# Two counterexamples in strong Rayleigh theory [1]

**Abstract:** We discuss two negative dependence properties of binary variables, both of which imply pairwise negative correlation, namely the strong Rayleigh property (SR) and negative association (NA). A seminal paper of Borcea, Brändén and Liggett (2009) showed that SR $\Rightarrow$ NA. Natural examples in NA $\setminus$ SR, are rare; by "natural" we mean measures one cares about for reasons other than merely to disprove the converse implication. The primary goal of this note is to provide two examples of measures that are not SR, for which the weaker properties of NA and negative pairwise correlation were either known or conjectured. Our examples are small, elementary and specific; their significance lies in how they alter our understanding of the relation between NA and SR. In the case of the random cluster model with $q < 1$, for example, our counterexample removes strong Rayleigh theory as a possible attack on the conjectured negative correlation for the random cluster with $q < 1$.

Xiangying Huang, University of North Carolina at Chapel Hill, Department of Statistics and Operations Research, 305 Hanes Hall, Chapel Hill, NC 27514 zoehuang@unc.edu

Robin Pemantle, University of Pennsylvania, Department of Mathematics, 209 South 33rd Street, Philadelphia, PA 19104 pemantle@math.upenn.edu

Rémy Poudevigne-Auboiron, Sorbonne Université, LPSM, 4 place de Jussieu, 75014 Paris rpoudevigne@lpsm.paris

*Subject classification:* primary 62H20; secondary 05A20.

*Key words and phrases*: random cluster model, coalescing random walk, uniform forest, negative association.

[1] We gratefully acknowledge the $70 + \epsilon$ birthday conference for Geoffrey Grimmett, where the crucial part of this work took place.

# 1 Introduction

Collections of random variables taking values in $\{0, 1\}$ are ubiquitous in random graph theory, information theory, theoretical computer science, and statistical physics where often they are called *spin systems* (with states sometimes labeled $\{-1, +1\}$). We call the collection **pairwise positively correlated** if all pairs are nonnegatively correlated and **pairwise negatively correlated** if all pairs are nonpositively correlated.

These are relatively weak properties but sometimes difficult to prove. One way these have historically been proved is by proving positive or negative joint dependence properties that imply the pairwise property. A celebrated result in the spin system literature is the FKG inequailty [FKG71] saying that nonnegative correlation of any pair when you condition on the remaining $n - 2$ variables (equivalently supermodularity of the lattice of probabilities of individual configurations) implies **positive association** (PA). Positive association is defined as nonnegative correlation of every pair of functions that are increasing with respect to the natural partial order; in particular any two spins are nonnegatively correlated. Via FKG, one easily proves that the Ising model with nonnegative interactions is PA. This method also works to prove PA for the **random cluster model** with parameter $q \geq 1$. This model is defined in the next section.

Proving negative dependence properties is harder for a number of reasons. For a lengthy discussion one may look in [Pem00]; a brief list of intuitive reasons is as follows. First, there is much less room for negative covariances among $n$ binary variables than there is for positive covariance (the total covariance of distinct pairs must be between $-n/2$ and $n^2/2$); secondly pairwise negative correlations among any set of three variables frustrate each other; thirdly, even the definitions are more complicated. For example, because a function can't be negatively correlated with itself, one must do more than simply reverse the inequality in the definition of PA to arrive at the following definition of NA.

**Definition 1** (negative association)**.** *Let $\mathcal{B}_n$ be the Boolean lattice $\{0, 1\}^n$ with coordinate functions $X_j(\omega) := \omega_j$. Functions $f$ and $g$ on $\mathcal{B}_n$ are said to depend on disjoint sets of coordinates if for some set $A \subseteq [n]$, the value of $f(\omega)$ depends only on the values $\{\omega_j : j \in A\}$ and the value of $g(\omega)$ depends only on the values $\{\omega_j : j \notin A\}$. Say that a probability measure $\mathbb{P}$ on $\mathcal{B}_n$ is* **negatively associated** *(NA) if and only if every pair of functions $f, g : \mathcal{B}_n \to \mathbb{R}$ depending on disjoints sets of coordinates is negatively correlated.*

This condition is harder to check than is positive association. In fact the theory of joint negative dependence properties lagged far behind the theory of positive dependence until 2009, when the seminal paper [BBL09] defined a property called the **strong Rayleigh property** (SR) and proved that SR $\Rightarrow$ NA. The definition of the strong Rayleigh property is provided in the next section. For now, suffice it to say that, while at first the property may seem opaque, the paper in which it is defined gives several schemes for inferring SR for some measures when you know SR for others, the upshot of which is that many families of measures are known to be SR. A hit parade includes these four families of measures.

- Hermitian determinantal measures [BBL09]
- Time $t$ laws of the exclusion process on a general graph [BBL09]
- Product Bernoulli measures conditioned on the sum [BBL09]
- Many other sampling measures, inclucding pivot sampling, Sampford sampling, Pareto sampling and conditional Poisson sampling [BJ12].

As a tool for proving NA, the strong Rayleigh property is so effective that in fact we know relatively few cases where NA is proved via some other method. In this paper we explore NA \ SR, the set of measures that are negatively associated but not strong Rayleigh. In the examples we care about, often at least one of the two properties is not settled. We therefore separate the class of measures that are potentially in NA \ SR, with at least one of the two settled, into three knowledge classes. The following table lists some measures we will discuss below, broken into these categories, with boldface indicating that the categorization is new in this paper.

| **NA** | **SR** | name of measure |
|---|---|---|
| Yes | No | uniform basis of finite geometries |
| Yes | No | **time $t$ law of coalescing random walk** |
| Yes | ??? | Mohit Singh's Brownian sampler |
| ??? | No | **random cluster measure with $q < 1$** |
| ??? | No | **uniform random forest** |

Usually, when the strong Rayleigh property is established for a measure, it is established for a whole class of measures such as exclusion measures, determinantal measures, etc. The above table refers to *classes* of measures. In other words, failure of the coalescing random walk (CRW) law to be SR means that for some graph, some initial configuration, and some time $t$, the time $t$ joint law of occupation indicators of the CRW fails to be SR.

If there is one main result in the note, it is that the random cluster model with $q < 1$ is not SR (Corollary 7 below). Whether the random cluster model with $q < 1$ has even pairwise negative correlations is still unknown, the answer to which is of great interest. One of the authors of this note had conjectured this measure to be SR, and that this would be the way to prove negative correlations. We have now killed that hope. Taking a more positive view, this note saves everyone time and effort trying an attack on proving negative correlations in the random cluster model that is doomed to fail. The negative result for CRW is less consequential but is also part of the program to try to understand NA \ SR. One motivation to understand this class is to understand the limits on proving NA via SR and when one is likely to need to develop other methods for proving negative dependence properties.

We end the introduction with remarks on what was already known. We consider NA to have been generally known for CRW occupation measures. The published result of which we are aware is [CRS04, Lemma 2.4], drawing on [vdBK02]; neither states the result in full generality. We therefore sketch a short and general proof. Our main tool in disproving SR is another result from [BBL09], namely that SR implies **stochastically increasing levels**. We thank Tyler Helmuth for pointing out to us this avenue for disproving SR for uniform forests and Peter Wildemann for bringing the third author's work to the attention of the other two and explaining its significance. Finally, the limiting relation between the random cluster model and the uniform forest is well known; its significance here was pointed out to us by Geoffrey Grimmett at the same conference acknowledged on the title page of this note.

## 2 Formal statements

Given a probability measure $\mathbb{P}$ on $\mathcal{B}_n$, its generating function is the multi-affine $n$-variable polynomial $F = F_{\mathbb{P}}$ defined by

$$F(z_1, \ldots, z_n) = \sum_{\omega \in \mathcal{B}_n} \mathbb{P}(\omega) \mathbf{z}^\omega \tag{1}$$

where we use the monomial power notation $\mathbf{z}^\omega := \prod_{j=1}^n z_j^{\omega_j}$. The generating polynomial $F$ is said to be multi-affine because each of its monomials has degree 0 or 1 in each variable separately.

**Definition 2** (strong Rayleigh property)**.** *The measure $\mathbb{P}$ on $\mathcal{B}_n$ is said to have the strong Rayleigh property if its generating function $F$ has no zeros on the product upper complex half-plane. More formally $\mathbb{P} \in SR$ if and only if $F(\mathbf{z}) \neq 0$ whenever $\min_{1 \leq j \leq n} \Im\{z_j\} > 0$.*

For $\omega \in \mathcal{B}_n$ let $|\omega| := \sum_{j=1}^n \omega_j$ denote the level of $\omega$ when $\mathcal{B}_n$ is viewed as a graded poset in the usual way. Given $\mathbb{P}$ on $\mathcal{B}_n$ and $j$ such that $\mathbb{P}(|\omega| = j) > 0$ we define the level-$j$ law of $\mathbb{P}$ as the conditional law $\mu_j := \mathbb{P}\big( \cdot \,\big|\, |\omega| = j\big)$.

**Definition 3** (stochastically increasing levels)**.** *The measure $\mathbb{P}$ on $\mathcal{B}_n$ is said to have stochastically increasing levels (SIL) if the set of $k$ such that $\mathbb{P}(|\omega| = k) \neq 0$ forms an interval $I$, and the level-$j$ laws $\{\mu_j : j \in I\}$ form a stochastically increasing sequence: $\mu_j \preceq \mu_{j+1}$ when $j, j+1 \in I$.*

The next proposition states useful implications found in [BBL09], including those that we have quoted in the introduction. The notion of *reweighting* a measure will be useful. If $\mathbb{P}$ is a probability measure on $\mathcal{B}_n$ and $\{w_j : 1 \leq j \leq n\}$ are positive weights (real numbers) then the measure $\mathbb{P}_{\mathbf{w}}$ defined by

$$\mathbb{P}_{\mathbf{w}}(\omega) = C\, \mathbb{P}(\omega) \prod_{j=1}^n w_j^{\omega_j}$$

is called the reweighting of $\mathbb{P}$ by weights $\mathbf{w}$; here $C$ is the normalizing constant. When the coordinates are magnetic spins, this is equivalent to the imposition of an external field with strength $\log w_j$ on spin $j$.

**Proposition 4.**

*(i) SR implies NA. [BBL09, Theorem 4.9]*

*(ii) SR implies SIL. [BBL09, Theorem 4.19]*

*(iii) Hermitian determinantal measures are SR. [BBL09, Proposition 3.5]*

*(iv) The SR property is closed under reweighting / external fields.*

We next define some classical random graph models from the statistical mechanics literature. Let $G = (V, E)$ be a finite connected graph. A subset $W \subseteq E$ containing no cycles is said to be **acyclic**. The maximum size of an acyclic set of edges is $|V| - 1$. If $E'$ is acyclic and $|E'| = |V| - 1$ then the induced subgraph $(V, E')$ is connected and is said to be a **spanning tree**. For general acyclic sets $W$, the induced graph $(V, W)$ is said to be a **forest**. Random subsets of the edge sets of a graph are identified with probability measures on $\mathcal{B}_n$ in a straightforward manner. Letting $n := |E|$ and choosing an aribtrary bijection between $E$ and $[n]$, we identify subsets $W \subseteq E$ of edges with elements of $\mathcal{B}_n$ via $\omega_j := \mathbf{1}_{\{j \in W\}}$.

Next, we define the random cluster model on a finite graph. Historically, this was devised as a probability measure on subgraphs that allows one to sample configurations of the Potts model by requiring the spin to be constant on components of the random subgraph and then assigning the component spins by independent fair coin-flips; this coupling of bonds and spins is called the Edwards-Sokal representation. The marginal on bonds can be defined as follows, with the (positive or negative) interaction strength of the Ising model encoded as a positive real parameter $p$ of the random cluster model, along with a Potts parameter $q > 0$ which need not be an integer. Given a graph $G = (V, E)$ and a subset $W \subseteq E$ of the edges, let $N(W)$ denote the number of components of the induced subgraph $(V, W)$.

**Definition 5** (random cluster model)**.** *Let $G = (V, E)$ be a finite connected graph. Fix parameters $p \in (0, 1)$ and $q > 0$. The random cluster measure with parameters $p$ and $q$ is the probability measure $\mathbb{P} = \mathbb{P}_{p,q}$ on subsets of $E$ defined by*

$$\mathbb{P}(W) = \frac{1}{Z_{p,q}} p^{|W|}(1-p)^{|E\setminus W|} q^{N(W)}$$

*where $Z_{p,q} := \sum_{W\subseteq E} p^{|W|}(1-p)^{|E\setminus W|} q^{N(W)}$ is the partition function. If $G = (V, E, \mathbf{w})$ is a weighted graph, then the weighted random cluster measure with parameter $q$ is obtained by replacing $p^{|W|}(1-p)^{|E\setminus W|}$ by* $\displaystyle\prod_{e\in W} \frac{w_e}{1+w_e} \prod_{e\in E\setminus W} \frac{1}{1+w_e}$.

The uniform measure of spanning trees of $G$ is Hermitian determinantal, hence SR. Our first main result is that this does not extend to the uniform measure on forests.

**Theorem 6.** *There is a finite connected graph $G = (V, E)$ for which the uniform measure $\mathbb{P}_{UF}$ on forests of $G$ is not SR.*

**Corollary 7.** *There is a finite connected graph $G = (V, E)$ and values $p \in (0, 1)$ and $q < 1$ such that the random cluster on $G$ with parameters $p$ and $q$ is not SR.*

Changing topics, let $V = [n]$ and let $\rho : V \times V \times \mathbb{R}^+ \to \mathbb{R}^+$ be a bounded continuous function, which we interpret as a set of jump rates. The natural configuration space is $\Xi := \{0, 1\}^V$. Let $\xi_0 \in \Xi$ be any initial configuration, where we interpret $v \in V$ to be initially occupied if and only if $\xi_0(v) = 1$. For any configuration $\xi \in \Xi$ and any distinct $v, w \in V$, let $\xi_{v\to w}$ denote the configuration agreeing with $\xi$ at vertices other than $v$ and $w$ and satisfying $\xi_{v\to w}(v) = 0$ and $\xi_{v\to w}(w) = \max\{\xi(v), \xi(w)\}$. Informally, if $v$ is occupied at time $t$ the particle at $v$ jumps to $w$, where it coalesces with any existing particle at $w$.

The natural path space on which to define the coalescing random walk measure $\mathbb{P} = \mathbb{P}_{\xi_0,\rho}$ is the set $\Gamma$ of right-continuous paths $\gamma : \mathbb{R}^+ \to \Xi$. Because the jump rates are bounded and continuous, the validity of the following construction is immediate.

**Definition 8** (coalescing random walk)**.** *Given $G, V, D, \rho$ and $\pi$ as above, there is a unique measure $\mathbb{P}$ on $\Gamma$ that defines a Markov process $\{\xi_t : t \geq 0\}$ with starting configuration $\xi_0$ and a jump kernel $\{p_t(\xi, \cdot) : t \geq 0, \xi \in \Xi\}$ that jumps from $\xi$ to $\xi_{v\to w}$ at rate $\rho(v, w, t)$.*

Rather than require our coalescing random walk always to start at a deterministic state, we can prove negative association for a bigger class of initial laws, though not for absolutely any law that itself is NA. Say that a law on $[n]$ is **admissible** if it can be reached by starting from a configuration on some set $U$ disjoint from $[n]$ and letting each particle in $v \in U$ at time 0 jump simultaneously according to probabilities $p_{v,j}$ to a point $j \in [n+1]$, coalescing where more than one particle arrives and eliminating the particle at $n+1$ if there is one. For example product measure on $[n]$ with probabilities $p_1, \ldots, p_n$ is admissible because one can take $U = \{-1, -2, \ldots, -n\}$ and let $-j$ jump to $j$ with probability $p_j$ and $n+1$ with probability $1 - p_j$.

**Theorem 9.**

*(i) Given any CRW started from an admissible law and any time $T > 0$, the law of $\xi_T$ is NA.*

*(ii) There is a time-homogeneous CRW started from a deterministic initial state and a time $T > 0$ such that the law of $\xi_T$ is not SR.*

# 3 Proofs

## 3.1 Proof of Theorem 6 and Corollary 7

By part (iv) of Proposition 4, to prove Theorem 6 it suffices to find a weighted graph $G$ for which the **weighted forest** measure (also call the **arboreal gas** measure) $\mathbb{P}_{\text{WSF}}$ is not SR. Here $\mathbb{P}_{\text{WSF}}(\omega) = C\,\mathbb{P}_{\text{UF}}(\omega) \prod_{j=1}^n w_j^{\omega_j}$. Define the **weighted spanning tree** measure $\mathbb{P}_{\text{WST}}$ for $G$ as the reweighting of the uniform spanning tree measure $\mathbb{P}_{\text{UT}}$ on $G$. For any weighted graph $G$, if $\mathbb{P}_{\text{WST}}$ on $G$ is SR then part (ii) of Proposition 4 implies the level measures $\{\mu_k\}$ stochastically increase. In particular, letting $n := |V| - 1$, the highest level measure $\mu_n = \mathbb{P}_{\text{WST}}$ stochasically dominates all others, hence any mixture of the level measures, hence $\mathbb{P}_{\text{WST}} \succeq \mathbb{P}_{\text{WSF}}$.

Figure 1 shows a graph with 7 vertices, with one edge weighted $x$ and the other 10 edges weighted $y$. We will show that the weighted graph $G$ in the figure is a couterexample whenever $y > 2$ because the edge with weight $x$, call it $e$, is more likely to be in the WSF than it is to be in the WST, violating $\mathbb{P}_{\text{WST}} \succeq \mathbb{P}_{\text{WSF}}$.

The total weight of WST configurations containing the edge $e$ is $x(2y)^5$ because they all contain one edge of weight $x$, five edges of weight $y$ and there are $2^5$ such configurations (choose precisely one of the two incident edges at the five "middle" vertices not adjacent to $e$). The total weight of WST configurations not containing $e$ is $5 \cdot 2^4 y^6$: each contains precisely six edges, all with weight $y$, and to count them, choose one of the five middle vertices to have two adjacent edges and choose precisely one of the two adjacent edges at each other middle vertex. Summarizing,

$$\mathbb{P}_{\text{WST}}(\omega_e = 1) = \frac{32xy^5}{32xy^5 + 80y^6} = \frac{1}{1 + \frac{1}{x}\left(\frac{5y}{2}\right)} \,. \tag{2}$$

The total weight of WSF configurations containing $e$ is equal to $x(1+2y)^5$ because once $e$ is in the forest, each middle vertex independently may have at most one of the two adjacent edges present, for a weight of 1 (neither edge) plus $2y$ (one edge or the other). The total weight of WSF configurations not containing $e$ is only slightly messier. The number of middle vertices with both adjacent edges present is either zero or one. In the former case, choosing independently at each of these yields a total weight of $(1+2y)^5$. In the latter case there are 5 ways to choose which middle vertex has two adjacent edges, leading to weight $5y^2(1+2y)^4$. Thus,

$$\mathbb{P}_{\text{WSF}}(\omega_e = 1) = \frac{x(1+2y)^5}{x(1+2y)^5 + (1+2y)^5 + 5y^2(1+2y)^4} = \frac{1}{1 + \frac{1}{x}\left(1 + \frac{5y^2}{1+2y}\right)} \,. \tag{3}$$

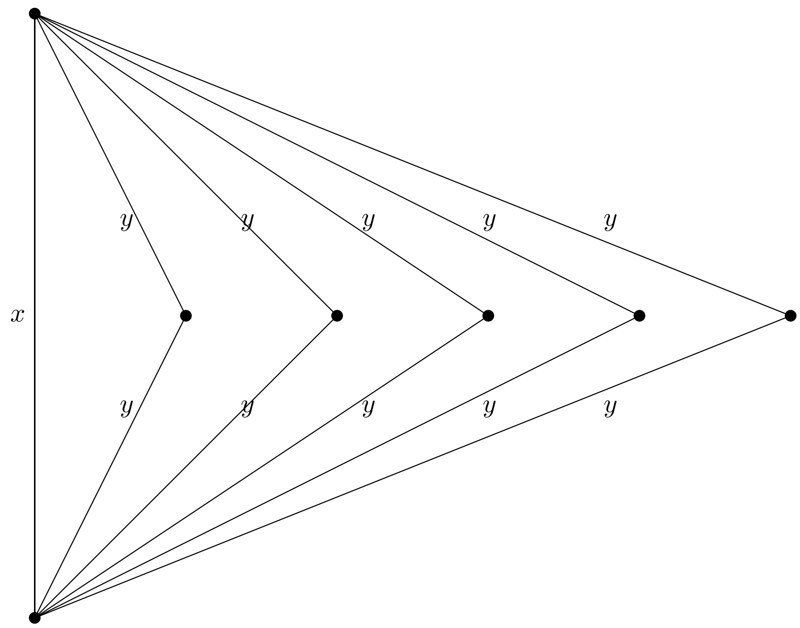


Figure 1:

It is easily checked that $1+\dfrac{5y^2}{1+2y} < \dfrac{5y}{2}$ if and only if $y > 2$, hence $\mathbb{P}_{\mathrm{WSF}}(\omega_e = 1) > \mathbb{P}_{\mathrm{WST}}(\omega_e = 1)$ precisely when $y > 2$. Thus, for $y > 2$, the graph in Figure 1 is a counterexample to the SR property for the WSF, finishing the proof of Theorem 6. $\square$

*Remark.* When the five middle points are replaced by $n$ middle points, the crossover point where $\mathbb{P}_{\mathrm{WSF}}(\omega_e = 1)$ becomes greater than $\mathbb{P}_{\mathrm{WST}}(\omega_e = 1)$ is $y = 2/(n-4)$. Thus, five middle points is the least number that works in this construction.

*Proof of Corollary 7.* Let $G = (V, E)$ be a finite connected graph for which the uniform measure $\mathbb{P}_{\mathrm{UF}}$ on forests of $G$ is not SR, whose existence is guaranteed by Theorem 6. Let $\mathbb{P}_{p,q}$ be the random cluster measure on $G$ with parameters $p$ and $q$. By [Gri10, Theorem (1.23)], letting $p, q \downarrow 0$ with $p = q$, we have

$$\lim_{p=q\downarrow 0} \mathbb{P}_{p,q} = \mathbb{P}_{\mathrm{UF}} .$$

The SR property for measures on a given finite set is a closed property, meaning that it is preserved under weak limits. Consequently, for all $p = q$ sufficiently small, $\mathbb{P}_{p,q}$ fails to be SR, proving Corollary 7. $\square$

## 3.2 Proof of Theorem 9, part $(i)$

As we have said, this is essentially known, but we sketch a proof of the general result that we could not find stated anywhere. Two configurations $\xi, \eta \in \Xi$ satisfy the partial order $\xi \leq \eta$ if $\xi(v) \leq \eta(v)$ for all $v \in V$. A set $A \subseteq \Xi$ is called an *upwardly closed set* if $\xi \in A$ and $\xi \leq \eta$ imply $\eta \in A$.

1. **Define a general discrete time CRW on $V$.** Let $p : V \times V \times \mathbb{Z}^+ \to [0, 1]$ be jump probabilities, that is, $\sum_{w\in V} p(v, w, t) = 1$ for all $v, t$. Let $\{Z(v, t) : v \in V, t \in \mathbb{Z}^+\}$ be independent random variables

with $\mathbb{P}(Z(v,t) = w) = p(v,w,t)$. These may be thought of arrows that give a graphical representation of a general discrete time CRW $\{\xi_t : t \in \mathbb{Z}^+\}$ on $V$ with particles following the arrows and coalescing if two land in the same place. To make this well defined, we stipulate that particles at $v$ and $w$ can switch places without coalescing.

2. **Witnesses for upwardly closed sets.** Fixing an initial configuration and a time $T \geq 1$, a set $S \subseteq V \times [0, T-1]$ witnesses an upwardly closed set $A \subseteq \Xi$ if the values of the random variables $\{Z(v,t) : (v,t) \in S\}$ allow one to conclude that $\xi_T \in A$.

3. **Disjoint witnesses.** Let $W$ be a proper subset of $V$. Let $A$ be an upwardly closed set of $\Xi$ such $\mathbf{1}_A(\omega)$ is measurable with respect to $\{\omega_j : j \in W\}$ and let $B$ be an upwardly closed set of $\Xi$ such $\mathbf{1}_B(\omega)$ is measurable with respect to $\{\omega_j : j \notin W\}$. Then $A \cap B = A \Box B$ where $A \Box B$ is the event that there is a set $S$ of arrows such that $S$ witnesses $\{\xi_t \in A\}$ while the complementary set $S^c = V \times [0, T-1] \setminus S$ witnesses $B$. This is easily checked because one can take $S$ to be the arrows with terminus in $W$ and $S^c$ to be the arrows with terminus in $W^c$, where the terminus of an arrow is the vertex one arrives at by following that arrow and each subsequent arrow beginning where the previous arrow has led.

4. **Use BKR.** By the van den Berg-Kesten-Reimer inequality applied to the independent variables $\{Z(v,t) : v \in V, t \in \mathbb{Z}^+\}$, we know that
$$\mathbb{P}(A \Box B) \leq \mathbb{P}(A)\mathbb{P}(B) ,$$
therefore we have established that any two upwardly closed indicator functions of $\xi_T$ measurable with respect to disjoint subsets of $V$ are nonpositively correlated. Any increasing function $f$ on $\mathcal{B}_n$ measurable with respect to $\{\omega_j : j \in W\}$ is a nonnegative sum $\sum_i a_i \mathbf{1}_{A_i}$ of indicator functions of upwardly closed sets measurable with respect to $\{\omega_j : j \in W\}$. If also $g : \mathcal{B}_n \to \mathbb{R}$ is increasing and measurable with respect to $\{\omega_j : j \in W^c\}$ then also writing $g = \sum_j b_j \mathbf{1}_{B_j}$,
$$\begin{aligned} \mathbb{E} fg &= \sum_{i,j} a_i b_j \mathbb{P}(A_i \cap B_j) \\ &\leq \sum_{i,j} a_i b_j \mathbb{P}(A_i)\mathbb{P}(B_j) \\ &= (\mathbb{E} f)(\mathbb{E} g) \end{aligned}$$
proving negative association.

5. **Pass to continuous time.** Couple a continuous time CRW on the time interval $[0, T]$ to discrete time processes $\mathbf{CRW}_n$ with time step $\varepsilon = T/n$. Because the Bernoulli with parameter $\varepsilon$ and the Poisson with parameter $\varepsilon$ differ in total variation by $O(\varepsilon^2)$ we can couple these two processes so that the coupling fails at each time and place with probability $O(\varepsilon^2)$, hence the processes agree up to time $T$ with probability $1 - O(\varepsilon T)$. This establishes that the continuous CRW law of $\xi_T$ is the weak limit of the laws of the discrete process, hence also NA.

6. **Start from an admissible law.** Add a single timestep populating the time 0 state from a deterministic configuration at time $-1$. This time step never rescales as the others do. This reduces admissible starting measures to deterministic starting states. $\square$

We remark that [CRS04] prove something called *negative upward orthant dependence* which is weaker than NA and they use only the case of the BKR inequality proved earlier in [vdBF87]. Our result requires the full BKR proved by Reimer [Rei00].

### 3.3 Proof of Theorem 9, part *(ii)*

Let $V = [4]$. Consider the admissible state where precisely one of the sites 1 and 2 is populated and precisely one of the sites 3 and 4 is populated. Let the only nonzero jump rate be a constant jump rate from 1 to 3. At time infinity, any particle that was at 1 has jumped to 3. The infinite time evolution maps the four equally likely initial occupied pairs $\{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}$ to the uniform measure $Q$ on four equally likely final occupied sets $\{3\}, \{3,4\}, \{2,3\}, \{2,4\}$. This fails to have the SIL property because $\mu_1 = \delta_{\{3\}}$ whereas $\mu_2$ has the occupation probability of 3 at 2/3 rather than 1.

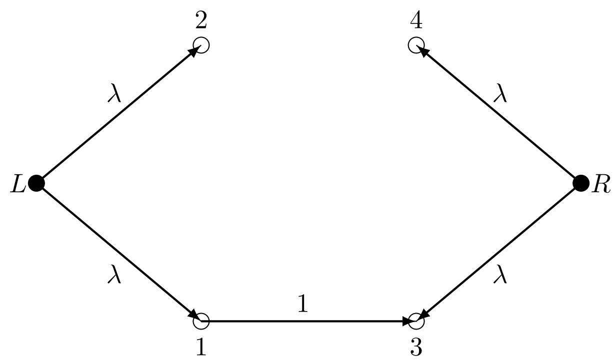


Figure 2: Rates for CRW; start with $L$ and $R$ occupied

Now consider $V = \{1, 2, 3, 4, L, R\}$ with jump rates $p(L, 1) = p(L, 2) = p(R, 3) = p(R, 4) = \lambda$, $p(1, 3) = 1$, and all other rates zero. See Figure 2. As $\lambda, T \to \infty$ the law of $\xi_T$ started in the initial configuration with $L$ and $R$ occupied converges to $Q$. The set of NA laws is closed and its complement is open, therefore for sufficiently large $\lambda$ and $T$, the time-$T$ law of the time-homogeneous CRW in Figure 2 with $L$ and $R$ occupied initially fails to be SR. $\square$

## References


[BBL09] J. Borcea, P. Brändén, and T. Liggett. Negative dependence and the geometry of polynomials. *J. AMS*, 22:521–567, 2009.

[BJ12] P. Brändén and J. Jonasson. Negative dependence in sampling. *Scand. J. Stat.*, 39:830–838, 2012.

[CRS04] E. Cśaki, P. Révész, and Z. Shi. Large void zones and occupation times for coalescing random walks. *Stoch. Proc. Appl.*, 111:97–218, 2004.

[FKG71] C. Fortuin, P. Kasteleyn, and P. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22:89–103, 1971.

[Gri10] G. Grimmett. *The Random-Cluster Model*, volume 333 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York, 2010.

[Pem00] R. Pemantle. Toward a theory of negative dependence. *J. Math. Phys.*, 41:1371–1390, 2000.

[Rei00] D. Reimer. Proof of the van den Berg-Kesten conjecture. *Combin. Probab. Comput.*, 9:27–32, 2000.

[vdBF87] R. van den Berg and U. Fiebig. On a combinatorial conjecture concerning disjoint occurrences of events. *Ann. Probab.*, 15:354–374, 1987.

[vdBK02] R. van den Berg and H. Kesten. Randomly coalescing random walk in dimension $\geq 3$. In V. Sidoravicius, editor, *In and Out of Equilibrium: Probability with a Physics Flavor*, pages 1–45. Birkhäuser, Boston, 2002.