependence. *Communications in Statistics-Theory and Methods*, 10(4):307–337, 1981.

[18] J Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. Determinantal Processes and Independence. *Probab. Surv.*, 3:206–229, 2006.

[19] T Hu and Jinjin Hu. Sufficient conditions for negative association of random variables. *Stat. Probab. Lett.*, 45:167–173, 1999.

[20] Taizhong Hu and Chaode Xie. Negative dependence in the balls and bins experiment with applications to order statistics. *Journal of Multivariate Analysis*, 97(6):1342–1354, 2006.

[21] Stefen Hui and CJ Park. The representation of hypergeometric random variables using independent bernoulli random variables. *Communications in Statistics-Theory and Methods*, 43(19):4103–4108, 2014.

[22] K Joag-Dev and F Proschan. Negative association of random variables with applications. *Ann. Stat.*, 11:286–295, 1983.

[23] Harry Joe. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

[24] Samuel Karlin and Yosef Rinott. Classes of orderings of measures and related correlation inequalities ii. multivariate reverse rule distributions. *Journal of Multivariate Analysis*, 10(4):499–516, 1980.

[25] Alex Kulesza, Ben Taskar, et al. Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3):123–286, 2012.

[26] Rafał Kulik and Ryszard Szekli. Dependence orderings for some functionals of multivariate point processes. *Journal of multivariate analysis*, 92(1):145–173, 2005.

[27] Günter Last and Mathew Penrose. *Lectures on the poisson process*, volume 7. Cambridge University Press, 2017.

[28] Chengtao Li, Stefanie Jegelka, and Suvrit Sra. Efficient sampling for k-determinantal point processes. *arXiv preprint arXiv:1509.01618*, 2015.

[29] Chengtao Li, Suvrit Sra, and Stefanie Jegelka. Fast mixing markov chains for strongly rayleigh measures, dpps, and constrained sampling. In *Advances in Neural Information Processing Systems*, pages 4188–4196, 2016.

[30] Thomas M Liggett. Ultra logconcave sequences and negative dependence. *journal of combinatorial theory, Series A*, 79(2):315–325, 1997.

[31] Thomas M Liggett. Negative correlations and particle systems. *Markov Proc. Related Fields*, 8:547–564, 2002.

[32] Thomas M Liggett. Distributional limits for the symmetric exclusion process. *Stochastic Processes and their Applications*, 119(1):1–15, 2009.

[33] Russell Lyons. Determinantal probability measures. *Publications mathématiques de l'IHÉS*, 98(1):167–212, 2003.

[34] Russell Lyons. Determinantal probability: basic properties and conjectures. *arXiv preprint arXiv:1406.2707*, 2014.

[35] C.L. Mallows. An inequality involving multinomial probabilities. *Biometrika*, 55:422–424, 1968.

[36] A Müller and D Stoyan. Comparison methods for stochastic models and risks. *John Wiley&Sons Ltd., Chichester*, 2002.

[37] Charles M. Newman. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Inequalities in Statistics and Probability*, pages 127–140. Institute of Mathematical Statistics, 1984.

[38] Robin Pemantle. Towards a theory of negative dependence. *J. Math. Phys.*, 41:1371–1390, 2000.

[39] Robin Pemantle and Yuval Peres. Concentration of lipschitz functionals of determinantal and other strong rayleigh measures. *Combinatorics, Probability and Computing*, 23(1):140–160, 2014.

[40] Yuval Peres, Mohit Singh, and Nisheeth K Vishnoi. Random walks in polytopes and negative dependence. *Innovations in Theoretical Computer Science (ITCS)*, 2017.

[41] Arnaud Poinas, Bernard Delyon, and Frédéric Lavancier. Mixing properties and central limit theorem for associated point processes. *arXiv preprint arXiv:1705.02276*, 2017.

[42] Ludger Rüschendorf. Comparison of multivariate risks and positive dependence. *Journal of Applied Probability*, 41(2):391–406, 2004.

[43] Alexander Soshnikov. Determinantal random point fields. *Russian Mathematical Surveys*, 55(5):923, 2000.

[44] Ryszard Szekli. *Stochastic Ordering and Dependence in Applied Probability*, volume 97 of *Lecture Notes in Statistics*. Springer-Verlag, 1995.

[45] David G Wagner. Negatively correlated random variables and mason's conjecture for independent sets in matroids. *Annals of Combinatorics*, 12(2):211–239, 2008.

[46] Ward Whitt. Uniform conditional variability ordering of probability distributions. *Journal of Applied Probability*, 22(3):619–633, 1985.

[47] Dhandapani Yogeshwaran. Negative association of point processes. *https://sites.google.com/site/yogeshacademics/home/publications*, 2013.