Stein's method for positively associated random variables with applications to the Ising and voter models, bond percolation, and contact process

Larry Goldstein[1] and Nathakhun Wiroonsri

University of Southern California, Kaprielian Hall, Room 108, 3620 Vermont Avenue, Los Angeles, CA 90089-2532, USA.
E-mail: larry@math.usc.edu; wiroonsr@usc.edu

Abstract. We provide nonasymptotic L^1 bounds to the normal for four well-known models in statistical physics and particle systems in $\mathbb{Z}^d$; the ferromagnetic nearest-neighbor Ising model, the supercritical bond percolation model, the voter model and the contact process. In the Ising model, we obtain an L^1 distance bound between the total magnetization and the normal distribution at any temperature when the magnetic moment parameter is nonzero, and when the inverse temperature is below critical and the magnetic moment parameter is zero. In the percolation model we obtain such a bound for the total number of points in a finite region belonging to an infinite cluster in dimensions $d \geq 2$, in the voter model for the occupation time of the origin in dimensions $d \geq 7$, and for finite time integrals of nonconstant increasing cylindrical functions evaluated on the one dimensional supercritical contact process started in its unique invariant distribution.

The tool developed for these purposes is a version of Stein's method adapted to positively associated random variables. In one dimension, letting $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ be a positively associated mean zero random vector with components that obey the bound $|\xi_i| \leq B, i = 1, \ldots, m$, and whose sum $W = \sum_{i=1}^m \xi_i$ has variance 1, it holds that

$$d_1\big(\mathcal{L}(W), \mathcal{L}(Z)\big) \leq 5B + \sqrt{\frac{8}{\pi}} \sum_{i \neq j} \mathbb{E}[\xi_i \xi_j],$$

where Z has the standard normal distribution and $d_1(\cdot, \cdot)$ is the L^1 metric. Our methods apply in the multidimensional case with the L^1 metric replaced by a smooth function metric.

Résumé. Nous obtenons des bornes L^1 non asymptotiques montrant l'approximation gaussienne pour quatre modèles classiques de physique statistique et de systèmes de particules dans $\mathbb{Z}^d$: le modèle d'Ising ferromagnétique au plus proche voisin, la percolation par arêtes sur-critique, le modèle du votant et le processus de contact. Pour le modèle d'Ising, nous obtenons une borne en distance L^1 entre la magnétisation totale et la loi normale, soit à toute température lorsque le paramètre de moment magnétique est non nul, soit, dans la phase sous-critique, si ce paramètre est zéro. Pour le modèle de percolation, nous obtenons une telle borne pour le nombre total de points dans une région finie qui appartiennent à un cluster infini, en dimensions $d \geq 2$. Pour le modèle du votant nous considérons le temps d'occupation à l'origine en dimensions $d \geq 7$, et pour le processus de contact, nous nous intéressons à des intégrales temporelles à horizon fini de fonctions cylindriques croissantes, évaluées en le processus de contact unidimensionnel sous l'unique mesure invariante.

L'outil que nous développons à cette fin est une version de la méthode de Stein adaptée à des variables aléatoires associées positivement. En dimension 1, si $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ désigne un vecteur aléatoire centré et positivement associé dont les composantes

satisfont la borne $|\xi_i| \leq B$, $i = 1, \ldots, m$, et dont la somme $W = \sum_{i=1}^{m} \xi_i$ est de variance 1, alors

$$d_1\big(\mathcal{L}(W), \mathcal{L}(Z)\big) \leq 5B + \sqrt{\frac{8}{\pi}} \sum_{i \neq j} \mathbb{E}[\xi_i \xi_j]$$

où Z est une variable aléatoire gaussienne centrée réduite et $d_1(\cdot, \cdot)$ est la distance L^1. Nos méthodes s'appliquent au cas multidimensionnel si la distance L^1 est remplacée par une distance plus faible.

MSC: 60F05; 82B30; 60G60

Keywords: Random fields; Block dependence; Correlation inequality; Positive dependence

---

1. Introduction

We provide nonasymptotic L^1 bounds to the normal for four well-known models in statistical physics and particle systems in $\mathbb{Z}^d$; the ferromagnetic nearest-neighbor Ising model, supercritical bond percolation, the voter model and the contact process. We recall that the L^1, or Wasserstein, distance between the distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ of real valued random variables X and Y is given by

$$d_1\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \int_{-\infty}^{\infty} \big| P(X \leq t) - P(Y \leq t) \big| \, dt. \tag{1}$$

Taking the Ising model first, questions regarding the distribution of total magnetization in statistical mechanical models have attracted attention for quite some time. To fix terms, with $\mathbb{Z}$ the set of integers and $k \in \mathbb{Z}$ let $\mathbb{N}_k = [k, \infty) \cap \mathbb{Z}$, and for d a positive integer let $\mathbb{Z}^d$ denote the d-dimensional lattice composed of vectors $\mathbf{i} = (i_1, \ldots, i_d)$ having integer components, and let Λ be a symmetric hypercube, that is,

$$\Lambda = \big\{ \mathbf{i} \in \mathbb{Z}^d : |\mathbf{i}|_\infty \leq k \big\} \quad \text{for some } k \in \mathbb{N}_0,$$

where $|\mathbf{i}|_\infty = \sup_{1 \leq j \leq d} |i_j|$, the supremum norm. We say $\omega = \{\omega_\mathbf{i} : \mathbf{i} \in \Lambda\}$ is a configuration on Λ when $\omega_\mathbf{i}$, denoting the magnetic spin at site $\mathbf{i}$, takes values in the set $\{-1, +1\}$ for each $\mathbf{i} \in \Lambda$.

We define a probability measure on configurations on Λ, corresponding to the nearest-neighbor ferromagnetic Ising model. First, fix a configuration $\widetilde{\omega} = \{\widetilde{\omega}_\mathbf{i} : \mathbf{i} \in \mathbb{Z}^d\}$ on all of $\mathbb{Z}^d$, and the interaction strength $J \geq 0$. Now define the energy function, or Hamiltonian $H_{\Lambda, h, \widetilde{\omega}}$, by

$$H_{\Lambda, h, \widetilde{\omega}}(\omega) = -J \bigg(\sum_{\substack{\mathbf{i}, \mathbf{j} \in \Lambda \\ |\mathbf{i} - \mathbf{j}|_1 = 1}} \omega_\mathbf{i} \omega_\mathbf{j} + \sum_{\substack{\mathbf{i} \in \Lambda, \mathbf{j} \notin \Lambda \\ |\mathbf{i} - \mathbf{j}|_1 = 1}} \omega_\mathbf{i} \widetilde{\omega}_\mathbf{j} \bigg) - h \sum_{\mathbf{i} \in \Lambda} \omega_\mathbf{i}, \tag{2}$$

where h is the magnetic moment parameter and $|\mathbf{x}|_1 = \sum_{i=1}^{d} |x_i|$ is the L^1 norm of $\mathbf{x} \in \mathbb{R}^d$. Given an inverse temperature $\beta > 0$, the finite volume Ising probability measure on Λ is given by

$$\mu_{\Lambda, \beta, h, \widetilde{\omega}}(\omega) = \exp\big(-\beta H_{\Lambda, h, \widetilde{\omega}}(\omega)\big) / Z_{\Lambda, \beta, h, \widetilde{\omega}},$$

for a normalizing constant $Z_{\Lambda, \beta, h, \widetilde{\omega}}$. The cases where $\widetilde{\omega}_\mathbf{j}$ takes the value $0, +1$ and -1 for all $\mathbf{j} \in \mathbb{Z}^d$ are typically referred to as *free, positive, and negative boundary conditions*, respectively, see [20], for example.

Now let $\{\Lambda_n : n \in \mathbb{N}_1\}$ be any increasing sequence of symmetric hypercubes of $\mathbb{Z}^d$ whose union is $\mathbb{Z}^d$. For fixed β and h, we say that $\mu_{\Lambda_n, \beta, h, \widetilde{\omega}}$ converges *weakly* to $\mu_{\beta, h, \widetilde{\omega}}$ as $n \to \infty$ if

$$\lim_{n \to \infty} \int f \, d\mu_{\Lambda_n, \beta, h, \widetilde{\omega}} = \int f \, d\mu_{\beta, h, \widetilde{\omega}}$$

for all local functions $f : \mathbb{R}^d \to \mathbb{R}$, where we recall that a function f is said to be *local* if it depends only on Λ_n for some n. Each such weak limit is a probability measure on $\{-1, 1\}^{\mathbb{Z}^d}$. The set of infinite volume Gibbs measures, denoted $\mathcal{G}_{\beta, h}$, is the closed convex hull of all weak limits of $\mu_{\Lambda_n, \beta, h, \widetilde{\omega}}$ for $\widetilde{\omega}$ any configuration on $\mathbb{Z}^d$.

The structure of $\mathcal{G}_{\beta,h}$ depends on parameters β and h. In particular, by [20], there exists a critical inverse temperature β_c such that the limit measures depend, or do not depend, on the boundary conditions when $\beta > \beta_c$ or $\beta < \beta_c$, respectively. More precisely, for $\beta > 0, h \neq 0$ and $0 < \beta < \beta_c, h = 0$, the finite volume measures $\mu_{\Lambda_n,\beta,h,\widetilde{\omega}}$ have a unique weak limit for any choice of boundary conditions $\widetilde{\omega}$. Thus, in these cases, $\mathcal{G}_{\beta,h}$ consists of a unique measure, which we denote $\mu_{\beta,h}$. In the case $\beta > \beta_c, h = 0$, the set $\mathcal{G}_{\beta,0}$ contains two *pure phases*, $\mu_{\beta,0}^+$ and $\mu_{\beta,0}^-$, which are the infinite volume measures arising as weak limits under positive and negative boundary conditions respectively. Hence, by definition, $\mathcal{G}_{\beta,0}$ also contains the convex combinations $\mu_{\beta,0}^{(\alpha)} = \alpha\mu_{\beta,0}^+ + (1-\alpha)\mu_{\beta,0}^-, 0 < \alpha < 1$. These measures are called *mixed phases*. For $d = 2$, $\mathcal{G}_{\beta,0}$ consists of only pure and mixed phases. For $d \geq 3$, the set $\mathcal{G}_{\beta,0}$ contains nontranslation invariant measures for any sufficiently large β, which are not considered in this paper. For additional detail about finite and infinite volume Gibbs measures, see [20].

In this work, we focus on the Ising models on $\mathbb{Z}^d$ in the cases where the weak limit is unique, precisely, for measures in the set

$$\mathcal{M}_1 = \left\{ \mu_{\beta,h} : (\beta, h) \in \mathbb{R}^+ \times \mathbb{R} \setminus \{0\} \cup (0, \beta_c) \times \{0\} \right\}. \tag{3}$$

For these measures, we obtain L^1 bounds to the normal for the distribution of the total magnetization

$$M_{\mathbf{k}} = \sum_{\mathbf{i} \in B_{\mathbf{k}}^n} \omega_{\mathbf{i}} \tag{4}$$

over the finite block $B_{\mathbf{k}}^n \subset \mathbb{Z}^d$ with side length n "anchored" at $\mathbf{k} \in \mathbb{Z}^d$, defined in (15). See Remark 2.8 for a discussion regarding the applicability of our results to the pure or mixed phases when the inverse temperature is above critical, that is, for measures in the set

$$\mathcal{M}_2 = \begin{cases} \mu_{\beta,0}^+, \mu_{\beta,0}^- & : & \beta > \beta_c, h = 0 \\ \mu_{\beta,0}^{(\alpha)} & : & 0 < \alpha < 1, \beta > \beta_c, h = 0 \end{cases}. \tag{5}$$

The second model we consider is bond percolation on $\mathbb{Z}^d$ for $d \geq 2$ (see [24], for example). In graphical terms, nearest neighbor bonds are the edges between vertices $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{Z}^d$ satisfying $|\mathbf{x} - \mathbf{y}|_1 = 1$; we denote the collection of all such bonds by $\mathbb{E}^d$. For a given $\theta \in [0, 1]$, we declare each bond in $\mathbb{E}^d$ to be *open* with probability θ and *closed* otherwise, independently of all other bonds. More formally, we take $\mathbf{G} = \{0, 1\}^{\mathbb{E}^d}$, elements of which are represented as $\mathbf{g} = \{g(\mathbf{e}) : \mathbf{e} \in \mathbb{E}^d\}$, and called configurations. The value $g(\mathbf{e}) = 0$ corresponds to the bond $\mathbf{e}$ being closed and $g(\mathbf{e}) = 1$ corresponds to $\mathbf{e}$ being open.

Given a bond configuration $\mathbf{g}$ and $\mathbf{x} \in \mathbb{Z}^d$, let $\mathcal{C}(\mathbf{x})$ be the set of vertices connected to $\mathbf{x}$ through open bonds, and let $|\mathcal{C}(\mathbf{x})|$ denote the number of vertices in $\mathcal{C}(\mathbf{x})$, which may be infinite. Since the probability θ that a bond is open is the same for every $\mathbf{e} \in \mathbb{E}^d$ and each bond is independent of all others, $\mathcal{C}(\mathbf{x})$ has the same distribution for all $\mathbf{x} \in \mathbb{R}^d$. We define

$$\rho(\theta) = P\left(|\mathcal{C}(\mathbf{0})| = \infty \right) \quad \text{and} \quad \theta_c = \sup\{\theta : \rho(\theta) = 0\},$$

respectively representing the probability that for a given connectivity θ a vertex belongs to an infinite cluster, and its threshold, the *critical probability*. Our interest is in obtaining L^1 bounds to the normal, in the supercritical case $\theta > \theta_c$, for the distribution of the total number of points in a finite block that belong to an infinite cluster, specifically for

$$U_{\mathbf{k}} = \sum_{\mathbf{x} \in B_{\mathbf{k}}^n} \mathbf{1}_{\{|\mathcal{C}(\mathbf{x})|=\infty\}}, \tag{6}$$

where $B_{\mathbf{k}}^n \subset \mathbb{Z}^d$ is defined in (15).

The third application we consider is the d-dimensional voter model introduced in [14], a Markov process $\{\eta_t : t \geq 0\}$ taking values in $\{0, 1\}^{\mathbb{Z}^d}$, and one of the simplest interacting particle systems, see also [12]. With a precise formulation deferred to Section 2.3, the process is specified by taking the initial distribution $\{\eta_0(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\}$ to be

a family of independent, identically distributed Bernoulli random variables with parameter $\theta \in (0, 1)$, and having transitions

$$\eta_t(\mathbf{x}) \to 1 - \eta_t(\mathbf{x}) \quad \text{at rate } \frac{1}{2d} \big| \{\mathbf{y} : |\mathbf{y} - \mathbf{x}|_1 = 1 \text{ and } \eta_t(\mathbf{x}) \neq \eta_t(\mathbf{y})\} \big|.$$

That is, a site $\mathbf{x} \in \mathbb{Z}^d$ changes its state with rate equal to the fraction of its $2d$ nearest neighbors that are in the opposite state. We study the occupation time

$$T_s^t = \int_s^{s+t} \eta_u(\mathbf{0}) \, du \quad \text{for } t > 0, \, s \geq 0, \tag{7}$$

that is, the amount of time in $(s, s + t]$ that the origin $\mathbf{0}$ spends in state 1.

The last model we consider is the one dimensional contact process $\{\zeta^{\nu_\lambda}(t) : t \geq 0\}$ introduced in [26], a continuous time Markov process with state space $\mathscr{P}(\mathbb{Z})$, the set of all subsets of $\mathbb{Z}$. The subset $\zeta^{\nu_\lambda}(t)$ models the collection of "infected" individuals at time t, where each infected site infects healthy neighbors at rate λ, and infected sites recover at rate 1. The distribution of $\zeta^{\nu_\lambda}(0)$ is ν_λ, the unique invariant measure defined in [16] as the limiting distribution of $\zeta^{\mathbb{Z}}(t)$,

$$\zeta^{\mathbb{Z}}(t) \xrightarrow{d} \nu_\lambda \quad \text{as } t \to \infty,$$

where $\zeta^{\mathbb{Z}}(t)$ has initial distribution putting mass one on all of $\mathbb{Z}$.

Define

$$\lambda_* = \sup\{\lambda > 0 : \zeta^{\mathbb{Z}}(t) \to \delta_\varnothing \text{ weakly as } t \to \infty\}. \tag{8}$$

The contact process is said to be supercritical when $\lambda > \lambda_*$, and, by [16], in this case ν_λ is the nontrivial, unique invariant measure of the process; $\nu_\lambda = \delta_\varnothing$ when $\lambda < \lambda_*$.

Recall that a function $f : \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$ is said to be *cylindrical* if it depends on only finitely many sites and *increasing* if $f(A) \leq f(B)$ whenever $A \subset B \subset \mathbb{Z}$. In the supercritical case, we study the cumulative value of a cylindrical f evaluated on the process over the interval $(s, s + t]$,

$$D_{s,f}^t = \int_s^{s+t} f(\zeta^{\nu_\lambda}(u)) \, du \quad \text{for } s \geq 0, t > 0. \tag{9}$$

For instance, for B a finite subset of $\mathbb{Z}$, letting $I_{\overline{B}}(\cdot)$ be the cylindrical, increasing indicator function given by

$$I_{\overline{B}}(\eta) = \mathbf{1}(\eta \cap B \neq \varnothing) \quad \text{for } \eta \subset \mathbb{Z}, \tag{10}$$

the value $D_{s, I_{\overline{B}}}^t$ yields the amount of time in $(s, s + t]$ that B contains at least one infected site.

The quantities (4), (6), (7) and (9) that we consider in these four models are obtained by summing or integrating positively associated variables, and we obtain our results by developing an L^1 version of Stein's method for such instances. We recall that a random vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^m$ is said to be *positively associated* whenever

$$\text{Cov}(\psi(\boldsymbol{\xi}), \phi(\boldsymbol{\xi})) \geq 0 \tag{11}$$

for all real valued coordinate-wise nondecreasing functions ψ and ϕ on $\mathbb{R}^m$ such that $\psi(\boldsymbol{\xi})$ and $\phi(\boldsymbol{\xi})$ posses second moments. In general, a collection $\{\xi_\alpha : \alpha \in I\}$ of real valued random variables indexed by a set I is said to be positively associated if all finite subcollections are positively associated. Positive association was introduced in [21] and has been found frequently in probabilistic models in several areas, especially statistical physics. In some literature (11) is termed the FKG-inequality, see [33] for example.

For the Ising model with measure in $\mathcal{M}_1 \cup \mathcal{M}_2$ as given in (3) and (5), it is well-known that the variables making up the configuration $\{\omega_{\mathbf{i}} : \mathbf{i} \in \mathbb{Z}^d\}$ are positively associated whenever $J \geq 0$. For the percolation model, it is also well-known that $\{\mathbf{1}_{\{|\mathcal{C}(\mathbf{x})| = \infty\}} : \mathbf{x} \in \mathbb{Z}^d\}$ is positively associated (see [33] for both models). For the voter model, taking $t > 0$,

$m \in \mathbb{N}_1$ and

$$X_{s,i}^t = \int_{s+(i-1)t/m}^{s+it/m} \eta_u(\mathbf{0}) \, du \quad \text{for } i \in [m], \text{ we have } T_s^t = \sum_{i=1}^m X_{s,i}^t, \tag{12}$$

where $[m] = \{1, 2, \ldots, m\}$. It was demonstrated in [13] that the family $\{X_{s,i}^t : i \in [m]\}$ is positively associated. For the contact process, similarly, with $t > 0$, $m \in \mathbb{N}_1$ and

$$Y_{s,i}^{t,m} = \int_{s+(i-1)t/m}^{s+it/m} f\big(\zeta^{v_\lambda}(u)\big) \, du \quad \text{for } i \in [m], \text{ we have } D_{s,f}^t = \sum_{i=1}^m Y_{s,i}^{t,m}. \tag{13}$$

We show in Lemma 2.18, using a simple consequence from arguments in the proof of Lemma 1 of [38], that the family $\{Y_{s,i}^{t,m} : i \in [m]\}$ is positively associated when f is increasing. Hence in each of the models we consider the quantity we study can be represented as the sum of positively associated random variables.

Over the last few decades, several papers established central limit theorems and rates of convergence for sums of positively associated random variables under various assumptions. To state some of these results, recall that a random field $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ is called *second order stationary* when $E X_{\mathbf{j}}^2 < \infty$ for all $\mathbf{j} \in \mathbb{Z}^d$ and the covariance $\text{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}})$ depends only on $\mathbf{j} - \mathbf{i}$, that is, $\text{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}}) = R(\mathbf{j} - \mathbf{i})$ for all $\mathbf{i}, \mathbf{j} \in \mathbb{R}^d$, with $R(\cdot)$ necessarily given by

$$R(\mathbf{k}) = \text{Cov}(X_{\mathbf{0}}, X_{\mathbf{k}}). \tag{14}$$

We say the field is *strictly stationary* if for all $m \in \mathbb{N}_1$ and $\mathbf{k}, \mathbf{j}_1, \ldots, \mathbf{j}_m \in \mathbb{Z}^d$, the vector $(X_{\mathbf{j}_1}, \ldots, X_{\mathbf{j}_m})$ has the same distribution as $(X_{\mathbf{j}_1+\mathbf{k}}, \ldots, X_{\mathbf{j}_m+\mathbf{k}})$. From the definitions, when second moments exist, strict stationarity implies second order stationarity, and if a second order stationary field is positively associated, then $R(\mathbf{k}) \geq 0$ for all $\mathbf{k} \in \mathbb{Z}^d$.

We let $\mathbf{1} \in \mathbb{Z}^d$ denote the vector with all components 1, and write inequalities such as $\mathbf{a} < \mathbf{b}$ for vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ when they hold componentwise. For $\mathbf{k} \in \mathbb{Z}^d$, $n \in \mathbb{N}_1$ and a random field $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$, define the "block sum" variables, over a block with side length n, by

$$S_{\mathbf{k}}^n = \sum_{\mathbf{j} \in B_{\mathbf{k}}^n} X_{\mathbf{j}} \quad \text{where } B_{\mathbf{k}}^n = \big\{\mathbf{j} \in \mathbb{Z}^d : \mathbf{k} \leq \mathbf{j} < \mathbf{k} + n\mathbf{1}\big\}. \tag{15}$$

Note that $B_{\mathbf{k}}^n = B_{\mathbf{0}}^n + \mathbf{k}$.

For a second order stationary field with $R(\cdot)$ given by (14), we have

$$\text{Var}\big(S_{\mathbf{k}}^n\big) = \sum_{\mathbf{i}, \mathbf{j} \in B_{\mathbf{k}}^n} \text{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}}) = n^d A_n \quad \text{where } A_n = \frac{1}{n^d} \sum_{\mathbf{i}, \mathbf{j} \in B_{\mathbf{1}}^n} R(\mathbf{i} - \mathbf{j}). \tag{16}$$

With

$$A = \sum_{\mathbf{k} \in \mathbb{Z}^d} R(\mathbf{k}), \tag{17}$$

Lemma 4 of [33] shows $R(\mathbf{0}) < \infty$, $R(\mathbf{k}) \geq 0$ and $A < \infty$ imply $\lim_{n \to \infty} A_n = A$. As A_n or A equals zero only if the field is trivial, we assume without loss of generality in the following that $A_n > 0$ for all $n \in \mathbb{N}_1$ and $A > 0$.

Theorem 1.1 of [33] shows that the collection of properly standardized block sums $S_{\mathbf{k}}^n$ of a strictly stationary, positively associated random field with covariance given by (14) converges jointly to independent normal variables when $A < \infty$.

Theorem 1.1 ([33]). *Let $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ be a positively associated strictly stationary random field with finite second moments and covariance $R(\mathbf{k})$ given by* (14) *that satisfies $A < \infty$ where A is given in* (17). *Then the standardized block sums*

$$\left\{ \frac{S_{n\mathbf{k}}^n - \mathbb{E}S_{n\mathbf{k}}^n}{n^{d/2}} : \mathbf{k} \in \mathbb{Z}^d \right\} \tag{18}$$

converge to independent mean zero normal distributions with variance A as $n \to \infty$. That is, the expectations of bounded continuous functions of (18) *that depend on only finitely many coordinates converge to the expectation of that same number of independent mean zero, variance A normal variables.*

We note that the blocks $B_{n\mathbf{k}}^n$ over which sums are taken in Theorem 1.1 are disjoint, and that the limiting distribution has independent coordinates. As an application of our Theorem 1.4, in Theorem 2.2 we obtain bounds for the sums over overlapping blocks to an approximating dependent multivariate normal distribution. Rectangular blocks, that is, blocks with varying side lengths, can also be accommodated by our methods, at the cost of some additional complexity in our computations.

The stationarity assumption was relaxed in [13], where it was shown that the block sums of a positively associated random field $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ converge to the normal whenever $X_{\mathbf{j}}$ has finite third moment for every $\mathbf{j} \in \mathbb{Z}^d$ and $u(n)$ converges to zero as $n \to \infty$ where

$$u(n) = \sup_{\mathbf{j} \in \mathbb{Z}^d} \sum_{\mathbf{i} \in \mathbb{Z}^d : |\mathbf{j} - \mathbf{i}|_\infty \geq n} \mathrm{Cov}(X_{\mathbf{i}}, X_{\mathbf{j}}). \tag{19}$$

The rate of convergence for sums of positively associated variables was first obtained in [6], which achieved the rate $\log(n)/\sqrt{n}$ in the Kolmogorov metric in the one dimensional case, assuming uniformly finite $3 + \varepsilon$-moments for some $\varepsilon > 0$, and that $u(n)$ decays at an exponential rate. Theorem 1.2 of [8] generalizes the results in [6]. In the following C will denote a constant whose value may change from line to line. Recall that the Kolmogorov, or L^∞ distance between distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ is given by

$$d_K\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \sup_{t \in \mathbb{R}} \big| P(X \leq t) - P(Y \leq t) \big|.$$

Theorem 1.2 ([8]). *Let $\varepsilon > 0$ and $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ be a mean zero, positively associated random field whose elements have uniformly bounded $3 + \varepsilon$ moments. Assume that there exists $\lambda > 0$ such that $u(n)$ as defined in* (19) *satisfies*

$$u(n) \leq \kappa_0 e^{-\lambda n} \quad \text{for some } \kappa_0 > 0.$$

Then for any finite subset $V \subset \mathbb{Z}^d$ and

$$S(V) = \sum_{\mathbf{j} \in V} X_{\mathbf{j}} \quad \text{with } \sigma^2(V) = \mathrm{Var}\big(S(V)\big),$$

one has

$$d_K\big(\mathcal{L}\big(S(V)/\sigma(V)\big), \mathcal{L}(Z)\big) \leq C|V|\sigma(V)^{-3}\big(\log(|V| + 1)\big)^d,$$

where Z is a standard normal random variable, $|V|$ denotes the size of V and $C > 0$ is a constant depending only on ε, d, κ_0 and λ.

In the present work we first provide bounds in the L^1 distance to the normal for a sum of elements of a positively associated random field in $\mathbb{Z}^d$, under the condition that covariances decay at an exponential rate. We apply these bounds to the total magnetization of the d-dimensional Ising model as defined in (4) and the total number of points belonging to an infinite cluster in the bond percolation model as defined in (6).

We also consider the occupation time of the voter model, where the assumption that the covariance decays exponentially is not satisfied. Without using positive association, [12] showed that the occupation time T_0^t of the voter model as defined in (7) satisfies the CLT for $d \geq 2$. In later work, [13] used the fact that T_0^t is the sum of positively associated random variables and that $u(n)$ as defined in (19) converges to zero as $n \to \infty$ to prove the CLT for $d \geq 5$. However, error bounds and the rate of convergence appear yet unaddressed in the literature. Bounds to the limit in the L^1 metric for $d \geq 7$ are obtained using the technique provided in Theorem 1.3 below.

Next, with definitions as in the Introduction, we provide L^1 bounds for the quantity $D_{s,f}^t$ in (9), an integral over the time interval $(s, s + t]$ of a nonconstant increasing cylindrical function f evaluated on the supercritical one dimensional contact process $\zeta^{\nu_\lambda}(t), t \geq 0$ with initial state having distribution ν_λ. Our argument exploits the exponential

decay of the covariance between $f(\zeta^{\nu_\lambda}(r))$ and $f(\zeta^{\nu_\lambda}(s))$ for any cylindrical f. Reference [38] applied the result in [34] along with the fact that $D_{0,f}^t$ is the sum of positively associated random variables to prove a CLT for $D_{0,f}^t$. Error bounds in Kolmogorov distance, without explicit constants, can be obtained by applying Theorem 1.2 for our applications to the Ising model, bond percolation and the contact process, contingent upon showing that $u(n)$ in (19) decays exponentially.

We obtain L^1 results on the four models considered by developing a version of Stein's method for sums of bounded positively associated random variables. Stein's method was first introduced by Charles Stein in his seminal paper [40] and has become one of the most powerful methods to prove convergence in distribution. Its main advantages are that it provides non asymptotic bounds on the distance between distributions, and that it can handle a number of situations involving dependence. For more detail about the method in general, see the text [11] and the introductory notes [37].

Stein's method has been used previously in several papers in statistical physics. In [19], a version of Stein's method using size bias couplings was developed for a class of discrete Gibbs measures, that is, probability measures on $\mathbb{N}_0$ proportional to $e^{V(k)}$ for some function $V : \mathbb{N}_0 \to \mathbb{R}$. This work provided bounds in the total variation distance between the distribution of certain sums of strongly correlated random variables and discrete Gibbs measures, with applications to interacting particle systems. Stein's method of exchangeable pairs was applied to the classical Curie–Weiss model for high temperatures and at the critical temperature in [17], where optimal Kolmogorov distance bounds were produced. The same result at the critical temperature was also obtained in [9]. These results were extended to the Curie–Weiss–Potts model in [18].

Our main result in the one dimensional case is the following.

Theorem 1.3. *Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m)$ be a positively associated mean zero random vector with components obeying the bound $|\xi_i| \leq B$, and whose sum $W = \sum_{i=1}^m \xi_i$ has variance 1. Let Z be a standard normal random variable. Then*

$$d_1\big(\mathcal{L}(W), \mathcal{L}(Z)\big) \leq 5B + \sqrt{\frac{8}{\pi}} \sum_{i \neq j} \sigma_{ij} \quad \text{where } \sigma_{ij} = \mathbb{E}[\xi_i \xi_j]. \tag{20}$$

We also consider a version of Theorem 1.3 adapted to the multidimensional case, with the L^1 metric replaced by a smooth functions metric, following the development of Chapter 12 of [11]. For a real valued function $\varphi(u)$ defined on the domain $\mathcal{D}$, let $|\varphi|_\infty = \sup_{x \in \mathcal{D}} |\varphi(x)|$. We include in this definition the $|\cdot|_\infty$ norm of vectors and matrices, for instance, by considering them as real valued functions of their indices.

For $m \in \mathbb{N}_0$, let $L_m^\infty(\mathbb{R}^p)$ be the collection of all functions $h : \mathbb{R}^p \to \mathbb{R}$ such that for all $\mathbf{k} = (k_1, \ldots, k_p) \in \mathbb{N}_0^p$ with $|\mathbf{k}|_1 \leq m$, the partial derivative

$$h^{(\mathbf{k})}(\mathbf{x}) = \frac{\partial^{|\mathbf{k}|_1} h}{\partial^{k_1} x_1 \cdots \partial^{k_p} x_p}$$

exists, and

$$|h|_{L_m^\infty(\mathbb{R}^p)} := \max_{0 \leq |\mathbf{k}|_1 \leq m} \left|h^{(\mathbf{k})}\right|_\infty \quad \text{is finite.}$$

Let

$$\mathcal{H}_{m,\infty,p} = \left\{ h \in L_m^\infty(\mathbb{R}^p) : |h|_{L_m^\infty(\mathbb{R}^p)} \leq 1 \right\},$$

and for random vectors $\mathbf{X}$ and $\mathbf{Y}$ in $\mathbb{R}^p$, define the smooth functions metric

$$d_{\mathcal{H}_{m,\infty,p}}\big(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y})\big) = \sup_{h \in \mathcal{H}_{m,\infty,p}} \big|\mathbb{E}h(\mathbf{X}) - \mathbb{E}h(\mathbf{Y})\big|. \tag{21}$$

For a positive semidefinite matrix H we let $H^{1/2}$ denote the unique positive semidefinite square root of H. When H is positive definite, we write $H^{-1/2} = (H^{1/2})^{-1}$. Our main multidimensional result is the following theorem:

Theorem 1.4. *With* $m, p \in \mathbb{N}_1$, *let* $\{\xi_{i,j} : i \in [m], j \in [p]\}$ *be positively associated mean zero random variables bounded in absolute value by some positive constant* B. *Let* $\mathbf{S} = (S_1, S_2, \ldots, S_p)$ *where* $S_j = \sum_{1 \le i \le m} \xi_{i,j}$ *for* $j \in [p]$ *and assume that* $\Sigma = \mathrm{Var}(\mathbf{S})$ *is positive definite. Then*

$$d_{\mathcal{H}_{3,\infty,p}}\big(\mathcal{L}\big(\Sigma^{-1/2}\mathbf{S}\big), \mathcal{L}(\mathbf{Z})\big) \le \left(\frac{1}{6} + 2\sqrt{2}\right) p^3 B \big|\Sigma^{-1/2}\big|_{\infty}^3 \sum_{j=1}^{p} \Sigma_{j,j}$$

$$+ \left(\frac{3}{\sqrt{2}} + \frac{1}{2}\right) p^2 \big|\Sigma^{-1/2}\big|_{\infty}^2 \sum_{j=1}^{p} \sum_{i,k \in [m], i \neq k} \mathrm{Cov}(\xi_{i,j}, \xi_{k,j})$$

$$+ \left(2\sqrt{2} p^3 B \big|\Sigma^{-1/2}\big|_{\infty}^3 + \left(\frac{3}{\sqrt{2}} + \frac{1}{2}\right) p^2 \big|\Sigma^{-1/2}\big|_{\infty}^2\right) \sum_{j,l \in [p], j \neq l} \Sigma_{j,l}, \qquad (22)$$

where $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I_p)$, *a standard normal vector in* $\mathbb{R}^p$.

The Ising and percolation models are special cases of random cluster models, see [25], for which positive association is well-known. These two cases, however, are the only ones for which covariances are known to decay exponentially. However, as in some generality the boundedness hypothesis of our main theorems hold for random cluster models, our results apply there whenever covariances can be shown to be sufficiently small. In particular, under known exponential decay, Corollary 2.7 may be applied; see Remark 2.8.

On the other hand, our results for the voter model in Section 2.3 illustrate that our methods apply in the absence of exponential decay. For Potts models, corresponding to the parameters $q \in \mathbb{N}_2$ of the random cluster model, it would therefore be of interest to determine if the covariance formula given in Theorem 9.11 of [25] admits sufficiently rapid covariance decay that would allow our methods to apply. We note that neither the voter model nor the contact process studied in Sections 2.3 and 2.4, respectively, are instances of the random cluster model; we apply specific techniques to handle their covariance terms.

The remainder of this work is organized as follows. In Section 2.1 we state general results for positively associated random fields whose covariance decays exponentially. These results are then applied to the Ising and percolation models in Section 2.2. The two main general results of this section, Theorems 1.3 and 1.4, are applied directly to the voter model and the contact process in Sections 2.3 and 2.4. We then prove the general results of this section in Section 3. One advantage of these latter results are that, unlike many results based on Stein's method, they may be applied without the need for coupling constructions.

2. Applications

We begin this section with a general result for a certain class of positively associated random fields.

2.1. *Second order stationary positively associated fields with exponential covariance decay*

Let $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ be a positively associated random field on the d-dimensional integer lattice $\mathbb{Z}^d$. Assume that the field is second order stationary, that is, for all $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$ the covariance $\mathrm{Cov}(X_{\mathbf{j}}, X_{\mathbf{k}})$ exists and equals $R(\mathbf{j} - \mathbf{k})$ for some function R. With $S_{\mathbf{k}}^n$ defined in (15), consider the standardized variables

$$W_{\mathbf{k}}^n = \frac{S_{\mathbf{k}}^n - \mathbb{E}S_{\mathbf{k}}^n}{\sqrt{n^d A_n}}, \qquad \mathbf{k} \in \mathbb{Z}^d, n \in \mathbb{N}_1, \qquad\qquad (23)$$

that have mean zero and variance 1. The following theorem provides a bound of order $n^{-d/(2d+2)}$ with an explicit constant on the L^1 distance between the distribution of $W_{\mathbf{k}}^n$ and the normal under the condition that the covariance function $R(\cdot)$ decays exponentially in the L^1 norm in $\mathbb{R}^d$. We have chosen to bound the covariance decay in the L^1 norm for convenience in the proof of Lemma 2.4; as all norms in $\mathbb{R}^d$ are equivalent, when inequality (24) holds for any norm with some $\lambda > 0$ then it holds for any other norm with λ replaced by some positive constant multiple.

Theorem 2.1. *Let* $d \in \mathbb{N}_1$ *and* $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ *be a positively associated second order stationary random field with covariance function* $R(\mathbf{k}) = \mathrm{Cov}(X_{\mathbf{j}}, X_{\mathbf{j}+\mathbf{k}})$ *for all* $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$, *and suppose that for some* $K > 0$ *it holds that* $|X_{\mathbf{j}}| \leq K$ *a.s. for all* $\mathbf{j} \in \mathbb{Z}^d$. *Assume that there exist* $\lambda > 0$ *and* $\kappa_0 > 0$ *such that*

$$R(\mathbf{k}) \leq \kappa_0 e^{-\lambda |\mathbf{k}|_1} \quad \text{for all } \mathbf{k} \in \mathbb{Z}^d \tag{24}$$

and let

$$\mu_\lambda = \frac{e^\lambda}{(e^\lambda - 1)^2}, \qquad \upsilon_\lambda = \frac{e^{2\lambda}}{(e^\lambda - 1)^2} \quad \text{and} \quad \gamma_{\lambda,d} = (4\mu_\lambda + 2\upsilon_\lambda)^d - (2\upsilon_\lambda)^d \tag{25}$$

and

$$C_{\lambda,\kappa_0,d} = \frac{5Kd\sqrt{\pi A_n}}{\sqrt{2}\kappa_0 \gamma_{\lambda,d}}.$$

Then, for any $\mathbf{k} \in \mathbb{Z}^d$, *with* $W_{\mathbf{k}}^n$ *as given in* (23) *and* Z *a standard normal random variable,*

$$d_1\big(\mathcal{L}(W_{\mathbf{k}}^n), \mathcal{L}(Z)\big) \leq \frac{\kappa_1}{n^{d/(2d+2)}} \quad \text{for all } n \geq \max\{C_{\lambda,\kappa_0,d}^{2/d}, C_{\lambda,\kappa_0,d}^{-2/(d+2)}\},$$

where, with A_n *is as in* (16),

$$\kappa_1 = \left(\frac{10K\kappa_0^d \gamma_{\lambda,d}^d 2^{3d/2}}{\pi^{d/2} A_n^{d+1/2}}\right)^{1/(d+1)} \left(\frac{1}{d^{\frac{d}{d+1}}} + 2d^{\frac{1}{d+1}}\right). \tag{26}$$

In the following we will make use of the identities

$$\sum_{k=1}^{n-1} (n-k)w^k = \frac{w((n-1) - nw + w^n)}{(w-1)^2} \quad \text{for } w \neq 1, \tag{27}$$

$$n + \sum_{a=1}^{n-1} (n-a)\big(v^a + v^{-a}\big) = \frac{v^{1-n}(v^n - 1)^2}{(v-1)^2} \quad \text{for } v \neq 1, \tag{28}$$

and

$$n + 2\sum_{b=1}^{n-1} (n-b)u^b = \frac{(1-u^2)n - 2u + 2u^{n+1}}{(u-1)^2} \quad \text{for } u \neq 1. \tag{29}$$

We have already taken A in (17), without loss of generality, to be positive, and the exponential decay condition (24) implies that it is also finite. Lemma 4 of [33] can now be invoked to yield that $\lim_{n\to\infty} A_n = A$. Hence the values A_n are bounded away from zero and infinity as a function of n, and replacing A_n by its limiting value A in (26) only affects the bound by a constant and does not alter the rate of convergence.

In certain instances the values of A_n and A will be close even for moderate the values n, for instance if equality holds in (24), then using (29) with $u = e^{-\lambda}$, which is not equal to one since $\lambda > 0$, as $n \to \infty$ we obtain

$$\begin{aligned}
A_n &= \frac{\kappa_0}{n^d} \sum_{\mathbf{i},\mathbf{j} \in B_{\mathbf{1}}^n} e^{-\lambda|\mathbf{i}-\mathbf{j}|_1} = \frac{\kappa_0}{n^d} \sum_{\substack{i_1,\dots,i_d=1 \\ j_1,\dots,j_d=1}}^{n} \prod_{k=1}^{d} e^{-\lambda|i_k - i_k|} \\
&= \frac{\kappa_0}{n^d} \prod_{k=1}^{d} \sum_{i_k,j_k=1}^{n} e^{-\lambda|i_k - i_k|}
\end{aligned}$$

$$= \frac{\kappa_0}{n^d} \prod_{k=1}^{d} \sum_{a_k=-n+1}^{n-1} \left(n - |a_k|\right) e^{-\lambda|a_k|} = \frac{\kappa_0}{n^d} \left(n + 2 \sum_{b=1}^{n-1} (n-b) e^{-\lambda b}\right)^d$$

$$= \kappa_0 \left(\frac{1 - e^{-2\lambda} - \frac{2e^{-\lambda}}{n} + \frac{2e^{-\lambda(n+1)}}{n}}{(1 - e^{-\lambda})^2}\right)^d \longrightarrow \kappa_0 \left(\frac{1 + e^{-\lambda}}{1 - e^{-\lambda}}\right)^d = \kappa_0 \coth^d(\lambda/2) := A.$$

One can apply Theorem 1.2 above, due to [8], to obtain a bound of order $\log^d n / \sqrt{n}$, without an explicit constant, on the Kolmogorov distance between $W_{\mathbf{k}}^n$ and the normal by showing that $u(n)$ as defined in (19) converges to zero at an exponential rate. Here we use Theorem 1.3 to obtain an L^1 result.

We note that the technique in [8], first introduced in [41], presents some challenges when applied to obtain an L^1 bound. We consider the argument in the one dimensional case, as in [6], to illustrate the difficulty. Letting f_n be the characteristic function of W_1^n and following the proof in [6], one obtains the bound

$$\left|f_n(t) - e^{-t^2/2}\right| \leq C\left(n^{-1/2} \log(n) t^3 e^{-t^2/4} + n^{-1} \log^{3/2}(n) t + n^{-3/2} t\right)$$

over the region $0 \leq t \leq \gamma n^{1/2} / \log(n)$, where γ is a constant defined in [6], and where here and in the following, C denotes a constant whose value may change from line to line. Applying Theorem 1.5.2 in [28] with $T = n^{1/2} / \log(n)$, the Kolmogorov distance of order $\log(n) / n^{1/2}$ was obtained in [6]. To get the bound in L^1, one may apply Theorem 1.5.4 in [28] instead, however the integrand of the term (3) in Theorem 1.5.4

$$\int_{-T}^{T} \left(\frac{d}{dt} \frac{f_n(t) - e^{-t^2/2}}{t}\right)^2 dt \quad \text{includes} \quad \frac{(f_n(t) - e^{-t^2/2})^2}{t^4},$$

which is not integrable.

We also extend Theorem 2.1 to the multidimensional case. For any $p \in \mathbb{N}_1$ and indices $\mathbf{k}_1, \ldots, \mathbf{k}_p$ in $\mathbb{Z}^d$ satisfying a separation condition to ensure nondegeneracy of the limiting distibution, Theorem 2.2 provides a bound in the metric $d_{\mathcal{H}_{3,\infty,p}}$ to the multivariate normal for $\mathbf{S}^n = (S_{\mathbf{k}_1}^n, \ldots, S_{\mathbf{k}_p}^n)$ under exponential decay of the covariance function and a condition limiting the amount of variables common to the coordinates making up the sums $\mathbf{S}^n$. Regarding multidimensional convergence, under strictly stationary Theorem 1.1 of [33] shows that $(W_{n\mathbf{k}_1}^n, \ldots, W_{n\mathbf{k}_p}^n)$ converges to a standard normal random vector in $\mathbb{R}^p$, but does not provide a bound. In the following result and its proof, constants will not be tracked with precision, but will be indexed by the set of variables on which it depends.

Theorem 2.2. *For $d \in \mathbb{N}_1$, let $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ be a positively associated second order stationary random field with covariance function $R(\mathbf{k}) = \mathrm{Cov}(X_{\mathbf{j}}, X_{\mathbf{j}+\mathbf{k}})$ for all $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$, and suppose that there exists constants $K > 0, \kappa_0 > 0$ and $\lambda > 0$ such that $|X_{\mathbf{j}}| \leq K$ a.s. for all $\mathbf{j} \in \mathbb{Z}^d$ and*

$$R(\mathbf{k}) \leq \kappa_0 e^{-\lambda|\mathbf{k}|_1} \quad \text{for all } \mathbf{k} \in \mathbb{Z}^d.$$

For $p \in \mathbb{N}_1$ let $\mathbf{k}_1, \ldots, \mathbf{k}_p \in \mathbb{Z}^d$ be such that

$$\min_{q,s \in [p], q \neq s} |\mathbf{k}_q - \mathbf{k}_s|_\infty \geq (1 - \alpha)n, \tag{30}$$

for some $0 < \alpha < 1$ with αn an integer. Let $\mathbf{S}^n = (S_{\mathbf{k}_1}^n, \ldots, S_{\mathbf{k}_p}^n)$, where $S_{\mathbf{k}}^n$ is defined as in (15) and assume that the covariance matrix Σ of $\mathbf{S}^n$ is invertible, and let

$$\psi_n = n^{d/2} |\Sigma^{-1/2}|_\infty.$$

Then, there exists a constant $C_{\lambda, \kappa_0, d, p, K}$ such that with $\mathbf{Z}$ a standard normal random vector in $\mathbb{R}^p$,

$$d_{\mathcal{H}_{3,\infty,p}}\left(\mathcal{L}\left(\Sigma^{-1/2}(\mathbf{S}^n - \mathbb{E}\mathbf{S}^n)\right), \mathcal{L}(\mathbf{Z})\right)$$

$$\leq C_{\lambda,\kappa_0,d,p,K}\left((A_n+\alpha)^{\frac{1}{d+1}}\psi_n^{\frac{2d+3}{d+1}}\left(\frac{1}{d^{\frac{d}{d+1}}}+2d^{\frac{1}{d+1}}\right)n^{-\frac{d}{2(d+1)}}+\alpha\psi_n^2\right)$$

$$\text{for all } n \geq \max\left\{B_{n,d,\alpha}^{2/d},\, B_{n,d,\alpha}^{-2/(d+2)}\right\} \text{ where } B_{n,d,\alpha}=d\psi_n(A_n+\alpha). \tag{31}$$

For checking the hypothesis that the covariance matrix of $\mathbf{S}^n$ is invertible, and that the quantity ψ_n is of order one, and so does not affect the rate in (31) and also so that $B_{n,d,\alpha}$ is order one, we present the following condition, proved at the end of this section. We note in addition that if $\psi_n = O(1)$ and $\alpha = O(n^{-\frac{d}{2(d+1)}})$ then the bound in (31) will also be of order $O(n^{-\frac{d}{2(d+1)}})$.

Lemma 2.3. *Let $\{X_\mathbf{j} : \mathbf{j} \in \mathbb{Z}^d\}$ be a second order stationary random field with covariance function $R(\mathbf{k}) = \mathrm{Cov}(X_\mathbf{j}, X_{\mathbf{j}+\mathbf{k}})$ for all $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$ where $R(\cdot)$ satisfies* (24). *With $n \in \mathbb{N}_2$, the covariance matrix Σ of the random vector $\mathbf{S}^n \in \mathbb{R}^p$, $p \geq 2$ as given in Theorem* 2.2 *is invertible if for some $b > 0$*

$$|\mathbf{k}_q - \mathbf{k}_s|_\infty \geq n - b \quad \text{for all } q \neq s, q, s \in [p], \text{ and } b < \frac{nA_n}{(p-1)\kappa_0 \upsilon_\lambda^d}, \tag{32}$$

with λ and κ_0 as in (24), *and υ_λ as in* (25). *If $b = \alpha n$ for α satisfying*

$$0 < \alpha < \min\left\{1,\, \frac{A_n}{(p-1)\kappa_0 \upsilon_\lambda^d}\right\} \tag{33}$$

then the matrix Σ is invertible and

$$\left|\Sigma^{-1}\right|_\infty \leq \frac{1}{n^d(A_n-(p-1)\kappa_0 \upsilon_\lambda^d \alpha)}. \tag{34}$$

The proofs of Theorems 2.1 and 2.2 proceed by decomposing the sum $S_\mathbf{k}^n$ over the block $B_\mathbf{k}^n$ into sums over smaller, disjoint blocks whose side lengths are at most some integer l. In particular, given $1 \leq l \leq n$ uniquely write $n = (m-1)l + r$ for $m \geq 1$ and $1 \leq r \leq l$. We correspondingly decompose $B_\mathbf{k}^n$ into m^d disjoint blocks $D_{\mathbf{i},\mathbf{k}}^l$, $\mathbf{i} \in [m]^d$, where there are $(m-1)^d$ "main" blocks having all sides of length l, and $m^d - (m-1)^d$ remainder blocks having all sides of length r or l, with at least one side of length r.

In detail, for $\mathbf{k} \in \mathbb{Z}^d$ and $\mathbf{i} \in [m]^d$ set $D_{\mathbf{i},\mathbf{k}}^l = D_\mathbf{i}^l + \mathbf{k} - \mathbf{1}$ where

$$D_\mathbf{i}^l = \left\{\mathbf{j} \in \mathbb{Z}^d : (i_s-1)l+1 \leq j_s \leq i_s l, i_s \neq m, (m-1)l+1 \leq j_s \leq (m-1)l+r, i_s = m\right\}.$$

It is straightforward to verify that for all $\mathbf{i} \in [m]^d$ with no component equal to m, which are the indices of the "main blocks", that with $B_\mathbf{1}^l$ as in (15), we have

$$D_\mathbf{i}^l = B_\mathbf{1}^l + (\mathbf{i}-\mathbf{1})l \quad \text{for } \mathbf{i} \in [m-1]^d, \tag{35}$$

and if $r = l$ then $D_\mathbf{i}^l$ is given by (35) for all $i \in [m]^d$. Further, it is easy to see that the elements of the collection $\{D_{\mathbf{i},\mathbf{k}}^l, \mathbf{i} \in [m]^d\}$ are disjoint, and union to $B_\mathbf{k}^n$.

Letting

$$\xi_{\mathbf{i},\mathbf{k}}^l = \sum_{\mathbf{t} \in D_{\mathbf{i},\mathbf{k}}^l}(X_\mathbf{t} - \mathbb{E}X_\mathbf{t}) \quad \text{and} \quad W_\mathbf{k}^n = \sum_{\mathbf{i} \in [m]^d}\frac{\xi_{\mathbf{i},\mathbf{k}}^l}{\sqrt{n^d A_n}} \quad \text{for } \mathbf{i} \in [m]^d, \tag{36}$$

we see that $\xi_{\mathbf{i},\mathbf{k}}^l$ has mean zero, and $W_\mathbf{k}^n$ as in (36) agrees with its representation as given in (23), and has mean zero and variance one. For simplicity we will drop the index $\mathbf{k}$ in $\xi_{\mathbf{i},\mathbf{k}}^l$ when $\mathbf{k} = \mathbf{1}$, as we do also for $D_{\mathbf{i},\mathbf{k}}$, and may also suppress l.

By [21], when $\mathbf{X} = (X_1, \ldots, X_p)$ is positively associated and if $h_k : \mathbb{R}^p \to \mathbb{R}$, $k \in [q]$ are coordinate-wise non-decreasing functions, then $h_1(\mathbf{X}), \ldots, h_q(\mathbf{X})$ are positively associated. Hence, as the elements of $\{\xi_{\mathbf{i},\mathbf{k}} : \mathbf{i} \in [m]^d\}$ are increasing functions of $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ they are positively associated.

We prove Theorems 2.1 and 2.2 with the help of two lemmas that follow. The first, Lemma 2.4 bounds the sum of the covariances between $\xi_{\mathbf{i},\mathbf{k}}^l$ and $\xi_{\mathbf{j},\mathbf{k}}^l$, defined in (36), over $\mathbf{i}, \mathbf{j} \in [m]^d$. Next, Lemma 2.6 bounds the covariance between two block sums of size n^d whose centers are at least $n - b$ apart in the L_∞ distance for some integer $1 \le b \le n$.

Lemma 2.4. *Let $\{X_{\mathbf{j}} : \mathbf{j} \in \mathbb{Z}^d\}$ be a second order, positively associated stationary random field with covariance function $R(\mathbf{k}) = \mathrm{Cov}(X_{\mathbf{j}}, X_{\mathbf{j}+\mathbf{k}})$ for all $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$ where $R(\cdot)$ satisfies the exponential decay condition (24). For $n \ge 2$ and $1 \le l \le n$, let $n = (m-1)l + r$ for integers $m \in \mathbb{N}_1$ and $1 \le r \le l$. Then for $\mathbf{k} \in \mathbb{Z}^d$, with $\xi_{\mathbf{i},\mathbf{k}}^l$ given by (36) we have*

$$\sum_{\mathbf{i},\mathbf{j} \in [m]^d, \mathbf{i} \ne \mathbf{j}} \mathbb{E}\big[\xi_{\mathbf{i},\mathbf{k}}^l \xi_{\mathbf{j},\mathbf{k}}^l\big] \le \frac{\kappa_0 \gamma_{\lambda,d} n^d}{l},$$

where κ_0 is given in (24), and $\gamma_{\lambda,d}$ in (25).

Proof. By second order stationarity, it suffices to consider the case $\mathbf{k} = \mathbf{1}$. For convenience write $\sigma_{\mathbf{i},\mathbf{j}} = \mathbb{E}[\xi_{\mathbf{i}} \xi_{\mathbf{j}}]$. Since $D_{\mathbf{i}}^l \subset B_{\mathbf{1}}^l + (\mathbf{i} - \mathbf{1})l$ for all $\mathbf{i} \in [m]^d$ and $\mathrm{Cov}(X_{\mathbf{t}}, X_{\mathbf{s}}) \ge 0$ for all $\mathbf{t}, \mathbf{s} \in \mathbb{Z}^d$ by positive association, with $S_{\mathbf{i}}^l$ as given in (15), we have

$$\sigma_{\mathbf{i},\mathbf{j}} = \mathrm{Cov}\left(\sum_{\mathbf{t} \in D_{\mathbf{i}}^l} X_{\mathbf{t}}, \sum_{\mathbf{s} \in D_{\mathbf{j}}^l} X_{\mathbf{s}} \right) \le \mathrm{Cov}\big(S_{\mathbf{1}+(\mathbf{i}-\mathbf{1})l}^l, S_{\mathbf{1}+(\mathbf{j}-\mathbf{1})l}^l\big) \quad \text{for all } \mathbf{i} \ne \mathbf{j} \in [m]^d. \tag{37}$$

Note that when $\mathbf{i} \ne \mathbf{j}$ in $[m]^d$ the pair $(\mathbf{i}, \mathbf{j})$ must lie in exactly one of the sets E_s^m for $s \in [d]$, given by

$$E_s^m = \big\{(\mathbf{i}, \mathbf{j}) \in [m]^d \times [m]^d : \big|\{k \in [d] : i_k \ne j_k\}\big| = s\big\} \quad \text{for } s \in [d].$$

Hence, using (37),

$$\sum_{\mathbf{i},\mathbf{j} \in [m]^d, \mathbf{i} \ne \mathbf{j}} \sigma_{\mathbf{i},\mathbf{j}} = \sum_{s=1}^d \sum_{(\mathbf{i},\mathbf{j}) \in E_s^m} \sigma_{\mathbf{i},\mathbf{j}} \le \sum_{s=1}^d \sum_{(\mathbf{i},\mathbf{j}) \in E_s^m} \mathrm{Cov}\big(S_{\mathbf{1}+(\mathbf{i}-\mathbf{1})l}^l, S_{\mathbf{1}+(\mathbf{j}-\mathbf{1})l}^l\big). \tag{38}$$

Recalling the definition of the block sums in (15) and using second order stationarity, for $(\mathbf{i}, \mathbf{j}) \in E_s^m$ we have

$$
\begin{aligned}
\mathrm{Cov}\big(S_{\mathbf{1}+(\mathbf{i}-\mathbf{1})l}^l, S_{\mathbf{1}+(\mathbf{j}-\mathbf{1})l}^l\big) &= \sum_{\substack{p_1,\ldots,p_d=1 \\ q_1,\ldots,q_d=1}}^l R\left(\begin{bmatrix} (p_1 + (j_1-1)l) - (q_1 + (i_1-1)l) \\ \vdots \\ (p_d + (j_d-1)l) - (q_d + (i_d-1)l) \end{bmatrix} \right) \\
&= \sum_{\substack{p_1,\ldots,p_d=1 \\ q_1,\ldots,q_d=1}}^l R\left(\begin{bmatrix} p_1 - q_1 + (j_1 - i_1)l \\ \vdots \\ p_d - q_d + (j_d - i_d)l \end{bmatrix} \right) \\
&\le \kappa_0 \sum_{a_1,\ldots,a_d=-l+1}^{l-1} \big(l - |a_1|\big) \cdots \big(l - |a_d|\big) \exp\left(-\lambda \Big| \begin{bmatrix} a_1 + (j_1 - i_1)l \\ \vdots \\ a_d + (j_d - i_d)l \end{bmatrix} \Big|_1 \right), \tag{39}
\end{aligned}
$$

where we have applied (24) in the final inequality.

We now apply some invariance properties of the norm $|\mathbf{x}|_1$ in order to simplify the sum of expression (39) when taken over pairs of indices in E_s^m. We say that the pairs of indices $(\mathbf{i}, \mathbf{j})$ and $(\mathbf{i}', \mathbf{j}')$ in E_s^m are equivalent if $(\mathbf{i}, \mathbf{j})$ can be made to agree with $(\mathbf{i}', \mathbf{j}')$ by interchanging i_k and j_k for any values of $k \in [d]$ and then applying the same permutation to the resulting set of vectors. Note that as a_k and $-a_k$ appear symmetrically in the sum (39), and that $|\cdot|_1$ is invariant under sign changes and permutations of coordinates, equivalent vectors yield the same value of this sum.

Call an index pair $(\mathbf{i}, \mathbf{j}) \in E_s^m$ canonical if $i_k < j_k$ for all $k = 1, \ldots, s$. Note that as pairs in E_s^m have exactly s inequalities among their coordinates that canonical indices must agree in their remaining, and last, $d - s$ coordinates. Now associate to every pair $(\mathbf{i}, \mathbf{j}) \in E_s^m$ the equivalent, unique canonical vector that is obtained from $(\mathbf{i}, \mathbf{j})$ by interchanging all coordinates $k \in [d]$ for which $i_k > j_k$, and then applying the permutation that maps the s unequal coordinates of the pair to positions $1, 2, \ldots, s$, thus mapping the equal pairs to the last $d - s$ positions, leaving the relative order of each group unchanged. Partitioning the sum over E_s^m into smaller sums over all vectors that are equivalent to the same canonical vector yields a sum over all canonical vectors, with a factor of 2^s to account for the number arrangements of inequalities over the unequal pairs, and a factor of $\binom{d}{s}$ to account for the positions in which the unequal coordinates may occur. To sum (39) over E_s^m we apply identities of the form

$$
\sum_{a_1 \in I_1, \ldots, a_d \in I_d} \prod_{q=1}^{d} f_q(a_q) = \prod_{q=1}^{d} \sum_{a_q \in I_q} f_q(a_q)
$$

that hold for any index sets $I_1, \ldots, I_d$ and any functions f_q; these identities will be used to interchange summations and products. In order to be consistent with the formula used above, we notate the index set as

$$
I_d^{s,m} = \left\{ (\mathbf{i}, \mathbf{j}) \in \mathbb{Z}^d \times \mathbb{Z}^d : 1 \le i_k < j_k \le m, k = 1, \ldots, s, 1 \le i_k = j_k \le m, k = s+1, \ldots, d \right\},
$$

and thus obtain

$$
\sum_{(\mathbf{i}, \mathbf{j}) \in E_s^m} \mathrm{Cov}\big(S_{\mathbf{1}+(\mathbf{i}-\mathbf{1})l}^l, S_{\mathbf{1}+(\mathbf{j}-\mathbf{1})l}^l\big)
$$

$$
\le \kappa_0 \binom{d}{s} 2^s \sum_{(\mathbf{i}, \mathbf{j}) \in I_d^{s,m}} \sum_{a_1, \ldots, a_d = -l+1}^{l-1} (l - |a_1|) \cdots (l - |a_d|) \exp\left(-\lambda \left| \begin{bmatrix} a_1 + (j_1 - i_1)l \\ \vdots \\ a_d + (j_d - i_d)l \end{bmatrix} \right|_1 \right)
$$

$$
= \kappa_0 \binom{d}{s} 2^s \sum_{(\mathbf{i}, \mathbf{j}) \in I_d^{s,m}} \sum_{a_1, \ldots, a_d = -l+1}^{l-1} \prod_{q=1}^{d} (l - |a_q|) e^{-\lambda |(j_q - i_q)l + a_q|}
$$

$$
= \kappa_0 \binom{d}{s} 2^s \sum_{(\mathbf{i}, \mathbf{j}) \in I_d^{s,m}} \prod_{q=1}^{d} \sum_{a_q = -l+1}^{l-1} (l - |a_q|) e^{-\lambda |(j_q - i_q)l + a_q|}
$$

$$
= \kappa_0 \binom{d}{s} 2^s \sum_{(\mathbf{i}, \mathbf{j}) \in I_d^{s,m}} \left(\prod_{q=1}^{s} \sum_{a_q = -l+1}^{l-1} (l - |a_q|) e^{-\lambda(j_q - i_q)l - \lambda a_q} \right)
$$

$$
\times \left(\prod_{q=s+1}^{d} \sum_{a_q = -l+1}^{l-1} (l - |a_q|) e^{-\lambda |a_q|} \right),
$$

where in the final equality we have expressed the product in q as two separate products, and have noted in the first of these, as $j_k - i_k \ge 1$ and a_k ranges from $-l+1$ to $l-1$, the terms $(j_k - i_k)l + a_k$ appearing in the absolute value of the exponent are always positive. Now expanding the first multiple summation as an inner summation on the first s indices and outer sum over the final $d - s$ indices, separating out the exponential terms in the first product and noting

that the remaining product terms are purely mulitplicative factors, we obtain

$$
\kappa_0 \binom{d}{s} 2^s \sum_{\substack{1 \le i_k = j_k \le m \\ k=s+1,\ldots,d}} \sum_{\substack{1 \le i_k < j_k \le m \\ k=1,\ldots,s}} \prod_{q=1}^{s} e^{-\lambda(j_q - i_q)l}
$$

$$
\times \left(\prod_{q=1}^{s} \sum_{a_q=-l+1}^{l-1} (l - |a_q|) e^{-\lambda a_q} \right) \left(\prod_{q=s+1}^{d} \sum_{a_q=-l+1}^{l-1} (l - |a_q|) e^{-\lambda |a_q|} \right)
$$

$$
= \kappa_0 \binom{d}{s} 2^s \left(\sum_{i=1}^{m} 1 \right)^{d-s} \left(\prod_{q=1}^{s} \sum_{1 \le i_q < j_q \le m} e^{-\lambda(j_q - i_q)l} \right)
$$

$$
\times \left(\prod_{q=1}^{s} \sum_{a_q=-l+1}^{l-1} (l - |a_q|) e^{-\lambda a_q} \right) \left(\prod_{q=s+1}^{d} \sum_{a_q=-l+1}^{l-1} (l - |a_q|) e^{-\lambda |a_q|} \right)
$$

$$
= \kappa_0 \binom{d}{s} 2^s m^{d-s} \left(\sum_{k=1}^{m-1} (m - k) e^{-\lambda k l} \right)^s
$$

$$
\times \left(l + \sum_{a=1}^{l-1} (l - a) \left(e^{\lambda a} + e^{-\lambda a} \right) \right)^s \left(l + 2 \sum_{b=1}^{l-1} (l - b) e^{-\lambda b} \right)^{d-s}.
$$

By applying identity (27) with $n = m$ and $w = e^{-\lambda l}$ and identities (28) and (29) with $n = l$, $v = e^{\lambda}$ and $u = e^{-\lambda}$, noting that u, v and w are not one as $\lambda > 0$, the term above equals

$$
\kappa_0 \binom{d}{s} 2^s m^{d-s} \left(\frac{e^{\lambda} e^{-2\lambda l} (e^{\lambda l} - 1)^2 (m - 1 - m e^{-\lambda l} + e^{-\lambda l m})}{(e^{\lambda} - 1)^2 (1 - e^{-\lambda l})^2} \right)^s
$$

$$
\times \left(\frac{(1 - e^{-2\lambda}) l - 2 e^{-\lambda} + 2 e^{-\lambda(1+l)}}{(1 - e^{-\lambda})^2} \right)^{d-s}
$$

$$
\le \kappa_0 \binom{d}{s} 2^s m^d \left(\frac{e^{\lambda}}{(e^{\lambda} - 1)^2} \right)^s \left(\frac{l}{(1 - e^{-\lambda})^2} \right)^{d-s}
$$

$$
= \kappa_0 m^d \binom{d}{s} (2\mu_\lambda)^s (l\upsilon_\lambda)^{d-s}, \tag{40}
$$

where μ_λ and υ_λ are given in (25). By (38) and (40), we have

$$
\sum_{\mathbf{i},\mathbf{j} \in [m]^d, \mathbf{i} \ne \mathbf{j}} \sigma_{\mathbf{i},\mathbf{j}} \le \kappa_0 m^d \sum_{s=1}^{d} \binom{d}{s} (2\mu_\lambda)^s (l\upsilon_\lambda)^{d-s}
$$

$$
\le \frac{\kappa_0 n^d}{l} \sum_{s=1}^{d} \binom{d}{s} (4\mu_\lambda)^s (2\upsilon_\lambda)^{d-s}
$$

$$
= \frac{\kappa_0 n^d ((4\mu_\lambda + 2\upsilon_\lambda)^d - (2\upsilon_\lambda)^d)}{l},
$$

where we have used the bounds $m \le 2n/l$ and $1/l^s \le 1/l$ for all $s \in [d]$ in the second inequality. $\square$

Lemma 2.5. *For all $n \in \mathbb{N}_2$ and $\lambda > 0$,*

$$\sum_{a=-n+1}^{n-1} \left(n - |a|\right)e^{-\lambda|q+a|} \quad \text{is decreasing as a function of } |q| \in \mathbb{N}_0.$$

Proof. We note that the sequence $f(a) = (n - |a|)\mathbf{1}(|a| \leq n)$ is unimodal, and $g(a) = e^{-\lambda|a|}$ is log-concave, and hence their convolution

$$h(q) = \sum_{a=-\infty}^{\infty} f(a)g(q - a) = \sum_{a=-n+1}^{n-1} \left(n - |a|\right)e^{-\lambda|q-a|} = \sum_{a=-n+1}^{n-1} \left(n - |a|\right)e^{-\lambda|q+a|}$$

is unimodal, see [7]. Since $g(a)$ is symmetric unimodal, one can also use the fact from [39] that the convolution of two symmetric unimodal sequences is symmetric unimodal to prove the result. $\qquad \square$

Below we index the coordinates of vectors $\mathbf{k}_j \in \mathbb{Z}^d$ as $(k_1^j, \ldots, k_d^j)$.

Lemma 2.6. *Let $\{X_\mathbf{j} : \mathbf{j} \in \mathbb{Z}^d\}$ be a second order stationary random field with covariance function $R(\mathbf{k}) = \mathrm{Cov}(X_\mathbf{j}, X_{\mathbf{j}+\mathbf{k}})$ for all $\mathbf{j}, \mathbf{k} \in \mathbb{Z}^d$ where $R(\cdot)$ satisfies (24). Let $n \in \mathbb{N}_2$ and let b be an integer such that $1 \leq b \leq n$. Then if $\mathbf{k}_1$ and $\mathbf{k}_2$ are vectors in $\mathbb{Z}^d$ such that*

$$|\mathbf{k}_1 - \mathbf{k}_2|_\infty \geq n - b,$$

then with λ and κ_0 as in (24), and υ_λ as in (25),

$$\mathrm{Cov}\left(S_{\mathbf{k}_1}^n, S_{\mathbf{k}_2}^n\right) \leq \kappa_0 \upsilon_\lambda^d b n^{d-1}.$$

Proof. Arguing as in the proof of Lemma 2.4 we have

$$\mathrm{Cov}\left(S_{\mathbf{k}_1}^n, S_{\mathbf{k}_2}^n\right) = \sum_{\substack{p_1,\ldots,p_d=0 \\ q_1,\ldots,q_d=0}}^{n-1} R\left(\begin{bmatrix} (p_1 + k_1^2) - (q_1 + k_1^1) \\ \vdots \\ (p_d + k_d^2) - (q_d + k_d^1) \end{bmatrix}\right)$$

$$\leq \kappa_0 \sum_{a_1,\ldots,a_d=-n+1}^{n-1} \left(n - |a_1|\right) \ldots \left(n - |a_d|\right) \exp\left(-\lambda \left|\begin{bmatrix} a_1 + (k_1^2 - k_1^1) \\ \vdots \\ a_d + (k_d^2 - k_d^1) \end{bmatrix}\right|_1\right)$$

$$= \kappa_0 \prod_{i=1}^{d} \sum_{a_i=-n+1}^{n-1} \left(n - |a_i|\right)e^{-\lambda|(k_i^2 - k_i^1)+a_i|}. \tag{41}$$

Lemma 2.5 yields that

$$\sum_{a_i=-n+1}^{n-1} \left(n - |a_i|\right)e^{-\lambda|(k_i^2 - k_i^1)+a_i|} \quad \text{is a decreasing function of } \left|k_i^1 - k_i^2\right|. \tag{42}$$

In particular the ith sum appearing in the product (41) is bounded by its value when $k_i^1 = k_i^2$. As $|\mathbf{k}_1 - \mathbf{k}_2|_\infty \geq n - b$, there must exist at least one i for which $|k_i^2 - k_i^1| \geq n - b$, and whose corresponding sum is maximized by its value when equality to $n - b$ is achieved, again using (42). The product of these sums, by (42) again, is maximized when there is just a single coordinate achieving $n - b$ as its absolute difference, and where this difference in all other terms

achieve equality to zero. Hence, by symmetry the term above is bounded by the case where $k_i^1 = k_i^2$ for $i \in [d-1]$ and $k_d^2 - k_d^1 = n - b$ and thus

$$\text{Cov}\big(S_{\mathbf{k}_1}^n, S_{\mathbf{k}_2}^n\big) \le \kappa_0 \prod_{i=1}^{d-1} \sum_{a_i=-n+1}^{n-1} \big(n - |a_i|\big) e^{-\lambda|a_i|} \sum_{a_d=-n+1}^{n-1} \big(n - |a_d|\big) e^{-\lambda|a_d+n-b|}$$

$$\le \kappa_0 (n\upsilon_\lambda)^{d-1} \sum_{a_d=-n+1}^{n-1} \big(n - |a_d|\big) e^{-\lambda|a_d+n-b|}, \tag{43}$$

where we have applied (29) in the final inequality, and υ_λ as given in (25).

Now considering the sum in (43) for $2 \le b \le n$, we obtain

$$\sum_{a=-n+1}^{n-1} \big(n - |a|\big) e^{-\lambda|a+n-b|} = \sum_{a=-n+1}^{-n+b-1} \big(n - |a|\big) e^{\lambda(a+n-b)} + \sum_{a=-n+b}^{n-1} \big(n - |a|\big) e^{-\lambda(a+n-b)}$$

$$= \sum_{a=n-b+1}^{n-1} (n-a) e^{\lambda(-a+n-b)} + \sum_{a=1}^{n-b} (n-a) e^{-\lambda(-a+n-b)}$$

$$+ \sum_{a=0}^{n-1} (n-a) e^{-\lambda(a+n-b)}.$$

For the first sum, making a change of variable and applying (27), we obtain

$$\sum_{a=n-b+1}^{n-1} (n-a) e^{\lambda(-a+n-b)} = e^{-\lambda b} \sum_{a=1}^{b-1} a e^{\lambda a} = e^{-\lambda b} \left(b \sum_{a=1}^{b-1} e^{\lambda a} - \sum_{a=1}^{b-1} (b-a) e^{\lambda a} \right)$$

$$= e^{-\lambda b} \left(b \frac{e^{\lambda b} - e^{\lambda}}{e^{\lambda} - 1} - \frac{e^{\lambda}((b-1) - b e^{\lambda} + e^{\lambda b})}{(e^{\lambda} - 1)^2} \right)$$

$$= \frac{b(e^{\lambda} - 1) + e^{\lambda(1-b)} - e^{\lambda}}{(e^{\lambda} - 1)^2},$$

and similarly we may obtain

$$\sum_{a=1}^{n-b} (n-a) e^{-\lambda(-a+n-b)} = \frac{b e^{\lambda}(e^{\lambda} - 1) + e^{\lambda} + n(1 - e^{\lambda}) e^{\lambda(1+b-n)} - e^{\lambda(1+b-n)}}{(e^{\lambda} - 1)^2}$$

and

$$\sum_{a=0}^{n-1} (n-a) e^{-\lambda(a+n-b)} = \frac{e^{\lambda(1+b-2n)} - e^{\lambda(1+b-n)} + n(e^{\lambda} - 1) e^{\lambda(1+b-n)}}{(e^{\lambda} - 1)^2}.$$

Summing these three terms yields

$$\sum_{a=-n+1}^{n-1} \big(n - |a|\big) e^{-\lambda|a+n-b|} = \frac{b(e^{2\lambda} - 1) + e^{(1-b)\lambda} - e^{-\lambda(n-b-1)}(2 - e^{-\lambda n})}{(e^{\lambda} - 1)^2} \le \frac{b e^{2\lambda}}{(e^{\lambda} - 1)^2} = b\upsilon_\lambda,$$

where we lastly note that the equality holds also for $b = 1$. $\qquad \square$

Now we use Theorem 1.3 and Lemma 2.4 to prove Theorem 2.1. In the following, for positive numbers a and b we will seek to minimize a quantity of the form $al^d + b/l$ over nonnegative integers l. Over real values, it is easy to

verify that the minimum is achieved at $l_0 = (b/ad)^{1/(d+1)}$. Taking $l = \lfloor l_0 \rfloor$ when $l_0 \geq 1$ and using that $l_0/2 \leq l \leq l_0$ yields

$$\min_{l \in \mathbb{N}_1} \left(al^d + \frac{b}{l} \right) \leq a \left(\frac{b}{ad} \right)^{\frac{d}{d+1}} + 2b \left(\frac{ad}{b} \right)^{\frac{1}{d+1}} = a^{\frac{1}{d+1}} b^{\frac{d}{d+1}} \left(\frac{1}{d^{\frac{d}{d+1}}} + 2d^{\frac{1}{d+1}} \right). \tag{44}$$

Proof of Theorem 2.1. By second order stationarity, it suffices to prove the case $\mathbf{k} = \mathbf{1}$. Let $n \geq 2$, $B_{\mathbf{1}}^n$ the block of size n^d as given in (15), and $W_{\mathbf{1}}^n$ the standardized sum over that block, as in (23). For any $1 \leq l \leq n$ write $n = (m-1)l + r$, $1 \leq r \leq l$, and decompose $W_{\mathbf{1}}^n$ as the sum of $\xi_{\mathbf{i}}/\sqrt{n^d A_n}$ over $\mathbf{i} \in [m]$, as in (36).

We apply Theorem 1.3, handling the two terms on the right-hand side of (20). For the first term, using $|X_{\mathbf{j}}| \leq K$, the definition (36) of $\xi_{\mathbf{i}}$, and the fact that the side lengths of all blocks $D_{\mathbf{i}}^l$ are at most l, we have

$$\left| \frac{\xi_{\mathbf{i}}}{\sqrt{n^d A_n}} \right| \leq B \quad \text{with } B = \frac{2Kl^d}{\sqrt{n^d A_n}} \text{ for all } \mathbf{i} \in [m]^d.$$

Applying Lemma 2.4 for the second term, invoking Theorem 1.3 now yields

$$d_1 \big(\mathcal{L}(W_{\mathbf{1}}), \mathcal{L}(Z) \big) \leq \frac{10Kl^d}{\sqrt{n^d A_n}} + \frac{2\sqrt{2}\kappa_0 \gamma_{\lambda,d}}{\sqrt{\pi} l A_n}.$$

Recalling that $A_n > 0$ for all $n \in \mathbb{N}_1$, applying the bound (44) yields the result, noting that $l_0 = C_{\lambda,\kappa_0,d}^{-1/(d+1)} n^{d/(2d+2)}$ satisfies $1 \leq l_0 \leq n$ for $n \geq \max\{C_{\lambda,\kappa_0,d}^{2/d}, C_{\lambda,\kappa_0,d}^{-2/(d+2)}\}$. $\qquad \square$

To prove Theorem 2.2 we apply Theorem 1.4 and use the same techniques as for Theorem 2.1. We remind the reader that for this result we do not explicitly compute the constants, but index them by the parameters on which they depend.

Proof of Theorem 2.2. We proceed as in the one dimensional case. For $n \geq 2$, and $1 \leq l \leq n$ we write $n = (m-1)l + r$ with $m \geq 1$ and $1 \leq r \leq l$, and decompose $S_{\mathbf{k}_q}^n - \mathbb{E}S_{\mathbf{k}_q}^n$ for $q \in [p]$ as the sum over $\mathbf{i} \in [m]^d$ of the variables $\xi_{\mathbf{i},\mathbf{k}_q}$ given in (36).

Applying Theorem 1.4, we handle the three terms on the right-hand side of (22). For the first term, using the definition (36) of $\xi_{\mathbf{i},\mathbf{k}_q}$ and $|X_{\mathbf{t}}| \leq K$, we have

$$|\xi_{\mathbf{i},\mathbf{k}_q}| \leq B \quad \text{where } B = 2Kl^d \text{ for all } \mathbf{i} \in [m]^d, q \in [p]$$

and thus noting that $|\Sigma^{-1/2}|_\infty = n^{-d/2}\psi_n$ and that $\Sigma_{jj} = n^d A_n$, we may bound the first term as

$$\left(\frac{1}{6} + 2\sqrt{2} \right) p^3 B n^{-3d/2} \psi_n^3 \sum_{q=1}^p \Sigma_{q,q} \leq \frac{C_{p,K} l^d A_n \psi_n^3}{n^{d/2}}. \tag{45}$$

For the second term, by Lemma 2.4 we have

$$\left(\frac{3}{\sqrt{2}} + \frac{1}{2} \right) p^2 n^{-d} \psi_n^2 \sum_{q=1}^p \sum_{\mathbf{i},\mathbf{j} \in [m]^d, \mathbf{i} \neq \mathbf{j}} \mathbb{E}(\xi_{\mathbf{i},\mathbf{k}_q} \xi_{\mathbf{j},\mathbf{k}_q})$$

$$= C_p \psi_n^2 \sum_{q=1}^p \sum_{\mathbf{i},\mathbf{j} \in [m]^d, \mathbf{i} \neq \mathbf{j}} \mathbb{E}\left(\frac{\xi_{\mathbf{i},\mathbf{k}_q} \xi_{\mathbf{j},\mathbf{k}_q}}{n^d} \right) \leq \frac{C_{\lambda,\kappa_0,p,d} \psi_n^2}{l}. \tag{46}$$

Invoking Lemma 2.6 and assumption (30) we have

$$\Sigma_{q,s} = \text{Cov}\big(S_{\mathbf{k}_q}^n, S_{\mathbf{k}_s}^n \big) \leq \kappa_0 \upsilon_\lambda^d \alpha n^d \quad \text{for } q \neq s \in [p],$$

and hence we may bound the last term as

$$
\left(2\sqrt{2} p^3 B n^{-3d/2} \psi_n^3 + \left(\frac{3}{\sqrt{2}} + \frac{1}{2} \right) p^2 n^{-d} \psi_n^2 \right) \sum_{q,s \in [p], q \neq s} \Sigma_{q,s}
$$

$$
\leq \left(\frac{C_{p,K} l^d \psi_n^3}{n^{3d/2}} + \frac{C_p \psi_n^2}{n^d} \right) \kappa_0 \upsilon_\lambda^d \alpha n^d
$$

$$
\leq \frac{C_{\lambda, \kappa_0 d, p, K} \alpha l^d \psi_n^3}{n^{d/2}} + C_{\lambda, \kappa_0, d, p} \alpha \psi_n^2. \tag{47}
$$

By Theorem 1.4 and (45), (46) and (47),

$$
d_{\mathcal{H}_{3,\infty,p}} \left(\mathcal{L}(\Sigma^{-1/2} \mathbf{W}), \mathcal{L}(\mathbf{Z}) \right) \leq (C_{p,K} A_n + C_{\lambda, \kappa_0, d, p, K} \alpha) \frac{\psi_n^3 l^d}{n^{d/2}} + \frac{C_{\lambda, \kappa_0, d, p} \psi_n^2}{l} + C_{\lambda, \kappa_0, d, p} \alpha \psi_n^2
$$

$$
\leq C_{\lambda, \kappa_0, d, p, K} \left((A_n + \alpha) \frac{\psi_n^3 l^d}{n^{d/2}} + \frac{\psi_n^2}{l} + \alpha \psi_n^2 \right).
$$

Applying (44) to the first two terms in parenthesis, we obtain

$$
l_0 = \left(\frac{1}{d \psi_n (A_n + \alpha)} \right)^{\frac{1}{d+1}} n^{\frac{d}{2d+2}},
$$

which satisfies $1 \leq l_0 \leq n$ for the range of n given in (31), and applying (44) yields the result.   □

We now prove the sufficient condition given above for the invertibility of the covariance matrix of $\mathbf{S}^n$, and the bound on the norm of its inverse. Recall that a matrix $A \in \mathbb{R}^{p \times p}$ is said to be *strictly diagonally dominant* if $|a_{ii}| - \sum_{j \neq i} |a_{ij}| > 0$ for all $i \in [p]$.

Proof of Lemma 2.3. By Lemma 2.6 and the upper bound on b given in (32), for all $q \in [p]$ we have

$$
\Sigma_{q,q} - \sum_{1 \leq s \leq p, s \neq q} |\Sigma_{q,s}| \geq n^{d-1} \left(n A_n - (p-1) \kappa_0 \upsilon_\lambda^d b \right) > 0.
$$

Hence, Σ is a strictly diagonally dominant matrix, and is therefore invertible by the Gershgorin circle theorem, see for instance Theorem 15.10 of [4]. The final claim of the lemma follows from [1], where it is shown that the bound (34) holds for the norm $\|C\|_\infty = \max_i \sum_{j=1}^p |c_{ij}|$, which dominates $|C|_\infty$.   □

2.2. *The Ising and percolation models*

In this section we apply Theorems 2.1 and 2.2 over finite blocks to the net magenetization of the Ising model and the number of points in the supercritical bond percolation model that belong to an infinite cluster.

Corollary 2.7. *The bounds of Theorem 2.1 (Theorem 2.2) apply to the total magnetization $M_{\mathbf{k}}$ of (4) of the Ising model having infinite volume Gibbs measure in $\mathcal{M}_1$ of (3), and to the total number of points $U_{\mathbf{k}}$ in (6) that belong to an infinite cluster in the supercritical bond percolation model in dimensions $d \geq 2$ (when choosing α in (30) to satisfy (33)).*

Proof. We show that the fields in the two models in question posses the properties required for the application of Theorems 2.1 and 2.2. First, each must comprise almost surely uniformly bounded variables, must be second order stationary, positively associated and must have an exponentially decreasing covariance function. Clearly the uniformly bounded condition holds for both, implying the existence of second moments, so second order stationarity holds under strictly stationary.

By [20], any Ising model corresponding to an infinite volume Gibbs measure in $\mathcal{M}_1$ is strictly stationary, and by [33], positive association holds whenever the interaction J that appears in the Hamiltonian in (2) is nonnegative. By [29] and [2] (see also [15]), the covariance between spins decays at an exponential rate when $\beta > 0, h \neq 0$ and $\beta < \beta_c, h = 0$, respectively.

For the supercritical bond percolation model in dimension 2 or higher, strict stationarity follows from the fact that the vertex connection probability θ is the same for each bond $\mathbf{e} \in \mathbb{E}^d$, and that the status of each bond is independent of all others. Again by [33], the binary field $\{\mathbf{1}_{\{|\mathcal{C}(\mathbf{x})|=\infty\}} : \mathbf{x} \in \mathbb{Z}^d\}$ is positively associated. By [10], the covariance between $\mathbf{1}_{\{|\mathcal{C}(\mathbf{x})|=\infty\}}$ and $\mathbf{1}_{\{|\mathcal{C}(\mathbf{y})|=\infty\}}$ decays exponentially.

Lastly, for the application of Theorem 2.2, assumption (33) on α and Lemma 2.3 yield the required invertibility of the covariance matrix. $\qquad\square$

Remark 2.8. The earlier version [23] of this work depended on a result of the subsequently withdrawn manuscript [3] to handle the Ising model with infinite volume Gibbs measures in $\mathcal{M}_2$ of (5). In particular, [23] shows how the results of Corollary 2.7 apply to this case of the Ising model under the condition that its covariance function is exponentially decreasing.

2.3. *The voter model*

In this section, we consider the d-dimensional voter model $\{\eta_t : t \geq 0\}$ with initial state at time zero given by independent Bernoulli variables with parameter $\theta \in (0, 1)$ on each site in $\mathbb{Z}^d$, and the occupation time process $\{T_s^t : t \geq 0\}$, $s \geq 0$, defined in (7), that records the amount of time in $(s, s + t]$ that the origin spends in state one. The work [12] demonstrates that T_0^t satisfies the CLT for $d \geq 2$ and any $t > 0$. Here we refine the results for $d \geq 7$ in Theorem 2.9 by applying Theorem 1.3 to obtain a rate of convergence in the L^1 metric.

Though the authors are unaware of any results in the literature that yield distributional bounds for the voter model, Stein's method has been successful in obtaining bounds in the anti-voter model, see [36]. The anti-voter model differs from the voter model in that after a uniformly neighbor of a vertex is chosen the state of that vertex changes to the opposite state of that neighbor. The work of [36] used the exchangeable pair approach to Stein's method, and positive association was not an ingredient.

By [12], for $d \geq 5$ we have

$$\mathbb{E}T_0^t = \theta t \text{ and } \lim_{t \to \infty} A_t = \kappa_2 \text{ for some } \kappa_2 > 0, \text{ where } A_t = \frac{\text{Var}(T_0^t)}{t} \text{ for } t > 0. \tag{48}$$

Letting

$$A_s^t = \frac{\text{Var}(T_s^t)}{t}, \tag{49}$$

we will show in Lemma 2.14 that for all $s \geq 0$, $\lim_{t \to \infty} A_s^t = \kappa_2$, and for all $t > 0$ that $\mathbb{E}T_s^t = \theta t$ and $A_s^t > 0$.

Standardizing T_s^t to have mean zero and variance one we obtain

$$W_s^t = \frac{T_s^t - \theta t}{\sqrt{A_s^t t}}. \tag{50}$$

Now we define the last exit time of the origin in $[0, u]$ and $[0, \infty)$ by

$$L_u = \sup\{s \leq u : Y_s = \mathbf{0}\} \quad \text{for } u \geq 0 \tag{51}$$

and

$$L = \sup\{s < \infty : Y_s = \mathbf{0}\}, \tag{52}$$

respectively, where Y_s is a rate 1 simple symmetric random walk in $\mathbb{Z}^d$ starting at the origin at time zero. By definition, it is clear that L_u is increasing in u and $L_u \leq L$ for all $u \geq 0$. The following theorem provides L^1 bounds to the normal

for the standardized occupation time W_s^t, given in (50). Lemma 2.13 shows that the second moment of L_u, appearing in the bounds below, is finite.

Theorem 2.9. *For W_s^t as defined in (50), for all $d \geq 7$ and $s \geq 0$,*

$$d_1\big(\mathcal{L}\big(W_s^t\big), \mathcal{L}(Z)\big) \leq \left(\frac{180\sqrt{2}\theta(1-\theta)\mathbb{E}L_{2(s+t)}^2}{\sqrt{\pi}(A_s^t)^{3/2}} \right)^{1/2} t^{-1/4}$$

$$\text{for all } t \geq \left(\frac{\sqrt{2}\theta(1-\theta)\mathbb{E}L_{2(s+t)}^2}{5\sqrt{\pi A_s^t}} \right)^{2/3},$$

with $Z \sim \mathcal{N}(0, 1)$, A_s^t is as in (49) and L_u as in (51).

We also extend Theorem 2.9 to the multidimensional case, obtaining a bound in Theorem 2.10 in the metric $d_{\mathcal{H}_{3,\infty,p}}$ to the multivariate normal for $\mathbf{S}^t = (T_{s_1}^t, \ldots, T_{s_p}^t)$. The results could easily be extended to the case where the occupation times are measured over intervals of varying length, as recorded in the vector $(T_{s_1}^{t_1}, \ldots, T_{s_p}^{t_p})$.

Theorem 2.10. *Let $d \geq 7$ and $t > 0$. For $p \in \mathbb{N}_1$ let $s_1, \ldots, s_p \geq 0$ be such that*

$$\min_{k,l \in [p], k \neq l} |s_k - s_l| \geq (1 - \alpha)t, \tag{53}$$

for some $0 < \alpha < 1$. Let $\mathbf{S}^t = (T_{s_1}^t, \ldots, T_{s_p}^t)$ where T_s^t is defined as in (7), assume that the covariance matrix Σ of $\mathbf{S}^t$ is invertible and let

$$\psi_t = t^{1/2}\big|\Sigma^{-1/2}\big|_\infty.$$

Then, there exists a constant $C_{p,\theta}$ such that with $\mathbf{Z}$ a standard normal random vector in $\mathbb{R}^p$,

$$d_{\mathcal{H}_{3,\infty,p}}\big(\mathcal{L}\big(\Sigma^{-1/2}\big(\mathbf{S}^t - \theta t\big)\big), \mathcal{L}(\mathbf{Z})\big) \leq C_{p,\theta}\left(\left(\sum_{k=1}^p A_{s_k}^t + \alpha + \frac{1}{t} \right)^{1/2} \psi_t^{5/2} t^{-1/4} + \psi_t^2\left(\alpha + \frac{1}{t} \right) \right)$$

$$\text{for all } t \geq \left(\psi_t\left(\sum_{k=1}^p A_{s_k}^t + \alpha + \frac{1}{t} \right) \right)^{-2/3}. \tag{54}$$

From (49) we have $\text{Var}(T_s^t) = t A_s^t$ and we show in Lemma 2.14 that $A_s^t \to \kappa_2$ as $t \to \infty$. For checking the hypothesis that the covariance matrix of $\mathbf{S}^t$ is invertible and the quantity ψ_t is of order one, we present the following sufficient condition, proved at the end of this section. In addition, we see from (54) that if $\psi_t = O(1)$ and $\alpha = O(t^{-\frac{1}{4}})$ then the bound will also be of order $O(t^{-\frac{1}{4}})$.

Lemma 2.11. *With $t > 0$, Σ the covariance matrix of the random vector $\mathbf{S}^t \in \mathbb{R}^p$, $p \geq 2$ as given in Theorem 2.10 is invertible if for some $b \geq 0$*

$$|s_k - s_l| \geq t - b \quad \text{for all } k \neq l, k, l \in [p], \quad \text{and} \quad b < \frac{t \min_{k \in [p]} A_{s_k}^t - (p-1)\theta(1-\theta)\mathbb{E}[L^2]}{2(p-1)\theta(1-\theta)\mathbb{E}[L]}, \tag{55}$$

where A_s^t is given in (49) and L in (52). If $b = \alpha t$ for α satisfying

$$0 < \alpha < \min\left\{ 1, \frac{\min_{k \in [p]} A_{s_k}^t - (p-1)\theta(1-\theta)\mathbb{E}[L^2]/t}{2(p-1)\theta(1-\theta)\mathbb{E}[L]} \right\},$$

then the matrix Σ is invertible and

$$\left|\Sigma^{-1}\right|_{\infty} \leq \frac{1}{t\left(\min_{k\in[p]} A_{s_k}^t - (p-1)\theta(1-\theta)(\mathbb{E}[L^2]/t + 2\alpha\mathbb{E}[L])\right)}.$$

To prove Theorems 2.9 and 2.10, we apply the following result implied by (0.8) and (0.9) of [12], shown there using a duality that connects the voter model with a system of coalescing random walks constructed by a time reversal, and tracing the $\{0, 1\}$ "opinion" of every site back to its genesis at time zero.

Lemma 2.12 ([12]). *For $t \geq 0$ and $0 \leq u \leq v$, with L_u as in (51),*

$$\text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) = \theta(1-\theta)P(L_{u+v} > v - u). \tag{56}$$

We now use Lemma 2.12 to prove the following result which will be used in several places in this section.

Lemma 2.13. *For all $0 \leq r < s < t$,*

$$\int_r^s \int_s^t \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du \leq \frac{\theta(1-\theta)}{2}\mathbb{E}\big[L_{s+t}^2\big],$$

where L_u is as in (51). Moreover, with L as in (52) we have $\mathbb{E}[L_u^2] \leq \mathbb{E}[L^2] < \infty$ for all $u > 0$ and $d \geq 7$.

Proof. Applying covariance identity (56) from Lemma 2.12 and using the fact that L_u is increasing in u, we obtain

$$\int_r^s \int_s^t \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du = \theta(1-\theta) \int_r^s \int_s^t P(L_{u+v} > v - u) \, dv \, du$$

$$\leq \theta(1-\theta) \int_r^s \int_s^t P(L_{s+t} > v - u) \, dv \, du.$$

By a change of variables that preserves the difference $v - u$, we have

$$\int_r^s \int_s^t P(L_{s+t} > v - u) \, dv \, du = \int_0^{s-r} \int_{s-r}^{t-r} P(L_{s+t} > v - u) \, dv \, du$$

$$\leq \int_0^{s-r} \int_{s-r}^t P(L_{s+t} > v - u) \, dv \, du.$$

Hence,

$$\int_r^s \int_s^t \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du \leq \theta(1-\theta) \int_0^{s-r} \int_{s-r}^t P(L_{s+t} > v - u) \, dv \, du$$

$$= \theta(1-\theta) \int_0^{s-r} \int_{s-r-u}^{t-u} P(L_{s+t} > v) \, dv \, du$$

$$\leq \theta(1-\theta) \int_0^{s-r} \int_{s-r-u}^t P(L_{s+t} > v) \, dv \, du$$

$$= \theta(1-\theta) \int_0^{s-r} \int_u^t P(L_{s+t} > v) \, dv \, du$$

$$\leq \theta(1-\theta) \int_0^t \int_u^t P(L_{s+t} > v) \, dv \, du$$

$$= \theta(1-\theta) \int_0^t \int_0^v P(L_{s+t} > v) \, du \, dv$$

$$= \frac{\theta(1-\theta)}{2} \int_0^t 2v P(L_{s+t} > v) \, dv$$

$$\leq \frac{\theta(1-\theta)}{2} \mathbb{E}[L_{s+t}^2],$$

where change of variables are applied in first and second equality, Fubini's theorem in the third equality, and where in the final inequality we apply the standard fact, easily shown using Fubini's theorem, that if $\phi(y)$ is a differentiable function on $[0, \infty)$ satisfying $\phi(0) = 0$ then for $Y \geq 0$,

$$\mathbb{E}[\phi(Y)] = \int_0^\infty \phi'(y) P(Y > y) \, dy. \tag{57}$$

Since $L_u \leq L$ for all $u \geq 0$, to prove the second claim, it suffices to show that $\mathbb{E}[L^2] < \infty$ for $d \geq 7$. Returning to the integral expression for the second moment, and slightly modifying the calculation for the first moment in [12] following (0.11), we have

$$\mathbb{E}[L^2] = \int_0^\infty 2t P(L > t) \, dt = \int_0^\infty 2t \int_t^\infty P(Y_s = 0) \, ds \, \gamma_d \, dt, \tag{58}$$

where $\gamma_d = P(Y_t \neq 0 \; \forall t \geq 1)$, known to be positive and increasing for $d \geq 3$. Changing the order of integration in (58) yields

$$\mathbb{E}[L^2] = \gamma_d \int_0^\infty \int_0^s 2t P(Y_s = \mathbf{0}) \, dt \, ds = \gamma_d \int_0^\infty s^2 P(Y_s = \mathbf{0}) \, ds,$$

implying that $\mathbb{E}[L^2] < \infty$ for $d \geq 7$ as, see [12] for instance, $P(Y_s = \mathbf{0}) = O(s^{-d/2})$. □

The next result shows some important properties of the mean and the variance of T_s^t.

Lemma 2.14. *For $d \geq 7$, $t > 0$ and $s \geq 0$, with T_s^t as in (7) and A_s^t as in (49),*

$$\mathbb{E}T_s^t = \theta t, \quad A_s^t > 0 \quad and \quad \lim_{t \to \infty} A_s^t = \kappa_2,$$

where κ_2 is given in (48).

Proof. We verify the first claim using the definition of T_s^t and (48) to yield

$$\mathbb{E}T_s^t = \mathbb{E}[T_0^{t+s} - T_0^s] = (t+s)\theta - s\theta = t\theta.$$

The second claim follows from (2.1) in [12] and the fact that $P(Y_u = \mathbf{0})$ is strictly positive for all $u > 0$.

For the final claim, note that for $s = 0$, $A_0^t = A_t$ and the result reduces to (48). For $s > 0$, we have

$$A_s^t = \frac{\mathrm{Var}(T_s^t)}{t} = \frac{\mathrm{Var}(T_0^{t+s} - T_0^s)}{t}$$

$$= \frac{\mathrm{Var}(T_0^{t+s}) + \mathrm{Var}(T_0^s) - 2\,\mathrm{Cov}(T_0^s, T_0^{t+s})}{t}$$

$$= \frac{(t+s)A_{t+s}}{t} + \frac{sA_s}{t} - \frac{2\,\mathrm{Cov}(T_0^{t+s}, T_0^s)}{t}. \tag{59}$$

The first term converges to κ_2 and the second term tends to zero as $t \to \infty$. Hence it suffices to show that the last term also tends to zero.

Writing the covariance in integral form in two parts and applying Lemmas 2.12 and 2.13 to the first and second terms, respectively, we obtain

$$
\begin{aligned}
\mathrm{Cov}\big(T_0^s, T_0^{t+s}\big) &= \int_0^s \int_0^s \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du + \int_0^s \int_s^{s+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du \\
&\leq 2\theta(1-\theta) \int_0^s \int_u^s P(L_{u+v} > v - u)\,dv\,du + \frac{\theta(1-\theta)}{2}\mathbb{E}\big[L_{2s+t}^2\big] \\
&\leq s^2 + \mathbb{E}\big[L^2\big],
\end{aligned}
$$

and thus the last term in (59) tends to zero as $t \to \infty$ since $\mathbb{E}[L^2]$ is finite for $d \geq 7$ by Lemma 2.13. $\qquad\square$

Now we use Lemma 2.13 to prove Lemma 2.15, which bounds the sum of the covariances between $X_{s,i}^t$ and $X_{s,j}^t$ defined in (12) over $i \neq j \in [m]$, followed by Lemma 2.16, which bounds the covariance between occupation times starting at r, s satisfying $|r - s| \geq t - b$.

Lemma 2.15. *For $t > 0$, $s \geq 0$ and $m \in \mathbb{N}_1$, with $X_{s,i}^t$ as in (12) and L_u as in (51),*

$$
\sum_{i,j\in[m],i\neq j} \mathrm{Cov}\big(X_{s,i}^t, X_{s,j}^t\big) \leq \theta(1-\theta)(m-1)\mathbb{E}\big[L_{2(s+t)}^2\big].
$$

Proof. Using the definition of $X_{s,i}^t$ in (12), we have

$$
\begin{aligned}
\sum_{i,j\in[m],i\neq j} \mathrm{Cov}\big(X_{s,i}^t, X_{s,j}^t\big) &= 2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \mathrm{Cov}\big(X_{s,i}^t, X_{s,j}^t\big) \\
&= 2\sum_{i=1}^{m-1} \mathrm{Cov}\bigg(X_{s,i}^t, \sum_{j=i+1}^{m} X_{s,j}^t\bigg) \\
&= 2\sum_{i=1}^{m-1} \int_{s+(i-1)t/m}^{s+it/m} \int_{s+it/m}^{s+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du.
\end{aligned}
$$

Applying Lemma 2.13 to the integrals above, and using the monotonicity property of L_u in $u > 0$, we have

$$
\sum_{i,j\in[m],i\neq j} \mathrm{Cov}\big(X_{s,i}^t, X_{s,j}^t\big) \leq \sum_{i=1}^{m-1} \theta(1-\theta)\mathbb{E}\big[L_{2(s+t)}^2\big] = \theta(1-\theta)(m-1)\mathbb{E}\big[L_{2(s+t)}^2\big].
$$

$\qquad\square$

Lemma 2.16. *For $t > 0$, let $0 \leq b \leq t$ and let $r, s \geq 0$ satisfy $|r - s| \geq t - b$. Then, with T_s^t as in (7) and L as in (52),*

$$
\mathrm{Cov}\big(T_s^t, T_r^t\big) \leq \theta(1-\theta)\big(\mathbb{E}\big[L^2\big] + 2b\mathbb{E}[L]\big). \tag{60}
$$

Proof. It suffices to consider the case $r \geq s$. First assume that $r < s + t$. Using the definition of T_s^t in (7) and breaking the integral that expresses the covariance we wish to bound into three parts, we have

$$
\begin{aligned}
\mathrm{Cov}\big(T_s^t, T_r^t\big) &= \int_s^{s+t} \int_r^{r+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du \\
&= \int_s^r \int_r^{r+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du + \int_r^{s+t} \int_{s+t}^{r+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du \\
&\quad + \int_r^{s+t} \int_r^{s+t} \mathrm{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big)\,dv\,du.
\end{aligned}
$$

Applying Lemma 2.13 to the first two integrals, and using the fact that $L_u \leq L$, accounts for the first term in (60). For the last integral, using (56) from Lemma 2.12, we obtain

$$
\begin{aligned}
\int_r^{s+t} \int_r^{s+t} \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du &= 2\theta(1-\theta) \int_r^{s+t} \int_u^{s+t} P(L_{u+v} > v - u) \, dv \, du \\
&\leq 2\theta(1-\theta) \int_r^{s+t} \int_u^{s+t} P(L > v - u) \, dv \, du \\
&= 2\theta(1-\theta) \int_r^{s+t} \int_0^{s+t-u} P(L > v) \, dv \, du \\
&\leq 2\theta(1-\theta) \int_r^{s+t} \int_0^\infty P(L > v) \, dv \, du \\
&= 2\theta(1-\theta) \int_r^{s+t} \mathbb{E}[L] \, du \\
&= 2\theta(1-\theta)(t - r + s)\mathbb{E}[L] \\
&\leq 2\theta(1-\theta)b\mathbb{E}[L],
\end{aligned}
$$

thus accounting for the second term in (60), where a change of variables is applied in the second equality, (57) in the third equality, and the assumption that $r - s \geq t - b$ in the final inequality.

When $r \geq s + t$, by the positivity of the covariance due to assocation, we have

$$
\begin{aligned}
\text{Cov}\big(T_s^t, T_r^t\big) &= \int_s^{s+t} \int_r^{r+t} \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du \\
&\leq \int_s^r \int_r^{r+t} \text{Cov}\big(\eta_u(\mathbf{0}), \eta_v(\mathbf{0})\big) \, dv \, du.
\end{aligned}
$$

Now applying Lemma 2.13 to the integral above, it is bounded by $\theta(1-\theta)\mathbb{E}[L^2]/2$, hence the claim of the lemma is true in this case as it holds with b replaced by zero. $\qquad\square$

Now we use Theorem 1.3 and Lemmas 2.13 and 2.15 to prove Theorem 2.9.

Proof of Theorem 2.9. Let m be a positive integer and

$$
\xi_{s,i}^t = \frac{X_{s,i}^t - (t/m)\theta}{\sqrt{A_s^t t}} \quad \text{for } i \in [m], s \geq 0, \tag{61}
$$

where $X_{s,i}^t$ is defined as in (12). We apply Theorem 1.3 to $W_s^t = \sum_{i=1}^m \xi_{s,i}^t$, having mean zero and variance one. By [21], the components of the vector $\boldsymbol{\xi}_t^s = (\xi_{s,1}^t, \ldots, \xi_{s,m}^t)$ are positively associated as they are increasing functions of the positively associated variables $\{X_{s,i}^t : i \in \mathbb{Z}\}$.

We now handle the terms on the right-hand side of the bound (20). For the first term, using the definition of $X_{s,i}^t$ in (12) we have

$$
0 \leq X_{s,i}^t = \int_{s+(i-1)t/m}^{s+it/m} \eta_u(\mathbf{0}) \, du \leq \int_{s+(i-1)t/m}^{s+it/m} 1 \, ds = \frac{t}{m},
$$

and thus, by using the definition of $\xi_{s,i}^t$ in (61),

$$
\big|\xi_{s,i}^t\big| \leq B \quad \text{where } B = \frac{2\sqrt{t}}{m\sqrt{A_s^t}} \text{ for all } i \in [m].
$$

Applying Lemma 2.15 for the second term, invoking Theorem 1.3 now yields

$$d_1\big(\mathcal{L}\big(W_s^t\big), \mathcal{L}(Z)\big) \le \frac{10\sqrt{t}}{m\sqrt{A_s^t}} + \frac{2\sqrt{2}\theta(1-\theta)m\mathbb{E}[L_{2(s+t)}^2]}{\sqrt{\pi}A_s^t t},$$

where $\mathbb{E}[L_{2(s+t)}^2] < \infty$ by Lemma 2.13. Applying (44) with $d = 1$ and with m in place of l now yields the result, noting that the lower bound on t in the theorem implies that

$$m_0 = \left(\frac{5\sqrt{\pi A_s^t}t^{3/2}}{\sqrt{2}\theta(1-\theta)\mathbb{E}L_{2(s+t)}^2}\right)^{1/2}$$

is at least one. $\qquad\square$

To prove Theorem 2.10 we apply Theorem 1.4 and use the same techniques as for Theorem 2.9. For this result we index constants by the parameters on which they depend, though we do not explicitly compute them.

Proof of Theorem 2.10. For $m \in \mathbb{N}_1$ decompose $T_{s_k}^t - \theta t$ for $k \in [p]$ as the sum over $i \in [m]$ of the variables $\xi_i^k = X_{s_k,i}^t - (t/m)\theta$ where $X_{s_k,i}^t$ is given in (12).

Applying Theorem 1.4, we handle the three terms on the right-hand side of (22). In the calculation below we use that $\mathbb{E}[L]$ and $\mathbb{E}[L^2]$ are finite for $d \ge 7$ by Lemma 2.13. For the first term, again using the definition of $X_{s_k,i}^t$, we have

$$\big|\xi_i^k\big| \le B \quad \text{where } B = \frac{2t}{m} \text{ for all } i \in [m] \text{ and } k \in [p].$$

Since $|\Sigma^{-1/2}|_\infty = t^{-1/2}\psi_t$, and $\Sigma_{k,k} = tA_{s_k}^t$ we may bound the first term as

$$\left(\frac{1}{6} + 2\sqrt{2}\right)p^3 B t^{-3/2}\psi_t^3 \sum_{k=1}^p \Sigma_{k,k} = \frac{C_p\psi_t^3\sqrt{t}}{m}\sum_{k=1}^p A_{s_k}^t. \tag{62}$$

For the second term, by Lemma 2.15 we have

$$\left(\frac{3}{\sqrt{2}} + \frac{1}{2}\right)p^2 t^{-1}\psi_t^2 \sum_{k=1}^p \sum_{i,j\in[m],i\neq j} \mathbb{E}\big(\xi_i^k\xi_j^k\big) = C_p\psi_t^2 \sum_{k=1}^p \sum_{i,j\in[m],i\neq j} \frac{\text{Cov}(X_{s_k,i}^t, X_{s_k,j}^t)}{t}$$

$$\le \frac{C_{p,\theta}\psi_t^2 m}{t}\sum_{k=1}^p \mathbb{E}[L_{s_k+t}^2] \le \frac{C_{p,\theta}\psi_t^2 m}{t}, \tag{63}$$

where we have applied Lemma 2.13 for the final inequality.

Invoking Lemma 2.16 and assumption (53) we have

$$\Sigma_{k,l} = \text{Cov}\big(T_{s_k}^t, T_{s_l}^t\big) \le \theta(1-\theta)\big(\mathbb{E}[L^2] + 2\alpha t\mathbb{E}[L]\big) \quad \text{for } k \neq l \in [p],$$

and hence we may bound the last term as

$$\left(2\sqrt{2}p^3 B t^{-3/2}\psi_t^3 + \left(\frac{3}{\sqrt{2}} + \frac{1}{2}\right)p^2 t^{-1}\psi_t^2\right) \sum_{k,l\in[p],k\neq l} \Sigma_{k,l}$$

$$\le \left(\frac{C_p t\psi_t^3}{mt^{3/2}} + \frac{C_p\psi_t^2}{t}\right)\theta(1-\theta)\big(\mathbb{E}[L^2] + 2\alpha t\mathbb{E}[L]\big)$$

$$\le \frac{C_{p,\theta}\psi_t^3}{m\sqrt{t}} + \frac{C_{p,\theta}\psi_t^2}{t} + \frac{C_{p,\theta}\alpha\sqrt{t}\psi_t^3}{m} + C_{p,\theta}\alpha\psi_t^2. \tag{64}$$

By Theorem 1.4 and (62), (63) and (64),

$$
d_{\mathcal{H}_{3,\infty,p}}\big(\mathcal{L}\big(\Sigma^{-1/2}(\mathbf{S}^t - \theta t)\big), \mathcal{L}(\mathbf{Z})\big) \leq \left(C_p \sqrt{t} \sum_{k=1}^{p} A_{s_k}^t + \frac{C_{p,\theta}}{\sqrt{t}} + C_{p,\theta}\alpha\sqrt{t} \right) \frac{\psi_t^3}{m}
$$

$$
+ \frac{C_{p,\theta}\psi_t^2 m}{t} + \frac{C_{p,\theta}\psi_t^2}{t} + C_{p,\theta}\alpha\psi_t^2
$$

$$
\leq C_{p,\theta}\left(\left(\sqrt{t} \sum_{k=1}^{p} A_{s_k}^t + \frac{1}{\sqrt{t}} + \alpha\sqrt{t} \right) \frac{\psi_t^3}{m} + \frac{\psi_t^2 m}{t} + \psi_t^2\left(\alpha + \frac{1}{t}\right) \right).
$$

Applying (44) to the first two terms in parenthesis now yields the result, noting that

$$
m_0 = \left(\psi_t \left(\sum_{k=1}^{p} A_{s_k}^t + \alpha + \frac{1}{t} \right) \right)^{\frac{1}{2}} t^{\frac{3}{4}}
$$

is at least one for the range of t given in (54). $\square$

Finally, we present the sufficient condition given above for the invertibility of the covariance matrix of $\mathbf{S}^t$.

Proof of Lemma 2.11. By Lemma 2.16 and the upper bound on b given in (55), for all $k \in [p]$ we have

$$
\Sigma_{k,k} - \sum_{l \in [p], l \neq k} |\Sigma_{k,l}| \geq t A_{s_k}^t - (p-1)\theta(1-\theta)\big(\mathbb{E}[L^2] + 2b\mathbb{E}[L]\big) > 0.
$$

Hence, Σ is a strictly diagonally dominant matrix, and the claims follow as in the proof of Lemma 2.3. $\square$

2.4. *The contact process*

As detailed in the Introduction, in this section we consider the functional

$$
D_{s,f}^t = \int_s^{s+t} f\big(\zeta^{\nu_\lambda}(u)\big)\,du \quad \text{for } s \geq 0, t > 0, \tag{65}
$$

where f is a nonconstant increasing cylindrical function for the supercritical one dimensional contact process $\{\zeta^{\nu_\lambda}(t) : t \geq 0\}$ with infection rate λ, recovery rate 1, and initial configuration having distribution ν_λ, the unique nontrivial invariant measure of the process. It was proved in [38] that in the supercritical case $D_{0,f}^t$ satisfies the CLT for any cylindrical f. In Theorems 2.19 and 2.20 we provide finite sample bounds in the L^1 metric, and a multivariate smooth function metric, for this asymptotic when f is increasing and nonconstant by applying Theorems 1.3 and 1.4, respectively.

In addition to the time dependent quantity (65) studied in this section, the results in Section 2.1 also allow us to obtain a version of Corollary 2.7 for the number of infected sites at any fixed time $t > 0$ in a bounded region for the supercritical multidimensional contact process $\{\zeta_t^{\mathbb{Z}^d}(\mathbf{x}) : \mathbf{x} \in \mathbb{Z}^d\}$ with intial state having mass one on $\mathbb{Z}^d$. In particular, the values of the process at t are positively associated by Theorem B17 of [32], and the covariance between $\zeta_t^{\mathbb{Z}^d}(\mathbf{x})$ and $\zeta_t^{\mathbb{Z}^d}(\mathbf{y})$ decays exponentially by Theorem 1.7 of [22]. Strictly stationary follows from the fact that at time zero all sites are infected.

For the main goal in this section, we apply Lemma 2.17 which provides exponential decay of the covariances of $f(\zeta^{\nu_\lambda}(\cdot))$ in time. We note the brief but important comment before the statement of Lemma 1 of [38], that the process considered there, and defined in an indirect manner, has the distribution of the contact process started in its unique invariant nontrivial distribution. We recall the definition of λ_* given in (8).

Lemma 2.17 (Lemma 2 of [38]). *For $\lambda > \lambda_*$ and for any cylindrical $f : \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$ there exist positive constants $\gamma = \gamma(f, \lambda)$ and $\kappa = \kappa(f, \lambda)$ depending on λ and f such that*

$$\left| \mathrm{Cov}\big(f\big(\zeta^{\nu_\lambda}(r)\big), f\big(\zeta^{\nu_\lambda}(s)\big)\big)\right| \leq \kappa e^{-\gamma|s-r|} \quad \text{for } \{r, s\} \subset (0, \infty). \tag{66}$$

The results of [38] make use of the fact that any cylindrical function f is a finite linear combination of indicators of the form $I_{\bar{B}}(\cdot)$ given in (10) with $|B| < \infty$. Hence, $M_f := |f|_\infty < \infty$, implying that the moments of f on the process always exist.

The next lemma verifies the positive association required for the application of our main results; let $Y_{s,i}^{t,m}$ be as in (13).

Lemma 2.18. *For $\lambda > \lambda_*$ and any increasing cylindrical function $f : \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$, the families $\{f(\zeta^{\nu_\lambda}(t)) : t \geq 0\}$, and $\{Y_{s,i}^{t,m} : i \in [m]\}$ for $s \geq 0, t > 0$ and $m \in \mathbb{N}_1$, are positively associated.*

Proof. By applying (2.21) in [31], a corollary of Harris' theorem, the proof of Lemma 1 of [38] demonstrates that $\{\zeta^{\nu_\lambda}(t) : t \geq 0\}$ is a positively associated family of random variables. Now both claims follow by the fact that both families in the statement are increasing functions of $\{\zeta^{\nu_\lambda}(t) : t \geq 0\}$. $\qquad \square$

Using the remark above Proposition 1 in [38], the process $\{\zeta^{\nu_\lambda}(t) : t \in \mathbb{R}\}$ is strictly stationary, hence we may let

$$\frac{\mathrm{Var}(D_{s,f}^t)}{t} = A_f^t \quad \text{for all } s \geq 0, t > 0. \tag{67}$$

Since f is nonconstant increasing and $\zeta^{\nu_\lambda}(u)$ is strictly stationary, by the Remark under Lemma 1 of [38], $\mathrm{Var}(f(\zeta^{\nu_\lambda}(u))) > 0$ and does not depend on u. Since $f(\zeta^{\nu_\lambda}(u))$, $u \geq 0$ is positively associated, we have that $A_f^t > 0$ for all $t > 0$. Using the definition of $D_{s,f}^t$ in (65), we also have $A_f^t \leq 2M_f^2 < \infty$.

Standardizing $D_{s,f}^t$ to have mean zero and variance one, we obtain

$$W_{s,f}^t = \frac{D_{s,f}^t - \mathbb{E}D_{s,f}^t}{\sqrt{tA_f^t}} = \frac{\sum_{i=1}^m (Y_{s,i}^{t,m} - \mathbb{E}Y_{s,i}^{t,m})}{\sqrt{tA_f^t}}. \tag{68}$$

The following lemma provides a bound on the L^1 distance between $W_{s,f}^t$ and the standard normal.

Theorem 2.19. *Let $\lambda > \lambda_*$, $s \geq 0$ and $t > 0$. Then for any nonconstant increasing cylindrical function $f : \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$, with $W_{s,f}^t$ as in (68) and $Z \sim \mathcal{N}(0, 1)$,*

$$d_1\big(\mathcal{L}\big(W_{s,f}^t\big), \mathcal{L}(Z)\big) \leq \left(\frac{360\sqrt{2}\kappa M_f}{\sqrt{\pi}(A_f^t)^{3/2}\gamma^2} \right)^{1/2} t^{-1/4}, \quad \text{for all } t \geq \left(\frac{2\sqrt{2}\kappa}{5\gamma^2 M_f \sqrt{\pi A_f^t}} \right)^{2/3},$$

where $M_f = |f|_\infty < \infty$, and κ and γ are as in (66).

We also prove the following multidimensional version of Theorem 2.19 for the p-vector

$$\mathbf{S}_f^t = \big(D_{s_1,f}^t, \ldots, D_{s_p,f}^t \big) \quad \text{with } s_1, \ldots, s_p \geq 0, \tag{69}$$

where $D_{s,f}^t$ is defined in (65).

Theorem 2.20. *Let $\lambda > \lambda_*$, $t > 0$ and for $p \in \mathbb{N}_1$ let $s_1, \ldots, s_p \geq 0$ satisfy*

$$\min_{k,l \in [p], k \neq l} |s_k - s_l| \geq (1 - \alpha)t$$

for some $0 < \alpha < 1$, and let $f : \mathscr{P}(\mathbb{Z}) \to \mathbb{R}$ be a nonconstant increasing cylindrical function. Assume that the covariance matrix Σ of $\mathbf{S}_f^t$ in (69) is invertible and let

$$\psi_f^t = t^{1/2} \big| \Sigma^{-1/2} \big|_\infty.$$

Then, there exists a constant $C_{p,f,\lambda}$ such that with $\mathbf{Z}$ a standard normal random vector in $\mathbb{R}^p$,

$$d_{\mathcal{H}_{3,\infty,p}}\big(\mathcal{L}\big(\Sigma^{-1/2}\big(\mathbf{S}_f^t - \mathbb{E}\mathbf{S}_f^t\big)\big), \mathcal{L}(\mathbf{Z})\big)$$

$$\leq C_{p,f,\lambda}\left(\left(A_f^t + \alpha + \frac{1}{t}\right)^{1/2}\big(\psi_f^t\big)^{5/2} t^{-1/4} + \big(\psi_f^t\big)^2\left(\alpha + \frac{1}{t}\right)\right)$$

$$\text{for all } t \geq \left(\psi_f^t\left(A_f^t + \alpha + \frac{1}{t}\right)\right)^{-2/3}. \tag{70}$$

We prove Theorems 2.19 and 2.20 using Theorems 1.3 and 1.4. Alternatively, one may apply Theorems 2.1 and 2.2 by breaking $D_{s,f}^t$ into a sum of integrals over intervals of length one and a "remainder integral" over an interval of length $t - \lfloor t \rfloor$. However, this approach will result in constants larger than the ones stated in Theorems 2.19 and 2.20 which breaks $D_{s,f}^t$ into integrals over intervals of equal, optimal length.

For checking the hypothesis that the covariance matrix of $\mathbf{S}_f^t$ is invertible and the quantity ψ_f^t is of order one, we present the following sufficient condition. We see from (70) that if $\psi_f^t = O(1)$ and $\alpha = O(t^{-\frac{1}{4}})$ then the bound will also be of order $O(t^{-\frac{1}{4}})$.

Lemma 2.21. *With $t > 0$, the covariance matrix Σ of the random vector $\mathbf{S}_f^t \in \mathbb{R}^p$, $p \geq 2$ in (69) is invertible if for some $b \geq 0$*

$$|s_k - s_l| \geq t - b \quad \text{for all } k \neq l, k, l \in [p], \quad \text{and} \quad b < \left(\frac{t A_f^t \gamma}{2(p-1)\kappa} - \frac{1}{\gamma}\right),$$

where A_f^t is given in (67) and κ, γ in (66). If $b = \alpha t$ for some α satisfying

$$0 < \alpha < \min\left\{1, \frac{A_f^t \gamma}{2(p-1)\kappa} - \frac{1}{\gamma t}\right\},$$

then the matrix Σ is invertible and

$$\big|\Sigma^{-1}\big|_\infty \leq \frac{1}{t(A_f^t - 2\kappa(p-1)(\alpha/\gamma + 1/(\gamma^2 t)))}.$$

Since Theorems 2.19 and 2.20 and Lemma 2.21 have the same conclusions as the voter model results in Theorems 2.9 and 2.10 and Lemma 2.11, respectively, only differing by constants depending on the parameters of the model, it suffices to prove "contact process" versions of Lemmas 2.15 and 2.16 and then the results in this section follow similarly as for their voter model counterparts. Proceeding in this manner, Lemma 2.22 bounds the sum of the covariances between $Y_{s,i}^{t,m}$ and $Y_{s,j}^{t,m}$, defined in (13), over distinct $i, j \in [m]$, and Lemma 2.23 bounds the covariance between $D_{r,f}^t$ and $D_{s,f}^t$ where r and s are at least $t - b$ apart for some $0 < b \leq t$.

Lemma 2.22. *For $t > 0$, $s \geq 0$ and $m \in \mathbb{N}_1$, with $Y_{s,i}^{t,m}$ as in (13),*

$$\sum_{i,j \in [m], i \neq j} \text{Cov}\big(Y_{s,i}^{t,m}, Y_{s,j}^{t,m}\big) \leq \frac{2\kappa m}{\gamma^2},$$

where κ and γ are as in (66).

Proof. Applying Lemma 2.17 and using the stationarity of $\zeta^{\nu_\lambda}(\cdot)$, for $1 \leq i < j \leq m$ we have

$$
\begin{aligned}
\operatorname{Cov}\big(Y_{s,i}^{t,m}, Y_{s,j}^{t,m}\big) &= \int_{(i-1)t/m}^{it/m} \int_{(j-1)t/m}^{jt/m} \operatorname{Cov}\big(f\big(\zeta^{\nu_\lambda}(u)\big), f\big(\zeta^{\nu_\lambda}(v)\big)\big) \, du \, dv \\
&\leq \kappa \int_{(i-1)t/m}^{it/m} \int_{(j-1)t/m}^{jt/m} e^{-\gamma(u-v)} \, du \, dv = \kappa \int_{(j-1)t/m}^{jt/m} e^{-\gamma u} \, du \int_{(i-1)t/m}^{it/m} e^{\gamma v} \, dv \\
&= \frac{\kappa}{\gamma^2} \big(e^{\gamma t/2m} - e^{-\gamma t/2m}\big)^2 e^{-(j-i)\gamma t/m}.
\end{aligned}
$$

Summing $\operatorname{Cov}(Y_{s,i}^{t,m}, Y_{s,j}^{t,m})$ over $i \neq j$, we obtain

$$
\begin{aligned}
\sum_{i,j\in[m], i\neq j} \operatorname{Cov}\big(Y_{s,i}^{t,m}, Y_{s,j}^{t,m}\big) &= 2 \sum_{1\leq i<j\leq m} \operatorname{Cov}\big(Y_{s,i}^{t,m}, Y_{s,j}^{t,m}\big) \\
&\leq \frac{2\kappa}{\gamma^2} \big(e^{\gamma t/2m} - e^{-\gamma t/2m}\big)^2 \sum_{1\leq i<j\leq m} e^{-(j-i)\gamma t/m} \\
&= \frac{2\kappa}{\gamma^2} \big(e^{\gamma t/2m} - e^{-\gamma t/2m}\big)^2 \sum_{k=1}^{m-1} (m-k) e^{-\gamma k t/m} \\
&= \frac{2\kappa}{\gamma^2} \big(e^{\gamma t/2m} - e^{-\gamma t/2m}\big)^2 \left(\frac{e^{-\gamma t/m}((m-1) - me^{-\gamma t/m} + e^{-\gamma t})}{(1 - e^{-\gamma t/m})^2} \right) \\
&= \frac{2\kappa}{\gamma^2} \big((m-1) - me^{-\gamma t/m} + e^{-\gamma t}\big) \\
&\leq \frac{2\kappa m}{\gamma^2},
\end{aligned}
$$

where in the third equality we apply the identity in (27) with $n = m$ and $w = e^{-\gamma t/m}$. $\qquad\square$

Lemma 2.23. *For $t > 0$, let $0 \leq b \leq t$ and let $r, s \geq 0$ satisfy $|s - r| \geq t - b$. Then, with $D_{s,f}^t$ as in (65),*

$$
\operatorname{Cov}\big(D_{r,f}^t, D_{s,f}^t\big) \leq 2\kappa \left(\frac{b}{\gamma} + \frac{1}{\gamma^2} \right),
$$

where κ and γ are as in (66).

Proof. It suffices to consider the case $s \geq r$. First assume that $s \leq r + t$. Using the definition of $D_{s,f}^t$ in (65) and breaking the integral that expresses the covariance we wish to bound into three parts, we have

$$
\begin{aligned}
\operatorname{Cov}\big(D_{s,f}^t, D_{r,f}^t\big) &= \int_s^{s+t} \int_r^{r+t} \operatorname{Cov}\big(f\big(\zeta^{\nu_\lambda}(u)\big), f\big(\zeta^{\nu_\lambda}(v)\big)\big) \, du \, dv \\
&= \int_s^{s+t} \int_r^s \operatorname{Cov}\big(f\big(\zeta^{\nu_\lambda}(u)\big), f\big(\zeta^{\nu_\lambda}(v)\big)\big) \, du \, dv \\
&\quad + \int_{r+t}^{s+t} \int_s^{r+t} \operatorname{Cov}\big(f\big(\zeta^{\nu_\lambda}(u)\big), f\big(\zeta^{\nu_\lambda}(v)\big)\big) \, du \, dv \\
&\quad + \int_s^{r+t} \int_s^{r+t} \operatorname{Cov}\big(f\big(\zeta^{\nu_\lambda}(u)\big), f\big(\zeta^{\nu_\lambda}(v)\big)\big) \, du \, dv.
\end{aligned}
$$

Now applying Lemma 2.17 we have

$$
\begin{aligned}
\mathrm{Cov}\big(D_{r,f}^t, D_{s,f}^t\big) &\le \kappa \int_s^{s+t} \int_r^s e^{-\gamma(v-u)}\,du\,dv + \kappa \int_{r+t}^{s+t} \int_s^{r+t} e^{-\gamma(v-u)}\,du\,dv \\
&\quad + \kappa \int_s^{r+t} \int_s^v e^{-\gamma(v-u)}\,du\,dv + \kappa \int_s^{r+t} \int_v^{r+t} e^{-\gamma(u-v)}\,du\,dv \\
&= \kappa \bigg(\frac{1 - e^{-\gamma(s-r)} - e^{-\gamma t} + e^{-\gamma(s-r+t)}}{\gamma^2} + \frac{1 - e^{-\gamma(r-s+t)} - e^{-\gamma(s-r)} + e^{-\gamma t}}{\gamma^2} \\
&\qquad + 2\bigg(\frac{t - (s-r)}{\gamma} - \frac{1 - e^{-\gamma(r-s+t)}}{\gamma^2} \bigg) \bigg) \\
&= \kappa \bigg(\frac{2(t - (s-r))}{\gamma} + \frac{e^{-\gamma(s-r+t)} + e^{-\gamma(r-s+t)} - 2e^{-\gamma(s-r)}}{\gamma^2} \bigg) \\
&\le \kappa \bigg(\frac{2(t - (s-r))}{\gamma} + \frac{2}{\gamma^2} \bigg) \\
&\le 2\kappa \bigg(\frac{b}{\gamma} + \frac{1}{\gamma^2} \bigg).
\end{aligned}
$$

When $s > r + t$ we have

$$
\begin{aligned}
\mathrm{Cov}\big(D_{s,f}^t, D_{r,f}^t\big) &\le \kappa \int_s^{s+t} \int_r^{r+t} e^{-\gamma(v-u)}\,du\,dv \\
&= \kappa \int_s^{s+t} e^{-\gamma v}\,dv \int_r^{r+t} e^{\gamma u}\,du \\
&= \frac{\kappa}{\gamma^2}\big(e^{-\gamma(s-r-t)} + e^{-\gamma(s-r+t)} - 2e^{-\gamma(s-r)}\big) \le \frac{2\kappa}{\gamma^2}
\end{aligned}
$$

hence the claim of the lemma is true in this case as it holds with b replaced by zero. $\qquad\square$

The proofs of Theorems 2.19 and 2.20 and Lemma 2.21 below closely follow those of Theorems 2.9 and 2.10 and Lemma 2.11 respectively, so only brief outlines are provided.

Proof of Theorem 2.19. Following the outline of the proof of Theorem 2.9, we apply Theorem 1.3 to the sum of

$$
\xi_{s,i}^t = \frac{Y_{s,i}^{t,m} - \mathbb{E}Y_{s,i}^{t,m}}{\sqrt{t A_f^t}}, \quad \text{which is absolutely bounded by } B = \frac{2\sqrt{t} M_f}{m\sqrt{A_f^t}}.
$$

Applying Lemma 2.22 for the second term of (20) yields

$$
d_1\big(\mathcal{L}\big(W_{s,f}^t\big), \mathcal{L}(Z)\big) \le \frac{10\sqrt{t} M_f}{m\sqrt{A_f^t}} + \frac{4\sqrt{2}\kappa m}{\sqrt{\pi}\gamma^2 t A_f^t}.
$$

Applying (44) with $d = 1$ and with m in place of l as in the proof of Theorem 2.9 now yields the result. $\qquad\square$

Proof of Theorem 2.20. We apply Theorem 1.4 as in the proof of Theorem 2.10. With $\xi_i^k = Y_{s_k,i}^{t,m} - \mathbb{E}Y_{s_k,i}^{t,m}$, we have $|\xi_i^k| \le B$ where $B = 2M_f t/m$. Then bounding the right-hand side of (22), using that $|\Sigma^{-1/2}|_\infty = t^{-1/2}\psi_f^t$ for the first term, applying Lemmas 2.22 and 2.23 for the second and the last terms respectively and invoking Theorem 1.4,

we obtain

$$
d_{\mathcal{H}_{3,\infty,p}}\big(\mathcal{L}\big(\Sigma^{-1/2}\big(\mathbf{S}^t - \mathbb{E}\mathbf{S}^t\big)\big), \mathcal{L}(\mathbf{Z})\big) \leq \left(C_{p,f}\sqrt{t}A_f^t + \frac{C_{p,f,\lambda}}{\sqrt{t}} + C_{p,f,\lambda}\alpha\sqrt{t} \right)\frac{(\psi_f^t)^3}{m}
$$

$$
+ \frac{C_{p,f,\lambda}(\psi_f^t)^2 m}{t} + \frac{C_{p,f,\lambda}(\psi_f^t)^2}{t} + C_{p,f,\lambda}\alpha\big(\psi_f^t\big)^2
$$

$$
\leq C_{p,f,\lambda}\left(\left(\sqrt{t}A_f^t + \frac{1}{\sqrt{t}} + \alpha\sqrt{t} \right)\frac{(\psi_f^t)^3}{m} + \frac{(\psi_f^t)^2 m}{t} + \big(\psi_f^t\big)^2\Big(\alpha + \frac{1}{t}\Big) \right).
$$

Applying (44) to the first two terms in parenthesis now yields the result. $\qquad\square$

Finally to verify Lemma 2.21, we follow the same idea as in the proof of Lemma 2.11 with the help of Lemma 2.23.

3. Proofs of main theorems

In this section we prove our main results, Theorems 1.3 and 1.4, stated in the Introduction. For this purpose we first recall that the pair of random variables (X, Y) is said to be *positive quadrant dependent*, or PQD, if for all $x, y \in \mathbb{R}$ we have

$$
H(x, y) \geq 0 \quad \text{where } H(x, y) = P(X > x, Y > y) - P(X > x)P(Y > y).
$$

It was shown in [27] (see also Lemma 2 of [30]), that if (X, Y) is PQD and if both X and Y have finite second moments then

$$
\mathrm{Cov}(X, Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} H(x, y)\, dx\, dy.
$$

In particular, if (X, Y) is PQD then $\mathrm{Cov}(X, Y) \geq 0$, and if $\mathrm{Cov}(X, Y) = 0$ then X and Y are independent.

The following two lemmas are invoked in the proofs of the main theorems. Lemma 3.1 is Lemma 3 of [33], and allows us to bound $\mathrm{Cov}(\phi(X), \psi(Y))$ by a constant times $\mathrm{Cov}(X, Y)$ for a PQD pair (X, Y) and ϕ and ψ sufficiently smooth.

Lemma 3.1 ([33]). *Let the random variables X, Y have finite second moments and be PQD. Then for any real valued functions ϕ and ψ that are absolutely continuous on all finite subintervals of $\mathbb{R}$,*

$$
\big|\mathrm{Cov}\big(\phi(X), \psi(Y)\big)\big| \leq \big|\phi'\big|_{\infty}\big|\psi'\big|_{\infty}\mathrm{Cov}(X, Y). \tag{71}
$$

The following result of [33], contained in the remark explaining (12) there, provides a version of (71) for vector valued functions.

Lemma 3.2 ([33]). *If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ are positively associated then for all real valued differentiable functions ϕ and ψ on $\mathbb{R}^n$,*

$$
\big|\mathrm{Cov}\big(\phi(\boldsymbol{\xi}), \psi(\boldsymbol{\xi})\big)\big| \leq 3\sqrt{2} \sum_{i,j=1}^{n} \left|\frac{\partial\phi}{\partial\xi_i}\right|_{\infty}\left|\frac{\partial\psi}{\partial\xi_j}\right|_{\infty}\mathrm{Cov}(\xi_i, \xi_j).
$$

As for any fixed u the indicator $\mathbf{1}_{X>u}$ is increasing in X, when X and Y are positively associated then they are PQD. As positive association is preserved by coordinate increasing functions, the following lemma is immediate.

Lemma 3.3. *The pair $(X, Y) = (\psi(\boldsymbol{\xi}), \phi(\boldsymbol{\xi}))$ is PQD whenever $\boldsymbol{\xi}$ is positively associated and $\psi(\boldsymbol{\xi})$ and $\phi(\boldsymbol{\xi})$ are coordinate wise increasing functions of $\boldsymbol{\xi}$.*

Though one form of the L^1, or Wasserstein distance $d_1(\mathcal{L}(X), \mathcal{L}(Y))$ is given in (1), in the proof that follows we apply the alternative characterization, see [35] for example:

$$d_1\big(\mathcal{L}(X), \mathcal{L}(Y)\big) = \sup_{h \in \mathbb{L}} \big|\mathbb{E}h(X) - \mathbb{E}h(Y)\big| \quad \text{where } \mathbb{L} = \big\{h : \big|h(y) - h(x)\big| \le |y - x|\big\}. \tag{72}$$

Proof of Theorem 1.3 For given $h \in \mathbb{L}$ let f be the unique bounded solution to the Stein equation

$$f'(w) - wf(w) = h(w) - Nh \quad \text{where } Nh = \mathbb{E}h(Z), \tag{73}$$

with $\mathcal{L}(Z)$ the standard normal distribution. Then, (see e.g. [11] Lemma 2.4),

$$\big|f'\big|_\infty \le \sqrt{\frac{2}{\pi}} \quad \text{and} \quad \big|f''\big|_\infty \le 2. \tag{74}$$

As $\text{Var}(W) = \sum_{i,j} \sigma_{ij} = 1$, we obtain

$$\begin{aligned}
\mathbb{E}\big[f'(W)\big] &= \mathbb{E}\left(\sum_{i=1}^{m} \sigma_i^2 f'(W) + \sum_{i \ne j} \sigma_{ij} f'(W)\right) \\
&= \mathbb{E}\left(\sum_{i=1}^{m} \xi_i^2 f'(W) + \sum_{i \ne j} \sigma_{ij} f'(W) + \sum_{i=1}^{m} (\sigma_i^2 - \xi_i^2) f'(W)\right).
\end{aligned}$$

Now letting $W^i = W - \xi_i$, write

$$\mathbb{E}\big[Wf(W)\big] = \mathbb{E}\sum_{i=1}^{m} \xi_i f(W) = \mathbb{E}\sum_{i=1}^{m} \xi_i f\big(W^i + \xi_i\big) = \mathbb{E}\sum_{i=1}^{m}\left[\xi_i f\big(W^i\big) + \xi_i^2 \int_0^1 f'\big(W^i + u\xi_i\big)\, du\right].$$

Recalling the Stein equation (73) and subtracting, we obtain

$$\begin{aligned}
\mathbb{E}\big[h(W) - Nh\big] &= \mathbb{E}\big[f'(W) - Wf(W)\big] \\
&= \mathbb{E}\left(\sum_{i=1}^{m} \xi_i^2\left(\int_0^1 \big(f'(W) - f'\big(W^i + u\xi_i\big)\big)\, du\right) + \sum_{i=1}^{m}(\sigma_i^2 - \xi_i^2) f'(W)\right. \\
&\quad \left. + \sum_{i \ne j} \sigma_{ij} f'(W) - \sum_{i=1}^{m} \xi_i f\big(W^i\big)\right).
\end{aligned} \tag{75}$$

Using the second inequality in (74), we bound the first term in (75) by

$$\begin{aligned}
&\left|\mathbb{E}\sum_{i=1}^{m} \xi_i^2 \int_0^1 \big(f'(W) - f'\big(W^i + u\xi_i\big)\big)\, du\right| \\
&= \left|\mathbb{E}\sum_{i=1}^{m} \xi_i^2 \int_0^1 \int_{u\xi_i}^{\xi_i} f''\big(W^i + t\big)\, dt\, du\right| \le 2\mathbb{E}\sum_{i=1}^{m} \xi_i^2\left(\int_0^1 \int_{u|\xi_i|}^{|\xi_i|} dt\, du\right) \\
&= \mathbb{E}\sum_{i=1}^{m} |\xi_i|^3 \le B\mathbb{E}\sum_{i=1}^{m} \xi_i^2 \le B,
\end{aligned} \tag{76}$$

using the almost sure bound on the variables ξ_i, and that their sum has mean zero and variance 1.

To handle the second term in (75), first note that W and ξ_i are coordinate wise increasing functions of $\boldsymbol{\xi}$, and hence PQD by Lemma 3.3. Now applying Lemma 3.1 with

$$
g(x) = \begin{cases} x^2, & |x| \le B, \\ B^2, & |x| > B, \end{cases}
$$

and again using the second inequality in (74), we have

$$
\left| \mathbb{E} \sum_{i=1}^{m} f'(W)\big(\sigma_i^2 - \xi_i^2\big) \right| = \left| \sum_{i=1}^{m} \mathrm{Cov}\big(f'(W), g(\xi_i)\big) \right|
$$
$$
\le 4B \sum_{i=1}^{m} \mathrm{Cov}(W, \xi_i) = 4B \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma_{ij} = 4B, \tag{77}
$$

using that $\mathrm{Var}(W) = \sum_{i,j} \sigma_{ij} = 1$.

For the third term in (75), using the nonnegativity of the covariances σ_{ij} and the first inequality in (74) we obtain

$$
\left| \mathbb{E} \sum_{i \ne j} \sigma_{ij} f'(W) \right| = \big| \mathbb{E} f'(W) \big| \sum_{i \ne j} \sigma_{ij} \le \sqrt{\frac{2}{\pi}} \sum_{i \ne j} \sigma_{ij}. \tag{78}
$$

For the final term in (75), we note that the variables W^i and ξ_i are coordinate wise increasing in $\boldsymbol{\xi}$, hence the pair (W^i, ξ_i) is PDQ by Lemma 3.3. Applying Lemma 3.1 and the first inequality in (74) now yields

$$
\left| \mathbb{E} \sum_{i=1}^{m} \xi_i f\big(W^i\big) \right| = \left| \sum_{i=1}^{m} \mathrm{Cov}\big(\xi_i, f(W^i)\big) \right| \le \sqrt{\frac{2}{\pi}} \sum_{i=1}^{m} \mathrm{Cov}\big(\xi_i, W^i\big) = \sqrt{\frac{2}{\pi}} \sum_{i \ne j} \sigma_{ij}. \tag{79}
$$

Summing the bounds (76)–(79) we find that $|\mathbb{E}h(W) - Nh|$ is bounded by the right-hand side of (20). Taking supremum over $h \in \mathbb{L}$ and using the characterization of the d_1 metric given in (72) completes the proof.

To prove Theorem 1.4 we apply the following result, which is a small variant of Lemma 2.6 of [11], due to [5]. Let $\mathbf{Z}$ be a standard normal random vector in $\mathbb{R}^p$. For $h : \mathbb{R}^p \to \mathbb{R}$ let $Nh = \mathbb{E}h(\mathbf{Z})$ and for $u \ge 0$ define

$$
(T_u h)(\mathbf{s}) = \mathbb{E}h\big(\mathbf{s}e^{-u} + \sqrt{1 - e^{-2u}}\,\mathbf{Z}\big).
$$

We write $D^2 h$ for the Hessian matrix of h when it exists.

Lemma 3.4. *For $m \ge 3$ and $h \in L_m^\infty(\mathbb{R}^p)$ the function*

$$
g(\mathbf{s}) = -\int_0^\infty \big[T_u h(\mathbf{s}) - Nh\big]\, du
$$

solves

$$
\mathrm{tr}\, D^2 g(\mathbf{s}) - \mathbf{s} \cdot \nabla g(\mathbf{s}) = h(\mathbf{s}) - Nh,
$$

and for any $0 \le |\mathbf{k}|_1 \le m$

$$
\big| g^{(\mathbf{k})} \big|_\infty \le \frac{1}{|\mathbf{k}|_1} \big| h^{(\mathbf{k})} \big|_\infty.
$$

Furthermore, for any $\boldsymbol{\lambda} \in \mathbb{R}^p$ and positive definite $p \times p$ matrix Σ, f defined by the change of variable

$$
f(\mathbf{s}) = g\big(\Sigma^{-1/2}(\mathbf{s} - \boldsymbol{\lambda})\big) \tag{80}
$$

solves

$$\operatorname{tr} \Sigma D^2 f(\mathbf{s}) - (\mathbf{s} - \boldsymbol{\lambda}) \cdot \nabla f(\mathbf{s}) = h\big(\Sigma^{-1/2}(\mathbf{s} - \boldsymbol{\lambda})\big) - Nh, \tag{81}$$

and satisfies

$$\big|f^{(\mathbf{k})}\big|_\infty \le \frac{p^{|\mathbf{k}|_1}}{|\mathbf{k}|_1} \big|\Sigma^{-1/2}\big|_\infty^{|\mathbf{k}|_1} \big|h^{(\mathbf{k})}\big|_\infty.$$

In particular, if $h \in \mathcal{H}_{m,\infty,p}$ then

$$\big|f^{(\mathbf{k})}\big|_\infty \le \frac{p^{|\mathbf{k}|_1}}{|\mathbf{k}|_1} \big|\Sigma^{-1/2}\big|_\infty^{|\mathbf{k}|_1} \quad \text{for all } 0 \le |\mathbf{k}|_1 \le m. \tag{82}$$

We apply the same technique as in the univariate case, along with Lemmas 3.2 and 3.4, to prove our main multivariate theorem.

Proof of Theorem 1.4. Given $h \in \mathcal{H}_{3,\infty,p}$, let f be the solution of (81) given by (80) with $\lambda = 0$. Writing out the expressions in (81) yields

$$\mathbb{E}\big[h\big(\Sigma^{-1/2}\mathbf{S}\big) - Nh\big] = \mathbb{E}\Bigg[\sum_{j=1}^{p} \sum_{l=1}^{p} \Sigma_{j,l} \frac{\partial^2}{\partial s_j\, \partial s_l} f(\mathbf{S}) - \sum_{j=1}^{p} S_j \frac{\partial}{\partial s_j} f(\mathbf{S}) \Bigg]$$

$$= \mathbb{E} \sum_{j=1}^{p} \Sigma_{j,j} \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) + \mathbb{E} \sum_{j,l \in [p],\, j \ne l} \Sigma_{j,l} \frac{\partial^2}{\partial s_j\, \partial s_l} f(\mathbf{S})$$

$$- \mathbb{E} \sum_{j=1}^{p} S_j \frac{\partial}{\partial s_j} f(\mathbf{S}). \tag{83}$$

We consider the first term of (83) and deal with each term under the sum separately for $j = 1, \ldots, p$. Letting $\sigma_{i,j}^2 = \operatorname{Var}(\xi_{i,j})$ and $\sigma_{i,j;k,l} = \operatorname{Cov}(\xi_{i,j}, \xi_{k,l})$, we have

$$\Sigma_{j,j} \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) = \sum_{i=1}^{m} \sigma_{i,j}^2 \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) + \sum_{i,k \in [m],\, i \ne k} \sigma_{i,j;k,j} \frac{\partial^2}{\partial s_j^2} f(\mathbf{S})$$

$$= \sum_{i=1}^{m} \xi_{i,j}^2 \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) + \sum_{i,k \in [m],\, i \ne k} \sigma_{i,j;k,j} \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) + \sum_{i=1}^{m} \big(\sigma_{i,j}^2 - \xi_{i,j}^2\big) \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}). \tag{84}$$

Now, with $S_{j*i} = S_j - \xi_{i,j}$ we may write the summands of the third term on the right-hand side of (83) as

$$S_j \frac{\partial}{\partial s_j} f(\mathbf{S}) = \sum_{i=1}^{m} \xi_{i,j} \frac{\partial}{\partial s_j} f(\mathbf{S})$$

$$= \sum_{i=1}^{m} \xi_{i,j} \frac{\partial}{\partial s_j} f(S_1, \ldots, S_{j*i}, \ldots, S_p)$$

$$+ \sum_{i=1}^{m} \xi_{i,j}^2 \int_0^1 \frac{\partial^2}{\partial s_j^2} f(S_1, \ldots, S_{j*i} + u\xi_{i,j}, \ldots, S_p)\, du. \tag{85}$$

Substituting (84) and (85) into (83), we obtain

$$
\begin{aligned}
\mathbb{E}\big[h\big(\Sigma^{-1/2}\mathbf{S}\big) - Nh\big] &= \mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}\xi_{i,j}^{2}\int_{0}^{1}\bigg(\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S}) - \frac{\partial^{2}}{\partial s_{j}^{2}}f(S_{1},\ldots,S_{j*i}+u\xi_{i,j},\ldots,S_{p})\bigg)\,du \\
&\quad + \mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}(\sigma_{i,j}^{2}-\xi_{i,j}^{2})\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S}) - \mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}\xi_{i,j}\frac{\partial}{\partial s_{j}}f(S_{1},\ldots,S_{j*i},\ldots,S_{p}) \\
&\quad + \mathbb{E}\sum_{j=1}^{p}\sum_{i,k\in[m],i\neq k}\sigma_{i,j;k,j}\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S}) + \mathbb{E}\sum_{j,l\in[p],j\neq l}\Sigma_{j,l}\frac{\partial^{2}}{\partial s_{j}\,\partial s_{l}}f(\mathbf{S}). \tag{86}
\end{aligned}
$$

We handle these five terms separately. For the first term in (86), using (82) we have

$$
\begin{aligned}
&\bigg|\mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}\xi_{i,j}^{2}\int_{0}^{1}\bigg(\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S}) - \frac{\partial^{2}}{\partial s_{j}^{2}}f(S_{1},\ldots,S_{j*i}+u\xi_{i,j},\ldots,S_{p})\bigg)\,du\bigg| \\
&= \bigg|\mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}\xi_{i,j}^{2}\int_{0}^{1}\int_{u\xi_{i,j}}^{\xi_{i,j}}\frac{\partial^{3}}{\partial s_{j}^{3}}f(S_{1},\ldots,S_{j*i}+t,\ldots,S_{p})\,dt\,du\bigg| \\
&\leq \frac{p^{3}}{3}\big|\Sigma^{-1/2}\big|_{\infty}^{3}\mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}\xi_{i,j}^{2}\int_{0}^{1}\int_{u|\xi_{i,j}|}^{|\xi_{i,j}|}\,dt\,du \\
&= \frac{p^{3}}{6}\big|\Sigma^{-1/2}\big|_{\infty}^{3}\sum_{j=1}^{p}\sum_{i=1}^{m}\mathbb{E}|\xi_{i,j}|^{3} \\
&\leq \frac{p^{3}}{6}\big|\Sigma^{-1/2}\big|_{\infty}^{3}B\sum_{j=1}^{p}\sum_{i=1}^{m}\mathbb{E}\xi_{i,j}^{2} \\
&\leq \frac{p^{3}}{6}\big|\Sigma^{-1/2}\big|_{\infty}^{3}B\sum_{j=1}^{p}\Sigma_{j,j}, \tag{87}
\end{aligned}
$$

using the almost sure bound on the variables $\xi_{i,j}$, and that their sum S_{j} over i from 1 to m has mean zero.

For the second term in (86), we have

$$
\begin{aligned}
\bigg|\mathbb{E}\sum_{j=1}^{p}\sum_{i=1}^{m}(\sigma_{i,j}^{2}-\xi_{i,j}^{2})\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S})\bigg| &= \bigg|\sum_{j=1}^{p}\sum_{i=1}^{m}\mathrm{Cov}\bigg(\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S}),\xi_{i,j}^{2}\bigg)\bigg| \\
&= \bigg|\sum_{j=1}^{p}\sum_{i=1}^{m}\mathrm{Cov}\bigg(\frac{\partial^{2}}{\partial s_{j}^{2}}f(\mathbf{S},\xi_{i,j}),g(\mathbf{S},\xi_{i,j})\bigg)\bigg|,
\end{aligned}
$$

where with some abuse of notation we let $f(\mathbf{s},x)=f(\mathbf{s})$, and define

$$
g(\mathbf{s},x) = \begin{cases} x^{2}, & |x|\leq B, \\ B^{2}, & |x| > B, \end{cases}
$$

for all $\mathbf{s} \in \mathbb{R}^p$ and $x \in \mathbb{R}$. Applying Lemma 3.2 and using the bound (82) and the fact that $(\mathbf{S}, \xi_{i,j})$ are positively associated for all i, j, we obtain

$$
\begin{aligned}
\left| \sum_{j=1}^p \sum_{i=1}^m \mathrm{Cov}\left(\frac{\partial^2}{\partial s_j^2} f(\mathbf{S}), \xi_{i,j}^2 \right) \right| &\le 3\sqrt{2} \left| \sum_{j=1}^p \sum_{i=1}^m \sum_{l=1}^p \left| \frac{\partial^3}{\partial s_l \, \partial s_j^2} f \right|_\infty \left| \frac{\partial g}{\partial x} \right|_\infty \mathrm{Cov}(S_l, \xi_{i,j}) \right| \\
&\le 2\sqrt{2} p^3 |\Sigma^{-1/2}|_\infty^3 B \sum_{j=1}^p \sum_{i=1}^m \sum_{l=1}^p \mathrm{Cov}(S_l, \xi_{i,j}) \\
&= 2\sqrt{2} p^3 |\Sigma^{-1/2}|_\infty^3 B \sum_{j,l=1}^p \sum_{i,k=1}^m \sigma_{i,j;k,l} \\
&= 2\sqrt{2} p^3 |\Sigma^{-1/2}|_\infty^3 B \sum_{j,l=1}^p \Sigma_{j,l}.
\end{aligned}
\tag{88}
$$

For the third term in (86), again applying Lemma 3.2 and arguing as for the second term, we have

$$
\begin{aligned}
\left| \mathbb{E} \sum_{j=1}^p \sum_{i=1}^m \xi_{i,j} \frac{\partial}{\partial s_j} f(S_1, \dots, S_{j*i}, \dots, S_p) \right| &\\
= \left| \sum_{j=1}^p \sum_{i=1}^m \mathrm{Cov}\left(\xi_{i,j}, \frac{\partial}{\partial s_j} f(S_1, \dots, S_{j*i}, \dots, S_p) \right) \right| &\\
\le 3\sqrt{2} \left| \sum_{j=1}^p \sum_{i=1}^m \left(\sum_{l \in [p]\setminus\{j\}} \left| \frac{\partial^2}{\partial s_l \, \partial s_j} f \right|_\infty \mathrm{Cov}(\xi_{i,j}, S_l) + \left| \frac{\partial^2}{\partial s_j^2} f \right|_\infty \mathrm{Cov}(\xi_{i,j}, S_{j*i}) \right) \right| &\\
\le \frac{3}{\sqrt{2}} p^2 |\Sigma^{-1/2}|_\infty^2 \left(\sum_{j,l \in [p], j \ne l} \sum_{i=1}^m \mathrm{Cov}(\xi_{i,j}, S_l) + \sum_{j=1}^p \sum_{i=1}^m \mathrm{Cov}(\xi_{i,j}, S_{j*i}) \right) &\\
= \frac{3}{\sqrt{2}} p^2 |\Sigma^{-1/2}|_\infty^2 \left(\sum_{j,l \in [p], j \ne l} \sum_{i,k=1}^m \sigma_{i,j;k,l} + \sum_{j=1}^p \sum_{i,k \in [m], i \ne k} \sigma_{i,j;k,j} \right) &\\
= \frac{3}{\sqrt{2}} p^2 |\Sigma^{-1/2}|_\infty^2 \left(\sum_{j,l \in [p], j \ne l} \Sigma_{j,l} + \sum_{j=1}^p \sum_{i,k \in [m], i \ne k} \sigma_{i,j;k,j} \right). &
\end{aligned}
\tag{89}
$$

For the fourth and the fifth terms in (86), again using (82) we have

$$
\left| \sum_{j=1}^p \sum_{i,k \in [m], i \ne k} \sigma_{i,j;k,j} \frac{\partial^2}{\partial s_j^2} f(\mathbf{S}) \right| \le \frac{p^2}{2} |\Sigma^{-1/2}|_\infty^2 \sum_{j=1}^p \sum_{i,k \in [m], i \ne k} \sigma_{i,j;k,j}
\tag{90}
$$

and

$$
\left| \mathbb{E} \sum_{j,l \in [p], j \ne l} \Sigma_{j,l} \frac{\partial^2}{\partial s_j \, \partial s_l} f(\mathbf{S}) \right| \le \frac{p^2}{2} |\Sigma^{-1/2}|_\infty^2 \sum_{j,l \in [p], j \ne l} \Sigma_{j,l}.
\tag{91}
$$

Summing the bounds (87)–(91) we find that $|\mathbb{E}[h(\Sigma^{-1/2}\mathbf{S}) - Nh]|$ is bounded by the right-hand side of (22). Taking supremum over $h \in \mathcal{H}_{3,\infty,p}$ and using the definition (21) of $d_{\mathcal{H}_{m,\infty,p}}$ completes the proof. $\square$

References

[1] J. H. Ahlberg and E. N. Nilson. Convergence properties of the spline fit. *J. SIAM.* **11** (1963) 95–104. MR0152788

[2] M. Aizenman, D. J. Barsky and R. Fernandez. The phase transition in a general class of Ising-type models is sharp. *J. Stat. Phys.* **47** (1987) 343–374. MR0894398

[3] M. Aizenman and H. Duminil-Copin. The truncated correlations of the Ising model in any dimension decay exponentially fast at all but the critical temperature, 2015. Available at arXiv:1506:00625.

[4] S. Banerjee and A. Roy. *Linear Algebra and Matrix Analysis for Statistics*. Chapman and Hall, New York, 2014. MR3222172

[5] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** (1990) 297–332. MR1035659

[6] T. Birkel. On the convergence rate in the central limit theorem for associated processes. *Ann. Probab.* **16** (1988) 1685–1698. MR0958210

[7] F. Brenti. Unimodal, log-concave, and Pólya frequency sequences in combinatorics. In *Memoirs Amer. Math. Soc.*, **413**. Rhode Island, 1989. MR0963833

[8] A. Bulinski. Rate of convergence in the central limit theorem for fields of associated random variables. *Theory Probab. Appl.* **40** (1995) 136–144. MR1346738

[9] S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** (2011) 464–483. MR2807964

[10] J. T. Chayes, L. Chayes, G. Grimmett, H. Kesten and R. H. Schonmann. The correlation length for the high-density phase of Bernoulli percolation. *Ann. Probab.* **17** (1989) 1277–1302. MR1048927

[11] L. H. Y. Chen, L. Goldstein and Q.-M. Shao. *Normal Approximation by Stein's Method*. Springer, New York, 2011. MR2732624

[12] T. Cox and D. Griffeath. Occupation time limit theorems for the voter model. *Ann. Probab.* **11** (1983) 876–893. MR0714952

[13] T. Cox and G. Grimmett. Central limit theorems for associated random variables and the percolation model. *Ann. Probab.* **12** (1984) 514–528. MR0735851

[14] P. Crifford and A. Subbery. A model for spatial conflict. *Biometrika* **60** (1973) 581–588. MR0343950

[15] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model, 2015. Available at arXiv:1502.03050v2.

[16] R. Durrett and D. Griffeath. Supercritical contact processes on $\mathbb{Z}$. *Ann. Probab.* **11** (1983) 1–15. MR0682796

[17] P. Eichelsbacher and G. Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **30** (2010) 962–988. MR2659754

[18] P. Eichelsbacher and B. Martschink. On rates of convergence in the Curie—Weiss—Potts model with an external field. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2015) 252–282. MR3300970

[19] P. Eichelsbacher and M. Reinert. Stein's method for discrete Gibbs measures. *Ann. Probab.* **18** (2008) 1588–1618. MR2434182

[20] R. S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics*. Springer, New York, 2006. MR2189669

[21] J. D. Esary, F. Proschan and D. W. Walkup. Association of random variables, with applications. *Ann. Math. Stat.* **38** (1967) 1466–1474. MR0217826

[22] M. Fiocco and W. R. Van Zwet. Decaying correlations for the supercritical contact process conditioned on survival. *Bernoulli* **9** (2003) 763–781. MR2047685

[23] L. Goldstein and N. Wiroonsri. Stein's method for positively associated random variables with applications to Ising, percolation and voter models, 2016. Available at arXiv:1603.05322v1.

[24] G. Grimmett. *Percolation*, 2nd edition. Springer, New York, 1999. MR1707339

[25] G. Grimmett. *The Random-Cluster Model*. Springer, New York, 2006. MR2243761

[26] T. E. Harris. Contact interactions on a lattice. *Ann. Probab.* **2** (1974) 969–988. MR0356292

[27] W. Hoeffding. Masstabinvariante Korrelations-theorie Schriften. *Math. Inst. Univ. Berlin* **5** (1940) 181–233. MR0004426

[28] I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen, 1971. MR0322926

[29] J. L. Lebowitz and O. Penrose. Analytic and clustering properties of thermodynamic functions and distribution functions for classical lattice and continuum systems. *Comm. Math. Phys.* **11** (1968) 99–124. MR0242450

[30] E. L. Lehmann. Some concepts of dependence. *Ann. Math. Stat.* **37** (1966) 1137–1153. MR0202228

[31] T. M. Liggett. *Interacting Particle Systems*. Springer, New York, 1985. MR0776231

[32] T. M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer, New York, 1999. MR1717346

[33] C. Newman. Normal fluctuations and the FKG inequality. *Comm. Math. Phys.* **74** (1980) 119–128. MR0576267

[34] C. Newman. A general central limit theorem for FKG systems. *Comm. Math. Phys.* **91** (1983) 75–80. MR0719811

[35] S. T. Rachev. The Monge-Kantorovich transference problem and its stochastic applications. *Theory Probab. Appl.* **29** (1984) 647–676.

[36] Y. Rinott and V. Rotar. On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** (4) (1997) 1080–1105. MR1484798

[37] N. Ross. Fundamentals of Stein's method. *Probab. Surv.* **8** (2011) 210–293. MR2861132

[38] R. H. Schonmann. Central limit theorem for the contact process. *Ann. Probab.* **14** (1986) 1291–1295. MR0866350

[39] R. P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. *Graph Theory and Its Applications: East and West* **576** (1989) 500–535. MR1110850

[40] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob. Univ. of California Press* **2** (1972) 210–293. MR0402873

[41] A. N. Tikhomirov. On the convergence rate in the central limit theorem for weakly dependent random variables. *Theory Probab. Appl.* **25** (1980) 790–809.