Chapter 23
Special Topic: Associated Random Fields, Positive Dependence, FKG Inequalities

The notion of *association* is a form of positive dependence among random variables independently introduced in reliability theory, percolation theory and statistical physics, where it is expressed in a form known as the "FKG-Inequalities." The main focus of this chapter is (i) a proof of Newman's central limit theorem for associated random fields with summable fast decay of correlations, and (ii) Pitt's characterization of association of multidimensional Gaussian distributions by non-negativity of covariances.

The notion of association as a form of positive dependence has proved to be of much interest in statistical physics,[1] but its potential importance goes beyond statistical physics applications. In 1980 C. M. Newman[2] announced a central limit theorem for associated random fields that will be the focus of this chapter. For stationary random fields the role of association in the asymptotic distribution of centered and scaled sums may be compared to that of martingales for stationary sequences, where only the finiteness of second moments come into play.

In their paper Esary et al. (1967), the notion is developed as a natural extension of weaker forms of positive dependence motivated by applications to reliability theory. It first appeared in Harris (1960), and was later generalized in Fortuin et al. (1971).

[1] A notion of *negative dependence* was explored by Pemantle (2000) from the perspective of statistical physics by way of stimulating examples and conjectures. However, the development of a comparable mathematical theory appears to be much less fruitful.

[2] The central limit theorem of Newman (1980) was extended to a functional central limit theorem for stationary associated sequences in Newman and Wright (1981).

We will restrict the exposition to *random fields* of real-valued random variables $\{X_x : x \in \mathbb{Z}^k\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and indexed by the k-dimensional integer lattice $\mathbb{Z}^k$. Here the natural extension of *stationarity* of sequences to that of random fields is as follows.

Definition 23.1. The random field $\mathbf{X} := \{X_x : x \in \mathbb{Z}^k\}$ is said to be *translation invariant* if for each fixed $z \in \mathbb{Z}^k$ the random field $\{X_{x+z} : x \in \mathbb{Z}^k\}$ is distributed as $\mathbf{X}$.

Definition 23.2. A finite set of random variables $X_1, \dots, X_m$ is said to be *associated* if

$$\text{Cov}(f(X_1, \dots, X_m), g(X_1, \dots, X_m))$$
$$\equiv \mathbb{E}f(X_1, \dots, X_m)g(X_1, \dots, X_m) - \mathbb{E}f(X_1, \dots, X_m)\mathbb{E}g(X_1, \dots, X_m) \geq 0$$

for any pair of bounded measurable coordinatewise non-decreasing functions f, g. An arbitrary collection $\{X_\lambda : \lambda \in \Lambda\}$ is said to be *associated* if every finite subcollection is associated.

The inequalities (23.1) are referred to as the *Fortuin–Kasteleyn–Ginbre (FKG) Inequalities*.[3] Let us begin with a useful formula for covariance in this context. The special case of this formula with $f(x) = x, g(y) = y$ was derived in Lehmann (1966) with attribution to Hoeffding (1940). Newman (1980) noticed the simple but significant extension presented here. (Recall the Definition 2.1 of the covariance of complex-valued random variables.)

Lemma 1 *(Hoeffding-Newman Covariance Formula).* Suppose that $f(X), g(Y) \in L^2(\Omega, \mathcal{F}, P)$ and assume f, g are continuously differentiable complex-valued functions on $\mathbb{R}$ having bounded derivatives. Then,

$$\text{Cov}(f(X), g(Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} f'(x)\overline{g}'(y)H_{X,Y}(x, y)dxdy,$$

where

$$H_{X,Y}(x, y) = \text{Cov}(\mathbf{1}_{[X>x]}, \mathbf{1}_{[Y>y]}) = P(X > x, Y > y)$$
$$- P(X > x)P(Y > y), \ x, y \in \mathbb{R}.$$

Proof. Let (X_1, Y_1) and (X_2, Y_2) be independent random vectors distributed as (X, Y). Note that $\mathbf{1}_{(u,\infty)}(X_1) - \mathbf{1}_{(u,\infty)}(X_2)$ is 1 if $X_2 < u < X_1$, -1 if $X_1 < u < X_2$, and 0 otherwise. Thus, by the fundamental theorem of calculus,

---

[3] Fortuin et al. (1971).

$$f(X_1) - f(X_2) = \int_{-\infty}^{\infty} f'(u)\{\mathbf{1}_{(u,\infty)}(X_1) - \mathbf{1}_{(u,\infty)}(X_2)\}du.$$

Similarly,

$$\overline{g}(X_1) - \overline{g}(X_2) = \int_{-\infty}^{\infty} \overline{g}'(u)\{\mathbf{1}_{(u,\infty)}(Y_1) - \mathbf{1}_{(u,\infty)}(Y_2)\}du.$$

Thus,

$$
\begin{aligned}
2&\mathrm{Cov}(f(X), g(Y)) \\
&= \mathbb{E}[f(X_1) - f(X_2)][\overline{g}(Y_1) - \overline{g}(Y_2)] \\
&= \mathbb{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\mathbf{1}_{[X_1>u]} - \mathbf{1}_{[X_2>u]}\right)\left(\mathbf{1}_{[Y_1>v]} - \mathbf{1}_{[Y_2>v]}\right) f'(u)\overline{g}'(v)dudv.
\end{aligned}
$$

The formula follows by an application of Fubini's theorem to interchange expected value with integrals, after canceling the factors of 2, since expanding the product of indicators one also has by independence and the specified common joint distributions of (X_i, Y_i), $i = 1, 2$, that

$$\mathbb{E}\left(\mathbf{1}_{[X_1>u]} - \mathbf{1}_{[X_2>u]}\right)\left(\mathbf{1}_{[Y_1>v]} - \mathbf{1}_{[Y_2>v]}\right) = 2\{P(X_1 > u, Y_1 > v) - P(X_1 > u)P(Y_1 > v)\}.$$

∎

Remark 23.1. Under the same conditions, the covariance formula may be expressed equivalently as

$$\mathrm{Cov}(f(X), g(Y)) = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{Cov}(\mathbf{1}_{[X>x]}, \mathbf{1}_{[Y>y]}) f'(x)\overline{g}'(y)dxdy.$$

Definition 23.3. A pair of real-valued random variables X, Y for which

$$P(X > u, Y > v) - P(X > u)P(Y > v) \geq 0 \quad \text{for all } u, v \in \mathbb{R},$$

is said to be *positive quadrant dependent*[4]

Proposition 23.1. Associated random variables are (pairwise) positive quadrant dependent.

Proof. Simply note that for any fixed number $a \in \mathbb{R}$, a function of the form $f(u) = \mathbf{1}_{[a,\infty)}(u)$ is non-decreasing. ∎

Newman's proof of the central limit theorem exploits the covariance formulae to compare characteristic functions of sums of random variables with the corre-

---

[4] The notion of positive quadrant dependence was introduced by Lehmann (1966).

sponding product of characteristic functions through the following key lemma. The non-negativity of the covariance is essential to this comparison.

Lemma 2 (Newman). Suppose that $f(X), g(Y) \in L^2(\Omega, \mathcal{F}, P)$ where X, Y are positive quadrant dependent and f, g are continuously differentiable complex-valued functions on $\mathbb{R}$ with bounded derivatives. Then

$$|\mathrm{Cov}\big(f(X), g(Y)\big)| \le ||f'||_\infty ||g'||_\infty \mathrm{Cov}(X, Y),$$

where $|| \cdot ||_\infty$ denotes the essential supremum norm. In particular,

$$|\mathbb{E}e^{irX+isY} - \mathbb{E}e^{irX}\mathbb{E}e^{isY}| \le |r||s|\mathrm{Cov}(X, Y), \quad r, s \in \mathbb{R}.$$

Proof. Using Lemma 1, the assertion follows from the triangle inequality, bounding the derivatives, and the positivity of $H(x, y)$. Specifically,

$$|\mathrm{Cov}\big(f(X), g(Y)\big)| \le ||f'||_\infty ||g'||_\infty \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y)dxdy = ||f'||_\infty ||g'||_\infty \mathrm{Cov}(X, Y).$$

This completes the proof of the general bound. The second bound is simply an application.   ∎

Let us say that a collection of functions $\mathcal{C}$ is *association determining* if one may restrict the FKG inequalities to $f, g \in \mathcal{C}$ to establish association. A proof of the following proposition is left to Exercise 8.

Proposition 23.2. The collections of coordinatewise non-decreasing binary $0 - 1$-valued functions, and of coordinatewise non-decreasing bounded continuous functions, respectively, are each association determining.

The following properties are useful in "tracking association" and/or building examples of associated families of random variables.

Proposition 23.3.

1. Any subcollection of associated random variables is associated.
2. The union of independent collections of associated random variables is associated.
3. Measurable coordinatewise non-decreasing or coordinatewise nonincreasing functions of associated random variables are associated.
4. If for each n, $X_1^{(n)}, \ldots, X_m^{(n)}$ is associated and if $(X_1^{(n)}, \ldots, X_m^{(n)})$ converges in distribution to $(X_1, \ldots, X_m)$, then $X_1, \ldots, X_m$ is associated.
5. A singleton $\{X_1\}$ is associated.
6. Independent random variables are associated.
7. If X, Y are binary random variables, the X, Y are associated if and only if $\mathrm{Cov}(X, Y) \ge 0$.

Proof. Part (1) follows directly from definition by considering functions whose values do not depend on variables not included in the subset. For (2), let $\mathbf{X} = (X_1, \ldots, X_m)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be two independent sequences of associated random variables. Let $\mathbf{Z} = (X_1, \ldots, X_m, Y_1, \ldots Y_n)$. For non-decreasing bounded measurable functions f, g of $m + n$ variables, since the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$ is a product measure by independence, one has

$$\mathrm{Cov}(f(\mathbf{Z}), g(\mathbf{Z})) = \mathbb{E}f(\mathbf{Z})g(\mathbf{Z}) - \mathbb{E}f(\mathbf{Z})\mathbb{E}g(\mathbf{Z})$$

$$= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})g(z_1, \ldots, z_{m+n})P_X(dz_1 \times \cdots \times dz_m)P_Y(dz_{m+1} \times \cdots \times dz_{m+n})$$

$$- \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})P_X(dz_1 \times \cdots \times dz_m)P_Y(dz_{m+1} \times \cdots \times dz_{m+n})$$

$$\times \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} g(z_1, \ldots, z_{m+n})P_X(dz_1 \times \cdots \times dz_m)P_Y(dz_{m+1} \times \cdots \times dz_{m+n})$$

$$= \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})g(z_1, \ldots, z_{m+n})dP_X \right.$$

$$- \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})dP_X \int_{\mathbb{R}^m} g(z_1, \ldots, z_{m+n})dP_X \left. \right\}dP_Y$$

$$+ \int_{\mathbb{R}^n} \left\{ \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})dP_X \int_{\mathbb{R}^m} g(z_1, \ldots, z_{m+n})dP_X \right\}dP_Y$$

$$- \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f(z_1, \ldots, z_{m+n})dP_XdP_Y \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} g(z_1, \ldots, z_{m+n})dP_XdP_Y$$

$$= \int_{\mathbb{R}^n} \mathrm{Cov}\left(f(X_1, \ldots, X_m, z_{m+1}, \ldots, z_{n+m}), g(X_1, \ldots, X_m, z_{m+1}, \ldots, z_{n+m}) \right)dP_Y$$

$$+ \mathrm{Cov}\left(\int_{\mathbb{R}^m} f(z_1, \ldots, z_m, Y_1, \ldots, Y_n)dP_X, \int_{\mathbb{R}^m} g(z_1, \ldots, z_m, Y_1, \ldots, Y_n)dP_X \right) \geq 0,$$

where $dP_X = P_X(dz_1 \times \cdots \times dz_m), dP_Y = P_X(dz_{m+1} \times \cdots \times dz_{m+n})$. The proof of part (3) follows directly from the definition since if $X_1, \ldots, X_m$ are associated and $Y_i = h_i(X_1, \ldots, X_m)$ for measurable coordinatewise non-decreasing functions $h_1, \ldots, h_m$, then $f(h_1, \ldots, h_m)$ and $g(h_1, \ldots, h_m)$ are bounded measurable coordinatewise non-decreasing whenever the same is true of f, g. For the coordinatewise nonincreasing case the composites $f(h_1, \ldots, h_m)$ and $g(h_1, \ldots, h_m)$ are bounded measurable coordinatewise nonincreasing for coordinatewise non-decreasing f, g. Now, $\mathrm{Cov}(f(h_1, \ldots, h_m), g(h_1, \ldots, h_m)) = \mathrm{Cov}(-f(h_1, \ldots, h_m), -g(h_1, \ldots, h_m))$ and $-f(h_1, \ldots, h_m)$ and $-g(h_1, \ldots, h_m)$ are bounded measurable coordinatewise non-decreasing. For part (4), by definition of weak convergence, $\mathrm{Cov}(f(\mathbf{X}), g(\mathbf{X})) = \lim_{n\to\infty} \mathrm{Cov}(f(\mathbf{X}^{(n)}), g(\mathbf{X}^{(n)}))$ for bounded continuous functions f, g. Therefore the result follows since, by the previous proposition, bounded continuous coordinatewise non-decreasing are association determining. To prove (5) restrict to the association determining class of non-decreasing binary functions, and observe that for non-decreasing binary functions f, g of a single variable one has either $f \leq g$ or $g \leq f$. Without loss of

generality consider the case $f \leq g$. Then $\text{Cov}(f(X_1), g(X_1)) = \mathbb{E}f(X_1)g(X_1) - \mathbb{E}f(X_1)\mathbb{E}g(X_1) = \mathbb{E}f(X_1) - \mathbb{E}f(X_1)\mathbb{E}g(X_1) = \mathbb{E}f(X_1)(1 - g(X_1)) \geq 0$. Property (6) follows by application of (2) and (5). Obviously $\text{Cov}(X, Y) \geq 0$ is necessary for X, Y to be associated. For binary $0 - 1$-valued, X, Y, suppose $\text{Cov}(X, Y) \geq 0$. Then the only binary coordinatewise non-decreasing $0 - 1$-valued functions of X, Y are $f_0 \equiv 0, f_1 \equiv 1, g_0(x, y) = x, g_1(x, y) = y$, $h_(x, y) = x \vee y, x \in \{0, 1\}$. Also, $f_0 \leq g_0, g_1 \leq h \leq f_1$. In particular, this ordering trivially implies that $\text{Cov}(f_j(X, Y), g_i(X, Y))] \geq 0, \text{Cov}(g_i(X, Y), h(X, Y)) \geq 0$, $\text{Cov}(f_j(X, Y), h(X, Y)) \geq 0$ for $i, j = 0, 1$. The case $\text{Cov}(g_0(X, Y), g_1(X, Y)) \geq 0$ is the hypothesis. Thus, X, Y is an associated pair proving part (7). ■

Remark 23.2. An alternative proof of the association of a single random variable by *coupling* is an Exercise 2 in Chapter 24.

Example 1 (*A Tendency to Align Under Associated Dependence*). The purpose of this example[5] is to illustrate the tendency for alignment under associated dependence. Consider identically distributed Bernoulli $0 - 1$-valued random variables Y_0, Y_1 with distribution specified by $P(Y_0 = j) = 1/2, P(Y_1 = j|Y_0 = j) = p, j = 0, 1$ for $p \in (0, 1)$. Association requires that Y_1 be most likely to align with the given value of Y_0. That is,

Proposition 23.4. Y_0, Y_1 is associated if and only if $p \geq 1/2$.

Proof. First observe that taking $f(i, j) = i$ and $g(i, j) = j, i, j = 0, 1$, one has that $\text{Cov}(f(Y_0, Y_1), g(Y_0, Y_1)) = \text{Cov}(Y_0, Y_1) = \frac{1}{2}p - \frac{1}{4} \geq 0$ if and only if $p \geq 1/2$. Thus $p \geq 1/2$ is necessary for association. Since Y_0, Y_1 are binary, it likewise follows from Proposition 23.3(g) that $p \geq 1/2$ is sufficient as well. ■

Remark 23.3. In the context of statistical physics association is often expressed as a property of the joint distribution μ of coordinate maps $X_x, x \in \Lambda$, on the product space $\Omega = \{-1, 1\}^\Lambda$ for some finite set Λ of integer lattice points connected to the origin; i.e., $X_x(\omega) = \omega_x, \omega \in \Omega$. The probability measure μ is said to satisfy the *FKG inequalities* if for any coordinatewise non-decreasing functions f, g on Ω one has

$$\int_\Omega f(X_x)g(X_y)d\mu \geq \int_\Omega f(X_x)d\mu \int_\Omega g(X_y)d\mu, \quad x, y \in \Lambda. \tag{23.1}$$

Equivalently, the FKG inequalities are the property that the collection of spin ± 1-valued random variables $X_x, x \in \Lambda$ have associated dependence. The ferromagnetic Ising model (see Chapter 13, Exercise 13) provides a well-known example in this context. The FKG inequalities for the ferromagnetic Ising model will be proved in Chapter 24, Proposition 24.12. The magnetic spin alignment reflected by association is a distinct feature of ferromagnets, responsible for their ability

---

[5] Also see Exercise 8 in Chapter 24 in this regard.

to be magnetized by placement in an external magnetic field. Other more general inequalities to be considered in Chapter 24 are available that imply association.[6]

The proof of the central limit theorem exploits the following basic inequality.

Lemma 3 *(Newman's Inequality).* Suppose that $X_1, \ldots, X_m$ are associated random variables having finite variance. Then for any $r_1, \ldots, r_m \in \mathbb{R}$ one has

$$|\mathbb{E} \exp\{i \sum_{j=1}^{m} r_j X_j\} - \prod_{j=1}^{m} \mathbb{E} e^{i r_j X_j}| \leq \sum_{1 \leq j < k \leq m} |r_j||r_k| \operatorname{Cov}(X_j, X_k).$$

Proof. The proof is by induction on m. The case $m = 1$ is obvious and the case $m = 2$ was proven in Lemma 2. Assume the inequality holds for all $m \leq M$ and rearrange the indices (if necessary) in such a way that $\operatorname{sgn}(r_j)$ is constant, say ϵ (either $+1$ or -1), for $1 \leq j \leq m_0$, and $\operatorname{sgn}(r_j)$ is also constant, say δ, for $m_0 + 1 \leq j \leq M$. Then $\epsilon r_j \geq 0, \delta r_j \geq 0$, so that each of $X = \sum_{j=1}^{m_0} \epsilon r_j X_j$ and $Y = \sum_{j=m_0+1}^{M+1} \delta r_j X_j$ is a non-decreasing function of associated variables $X_1, \ldots, X_{M+1}$ and therefore associated. Also $\sum_{j=1}^{M+1} r_j X_j = \epsilon X + \delta Y$. Thus, applying Lemma 2 and the induction hypothesis, one has

$$\left| \mathbb{E} \exp\left\{ i \sum_{j=1}^{M+1} r_j X_j \right\} - \prod_{j=1}^{M+1} \mathbb{E} e^{i r_j X_j} \right|$$

$$\leq \left| \mathbb{E} e^{i(\epsilon X + \delta Y)} - \mathbb{E} e^{i \epsilon X} \mathbb{E} e^{i \delta Y} \right| + \left| \mathbb{E} e^{i \epsilon X} \mathbb{E} e^{i \delta Y} - \mathbb{E} e^{i \epsilon X} \prod_{j=m_0+1}^{M+1} \mathbb{E} e^{i r_j X_j} \right|$$

$$+ \left| \mathbb{E} e^{i \epsilon X} \prod_{j=m_0+1}^{M+1} \mathbb{E} e^{i r_j X_j} - \left(\prod_{j=1}^{m_0} \mathbb{E} e^{i r_j X_j} \right) \prod_{j=m_0+1}^{M+1} \mathbb{E} e^{i r_j X_j} \right|$$

$$\leq |\epsilon||\delta| \operatorname{Cov}(X, Y) + \left| \mathbb{E} e^{i \delta Y} - \prod_{j=m_0+1}^{M+1} \mathbb{E} e^{i r_j X_j} \right| + \left| \mathbb{E} e^{i \epsilon X} - \prod_{j=1}^{m_0} \mathbb{E} e^{i r_j X_j} \right|$$

$$\leq \operatorname{Cov}\left(\sum_{j=1}^{m_0} \epsilon r_j X_j, \sum_{k=m_0+1}^{M+1} \delta r_k X_k \right) + \sum_{m_0+1 \leq j < k \leq M+1} |r_j||r_k| \operatorname{Cov}(X_j, X_k)$$

$$+ \sum_{1 \leq j < k \leq m_0} |r_j||r_k| \operatorname{Cov}(X_j, X_k)$$

$$= \sum_{1 \leq j < k \leq M+1} |r_j||r_k| \operatorname{Cov}(X_j, X_k).$$

∎

---

[6] See den Hollander and Keane (1986).

Of course, one may prefer the equivalent expression of Newman's bound as

$$\frac{1}{2} \sum_{1 \le j,k \le m, j \ne k} |r_j||r_k| \, \mathrm{Cov}(X_j, X_k) = \sum_{1 \le j < k \le M+1} |r_j||r_k| \, \mathrm{Cov}(X_j, X_k). \quad (23.2)$$

Lemma 4. Let $\mathbf{X} := \{X_x : x \in \mathbb{Z}^k\}$ be a translation invariant random field of associated random variables having finite second moments. Assume that

$$\gamma := \sum_{x \in \mathbb{Z}^k} \mathrm{Cov}(X_0, X_x) < \infty.$$

Let

$$B_x^{(N)} := \{y \in \mathbb{Z}^k : Nx_l \le y_l < N(x_l + 1), l = 1, \dots, k\}$$

denote a "block of lattice sites of length N located near Nx", $x = (x_1, \dots, x_k)$, and define a random field of centered and rescaled "block sum averages" by

$$A_x^{(N)} = N^{-\frac{k}{2}} \sum_{y \in B_x^{(N)}} (X_y - \mathbb{E}X_y), \qquad x \in \mathbb{Z}^k.$$

Then

$$\lim_{N \to \infty} \mathrm{Var}(A_x^{(N)}) = \gamma, \quad \text{and} \quad \lim_{N \to \infty} \mathrm{Cov}(A_x^{(N)}, A_y^{(N)}) = 0 \quad x \ne y.$$

Proof. By translation invariance it suffices to check the asserted limits for the case $x = 0$. Clearly

$$\mathrm{Var}(A_0^{(N)}) = N^{-k} \sum_{x \in B_0^{(N)}} \sum_{y \in B_0^{(N)}} \mathrm{Cov}(X_0, X_{y-x}) \le N^{-k} \sum_{x \in B_0^{(N)}} \sum_{y \in \mathbb{Z}^k} \mathrm{Cov}(X_0, X_{y-x}).$$

In particular, letting $N \to \infty$,

$$\limsup_{N \to \infty} \mathrm{Var}(A_0^{(N)}) \le \gamma.$$

For the reverse inequality let $0 < \epsilon < 1/2$ and define

$$B_0^{(N)}(\epsilon) := \{z = (z_1, \dots, z_k) : \epsilon N < z_i < (1 - \epsilon)N, i = 1, \dots, k\}.$$

Note that for $x \in B_0^{(N)}$, $y \notin B_0^{(N)}$, $|x - y| \ge \epsilon N$, so that

$$\text{Var}(A_0^{(N)}) \geq N^{-k} \sum_{x \in B_0^{(N)}(\epsilon)} \sum_{y \in B_0^{(N)}} \text{Cov}(X_0, X_{y-x})$$

$$\geq N^{-k} \sum_{x \in B_0^{(N)}(\epsilon)} \sum_{|y-x| \leq \epsilon N} \text{Cov}(X_0, X_{y-x}) = \frac{|B_0^{(N)}(\epsilon)|}{N^k} \sum_{|z| \leq \epsilon N} \text{Cov}(X_0, X_z),$$

where $|B|$ denotes cardinality of the set B. Choosing a sequence such that $\epsilon_N \downarrow 0$ and $\epsilon_N N \to \infty$, one obtains

$$\liminf_{N \to \infty} \text{Var}(A_0^{(N)}) \geq \gamma.$$

This proves the asserted asymptotic variance. For the covariance decay choose a sequence $M_N \leq N$ such that $M_N/N \to 1$ and $N - M_N \to \infty$ as $N \to \infty$.

$$\text{Var}(A_0^{(N)} - A_0^{(M_N)})$$

$$= \text{Var}(A_0^{(N)}) + \text{Var}(A_0^{(M_N)}) - 2(NM_N)^{-\frac{k}{2}} \text{Cov}\left(\sum_{y \in B_0^{(N)}} X_y, \sum_{y \in B_0^{(M_N)}} X_y \right)$$

$$\leq \text{Var}(A_0^{(N)}) + \text{Var}(A_0^{(M_N)}) - 2\left(\frac{M_N}{N}\right)^{\frac{k}{2}} \text{Var}(X_0^{(M_N)}) \to 0.$$

One has for $z \neq 0$,

$$\text{Cov}(A_0^{(N)}, A_z^{(N)}) = \text{Cov}(A_0^{(N)} - A_0^{(M_N)}, A_z^{(N)}) + \text{Cov}(A_0^{(M_N)}, A_z^{(N)})$$

$$\leq \sqrt{\text{Var}(A_0^{(N)} - A_0^{(M_N)})}\sqrt{\text{Var}(A_z^{(N)})} + \text{Cov}(A_0^{(M_N)}, A_z^{(N)}).$$

Thus the proof of covariance decay is therefore completed by the following calculation

$$\text{Cov}(A_0^{(M_N)}, A_z^{(N)}) = M_N^{-\frac{k}{2}} N^{-\frac{k}{2}} \sum_{x \in B_0^{(M_N)}} \sum_{y \in B_z^{(N)}} \text{Cov}(X_0, X_{x-y})$$

$$\leq \left(\frac{M_N}{N}\right)^{\frac{k}{2}} M_N^{-k} \sum_{x \in B_0^{(M_N)}} \sum_{|y-x| \geq N - M_N} \text{Cov}(X_0, X_{x-y})$$

$$= \left(\frac{M_N}{N}\right)^{\frac{k}{2}} \sum_{|y| \geq N - M_N} \text{Cov}(X_0, X_y) \to 0,$$

as $N \to \infty$.   ∎

To state Newman's central limit theorem it is helpful to have some extra notation. For $x = (x_1, \ldots, x_k) \in \mathbb{Z}^k$, a "block of lattice sites of length N located near Nx" is denoted

$$B_x^{(N)} := \{y \in \mathbb{Z}^k : Nx_l \leq y_l < N(x_l + 1), l = 1, \ldots, k\}.$$

Given a translation invariant random field with finite second moments $\mathbf{X} := \{X_x : x \in \mathbb{Z}^k\}$, the random field of centered and rescaled "block sum averages" is denoted

$$A_x^{(N)} = N^{-\frac{k}{2}} \sum_{y \in B_x^{(N)}} (X_y - \mathbb{E}X_y).$$

Theorem 23.5 *(Newman's Central Limit Theorem).* Let $\mathbf{X} := \{X_x : x \in \mathbb{Z}^k\}$ be a translation invariant random field of associated random variables having finite second moments. Assume that

$$\gamma := \sum_{x \in \mathbb{Z}^k} \mathrm{Cov}(X_0, X_x) < \infty.$$

Then for any finite number n of lattice sites $z_1, \ldots, z_n$, the (finite dimensional) distribution of $(A_{z_1}^{(N)}, A_{z_2}^{(N)}, \ldots, A_{z_n}^{(N)})$ converges weakly as $N \to \infty$ to the Gaussian distribution with mean zero and covariance matrix $\mathrm{diag}(\gamma, \ldots, \gamma)$.

Proof. By Newman's inequality and association inherited by the $A_z^{(N)}, z \in \mathbb{Z}^k$, and the previous lemma, it suffices to show convergence of $A_z^{(N)}$, i.e., $n = 1$, to obtain convergence for finite dimensional distributions of arbitrary size $n \geq 1$. More specifically, if one can show $\mathbb{E}e^{ir A_z^{(N)}} \to e^{-\frac{\gamma}{2}r^2}$ as $N \to \infty$, then

$$\lim_{N \to \infty} \left| \mathbb{E}e^{i \sum_{j=1}^n r_j A_{z_j}^{(N)}} - \prod_{j=1}^n e^{-\frac{\gamma}{2}r_j^2} \right| \leq \lim_{N \to \infty} \sum_{1 \leq m < j \leq n} |r_m||r_j| \mathrm{Cov}(A_{z_m}^{(N)}, A_{z_j}^{(N)}) = 0.$$

(23.3)

As noted earlier, by translation invariance it is sufficient to consider the case $z = 0$. For fixed $M = 1, 2, \ldots$, let $M_N = M[\frac{N}{M}] \leq N$, where $[\cdot]$ denotes integer-part. In the proof of the previous lemma it was shown that $\mathrm{Var}(A_0^{(N)} - A_0^{(M_N)}) \to 0$ as $N \to \infty$. Thus, one has

$$\left| \mathbb{E}e^{ir A_0^{(N)}} - \mathbb{E}e^{ir A_0^{(M_N)}} \right| \leq \mathbb{E} \left| e^{ir(A_0^{(N)} - A_0^{(M_N)})} - 1 \right|$$

$$\leq \mathbb{E} \left| A_0^{(N)} - A_0^{(M_N)} \right| \leq \sqrt{\mathrm{Var}(A_0^{(N)} - A_0^{(M_N)})} \to 0.$$

Next, using the simple property of the block averages that

$$A_0^{(N_1 N_2)} = N_1^{-\frac{k}{2}} \sum_{y \in B_0^{(N_1)}} A_0^{(N_2)}, \tag{23.4}$$

(for $M_N = M[\frac{N}{M}] = N_1 N_2$), one has by Newman's inequality (applied to $A_0^{(M)}$)

$$\left| \mathbb{E} e^{ir A_0^{(M[\frac{N}{M}])}} - \left(\mathbb{E} e^{ir[\frac{N}{M}]^{-\frac{k}{2}} A_0^{(M)}} \right)^{([\frac{N}{M}])^k} \right| \le \frac{1}{2} \sum_{\substack{x,y \in B_0^{([\frac{N}{M}])} \\ x \ne y}} r^2 \left(\left[\frac{N}{M}\right] \right)^{-k} \mathrm{Cov}\left(A_x^{(M)}, A_y^{(M)} \right).$$

This upper bound may be equivalently expressed using the block average property (23.4) as

$$\frac{r^2}{2} \left\{ \mathrm{Cov}\left(A_0^{(M[\frac{N}{M}])}, A_0^{(M[\frac{N}{M}])} \right) - \left[\frac{N}{M}\right]^{-k} \sum_{y \in B_0^{([\frac{N}{M}])}} \mathrm{Cov}\left(A_y^{(M)}, A_y^{(M)} \right) \right\}$$

$$= \frac{r^2}{2} \left\{ \mathrm{Var}\left(A_0^{(M[\frac{N}{M}])} \right) - \mathrm{Var}\left(A_0^{(M)} \right) \right\} \to \frac{r^2}{2} \left\{ \gamma - \mathrm{Var}\left(A_0^{(M)} \right) \right\}.$$

Letting $N \to \infty$ with M fixed, it follows that

$$\left(\mathbb{E} e^{ir[\frac{N}{M}]^{-\frac{k}{2}} A_0^{(M)}} \right)^{([\frac{N}{M}])^k} = \left(1 - \frac{r^2}{2} ([\frac{N}{M}])^{-k} \mathrm{Var}(A_0^{(M)}) + o([\frac{N}{M}]^{-k}) \right)$$

$$\to e^{-\frac{\mathrm{Var}(A_0^{(M)})}{2} r^2}. \tag{23.5}$$

Thus, combining these estimates, one has

$$\limsup_{N \to \infty} \left| \mathbb{E} e^{ir A_0^{(N)}} - e^{-\frac{\gamma}{2} r^2} \right| \le \frac{r^2}{2} \left\{ \gamma - \mathrm{Var}(A_0^{(M)}) \right\} + \left\{ e^{-\frac{\mathrm{Var}(A_0^{(M)})}{2} r^2} - e^{-\frac{\gamma}{2} r^2} \right\}.$$

Finally, letting $M \to \infty$ completes the proof. ∎

The following example[7] is a significant framework[8] in which association naturally occurs.

Example 2 (Two-dimensional Bond Percolation Model). The *independent bond percolation model*[9] on $\mathbb{Z}^2$ can be defined as follows: Each lattice site $x \in \mathbb{Z}^2$ has four *nearest neighbor sites* of the form $y = x \pm e$ where e is either $(1, 0)$ or $(0, 1)$. A pair of such nearest neighbor sites x, y, in turn, defines a (unoriented) *bond* $b = \{x, y\}$ of $\mathbf{Z}^2$. Let $\mathbb{L}^2$ denote the collection all such bonds of $\mathbb{Z}^2$. Let $\{Y_b : b \in \mathbb{L}^2\}$ be the i.i.d. random field of Bernoulli $0 - 1$ valued random variables with $p = P_p(Y_b = 1)$ defined by coordinate projections on the product probability space $\Omega = \{0, 1\}^{\mathbb{L}^2}$ equipped with the σ-field $\mathcal{F}$ generated by finite dimensional cylinder sets and product measure $P_p = \prod_{\mathbb{L}^2}(q\delta_{\{0\}} + p\delta_{\{1\}})$, where $q = 1 - p$. Declare the bonds b as *open* or *closed* according to whether the value of Y_b is 1 or 0, respectively. The usual interpretation of percolation is as a model for a disordered porous medium in which the open bonds permit fluid flow between nearest neighbor sites, while closed bonds block the passage of fluid. Two sites $x, z \in \mathbb{Z}^2$ are said to be *connected by an open path*, denoted $x \leftrightarrow z$ if there is a succession of sites in $\mathbb{Z}^2$, $x_0 = x, x_1, \ldots x_m = z, m \geq 1$, such that pairs x_i, x_{i+1} are nearest neighbor with $b_i = \{x_i, x_{i+1}\}$ open $(i = 0, \ldots, m - 1)$. A *cluster* $C(x)$ at site $x \in \mathbb{Z}^2$ is defined by the (random) set

$$C(x) := \{z \in \mathbb{Z}^2 : x \leftrightarrow z\}, \qquad x \in \mathbb{Z}^2.$$

The *cluster size* refers to the (possibly infinite) cardinality of $C(x)$ and is denoted by $|C(x)|$. The set $C(x)$ is referred to as a *percolation cluster*[10] at x if $|C(x)| = \infty$.

Definition 23.4. The existence of an infinite cluster that is the event $E :=$ $\cup_{x \in \mathbb{Z}^2}[|C(x)| = \infty]$ is referred to as the *percolation* event. Also, the *percolation probability* is defined by

$$\rho \equiv \rho(p) := P_p(E) = P_p(\cup_{x \in \mathbb{Z}^2}[|C(x)| = \infty]). \qquad (23.6)$$

---

[7] The survey article Last et al. (2020) is a source of a wide variety of additional examples of associated stochastic random fields. Extension of Newman's central limit theorem for Poisson cluster processes and random measures was developed in Burton and Waymire (1985), and independently by Evans (1989), provides an illustrative setting for applications of positive dependence.

[8] Also see Newman (1980) for examples in the context of mathematical physics.

[9] The mathematical interest in percolation models is usually traced to Broadbent and Hammersley (1957). Broadbent's work at the British Coal Utilization Research Association involved the design of porous gas masks for coal miners. The critical nature of pore size was empirically realized in this context, motivating the subsequent development of simpler models of such phenomena of wide interest in probability and mathematical physics.

[10] Uniqueness of such percolation clusters was originally established by Aizenmann et al. (1987). A widely recognized very simple proof of uniqueness was subsequently made by Burton and Keane (1989). This has become a standard approach to uniqueness.

Proposition 23.6. Define

$$\theta \equiv \theta(p) := P_p(|C(0)| = \infty). \tag{23.7}$$

Then $\rho(p) = 0$ or $\rho(p) = 1$ if and only if $\theta(p) = 0$ or $\theta(p) > 0$, respectively.

Proof. Note that the percolation event $E = \cup_{x \in \mathbb{Z}^2}[|C(x)| = \infty]$ is a tail event for a countable collection of i.i.d. random variables $\{Y_b : b \in \mathbb{L}^2\}$. The assertion follows immediately from subadditivity and Kolmogorov's zero-one law.[11] Namely, $\rho(p) = 0$ or 1, $\theta(p) \leq \rho(p)$, and $\rho(p) \leq \sum_{x \in \mathbb{Z}^2} \theta(p)$. So $\rho(p) = 0$ if and only if $\theta(p) = 0$, and $\theta(p) > 0$ if and only if $\rho(p) = 1$. ∎

Remark 23.4. A proof of the monotonicity of the percolation probability $p \to \theta(p)$ as a function of p by monotone coupling techniques is given for Proposition 24.3 in Chapter 24.

Definition 23.5. The *critical probability* for existence of an infinite cluster, i.e., percolation, is defined by

$$p_c = \sup\{p \in [0, 1] : \theta(p) = 0\}.$$

Remark 23.5. An important role for the FKG inequalities occurs in a simplified proof of the criticality of $p = 1/2$ for bond percolation by Bollabás and Riordan (2006). The original proof is the result of Kesten (1980), after completing the upper bound calculation from two-decades earlier by Harris (1960), that $p_c = 1/2$ for $2d$-bond percolation. The upper bound $p_c \leq 1/2$ had already involved inequalities, now known as *Harris inequalities*, that may be viewed as a special case of the FKG inequalities for product measure.

For probability measures μ_1 and μ_2 on the compact space (for product topology) $S = \{0, 1\}^\Lambda$, where Λ is a finite or countably infinite set, the *Holley inequalities*[12] are a generalization of associated dependence of the form

$$\int_S f d\mu_1 \geq \int_S f d\mu_2, \tag{23.8}$$

for coordinatewise non-decreasing functions f on S; equivalently it is non-decreasing with respect to the partial order $\preceq$ on S defined by $x \preceq y$ if and only if $x_j \leq y_j, j \in \Lambda$ for $x, y \in S$,

To see that (23.8) embodies association of a probability distribution μ on S, let f, g be nonnegative coordinatewise non-decreasing functions on S. Take $d\mu_1 = \frac{g d\mu}{\int_S g d\mu}$, $\mu_2 = \mu$. Holley's inequalities for μ_1 and μ_2 are then equivalent to the FKG

---

[11] See BCPT, p. 87.

[12] Holley (1974).

inequalities for μ. The so-called *log-convexity* type conditions[13] on μ, μ_1, μ_2 are available to ensure either FKG inequalities or Holley inequalities, respectively.

In the so-called *disordered phase*[14] defined by $0 < p < p_c$, the lattice a.s. consists of infinitely many disjoint *finite* random clusters of lattice sites connected by open bonds. The following simple path counting argument demonstrates the existence of a disordered phase.

Proposition 23.7. $p_c > 0$.

Proof. Consider the number N_n of open (self-avoiding) paths of length n starting at the origin. Clearly, noting that such a path can connect to any of the 4 neighbors of $(0, 0)$ and continue in $n - 1$ self-avoiding steps, $N_n \leq 4(3^{n-1})$. Thus for $p < 1/3$, applying a useful but very simple inequality for nonnegative integer-valued random variables,

$$P_p(N_n \geq 1) \leq \mathbb{E}_p N_n \leq 4(3^{n-1})p^n \to 0 \text{ as } n \to \infty.$$

In particular,[15] since $\theta(p) \leq P_p(N_n \geq 1)$ for all $n \geq 1$, one has $\theta(p) = 0$ for $p < 1/3$ and hence $p_c \geq 1/3$. ∎

Lemma 5 (*Harris' Lemma*[16]). Let $X_x = \mathbf{1}_{[C(x)\neq\emptyset]}$, $x \in \mathbb{Z}^2$. Then $\{X_x : x \in \mathbb{Z}^2\}$ is a translation invariant random field of associated random variables.

Proof. Translation invariance follows directly from the definition and the fact that the distribution of the underlying random field $\{Y_b : b \in \mathbb{L}^2\}$ is invariant under translation of the lattice $\mathbb{Z}^2$. Also each X_x, $x \in \mathbb{Z}^2$, is a (coordinatewise) non-decreasing function of $\mathbf{Y} \equiv \{Y_b : b \in \mathbb{L}^2\}$. Apply Proposition 23.3. ∎

For an application of the central limit theorem in this context we will establish the asymptotic normality of the cumulative size $\sum_{x \in B_0^{(N)}} |C(x)|$ of all clusters connected to points in the cube $B_0^{(N)}$, suitably centered and scaled for $0 < p < 1/3$. Additional applications[17] along these lines are given in the exercises. The conditions for the theorem will be checked in a sequence of simple lemmas, the first of which is a special case of an inequality known as the *BK Inequality* after its originators van den Berg and Kesten (1985).

---

[13] den Hollander and Keane (1986).

[14] Physicists often refer to the absence of long-range connectivities as "disorder."

[15] In his celebrated paper, Kesten (1980), it was proved that $p_c = 1/2$.

[16] A stronger version of this type result was first formulated and proven by Harris (1960) as a special case.

[17] The example given here serves the pedagogical purpose of simply illustrating the theorem. For more substantial applications, but requiring elements of percolation theory which are outside the scope of this exposition, consult the comprehensive text by Grimmett (1999), and numerous references therein.

To prepare for the BK inequality let us refer to a random variable X defined on Ω as *increasing* if $X(\omega_1) \leq X(\omega_2)$ whenever $\omega_1, \omega_2 \in \Omega$ satisfy $\omega_1(b) \leq \omega_2(b)$ for all $b \in \mathbb{L}^2$; the latter set of coordinatewise inequalities defines a partial order on Ω which we denote as $\omega_1 \preceq \omega_2$. Similarly we say that an event $A \in \mathcal{F}$ is an *increasing event* if $\mathbf{1}_A$ is an increasing random variable. Connectivity events of the form $[x \leftrightarrow y]$ are prototypical increasing events. From here out we restrict to this case.

Definition 23.6. The *disjoint occurrence* of two increasing events $A = [x \leftrightarrow y]$, $B = [z \leftrightarrow w]$ is an event denoted by $A \circ B$ and defined by

$$[x \leftrightarrow z] \circ [y \leftrightarrow w] = [x \leftrightarrow z, x \notin C(y), y \leftrightarrow w].$$

Lemma 6 (BK Inequality-Special Case). For $x, y, w, z \in \mathbb{Z}^2$

$$P_p([x \leftrightarrow z] \circ [y \leftrightarrow w]) \leq P_p(x \leftrightarrow z) P_p(y \leftrightarrow w).$$

Proof. Observe that

$$P_p([x \leftrightarrow z] \circ [y \leftrightarrow w]) = \mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[x \notin C(y)]} \mathbf{1}_{[y \leftrightarrow w]})$$

$$= \mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[x \notin C(y)]} \mathbf{1}_{[x \notin C(w)]} \mathbf{1}_{[y \leftrightarrow w]})$$

$$= P_p(x \leftrightarrow z, x \notin C(y), x \notin C(w), y \leftrightarrow w)$$

$$= P_p(y \leftrightarrow w | x \leftrightarrow z, x \notin C(y), x \notin C(w)) P_p(x \leftrightarrow z, x \notin C(y), x \notin C(w))$$

$$\leq P_p(y \leftrightarrow w | x \leftrightarrow z, x \notin C(y), x \notin C(w)) P_p(x \leftrightarrow z). \tag{23.9}$$

So it suffices to show that

$$P_p(y \leftrightarrow w | x \leftrightarrow z, x \notin C(y), x \notin C(w)) \leq P_p(y \leftrightarrow w). \tag{23.10}$$

Let A be an arbitrary but fixed finite connected subgraph of $\mathbb{L}^2$ with vertices x and z connected in A, but not connected to y nor w, i.e., having the properties of the conditioning. The graph A is referred to as a *lattice animal*. Denote the vertex and edge sets of A by A_v and A_e, respectively. Also define the *edge boundary* $\partial_e A$ as the set of (closed) edges which do not belong to A_e but have at least one endvertex in A_v. First consider the case in which y is "interior" to the lattice animal A and w is "exterior" to A in the sense that any path of bonds connecting y to w must include a bond from $\partial_e A$. Then, since on $[C(x) = A]$ the edges in $\partial_e A$ are all closed, one has for this case that

$$P_p(y \leftrightarrow w, C(x) = A | x \leftrightarrow z, x \notin C(y), x \notin C(w)) = 0.$$

On the other hand, for the case when $\partial_e A$ does not obstruct the existence of a path of open bonds connecting y to w, let us see that one may use association (FKG

inequalities) to establish that

$$P_p(y \leftrightarrow w, C(x) = A | x \leftrightarrow z, x \notin C(y), x \notin C(w)) \le P_p(y \leftrightarrow w). \qquad (23.11)$$

To prove (23.11) consider a lattice animal A such that $x \leftrightarrow z$ and $x \notin C(y), x \notin C(w)$ and for which y, w are not separated by $\partial_e A$ in the previous sense. Then $\mathbf{1}_{[y \leftrightarrow w]}$ and $\mathbf{1}_{[C(x)=A]}$ are, respectively, increasing and decreasing functions of independent random variables. Thus, by association (see Exercise 5),

$$P_p(y \leftrightarrow w, C(x) = A, x \leftrightarrow z, x \notin C(y), x \notin C(w)) = P_p(y \leftrightarrow w, C(x) = A)$$
$$\le P_p(y \leftrightarrow w) P_p(C(x) = A)$$
$$= P_p(y \leftrightarrow w) P_p(C(x) = A, x \leftrightarrow z, x \notin C(y), x \notin C(w)).$$

Divide by the common (positive) probability $P_p(C(x) = A, x \leftrightarrow z, x \notin C(y), x \notin C(w))$ to obtain the bound (23.11). Then summing over such lattice animals A completes the proof of (23.10) and thus the BK inequality follows. ∎

Lemma 7.

$$\mathrm{Cov}(\mathbf{1}_{[x \leftrightarrow z]}, \mathbf{1}_{[y \leftrightarrow w]}) \le \mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[x \leftrightarrow y]} \mathbf{1}_{[x \leftrightarrow w]}).$$

Proof. Let $\tau(x, z, y, w) := \mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[x \leftrightarrow y]} \mathbf{1}_{[x \leftrightarrow w]})$. Note that

$$\mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[y \leftrightarrow w]}) = \tau(x, z, y, w) + \mathbb{E}(\mathbf{1}_{[x \leftrightarrow z]} \mathbf{1}_{[x \notin C(y)]} \mathbf{1}_{[y \leftrightarrow w]})$$
$$= \tau(x, z, y, w) + P_p([x \leftrightarrow z] \circ [y \leftrightarrow w]). \qquad (23.12)$$

Now apply the BK inequality to the second term. Subtracting $\mathbb{E}\mathbf{1}_{[x \leftrightarrow z]} \mathbb{E}\mathbf{1}_{[y \leftrightarrow w]}$ from both sides establishes the assertion of the lemma. ∎

Lemma 8. Let $U_x = |C(x)| = \sum_z \mathbf{1}_{[x \leftrightarrow z]}$. Then

$$\gamma = \sum_{x \in \mathbb{Z}^2} \mathrm{Cov}(U_0, U_x) \le \mathbb{E}|C(0)|^3.$$

Proof. Using bi-linearity of covariance and the bound from the first lemma,

$$\mathrm{Cov}(\mathbf{1}_{[0 \leftrightarrow w]}, \mathbf{1}_{[y \leftrightarrow z]}) \le \tau(0, w, y, z) = \mathbb{E}(\mathbf{1}_{[0 \leftrightarrow w]} \mathbf{1}_{[w \leftrightarrow y]} \mathbf{1}_{[y \leftrightarrow z]}),$$

it follows that

$$\sum_{y \in \mathbb{Z}^2} \mathrm{Cov}(U_0, U_y) \le \sum_{y,w,z} \tau(0, w, y, z) = \sum_{y,w,z} \mathbb{E}(\mathbf{1}_{[0 \leftrightarrow w]} \mathbf{1}_{[0 \leftrightarrow y]} \mathbf{1}_{[0 \leftrightarrow z]})$$

$$= \mathbb{E} \left(\sum_{y,w,z} \mathbf{1}_{[0 \leftrightarrow w]} \mathbf{1}_{[0 \leftrightarrow y]} \mathbf{1}_{[0 \leftrightarrow z]} \right) = \mathbb{E}|C(0)|^3.$$

∎

Lemma 9. $\mathbb{E}|C(0)|^m < \infty$ for all $m \geq 0$.

Proof. Note from the proof of Proposition 23.7 that for $p < 1/3$, $\tau(0, x) = P_p(0 \leftrightarrow x) \leq \frac{4}{3} e^{-c|x|}$, where $c = -\ln(3p) > 0$. Thus, denoting by R_k the complimentary region to the (two-dimensional) square of side-lengths $2k + 1$ centered at 0, one has the tail probability bound

$$P_p(|C(0)| \geq (2k+1)^2) \leq \sum_{x \in R_k} \tau(0, x) \leq \frac{4}{3} \sum_{x \in R_k} e^{-c|x|}$$

$$\leq c' \sum_{j=k+1}^{\infty} j e^{-cj} \leq c'' k e^{-c''k},$$

for a suitable $c'' > 0$. The second to the last inequality is a consequence of summing over x on the perimeters at respective distances j from the origin, noting that the number of sites on the perimeter is linear in j. It now follows that $\sum_{k=1}^{\infty} k^{2m-1} P(\sqrt{|C(0)|} \geq k) < \infty$, and therefore $\mathbb{E}|C(0)|^m = \mathbb{E}(\sqrt{|C(0)|})^{2m} < \infty$. ∎

In view of Newman's central limit theorem this series of lemmas establishes the following fluctuation law.

Theorem 23.8. Consider two-dimensional bond percolation with $0 < p < 1/3$. Then the centered and rescaled cumulative size $\frac{1}{N} \sum_{x \in B_0^{(N)}} \{|C(x)| - \mathbb{E}|C(0)|\}$ of all clusters connected to points in the cube $B_0^{(N)}$ is asymptotically normal with mean zero and variance $0 < \gamma = \sum_{x \in \mathbb{Z}^2} \text{Cov}(|C(0)|, |C(x)|) \leq \mathbb{E}|C(0)|^3 < \infty$.

We close this chapter with a celebrated result of Loren Pitt on association of positively correlated normal random variables. We provide the essence of his[18] very clever proof leaving the technical details to exercises.

Theorem 23.9 (Pitt). Let $X = (X_1, \ldots, X_k)$ be a positively correlated normal random vector. Then $\{X_1, \ldots, X_k\}$ is an associated family.

Proof. First consider the case in which the covariance matrix $\Gamma = ((\gamma_{i,j}))$ is non-singular matrix with nonnegative entries. One may show that the collection of coordinatewise non-decreasing functions f, g on $\mathbb{R}^k$ that are continuously differentiable with bounded partials is association determining (Exercise 12). As

---

[18] Pitt (1982).

a result, one may restrict to such functions f, g. Let $Z = (Z_1, \ldots, Z_k)$ be an independent copy of X and define

$$Y(\lambda) = \lambda X + (1 - \lambda^2)^{\frac{1}{2}} Z.$$

Then, $Y(\lambda)$ is mean-zero normal with covariance matrix

$$\mathrm{Cov}(Y_i(\lambda), Y_j(\lambda)) = \lambda^2 \gamma_{i,j} + (1 - \lambda^2) \gamma_{i,j} = \gamma_{i,j}.$$

Also, $\mathrm{Cov}(X, Y(\lambda))_{i,j} = \lambda \gamma_{i,j}$. Consider

$$F(\lambda) = \mathbb{E} f(X) g(Y(\lambda)).$$

Then, $\mathrm{Cov}(f(X), g(X)) = F(1) - F(0)$. So it suffices to show $F'(\lambda)$ exists and is positive for $0 \le \lambda < 1$. This is where the analysis is required. Namely, writing $\Gamma^{-1} = ((c_{i,j}))$, let

$$\varphi(x) = (2\pi)^{-\frac{k}{2}} (\det \Gamma)^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \sum_{i,j=1}^{k} c_{i,j} x_i x_j \right\}$$

denote the Gaussian pdf[19] of X. Then the conditional pdf of $Y(\lambda)$ given $[X = x]$ is

$$p(\lambda; x, y) = (1 - \lambda^2)^{-\frac{k}{2}} \varphi((1 - \lambda^2)^{-\frac{1}{2}} (y - \lambda x)).$$

That is, $p(\lambda; x, y)$ is normal with covariance matrix $(1 - \lambda^2)\Gamma$ and mean vector λx. Thus,

$$F(\lambda) = \int_{\mathbb{R}^k} f(x) \varphi(x) g(\lambda, x) dx,$$

where

$$g(\lambda, x) = \int_{\mathbb{R}^k} g(y) p(\lambda; x, y) dy.$$

Observing that

$$g(\lambda, x) = \varphi(\lambda, \cdot) * g(\lambda x),$$

where $\varphi(\lambda, x) = (1 - \lambda^2)^{-\frac{k}{2}} \varphi((1 - \lambda^2)^{-\frac{1}{2}} x)$, one sees that $\partial g(\lambda, x)/\partial \lambda$ exists and is nonnegative, while $\partial g(\lambda, x)/\partial x_j$ exists and is bounded. To compute $\frac{\partial p}{\partial \lambda}$, let $h(t, y)$

---

[19] BCPT, p. 130.

denote the pdf of $\Gamma^{\frac{1}{2}} B_t$, where B is k-dimensional standard Brownian motion. Then one has

$$p(\lambda; x, y) = h(1 - \lambda^2, y - \lambda x).$$

Using the chain rule and an application of the heat equation for multivariate Brownian motion[20] one arrives at

$$\frac{\partial p}{d\lambda} = -\frac{1}{\lambda} \left\{ \sum_{i \neq j} \gamma_{i,j} \frac{\partial p}{\partial x_i, \partial x_j} - \sum_i x_i \frac{\partial p}{\partial x_i} \right\}.$$

Thus,

$$F'(\lambda) = -\frac{1}{\lambda} \int_{\mathbb{R}^k} f(x) \varphi(x) \left\{ \sum_{i \neq j} \gamma_{i,j} \frac{\partial g(\lambda, x)}{\partial x_i, \partial x_j} - \sum_i x_i \frac{\partial g(\lambda, x)}{\partial x_i} \right\} dx.$$

Finally, with an integration by parts one arrives at

$$F'(\lambda) = \frac{1}{\lambda} \int_{\mathbb{R}^k} \varphi(x) \left\{ \sum_{i \neq j} \gamma_{i,j} \frac{\partial f(x)}{\partial x_i} \frac{\partial g(\lambda, x)}{\partial x_j} \right\} dx \geq 0.$$

In the case Γ is singular one may replace Γ by the non-singular matrix $\Gamma + \epsilon \mathbf{1}_{k \times k}$, $\epsilon > 0$, and observe that for continuous f, g, $\mathrm{Cov}(f(X), g(X))$ depends continuously on Γ. Thus positivity is preserved in the limit as $\epsilon \to 0$. ∎

Exercises

1. Let $X_1, \ldots, X_n$ be associated random variables, and $Y_j = f_j(X_1, \ldots, X_n)$, $j = 1, \ldots, m$ where f_j is coordinatewise non-decreasing for $j = 1, \ldots, m$. Show that (a) $P(Y_1 \leq y_1, \ldots, Y_k \leq y_k) \geq \prod_{j=1}^k P(Y_j \leq y_j)$, and (b) $P(Y_1 > y_1, \ldots, Y_k > y_k) \geq \prod_{j=1}^k P(Y_j > y_j)$. [*Hint*: Define non-decreasing functions of Z_j's by $Z_j = \mathbf{1}_{[Y_j > y_j]}$, $j = 1, \ldots, m$, and note that $Z_1 \cdots Z_i$, and $Z_{i+1} \cdots Z_m$ are non-decreasing functions of Z_j's. Apply the FKG inequalities iteratively for $i = 1, \ldots, m$, noting $\mathbb{E} Z_j = P(Z_j = 1)$. (c) Suppose that $X_1, \ldots X_n$ are independent random variables and let $S_j = \sum_{i=1}^j X_i$, $j = 1, \ldots, n$. Show that $P(S_1 \leq s_1, \ldots, S_n \leq s_n) \geq \prod_{i=1}^n P(S_i \leq s_i)$ for

[20] For the heat equation connection see Bhattacharya and Waymire (2021), Chapter 6, Remark 6.2. Pitt uses a related computational device of Plackett (1954).

all $s_1, \ldots, s_n$. [*Hint*: S_j is a non-decreasing function of $X_1, \ldots, X_n$, and independent random variables are associated.]

2. Suppose f is a continuous function on $[0, 1]$ and consider the *Bernstein polynomial* defined by $f_n(x) = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} = \mathbb{E} f(\frac{S_n(x)}{n}), 0 \leq x \leq 1$, where $S_j(x) = X_1(x) + \cdots X_j(x)$, for i.i.d. Bernoulli $0 - 1$-valued random variables with $P(X_j(x) = 1) = x, j = 1, \ldots, n$. (a) Show that $f_n \to f$ uniformly on $[0, 1]$ as $n \to \infty$. (b)(*Seymour-Welsh*) Show[21] that for non-decreasing f, g on $[0, 1], (fg)_n(x) \geq f_n(x)g_n(x), 0 \leq x \leq 1$.

3. Show that binary $0 - 1$-valued random variables X, Y are associated if and only if they are positively quadrant dependent.

4. Complete the details for the extension of Hoeffding's lemma used in the generalization Lemma 2.

5. Suppose that $X = (X_1, \ldots, X_m)$ is a vector of associated random variables. Let f, g be, respectively, coordinatewise increasing and decreasing functions. Show that $\mathrm{Cov}(f(X), g(X)) \leq 0$. Extend this to countably many associated random variables.

6. Prove the alternative formula

$$H_{X,Y}(x, y) = P(X \leq x, Y \leq y) - P(X \leq x)P(Y \leq y).$$

7. Suppose $f(X), g(Y) \in L^2(\Omega, \mathcal{F}, P)$, where X, Y are real-valued random variables bounded below by a constant $b \in \mathbb{R}$, and f, g are continuously differentiable complex functions with bounded derivatives. Prove the Hoeffding-Newman formula in this case, starting from the familiar moment formulae

$$\mathbb{E}(f(X) - f(b)) = \mathbb{E} \int_b^X f'(x)dx = \int_b^\infty P(X > x)f'(x)dx,$$

and

$$\mathbb{E}(f(X) - f(b))(\overline{g(Y) - g(b)}) = \mathbb{E} \int_b^X \int_b^Y f'(x)\overline{g}'(y)dxdy.$$

8. Show that each of the collections of non-decreasing binary 0 or 1-valued functions and those of non-decreasing bounded continuous functions are association determining.

9. Consider the spatial intermittency of clusters as reflected in the density of isolated points and/or non-isolated points. A site $x \in \mathbb{Z}^2$ is *isolated* whenever $[C(x) = \emptyset]$. The numbers of isolated points and non-isolated points in a square $B_0^{(N)}$ are perfectly correlated since their total is (fixed) $|B_0^{(N)}|$. It is convenient to consider the number of non-isolated sites in the square $B_0^{(N)}$ (including surface sites for simplicity) as given by

---

[21] Seymour and Welsh (1975).

$$D_N = \sum_{x \in B_0^{(N)}} \mathbf{1}_{[C(x) \neq \emptyset]}.$$

Show that $\mathbb{E} D_N = (1 - q^4) N^2$ and compute the asymptotic distribution of D_N, suitably centered and scaled, as $N \to \infty$.

10. Fix a positive integer k and obtain the asymptotic fluctuation law for the numbers of sites x in $B_0^{(N)}$ belonging to a cluster of size at most k, i.e., such that $|C(x)| \leq k$.

11. Let $N = \{N(A) : A \in \mathcal{B}\}$ be a Poisson point process on $\mathbb{R}^n$. Show that N is an associated family of random variables.

12. Show that the collection of coordinatewise non-decreasing functions f, g on $\mathbb{R}^k$ that are continuously differentiable with bounded partial derivatives is association determining. [*Hint:* (a) Check that any measurable increasing functions is a pointwise a.s. limit of continuous increasing functions. (b) Check that if ρ_ϵ, $\epsilon > 0$ is a nonnegative C^∞-mollifier,[22] then $f * \rho_\epsilon$ is C^∞ increasing with bounded partials such that $f * \rho_\epsilon \to f$ pointwise, uniform boundedly.]

---

[22] See BCPT, p.77, for a mollifier example.