The Annals of Probability
1990, Vol. 18, No. 1, 359–377

POSITIVE DEPENDENCE PROPERTIES OF POINT PROCESSES

BY ROBERT M. BURTON, JR.[1] AND MARIE M. FRANZOSA

Oregon State University and Western Oregon State College

There are many ways of describing positive dependence, for example the strong FKG inequalities and association. It is known that for Bernoulli random variables the strong FKG inequalities are equivalent to all the conditional distributions being associated, which is in turn equivalent to all the conditional distributions having positively correlated marginals. These and similar definitions are extended to point processes on $\mathbb{R}^d$. Examples are given to show that, unlike the analogous Bernoulli random variable case, these conditions are no longer equivalent, although some are implied by others.

1. Introduction. In this paper we compare the various notions of positive dependence for point processes. Analogs of properties that are equivalent in the case of binary random variables turn out to be distinct in the point process case.

Section 2 describes the relationship between various positive dependence properties for 0–1 valued random variables. Such properties have been applied to many fields including reliability theory and statistical physics. In Section 3 point processes and their densities are defined. These densities are then used in Section 4 to define properties for point processes analogous to those defined in Section 2. In this section the theorems relating positive dependence properties for point processes are stated. Section 5 provides examples of point processes satisfying various positive dependence properties. In Section 6 the theorems of Section 4 are proven and counterexamples are given to show the limitations of these theorems.

2. Positive dependence for 0–1 valued random variables. We begin by considering positive dependence properties of Bernoulli random variables. Let P be a probability measure on $\Omega = \{0,1\}^N$ with the σ-algebra generated by points. Ω is a distributive lattice with the coordinatewise ordering $\alpha \wedge \beta = (\min(\alpha_1, \beta_1), \dots, \min(\alpha_N, \beta_N))$ and $\alpha \vee \beta = (\max(\alpha_1, \beta_1), \dots, \max(\alpha_N, \beta_N))$ for $\alpha, \beta \in \Omega$. Let $\mathbf{X} = (X_1, \dots, X_N)$ where $X_i \in \{0,1\}$ for each $i = 1, 2, \dots, N$, and X has distribution P. Positive dependence of such random variables can be described in many ways, several of which are given here. For more details as well as the proofs of statements given in this section the reader is referred to van den Berg and Burton [20].

DEFINITION 2.1. $\mathbf{X}$ satisfies the *strong FKG inequalities* if

$$P(\mathbf{X} = \alpha \wedge \beta)P(\mathbf{X} = \alpha \vee \beta) \geq P(\mathbf{X} = \alpha)P(\mathbf{X} = \beta) \quad \text{for all } \alpha, \beta \in \{0,1\}^N.$$

Received September 1987; revised December 1988.
[1]Supported in part by NSF Grant DMS-86-00021.
AMS 1980 *subject classifications*. Primary 60G55; secondary 60E15.
Key words and phrases. FKG, association, point process, product density.

DEFINITION 2.2. $\mathbf{X}$ is *associated* (or, $X_1, \ldots, X_N$ is an associated collection of random variables) if for all pairs of nondecreasing functions f and g, $\mathrm{Cov}(f(\mathbf{X}), g(\mathbf{X})) \geq 0$.

REMARK. The strong FKG inequalities are also known in the literature as the FKG lattice condition and the FKG property. That X satisfying the strong FKG inequalities implies X is associated is called the FKG theorem [7].

DEFINITION 2.3. $\mathbf{X}$ has *positively correlated increasing cylinder sets* ($\mathbf{X}$ has PCIC) if when I, K are disjoint subsets of $\{1, \ldots, N\}$ and A_I is the event that $X_i = 1$ for all $i \in I$, then $P(A_I \cap A_K) \geq P(A_I)P(A_K)$.

DEFINITION 2.4. $\mathbf{X}$ is *positively correlated* ($\mathbf{X}$ is PC) if

$$P(X_i = 1, X_j = 1) \geq P(X_i = 1)P(X_j = 1) \quad \text{for all } i, j \in \{1, \ldots, N\}.$$

Each of the above properties is implied by the preceding properties.

DEFINITION 2.5. $\mathbf{X}$ is *conditionally associated* if for each $J \subseteq \{1, \ldots, N\}$ and $\alpha_0 \in \{0,1\}^N$, $(\mathbf{X}|X_j = (\alpha_0)_j = (\alpha_0)_j \ \forall\ j \in J)$ is associated, provided that $P[X_j = (\alpha_0)_j \ \forall\ j \in J] > 0$.

DEFINITION 2.6. $\mathbf{X}$ has *conditionally positively correlated increasing cylinder sets* ($\mathbf{X}$ has CPCIC) when if I, J, K are all disjoint subsets of $\{1, \ldots, N\}$, $\alpha_0 \in \{0,1\}^N$ and if A_I is the event that $X_i = 1$ for all $i \in I$, then

$$P(A_I \cap A_K | X_j = (\alpha_0)_j \ \forall\, j \in J) \geq P(A_I | X_j = (\alpha_0)_j \ \forall\, j \in J) P(A_K | X_j = (\alpha_0)_j \ \forall\, j \in J),$$

provided that the event conditioned upon has positive probability.

DEFINITION 2.7. $\mathbf{X}$ is *conditionally positively correlated* ($\mathbf{X}$ is CPC) if for each $K \subset \{1, \ldots, N\}$ and $\alpha_0 \in \{0,1\}^N$,

$$P(X_i = 1, X_j = 1 | X_k = (\alpha_0)_k \ \forall\ k \in K) \geq P(X_i = 1 | X_k = (\alpha_0)_k \ \forall\ k \in K) P(X_j = 1 | X_k (\alpha_0)_k \ \forall\ k \in K)$$

for all $i, j \in \{1, \ldots, N\}$.

THEOREM 2.8 (van den Berg and Burton [20]; see also Kemperman [12]). *The following are equivalent*:

(a) $\mathbf{X}$ *satisfies the strong FKG inequalities.*
(b) $\mathbf{X}$ *is conditionally associated.*
(c) $\mathbf{X}$ *has conditionally positively correlated increasing cylinder sets* (*CPCIC*).

Furthermore, if P assigns positive probability to all configurations, then the above are also equivalent to:

(d) $\mathbf{X}$ *is conditionally positively correlated* (*CPC*).

For further information about related inequalities the reader can see [3, 4, 10, 12, 16]. For applications to probability and statistics see [3, 4, 15, 16, 19].

3. Preliminaries for the point processes. Let $\mathbb{R}^d$ be d-dimensional Euclidean space and $D \subseteq \mathbb{R}^d$ be a fixed, possibly infinite subrectangle. Let $\mathscr{B}^d$ be the collection of Borel subsets of D. Denote the subset of $\mathscr{B}^d$ consisting of bounded sets by $\hat{\mathscr{B}}^d$. Let N denote the set of all Radon counting measures on $(D, \mathscr{B}^d)$. Thus, $\mu \in N$ if and only if $\mu(B) \in \mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ and $\mu(B) < \infty$ for all $B \in \hat{\mathscr{B}}^d$. N is naturally identified with the set of all finite or infinite configurations of points (including multiplicities) in D without limit points.

Let $\mathscr{N}$ be the σ-algebra on N generated by sets of the form $\{\mu \in N | \mu(A) = k\}$ for all $A \in \hat{\mathscr{B}}^d$ and $0 \leq k < \infty$. $\mathscr{N}$ is the σ-algebra on N which allows us to count the points in bounded regions of D. N is a Polish space with the vague topology and its class of Borel sets is $\mathscr{N}$.

DEFINITION 3.1. A *point process* is a measurable mapping X from a probability space $(\Omega, \mathscr{F}, P)$ into $(N, \mathscr{N})$. The distribution of X is the induced measure on $(N, \mathscr{N})$ given by $P_X = P \circ X^{-1}$. If $A \in \mathscr{B}^d$, then we set $X(A)$ equal to the (random) number of occurrences in A and if f is a function with compact support, then we set $X(f)$ equal to the integral of f with respect to the random measure X.

For details on point processes the reader should consult Matthes, Kerstan and Mecke [17] and Kallenberg [9].

We now wish to define densities for point processes. That is, we want functions $p_n(x_1, \ldots, x_n)$ for each $n = 1, 2, \ldots$ such that $p_n(x_1 \cdots x_n)$ $|\Delta x_1| \cdots |\Delta x_n|$ approximates the probability of points occurring in the d-dimensional intervals $\Delta x_1 \cdots \Delta x_n$ about $x_1 \cdots x_n$ when the Lebesgue measure $|\Delta x_i|$ is sufficiently small. The nth-order moment measure M_n is defined by $M_n(A_1 \times \cdots \times A_n) = E[X(A_1) \cdots X(A_n)]$ for disjoint A_i. The nth-order factorial moment measure $M_{[n]}$ is defined by $M_{[n]}(A_1^{t_1} \times \cdots \times A_k^{t_k}) = E[X(A_1)^{[t_1]} \cdots X(A_k)^{[t_k]}]$, where $t_1 + \cdots + t_k = n$ and $s^{[t]} = s(s-1) \cdots (s-t+1)$. We will drop the subscript n or $[n]$ whenever it is clear which measure is meant. If the point process has a.s. no multiple occurrences and if the factorial moment measures are absolutely continuous with respect to Lebesgue measure, then their Radon–Nikodym derivatives are defined to be $p_n(x_1 \cdots x_n)$. These functions will be referred to as the *product densities* of X. We will usually write this as $p(x_1 \cdots x_n)$, dropping the subscript.

DEFINITION 3.2. Let S be the set of all real valued, bounded, measurable functions ξ on D satisfying:

(i) $0 \leq \xi(x) \leq 1$ for all $x \in D$.
(ii) $\xi(x) = 1$ on the complement of a bounded subset of D.

Then the *probability generating functional* corresponding to X is defined by

$$G(\xi) = E\left\{\exp\left(\int_{\mathbb{R}^d} \log \xi(x)\, dX(x)\right)\right\}, \qquad \xi \in S. \tag{3.1}$$

The factorial moment measures can be computed using the probability generating functional according to the following formula due to Moyal [18]:

$$\begin{aligned} M_{x_1,\dots,x_n} &\equiv M([0,x_1] \times \cdots \times [0,x_n]) \\ &= \lim_{\eta \uparrow 1} \left\{ \frac{\partial}{\partial \lambda_1 \cdots \partial \lambda_n} G\left(1 + \sum_{i=1}^{n} \lambda_i 1_{[a,x_i]}\right)\right\}_{\lambda_1 = \cdots = \lambda_n = 1}, \end{aligned} \tag{3.2}$$

where $1_{[0,x_i]}$ denotes the indicator function of the rectangle with diagonal corners the origin and x_i.

To calculate the product densities from this formula, we differentiate $M_{x_1,\dots,x_n}$, that is,

$$p(x_1,\dots,x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n}(M_{x_1,\dots,x_n}). \tag{3.3}$$

Given a compact subset of A of D and disjoint Borel subsets $A_1 \cdots A_n$ of A, let $R_A^n(A_1 \cdots A_n)$ be the probability of exactly one point occurring in each of the sets $A_1 \cdots A_n$ and no other points occurring in A. Assume that all the R_A^n are absolutely continuous with respect to Lebesgue measure on $A^n = A \times \cdots \times A$ and that a.s. X has no multiple occurrences. The *absolute product densities* $r_A^n(x_1,\dots,x_n)$ are defined to be $1/n!$ times the Radon–Nikodym derivative of R_A^n. $r_A^n(x_1,\dots,x_n)|\Delta x_1| \cdots |\Delta x_n|$ has the interpretation as an approximation to the probability that X has exactly n occurrences in the set A and that in each region Δx_i about x_i there is exactly one occurrence. Again we will usually drop the superscript and write r_A. Under very general conditions we have the following relations between these density functions which hold a.e.:

$$p(x_1,\dots,x_n) = \sum_{j=0}^{\infty} \frac{1}{j!} \int_{A^j} r_A(x_1,\dots,x_n,\theta_1,\dots,\theta_j)\, d\theta_1 \cdots d\theta_j \tag{3.4}$$

and

$$r_A(x_1,\dots,x_n) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{A^j} p(x_1,\dots,x_n,\theta_1,\dots,\theta_j)\, d\theta_1 \cdots d\theta_j. \tag{3.5}$$

The above results are given in Fisher [6], Macchi [14] and Moyal [18].

In order to consider conditional distributions we define conditional product densities the natural way.

DEFINITION 3.3. A point process X with product densities $p(x_1,\dots,x_n)$ has *conditional product densities* given by

$$p(x_1,\dots,x_n|y_1,\dots,y_m) = \frac{p(x_1,\dots,x_n,y_1,\dots,y_m)}{p(y_1,\dots,y_m)}$$

and its *conditional absolute product densities* are given by

$$r_A(x_1,\dots,x_n|y_1,\dots,y_m) = \frac{r_A(x_1,\dots,x_n,y_1,\dots,y_m)}{p(y_1,\dots,y_m)}$$

for $x_1,\dots,x_n,y_1,\dots,y_m \in A$.

The densities $p(x_1,\dots,x_n|y_1,\dots,y_m)$ correspond to the product densities of the reduced Palm measure of X (see Karr [11]).

If X has product densities one can also define *cumulant densities* $q(x_1,\dots,x_n)$ (also called the correlation function) corresponding to X inductively by the following relationship with $p(x_1,\dots,x_n)$:

$$\begin{aligned} p(x_1) &= q(x_1), \\ p(x_1,x_2) &= q(x_1,x_2) + q(x_1)q(x_2), \\ p(x_1,x_2,x_3) &= q(x_1,x_2,x_3) + q(x_1,x_2)q(x_3) + q(x_2,x_3)q(x_1) \\ &\quad + q(x_1,x_3)q(x_2) + q(x_1)q(x_2)q(x_3) \end{aligned} \tag{3.6}$$

and so on, so that $p(x_1,\dots,x_n)$ is written in terms of q by subdividing $(x_1,\dots,x_n)$ into all possible configurations of disjoint subsets and adding the corresponding product of q's.

4. Positive dependence properties of point processes. In this section, the definitions of positive dependence for Bernoulli random variables are extended to the point process case. The proofs of the theorems given are presented in Section 6. The first definition we give was originally stated by Burton and Waymire [1].

DEFINITION 4.1. X satisfies the *strong FKG inequalities* if for all sets A in $\hat{\mathscr{B}}^d$ there exists a version of the absolute product densities such that

$$r_A(x_1,\dots,x_n)r_A(x_i,\dots,x_j) \ge r_A(x_1,\dots,x_j)r_A(x_i,\dots,x_n) \tag{4.1}$$

for all $x_1,\dots,x_n \in A$, $1 \le i \le j \le n$.

There is the following restriction that the strong FKG inequalities put on possible configurations.

THEOREM 4.2. *Suppose that X satisfies the strong FKG inequalities. Then there is a subset $A \subset D$ so that $P[X(A) = 0] = 1$ and all finite configurations are possible on $D - A$, in the sense that if $B \subset D - A$ is a bounded Borel set,* *then there is a version of $r_B(x_1, \dots, x_n)$ that is strictly positive for all distinct $x_1, \dots, x_n$ in B.*

DEFINITION 4.3. X has *positively correlated increasing cylinder sets* (X has PCIC) if

$$p(x_1, \dots, x_n) \ge p(x_1, \dots, x_i)p(x_{i+1}, \dots, x_n). \tag{4.2}$$

DEFINITION 4.4. X has *conditionally positively correlated increasing cylinder sets* (X has CPCIC) if

$$p(x_1, \dots, x_n)p(x_i, \dots, x_j) \ge p(x_1, \dots, x_j)p(x_i, \dots, x_n) \tag{4.3}$$

for $1 \le i \le j \le n$.

REMARK. We define X to be *positively correlated* if $p(x, y) \ge p(x)p(y)$ for all $x, y \in A$. Similarly, X is *conditionally positively correlated* if for all $z_1, \dots, z_n \in A$ for which $p(z_1, \dots, z_n) > 0$ we have

$$p(x, y|z_1, \dots, z_n) \ge p(x|z_1, \dots, z_n)p(y|z_1, \dots, z_n).$$

If $p(z_1, \dots, z_n) \ge 0$ for all $z_1, \dots, z_n \in A$, then conditionally positively correlated is equivalent to CPCIC.

DEFINITION 4.5. X is *associated* if $\mathrm{Cov}(F(X), G(X)) \ge 0$ for all pairs of functions F, G: $N \to \mathbb{R}$ that are nondecreasing, measurable and bounded [where nondecreasing means nondecreasing with respect to the ordering on N given by $\mu \le \nu$ if $\mu(B) \le \nu(B)$ for all $B \in \mathscr{B}^d$]. X is *conditionally associated* (CA) if P_X-almost surely for $z_1, \dots, z_n$, the point process given by conditioning X on occurrences at $z_1, \dots, z_n$, is associated. If X has absolute product densities this means that for all compact $A \subseteq \mathbb{R}$, for all $z_1, \dots, z_n \in A$, the point process on A with absolute product densities $r_A(\cdot|z_1, \dots, z_n)$ is associated.

Burton and Waymire [1, 2] showed that Definition 4.4 is equivalent to the family of random variables $\{X(B)|B \in \hat{\mathscr{B}}^d\}$ being associated. That is, all finite subsets of $\{X(B)|B \in \hat{\mathscr{B}}^d\}$ are associated in the sense of Definition 2.2.

To see how these definitions are a natural extension of the definitions of positive dependence for Bernoulli random variables, consider the following method of approximating a points process X on the interval $[0, 1]$ whose densities are well defined. For each n and $k = 1, \dots, n$ define

$$X_k^{(n)} = \begin{cases} 1 & \text{if there is an occurrence of } X \text{ in } \left[\dfrac{k-1}{n}, \dfrac{k}{n}\right], \\ 0 & \text{if there is no occurrence of } X \text{ in } \left[\dfrac{k-1}{n}, \dfrac{k}{n}\right]. \end{cases} \tag{4.4}$$

That is, $X_k^{(n)} = \min(1, X[(k-1)/n, k/n])$ and each $X_k^{(n)}$ is a Bernoulli random variable. $\mathbf{X}^{(n)} = (X_1^{(n)}, \dots, X_k^{(n)})$ approximates X and $\mathbf{X}^{(n)}$ converges to X in distribution. Thus, for example, FKG for a point process X can be thought of as a limiting condition of FKG for the random vectors $\mathbf{X}^{(n)}$. In the limit we indicate just where the 1's are the 1's indicating the occurrence of points). Burton and Waymire used a similar approximation technique for point processes on $\mathbb{R}^d$ to prove the following theorem.

THEOREM 4.6 (Burton and Waymire [1]). *If a point process X with piecewise continuous absolute product densities satisfies the strong FKG inequalities, then X is associated.*

We must be careful with approximations such as the one suggested above. Based on the comparisons of the definitions one might expect a theorem for point processes analogous to Theorem 2.8. In particular, it seems reasonable to expect that strong FKG and CPCIC are equivalent. They are not, however, and the relationship between the positive dependence definitions becomes a bit more complicated.

THEOREM 4.7. *If X has piecewise continuous absolute product densities, then the following implications, and no others, hold:*

$$\swarrow \quad \textit{X satisfies the strong FKG inequalities} \quad \searrow$$

$$\textit{X is CPCIC} \qquad\qquad\qquad\qquad\qquad\qquad \textit{X is associated}$$

$$\searrow \qquad \textit{X has PCIC} \qquad \swarrow$$

REMARK. A conditional version of the FKG theorem shows that X satisfying the strong FKG inequalities implies X is conditionally associated. It is also true that X being conditionally associated implies both that X is associated and that X has CPCIC. It is an open problem to determine if X being conditionally associated is equivalent to X satisfying the strong FKG inequalities.

In the lattice case, Theorem 2.8 has been useful (see, for example, [13]). In a sense the main application of Theorem 4.7 is negative. CPCIC is a much easier condition to check in practice than strong FKG and it would have been nice if these conditions were equivalent. It is unknown whether an additional, easily verifiable assumption can be found making these equivalent. It is to be expected that CPCIC will be helpful in the future.

The cumulant densities defined in Section 3 also play a role in describing positive dependence properties.

THEOREM 4.8. *If a point process X has cumulant densities which are always nonnegative, then X has PCIC, but not conversely.*

5. Examples of point processes.

EXAMPLE 5.1. The most fundamental point process is the Poisson point process. Given a Radon measure Λ on $\mathbb{R}^d$, a *Poisson point process* X with intensity Λ is a point process with independent increments such that $X(B)$ is a Poisson random variable with parameter $\Lambda(B)$. When the measure Λ is taken to be a multiple of Lebesgue measure, i.e., $\Lambda(A) = \lambda|A|$ for some $0 < \lambda < \infty$, we obtain a *stationary Poisson process.*

The probability generating functional of a Poisson point process has the form

$$G(\xi) = \exp\left\{\int [\xi(x) - 1]\, d\Lambda(x)\right\}. \tag{5.1}$$

If Λ is absolutely continuous with respect to Lebesgue measure with density function $f(x)$, then the product densities of X are $p(x_1, \dots, x_n) = f(x_1) \cdots f(x_n)$ and the absolute product densities are $r_A(x_1, \dots, x_n) = f(x_1) \cdots f(x_n)\exp(-\int_A f(y)\, dy)$. In particular, X is strong FKG.

EXAMPLE 5.2 (Mixed Poisson process). Here X is conditionally stationary Poisson with random intensity I. We assume that all moments of I are finite. The absolute product densities of X are $r_B(x_1, \dots, x_n) = E[I^n e^{-I|B|}]$. The strong FKG inequalities hold because the moments of a nonnegative random variable are log-convex which follows from Hölder's inequality (see, e.g., Feller [5]).

EXAMPLE 5.3 (Mixed sample processes). This class of distributions is characterized by the property that the distributions are invariant under measure preserving transformations of D. The absolute product densities are independent of location, that is, there are functions $f_B(n)$ so that $r_B(x_1, \dots, x_n) = f_B(n)$. This means that, conditioned on the number of point occurrences in B, the points are distributed uniformly and independently. These models are analyzed in Kallenberg [8]. We shall see that such processes need not satisfy the strong FKG inequalities even if they have CPCIC.

EXAMPLE 5.4 (Poisson center cluster processes). Let U be a stationary Poisson point process on $\mathbb{R}^d$ with intensity ξ and let V be a point process so that $E[V(\mathbb{R}^d)] < \infty$. Suppose that the (random) occurrences of U are $\{u_i\}$ and that $V_1, V_2, \dots$ are iid and distributed as V. If X is defined by $X(A) = \sum V_i(A - u_i)$, then X is a well-defined point process (see Westcott [21]) and is said to be distributed as a Poisson center cluster process with centers U and clusters V. It is known (Burton and Waymire [1]) that X is associated, although it will be shown that such processes need not satisfy the strong FKG inequalities. The probability generating functional of X is given by

$$G(\xi) = \exp\left\{\int_D [G_V(T_x\xi) - 1]\, \xi\, dx\right\}, \tag{5.2}$$

where T_x is the translation operator, $(T_x\xi)(y) = \xi(x + y)$ and G_V is the probability generating functional of V.

6. Proofs of theorems.

PROOF OF THEOREM 4.2. We may, of course, assume that all the product densities and absolute product densities are Borel measurable. The relation

$$p(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_{B^n} r_B(x, \mathbf{y})\, d\mathbf{y} \tag{6.1}$$

of course holds only a.e. on B. If (6.1) does not hold for x, redefine $r_B(x, \mathbf{y}) = 0$ for all $\mathbf{y} \in B^n$ and all n. The strong FKG inequalities still hold and we may redefine $p(x) = 0$ so that (6.1) holds everywhere. Furthermore, this will change p only on a set of measure 0 (even as we vary B). This is because the expression (6.1) is essentially independent of B.

Now let $A = \{x \in D | p(x) = 0\}$ and let B be a bounded Borel subset. We show that if $(x_1, \dots, x_m) \in B^m$ has distinct coordinates, then $r_B(x_1, \dots, x_m) > 0$. Suppose otherwise. Repeated applications of the strong FKG inequalities give

$$0 = r_B(x_1, \dots, x_m) r_B(\varnothing)^{m-1} \geq r_B(x_1) r_B(x_2) \cdots r_B(x_m).$$

This means that there is an x_i so that $r_B(x_i) = 0$. We rename $x_i = x$. If $\mathbf{y}, \mathbf{z} \in B^n$ have distinct coordinates and if for each $i = 1, \dots, n$, $x \neq y_i \neq z_i \neq x$, then $0 = r_B(x, \mathbf{y}, \mathbf{z}) r_B(x) \geq r_B(x, \mathbf{y}) r_B(x, \mathbf{z})$. Thus at most one of $r_B(x, \mathbf{y})$ and $r_B(x, \mathbf{z})$ can be strictly positive. This implies $r_B(x, \mathbf{y}) = 0$ for all but at most one $\mathbf{y}$ on B^n. But in view of (6.1) this means that $p(x) = 0$ so $x \in A$, a contradiction. □

PROOF OF THEOREM 4.7. (a) That X satisfies the strong FKG inequalities implies all conditional distributions of X are associated by Theorem 4.6 (as noted previously).

(b) That X has CPCIC implies X has PCIC is immediate.

(c) We show that if X satisfies the strong FKG inequalities, then X has CPCIC. Let $\mathbf{x} = (x_1, \dots, x_n)$ and we suppose with no loss of generality that each x_i is not in $D - A$, that is, $p(x_i) > 0$. In fact, all the points discussed in this part of the proof are assumed to be in $D - A$. Let $r(\mathbf{x}|\mathbf{y})$ be the conditional absolute product density.

Define $\Phi_{\mathbf{z}}$ by $\Phi_{\mathbf{z}}(\mathbf{x}|\mathbf{y}) = r(\mathbf{x}, \mathbf{z}|\mathbf{y})/r(\mathbf{x}|\mathbf{y})$, where $\mathbf{z} = (z_1, \dots, z_j)$ [so that $(\mathbf{x}, \mathbf{z}) = (x_1, \dots, x_n, z_1, \dots, z_j)$]. Note that $\Phi_{\mathbf{z}}$ is an increasing function of $\mathbf{x}$ for fixed $\mathbf{y}$ since X satisfying the strong FKG inequalities implies $r(\mathbf{x}, \mathbf{y}, \mathbf{z}) r(\mathbf{y}) \geq r(\mathbf{x}, \mathbf{y}) r(\mathbf{y}, \mathbf{z}) \Rightarrow r(\mathbf{x}, \mathbf{z}|\mathbf{y}) r(\mathbf{y}|\mathbf{y}) \geq r(\mathbf{x}|\mathbf{y}) r(\mathbf{z}|\mathbf{y}) \Rightarrow r(\mathbf{x}, \mathbf{z}|\mathbf{y})/r(\mathbf{x}|\mathbf{y}) \geq r(\mathbf{z}|\mathbf{y})/r(\mathbf{y}|\mathbf{y})$. Then

$$\begin{aligned} E[\Phi_{\mathbf{z}}] &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int \Phi_{\mathbf{z}}(\mathbf{x}|\mathbf{y}) r(\mathbf{x}|\mathbf{y})\, d\mathbf{x} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int r(\mathbf{x}, \mathbf{z}|\mathbf{y})\, d\mathbf{x} = p(\mathbf{z}|\mathbf{y}) \end{aligned} \tag{6.2}$$

and

$$\begin{aligned} E[\Phi_{\mathbf{z}}\Phi_{\mathbf{w}}] &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int \Phi_{\mathbf{z}}(\mathbf{x}|\mathbf{y})\Phi_{\mathbf{w}}(\mathbf{x}|\mathbf{y}) r(\mathbf{x}|\mathbf{y})\, d\mathbf{x} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int \frac{r(\mathbf{x},\mathbf{z}|\mathbf{y}) r(\mathbf{x},\mathbf{w}|\mathbf{y})}{r(\mathbf{x}|\mathbf{y})}\, d\mathbf{x} \qquad (6.3) \\ &\le \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int r(\mathbf{x},\mathbf{z},\mathbf{w}|\mathbf{y})\, d\mathbf{x} \quad \text{(by the strong FKG inequalities)} \\ &= p(\mathbf{z},\mathbf{w}|\mathbf{y}). \end{aligned}$$

Since X satisfies the strong FKG inequalities its conditional distributions are associated. For fixed $\mathbf{y}$, $\Phi_{\mathbf{z}}$ and $\Phi_{\mathbf{w}}$ are increasing functions on N so $\operatorname{Cov}(\Phi_{\mathbf{z}}, \Phi_{\mathbf{w}}) \ge 0$, i.e., $E[\Phi_{\mathbf{z}}\Phi_{\mathbf{w}}] \ge E[\Phi_{\mathbf{z}}]E[\Phi_{\mathbf{w}}]$ or $p(\mathbf{z},\mathbf{w}|\mathbf{y}) \ge p(\mathbf{z}|\mathbf{y})p(\mathbf{w}|\mathbf{y})$ which is CPCIC.

(d) Now we show that if X is associated, then X has PCIC. If x is a point in the interior of D and Δ is a real number, then let $[x, \Delta]$ be the rectangle $\{z \in D | x_i \le z_i \le x_i + \Delta\}$ and note that $X([x, \Delta])$ is nondecreasing. For each $x_1, \dots, x_n, y_1, \dots, y_m$ by association we have that

$$\begin{aligned} &E[X([x_1, \Delta]) \cdots X([x_n, \Delta]) X([y_1, \Delta]) X([y_m, \Delta])] \\ &\quad - E[X([x_1, \Delta]) \dots X([x_n, \Delta])] E[X([y_1, \Delta]) X([y_m, \Delta])] \ge 0. \end{aligned}$$

Dividing by Δ^{n+m} and taking the limit as $\Delta \to 0$ gives

$$\begin{aligned} 0 &\le \frac{\partial^n}{\partial x_1 \cdots \partial x_n} \frac{\partial^m}{\partial y_1 \cdots \partial y_m} (M_{x_1, \dots, x_n, y_1, \dots, y_m}) \\ &\quad - \frac{\partial^n}{\partial x_1 \cdots \partial x_n} (M_{x_1, \dots, x_n}) \frac{\partial^m}{\partial y_1 \cdots \partial y_m} (M_{y_1, \dots, y_m}) \\ &= p(x_1, \dots, x_m, y_1, \dots, y_m) - p(x_1, \dots, x_m) p(y_1, \dots, y_m) \end{aligned}$$

as in Moyal's formula (3.3).

To complete the proof, we will find two examples, one of which has CPCIC but is not associated and the other which is associated but does not have CPCIC.

EXAMPLE 6.1. In this example, we construct a point process X on $D = [0, b] \subset \mathbb{R}$ that has CPCIC but is not associated. X will be a mixed sample process. That is, we will take Y to be a random variable with values in $\{0, 1, \dots\}$. Then, conditioned on $Y = k$, we let $Z_1, \dots, Z_k$ be uniformly and independently distributed on $[0, b]$ as occurrences of X. Thus the distribution of X depends only on the number of points and not on their locations.

For the actual construction of the densities, first note that such an X is a point process with product densities p and absolute product densities

$r_B(x_1,\dots,x_k) = f_B(k)$. In this case,

$$\begin{aligned} g(k) = p(x_1,\dots,x_k) &= \sum_{n=0}^{\infty} \frac{1}{n!} \int \cdots \int_{B\times\cdots\times B} f_B(k+n)\, dy_1 \cdots dy_n \\ &= \sum_{n=0}^{\infty} \frac{b^n}{n!} f_B(k+n). \end{aligned} \tag{6.4}$$

Conversely, if $p(x_1,\dots,x_n) = g(k)$ we find that

$$f_B(k) = r_B(x_1,\dots,x_k) = \sum_{n=0}^{\infty} (-1)^n \frac{b^n}{n!} g(k+n). \tag{6.5}$$

We will define the distribution of this process by finding values of the function $g(k)$ and then checking that they give a well-defined distribution. A given function $g(k)$ will represent product densities with corresponding absolute product densities given by (6.5) if the following conditions are satisfied:

$$f_B(k) \ge 0 \quad \text{for all } k. \tag{6.6)(a}$$

$$\sum_{n=0}^{\infty} \frac{b^n}{n!} f_B(n) = 1. \tag{6.6)(b}$$

$$g(k) = \sum_{n=0}^{\infty} \frac{b^n}{n!} f_B(k+n). \tag{6.6)(c}$$

LEMMA 6.2. *If g satisfies*

(a) $$\sum_{n=0}^{\infty} (-1)^n \frac{b^n}{n!} g(k+n) \ge 0 \quad \textit{for } k = 0,1,2,\dots,$$

(b) $$g(0) = 1,$$

(c) $$g(k) \ge 0 \quad \textit{for } k = 0,1,2,\dots,$$

(d) $$\sum_{L=0}^{\infty} \frac{(2b)^L}{L!} g(k+L) < \infty \quad \textit{for } k = 0,1,2,\dots,$$

then g determines product densities with corresponding absolute product densities given by (6.5).

PROOF. (a) gives us condition (6.6)(a), and, given condition (6.6)(c), (b) gives condition (6.6)(b). Thus we must only show that condition (6.6)(c) is satisfied:

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{b^n}{n!} f_B(k+n) &= \sum_{n=0}^{\infty} \frac{b^n}{n!} \sum_{j=0}^{\infty} (-1)^j \frac{b^j}{j!} g(j+k+n) \\ &= \sum_{n=0}^{\infty} \sum_{j=0}^{\infty} (-1)^j \frac{b^n}{n!} \frac{b^j}{j!} g(j+k+n). \end{aligned} \tag{6.7}$$

By (d), the above sum is absolutely convergent so we may rearrange (6.7), by summing over the diagonals, to get

$$\sum_{n=0}^{\infty}\sum_{j=0}^{\infty}(-1)^j\frac{b^n}{n!}\frac{b^j}{j!}g(j+k+n)$$

$$(6.8)\qquad =\sum_{L=0}^{\infty}\sum_{n=0}^{L}\frac{(-1)^{L-n}}{L!}\frac{L!}{n!\,(L-n)!}b^L g(L+k)$$

$$=\sum_{L=0}^{\infty}b^L\frac{g(L+k)}{L!}\sum_{n=0}^{L}\binom{L}{n}(-1)^{L-n}.$$

By the binomial theorem, the terms of this sum are all 0 except when $L = 0$. Thus the series becomes

$$\frac{b^0}{0!}g(k+0)\cdot 1 = g(k). \qquad \square$$

Any function g which is nonnegative and bounded by an exponential will satisfy condition (6.6)(c). For this example we must also choose g so that product densities satisfy the inequalities given that X has CPCIC. That is, we need $p(x_1,\dots,x_n)p(x_i,\dots,x_j) \geq p(x_1,\dots,x_j)p(x_i,\dots,x_n)$ or, in terms of g, $g(n)g(j-i) \geq g(j)g(n-i)$ for all $n \geq j \geq i$. This log-convexity condition is true if and only if

$$(6.9)\qquad g(n+1)g(n-1) \geq [g(n)]^2.$$

Let $g(0) = 1$, $g(1) = \alpha$ and $g(k) = \alpha^2\beta^{k-2}$ for $k \geq 3$, where $\alpha < \beta$ are to be determined later. It is easy to check that g satisfies (6.9), so that X has CPCIC.

To see that X can have a well-defined distribution, we set $\alpha = 1$ and $\beta = 2$, then calculate the absolute product densities:

$$f_B(0) = \sum_{n=0}^{\infty}(-1)^n\frac{b^n}{n!}g(0+n) = \frac{3}{4} - \frac{1}{2}b + \frac{1}{4}e^{-2b},$$

$$f_B(1) = \sum_{n=0}^{\infty}(-1)^n\frac{b^n}{n!}g(1+n) = 1 + \frac{1}{2}[e^{-2b} - 1],$$

$$f_B(k) = \sum_{n=0}^{\infty}(-1)^n\frac{b^n}{n!}g(k+n) = \frac{1}{4}2^k e^{-2b} \quad \text{for } k \geq 2.$$

Note that $f_B(k) \geq 0$ for all values of b if $k \geq 1$. In order to be certain that these functions determine absolute product densities [with corresponding product densities $g(k)$], it remains to choose b so that $f_B(0) \geq 0$. By the intermediate value theorem we may choose b so that $f_B(0) = 0$. X with the values of α, β and b thus chosen is well defined. Finally, we show that X is not associated. Notice that $P(X([0, b]) = 0) = f_B(0) = 0$, but $P(X([0, b/2]) = 0) = P(X([b/2, b]) = 0) = \sum_{k=0}^{\infty}(1/k!)(b/2)^k f_B(k) > 0$ since $f_B(0) = 0$ and $f_B(k) \geq 0$ for all $k \geq 1$. The indicators of the events that $X([0, b/2]) = 0$ and that

$X([b/2, b]) = 0$ are decreasing events and so would be positively correlated if X were associated, but the above shows that these events are negatively correlated. Thus this example has CPCIC but is not associated.

EXAMPLE 6.3. This example shows that X being associated does not imply that X has CPCIC. Consider a Poisson center cluster process on $\mathbb{R}$, where the Poisson process of cluster centers is stationary with intensity $\lambda = 1$. Assume also that there are exactly two points distributed at each cluster center according to a distance distribution F which has density f, continuous with the exception of a finite number of jump discontinuities. This process was shown to be associated by Burton and Waymire [1] but can be adjusted (by choosing an appropriate distance distribution) so that if points x_1, x_2 and x_3 are chosen properly the inequality $p(x_1, x_2, x_3)p(x_2) < p(x_1, x_2)p(x_2, x_3)$ holds, i.e., the process does not have CPCIC. This may be expected due to the fact that in this process one is more likely to observe points occurring in pairs than in triples or singles. Thus we would expect that both $p(x_1, x_2)$ and $p(x_2, x_3)$ are likely to be larger than $p(x_1, x_2, x_3)$.

In order to show that a process does not have CPCIC, we must first calculate some of the product densities. $p(x_1) = 2$, the intensity of the process. We calculate $p(x_1, x_2)$, $p(x_2, x_3)$ and $p(x_1, x_2, x_3)$ by making use of Moyal's formula [(3.2) and (3.3)]. In this case, from Example 5.4 of the previous section, the probability generating functional is given by

$$G(\xi) = \exp\int\left[\left(\int\xi(x + r)f(r)\,dr\right)^2 - 1\right]dx. \tag{6.10}$$

Thus

$$M_{x_1,\dots,x_k}$$

$$= \lim_{\eta\uparrow 1}\left\{\frac{\partial^k}{\partial x_1\cdots\partial x_k}\right. \tag{6.11}$$

$$\left.\times\exp\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\left[1 + \sum_{i=1}^{k}\lambda_i 1_{[0,x_i]}(x + r)\right]f(r)\,dr\right) - 1\,dx\right\}_{\lambda_1 = \cdots = \lambda_k = 0}.$$

Let

$$h_k(\lambda_1,\dots,\lambda_k) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\left[1 + \sum_{i=1}^{k}\lambda_i 1_{[0,x_i]}(x + r)\right]f(r)\,dr\right) - 1\,dx \tag{6.12}$$

and

$$\hat{G}_k(\lambda_1,\dots,\lambda_k) = \exp(h_k(\lambda_1,\dots,\lambda_k)). \tag{6.13}$$

Then

$$M_{x_1,\dots,x_k} = \frac{\partial^k}{\partial x_1\cdots\partial x_k}\{\hat{G}_k(\lambda_1,\dots,\lambda_k)\}_{\lambda_1 = \cdots = \lambda_k = 0}. \tag{6.14}$$

We use this to calculate $p(x_1, x_2)$:

$$M_{x_1, x_2} = \frac{\partial^2}{\partial\lambda_1\,\partial\lambda_2}\hat{G}_2(\lambda_1, \lambda_2)\bigg|_{\lambda_1=\lambda_2=0}, \tag{6.15}$$

$$\begin{aligned}\frac{\partial^2}{\partial\lambda_1\,\partial\lambda_2}\hat{G}_2(\lambda_1, \lambda_2) &= \exp(h_2(\lambda_1, \lambda_2))\frac{\partial^2 h_2}{\partial\lambda_1\,\partial\lambda_2} + \frac{\partial h_2}{\partial\lambda_1}\frac{\partial h_2}{\partial\lambda_2}\exp(h_2(\lambda_1, \lambda_2))\\ &= \exp(h_2(\lambda_1, \lambda_2))\left(\frac{\partial^2 h_2}{\partial\lambda_1\,\partial\lambda_2} + \frac{\partial h_2}{\partial\lambda_1}\frac{\partial h_2}{\partial\lambda_2}\right).\end{aligned} \tag{6.16}$$

Let

$$I_i = \int_{-\infty}^{\infty} 1_{[0, x_i]}(x + r) f(r)\, dr$$

but

$$\begin{aligned}\frac{\partial^2 h_2}{\partial\lambda_1\,\partial\lambda_2} &= \frac{\partial}{\partial\lambda_2}\left(\int_{-\infty}^{\infty} 2I_1\left(\int_{-\infty}^{\infty}\left[1 + \sum_{i=1}^{2}\lambda_i 1_{[0, x_i]}(x + r)\right] f(r)\, dr\right) dx\right)\\ &= \int_{-\infty}^{\infty} 2I_1\left(\int_{-\infty}^{\infty} 1_{[0, x_2]}(x + r) f(r)\, dr\right) dx = \int_{-\infty}^{\infty} 2I_1 I_2\, dx.\end{aligned} \tag{6.17}$$

Substituting into (6.16), we find that

$$\begin{aligned}&\frac{\partial^2}{\partial\lambda_1\,\partial\lambda_2}\hat{G}_2(\lambda_1, \lambda_2)\\ &\quad = \exp(h_2(\lambda_1, \lambda_2))\left(\int_{-\infty}^{\infty} 2I_1 I_2\, dx\right.\\ &\qquad + \left(\int_{-\infty}^{\infty} 2I_1\left(\int_{-\infty}^{\infty}\left[1 + \sum_{i=1}^{2}\lambda_i 1_{[0, x_i]}(x + r)\right] f(r)\, dr\right) dx\right)\\ &\qquad \left.\times\left(\int_{-\infty}^{\infty} 2I_1\left(\int_{-\infty}^{\infty}\left[1 + \sum_{i=1}^{2}\lambda_i 1_{[0, x_i]}(x + r)\right] f(r)\, dr\right) dx\right)\right).\end{aligned} \tag{6.18}$$

Evaluating (6.18) at $\lambda_1 = \lambda_2 = 0$, we obtain

$$\begin{aligned}M_{x_1, x_2} &= \int_{-\infty}^{\infty} 2I_1 I_2\, dx + \left(\int_{-\infty}^{\infty} 2I_1\, dx\right)\left(\int_{-\infty}^{\infty} 2I_2\, dx\right)\\ &= 4\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{[0, x_1]}(x + r) f(r)\, dr\, dx\right)\\ &\quad \times\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{[0, x_1]}(x + r) f(r)\, dr\, dx\right)\\ &\quad + 2\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} 1_{[0, x_1]}(x + r) f(r)\, dr\right)\\ &\quad \times\left(\int_{-\infty}^{\infty} 1_{[0, x_1]}(x + r) f(r)\, dr\right) dx.\end{aligned} \tag{6.19}$$

Differentiating, first with respect to x_1 and then with respect to x_2, we obtain $p(x_1, x_2)$. First, we change the order of integration in the integrals of the first term of (6.19):

$$M_{x_1,x_2} = 4\left(\int_{-\infty}^{\infty}\int_{r}^{x_1-r} f(r)\,dx\,dr\right)\left(\int_{-\infty}^{\infty}\int_{r}^{x_2-r} f(r)\,dx\,dr\right) \tag{6.20}$$
$$+ 2\int_{-\infty}^{\infty}\left(\int_{x}^{x_1-x} f(r)\,dr\right)\left(\int_{x}^{x_2-x} f(r)\,dr\right)dx$$
$$= 4x_1x_2 + 2\int_{-\infty}^{\infty}\left(\int_{x}^{x_1-x} f(r)\,dr\right)\left(\int_{x}^{x_2-x} f(r)\,dr\right)dx.$$

Differentiating (6.20) with respect to x_1 and x_2, we get

$$p(x_1, x_2) = 4 + 2\int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\,dx. \tag{6.21}$$

A similar calculation yields

$$p(x_1, x_2, x_3) = 32 + 4\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\,dx \tag{6.22}$$
$$+ 4\int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\,dx$$
$$+ 4\int_{-\infty}^{\infty} f(x_1 - x) f(x_3 - x)\,dx.$$

Having calculated $p(x_1)$, $p(x_1, x_2)$ and $p(x_1, x_2, x_3)$, we are able to show that X does not have CPCIC by making an appropriate choice of f so that

$$p(x_1, x_2, x_3)p(x_2) < p(x_1, x_2)p(x_2, x_3). \tag{6.23}$$

The left-hand side of (6.23) is

$$p(x_1, x_2, x_3)p(x_2) = 64 + 8\left(\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\,dx \right. \tag{6.24}$$
$$+ \int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\,dx$$
$$\left. + \int_{-\infty}^{\infty} f(x_1 - x) f(x_3 - x)\,dx\right),$$

while on the right-hand side we have

$$p(x_1, x_2)p(x_2, x_3) = 16 + 8\left(\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\,dx \right. \tag{6.25}$$
$$\left. + \int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\,dx\right)$$
$$+ 4\left(\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\,dx\right)$$
$$\times\left(\int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\,dx\right).$$

Cancelling like terms above, we are reduced to showing that

$$\begin{aligned}(6.26)\qquad & 12 + 2\int_{-\infty}^{\infty} f(x_1 - x) f(x_3 - x)\, dx \\ & < \left(\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\, dx\right)\left(\int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\, dx\right).\end{aligned}$$

We now determine an appropriate choice of f. For a fixed value of n let

$$(6.27)\qquad f_n(x) = \begin{cases} n & \text{if } 0 \le x \le n, \\ 0 & \text{otherwise,} \end{cases}$$

and let $x_1 = 0$, $x_2 = 1/2n$ and $x_3 = 1/n$. Then $f_n(x_1 - x) f_n(x_3 - x) = 0$. For $f = f_n(\cdot)$, (6.27) becomes

$$\begin{aligned} 12 &< \left(\int_{-\infty}^{\infty} f(x_2 - x) f(x_3 - x)\, dx\right)\left(\int_{-\infty}^{\infty} f(x_1 - x) f(x_2 - x)\, dx\right) \\ &= \left(\int_{0}^{1/2n} n \cdot n\, dx\right)\left(\int_{1/2n}^{1/n} n \cdot n\, dx\right) = \frac{n^2}{4}, \end{aligned}$$

which holds as long as $n^2 \ge 48$. So, for example, as long as $n \ge 7$ we get the desired result. This completes the proof of Theorem 4.7. □

PROOF OF THEOREM 4.8. In general $p(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+n})$ consists of the sum of products of $q(\cdot)$ terms where each such product is taken over a subdivision of $(x_1, \dots, x_{k+n})$ and all such subdivisions are represented once in the sum. On the other hand, $p(x_1, \dots, x_k) p(x_{k+1}, \dots, x_{k+n})$ is a product of similar sums for $p(x_1, \dots, x_k)$ and $p(x_{k+1}, \dots, x_{k+n})$. Clearly, every term in the resulting sum for $p(x_1, \dots, x_k) p(x_{k+1}, \dots, x_{k+n})$ is a term in the sum for $p(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+n})$. There are, however, in $p(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+n})$ additional terms in which the subdivisions allow for a combination of x_i's with $1 \le i \le k$ and x_i's with $k + 1 \le i \le k + n$. Since all of the cumulant densities are nonnegative, adding in these additional terms makes $p(x_1, \dots, x_k, x_{k+1}, \dots, x_{k+n})$ at least as large as $p(x_1, \dots, x_k) p(x_{k+1}, \dots, x_{k+n})$. Finally, we construct an example that has PCIC but also has some negative cumulants.

EXAMPLE 6.4. This example shows that X can have PCIC and still have cumulant densities that are negative. We again consider a mixed sample process, this time on the interval $[0, 1]$. In this case we take $p(x_1, \dots, x_n) = g(n)$, where

$$(6.28)\qquad g(n) = \begin{cases} c^{n/2} & \text{if } n \text{ is even,} \\ c^{(n-1)/2} & \text{if } n \text{ is odd.} \end{cases}$$

There are three things we must check:

(a) The product densities p given above determine a completely regular point process, that is, $g(n)$ satisfies the conditions of Lemma 6.2.

(b) The process has positively correlated increasing cylinder sets. That is, $p(x_1, \ldots, x_{k+n}) \geq p(x_1, \ldots, x_k)p(x_{k+1}, \ldots, x_{k+n})$ or $g(n + k) \geq g(n)g(k)$.

(c) The process has at least one negative cumulant density function.

We begin by checking (b) and (c). For (b) it is clear that we have equality when n and k are both even or when one is even and the other is odd. In the case when n and k are both odd, $g(n + k) \geq g(n)g(k)$ if and only if $c^{(n+k)/2} \geq c^{(n-1)/2}c^{(n-1)/2} = c^{(n+k-2)/2}$. That is, if and only if $1 \geq c^{-1/2}$. Thus, as long as we take $c \geq 1$, the process will have PCIC.

To guarantee that at least one cumulant density is negative, we calculate $q(x_1, x_2, x_3)$:

$$
\begin{aligned}
q(x_1, x_2, x_3) &= p(x_1, x_2, x_3) - [p(x_1, x_2) - p(x_1)p(x_2)]\,p(x_3) \\
&\quad - [p(x_1, x_3) - p(x_1)p(x_3)]\,p(x_2) \\
&\quad - [p(x_2, x_3) - p(x_2)p(x_3)]\,p(x_1) - p(x_1)p(x_2)p(x_3) \\
&= g(3) - 3g(2)g(1) + 2(g(1))^3 \\
&= 2 - 2c.
\end{aligned}
\tag{6.29}
$$

Thus $q(x_1, x_2, x_3) < 0$ as long as $c > 1$.

It remains to determine at least one value of c for which the conditions of Lemma 6.2 hold. It is only necessary to check that

$$f(k) = \sum_{n=0}^{\infty} (-1)^n (1/n!) g(k + n) \geq 0$$

for each value of k.

We consider separately the case where k is even and the case where k is odd:

$$f(2n) = \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!} g(j + 2n) = \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!} c^{\log_c g(j+2n)} \tag{6.30}$$

but

$$
\begin{aligned}
\log_c g(j + 2n) &= \begin{cases} n + \dfrac{j}{2} & \text{if } j \text{ is even,} \\[2mm] n + \dfrac{j-1}{2} & \text{if } j \text{ is odd} \end{cases} \\
&= n + \log_c g(j).
\end{aligned}
\tag{6.31}
$$

Thus

$$
\begin{aligned}
f(2n) &= \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!} c^{n + \log_c g(j)} \\
&= c^n f(0).
\end{aligned}
\tag{6.32}
$$

Likewise,

$$f(2n + 1) = c^n f(1). \tag{6.33}$$

Thus we need only check that $f(0) \geq 0$ and $f(1) \geq 0$:

$$f(0) = \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!} g(j) = \sum_{j \text{ even}} \frac{1}{j!} c^{j/2} - \sum_{j \text{ odd}} \frac{1}{j!} c^{(j-1)/2}. \tag{6.34}$$

We will compare the sum for j even with the sum for j odd by comparing them term by term. The nth term of the first sum in (6.34) (starting with $n = 0$) is $(1/2n!)c^n$, whereas the nth term of the second sum is (again starting with $n = 0$) $[1/(2n+1)!]c^n$. Since $c^n(1/2n! - [1/(2n+1)!]) \geq 0$ for all values of n and c, $f(0) \geq 0$.

Unfortunately, showing that $f(1) \geq 0$ is a little more complicated and requires some restrictions on c:

$$f(1) = \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!} g(j+1) = \sum_{j \text{ even}} \frac{1}{j!} c^{j/2} - \sum_{j \text{ odd}} \frac{1}{j!} c^{(j+1)/2}. \tag{6.35}$$

We again compare the last two sums term by term. In this case we compare the nth term of the sum over evens with the $n+1$st term of the sum over odds. Since the 0th term of the sum over the odds is $-c$, the remaining terms must sum to at least c.

Notice that

$$\frac{1}{(2n)!} c^n - \frac{1}{(2(n+1)+1)!} c^{n+2} = \frac{1}{(2n)!} c^n \left(1 - \frac{c^2}{(2n+3)(2n+2)(2n+1)}\right) \tag{6.36}$$

so that the pairings are nonnegative as long as $1 - [c^2/(2n+3)(2n+2)(2n+1)] \geq 0$. That is, as long as $c^2 \leq (2n+3)(2n+2)(2n+1)$. So we need $c^2 \leq 6$, i.e., $c \leq \sqrt{6}$.

It remains to check that we can also make these nonnegative terms add up to a quantity at least as large as c in order to ensure that $f(1) \geq 0$. We now choose a particular value of c, $1 < c \leq \sqrt{6}$, to ensure that X has PCIC, at least one negative cumulant density and g satisfies the conditions of Lemma 6.2.

Let $c = 1.1$. For $n = 0$, we get the term

$$1\left(1 - \frac{1.21}{6}\right) \approx 0.7983. \tag{6.37}$$

For $n = 1$, we get

$$\frac{1.1}{2}\left(1 - \frac{1.21}{69.88}\right) \approx 0.5405. \tag{6.38}$$

Summing (6.37) and (6.38), we get $1.3388 > c = 1.1$, guaranteeing that $f(1) \geq 0$. This completes the proof of Theorem 4.8. In closing we note that it may be checked that this X also does not have CPCIC. □

REFERENCES

[1] BURTON, R. M. and WAYMIRE, E. (1985). Classical scaling limits for associated random measures. *Ann. Probab.* **13** 1267–1278.

[2] BURTON, R. M. and WAYMIRE, E. (1986). The central limit problem for infinitely divisible random measures. In *Dependence in Probability and Statistics* (M. Taqqu and E. Eberlein, eds.) 383–395. Birkhäuser, Boston.

[3] DEN HOLLANDER, W. TH. F. and KEANE, M. (1986). Inequalities of the FKG type. *Phys. A* **138** 167–182.

[4] EATON, M. L. (1986). *Lectures on Topics in Probability Inequalities. CWI Tracts* **35.** Mathematical Centre, Amsterdam.

[5] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

[6] FISHER, L. (1972). A survey of the mathematical theory of multidimensional point processes. In *Stochastic Point Processes: Statistical Analysis, Theory, and Applications* (P. A. W. Lewis, ed.) 468–513. Wiley, New York.

[7] FORTUIN, C., KASTELYN, P. and GINIBRE, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89–103.

[8] KALLENBERG, O. (1975). On symmetrically distributed random measures. *Trans. Amer. Math. Soc.* **202** 105–121.

[9] KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Academic, New York.

[10] KARLIN, S. and RINOTT, J. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467–495.

[11] KARR, A. F. (1986). *Point Processes and Their Statistical Inference*. Dekker, New York.

[12] KEMPERMAN, J. H. B. (1977). On the FKG inequality for measures on a partially ordered space. *Indag. Math.* **39** 313–331.

[13] KESTEN, H. (1988). Asymptotics in high dimensions for the Fortuin–Kastelyn random cluster model. Unpublished manuscript.

[14] MACCHI, O. (1975). The coincidence approach to point processes. *Adv. in Appl. Probab.* **7** 83–122.

[15] MADRAS, N. (1986). A process in a randomly fluctuating environment. *Ann. Probab.* **14** 119–135.

[16] MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic, New York.

[17] MATTHES, K., KERSTAN, J. and MECKE, J. (1978). *Infinitely Divisible Point Processes*. Wiley, New York.

[18] MOYAL, J. E. (1962). The general theory of stochastic population processes. *Acta Math.* **108** 1–31.

[19] PERLMAN, M. and OLKIN, I. (1980). Unbiasedness for invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8** 1326–1341.

[20] VAN DEN BERG, J. and BURTON, R. M. (1989). FKG and equivalent conditions for binary random variables. Unpublished manuscript.

[21] WESTCOTT, M. (1972). The probability generating functional. *J. Austral. Math. Soc.* **14** 448–466.

DEPARTMENT OF MATHEMATICS
OREGON STATE UNIVERSITY
CORVALLIS, OREGON 97331

DEPARTMENT OF MATHEMATICS
WESTERN OREGON STATE COLLEGE
MONMOUTH, OREGON 97361