cemmap
centre for microdata methods and practice

Spatial circular matrices, with applications

Grant Hillier
Federico Martellosio

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP06/10

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

Spatial Circular Matrices, with Applications

Grant Hillier[a] and Federico Martellosio[b,*]

[a]CeMMAP and Department of Economics, University of Southampton,
Highfield, Southampton SO17 1BJ, UK
[b]School of Economics, University of Reading,
Whiteknights, Reading RG6 6UD, UK

March 5, 2010

Abstract

The cumulants of the quadratic forms associated to the so-called spatial design matrices are often needed for inference in the context of isotropic processes on uniform grids. Unfortunately, because the eigenvalues of the matrices involved are generally unknown, the computation of the cumulants may be very demanding if the grids are large. This paper constructs circular counterparts, with known eigenvalues, to the spatial design matrices. It then studies some of their properties, and analyzes their performance in a number of applications.

Keywords: circulant matrices; isotropic spatial processes; Moran correlogram; quadratic forms; spatial design matrices; variogram GLS fitting.
JEL Classification: C12, C21.

---

1 Introduction

When analyzing data observed on a uniform grid, it is often reasonable to assume that the underlying spatial stochastic process is isotropic. Isotropy requires that the variogram—or the covariance function, if it exists—does not depend on direction but only on distance; see, e.g., Cressie (1993), Ch. 2.3. The so-called spatial design matrices arise naturally in many inferential problems in the context of isotropic processes; see Genton (1998), Gorsich et al. (2002), Hillier and Martellosio (2006), Arellano-Valle and Genton (2010), and below. In particular, for testing or estimation purposes, one often needs the first few cumulants of quadratic forms, or ratios of quadratic forms, associated to the spatial design matrices. Unfortunately, the derivation of such cumulants may be computationally very demanding. This paper is concerned with an approximation aimed at alleviating the computational effort.

By d-dimensional uniform grid with n_i sites on the i-th axis we mean the set $\Gamma = \Gamma(n_1, ..., n_d)$ of the $N = \prod_{i=1}^{d} n_i$ sequences $\boldsymbol{\alpha} = (\alpha(1), ..., \alpha(d))$ of integers

*Corresponding author. Tel: +44 (0) 118 378 6033; fax: +44 (0) 118 378 4029
E-mail addresses: ghh@soton.ac.uk (G. Hillier), f.martellosio@reading.ac.uk (F. Martellosio).

$\alpha(i) = 0, ..., n_i - 1$. We call two elements $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \Gamma$ *h-neighbors* if the squared Euclidean distance $\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2$ is equal to h.[1] For convenience, we order the sequences in Γ lexicographically. An *off-diagonal spatial design matrix* is a matrix $\boldsymbol{A}_h$ indexed by $\Gamma \times \Gamma$ with entries

$$(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{cases} 1 & \text{if } \boldsymbol{\alpha} \text{ and } \boldsymbol{\beta} \text{ are } h\text{-neighbors} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

for $h = 1, 2, ..., \sum_{i=1}^{d}(n_i - 1)^2$. Given $\boldsymbol{A}_h$, we can define a *full spatial design matrix* as

$$\boldsymbol{L}_h = \boldsymbol{D}_h - \boldsymbol{A}_h, \tag{2}$$

where $\boldsymbol{D}_h$ denotes the diagonal matrix containing the row-sums of $\boldsymbol{A}_h$. Note that $(\boldsymbol{D}_h)_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$ is the number of h-neighbors of α. For $h = 0$, it proves convenient to define $\boldsymbol{A}_0 = \boldsymbol{D}_0 = 2\boldsymbol{I}_N$, where $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix.

Spatial design matrices appear in several important statistics related to stochastic processes on uniform grids. Let $\boldsymbol{z}$ denote the $N \times 1$ vector $(z(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in \Gamma)'$, where $z(\boldsymbol{\alpha})$ is the random variable observed at $\boldsymbol{\alpha}$. Let $N(h)$ denote the set of unordered pairs $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ that are h-neighbors, and let $N_h = |N(h)|$ be the number of h-neighbors on Γ. For any h such that $N_h \neq 0$, two fundamental quadratic forms associated to the spatial design matrices are

$$\widehat{\gamma}_h = \frac{1}{2N_h}\boldsymbol{z}'\boldsymbol{L}_h\boldsymbol{z} = \frac{1}{2N_h}\sum_{\boldsymbol{\alpha},\boldsymbol{\beta} \in N(h)} (z(\boldsymbol{\alpha}) - z(\boldsymbol{\beta}))^2, \tag{3}$$

and

$$\widehat{b}_h = \frac{1}{2N_h}\boldsymbol{z}'\boldsymbol{A}_h\boldsymbol{z} = \frac{1}{N_h}\sum_{\boldsymbol{\alpha},\boldsymbol{\beta} \in N(h)} z(\boldsymbol{\alpha})z(\boldsymbol{\beta}). \tag{4}$$

The quadratic form $\widehat{\gamma}_h$ is a sample equivalent of the semivariogram $\gamma(h)$ at distance $\sqrt{h}$ of an *intrinsically stationary and isotropic* process (e.g., Cressie, 1993). On the other hand, $\widehat{b}_h$ is a sample equivalent of the autocovariance $c(h)$ at distance $\sqrt{h}$ of a zero-mean *second-order stationary and isotropic* process. When $\mathrm{E}(z(\boldsymbol{\alpha}))$ is assumed to be constant over Γ, and unknown, $\widehat{b}_h$ can be generalized to

$$\widehat{c}_h = \frac{1}{2N_h}\boldsymbol{z}'\boldsymbol{M}\boldsymbol{A}_h\boldsymbol{M}\boldsymbol{z} = \frac{1}{N_h}\sum_{\boldsymbol{\alpha},\boldsymbol{\beta} \in N(h)} (z(\boldsymbol{\alpha}) - \bar{z})(z(\boldsymbol{\beta}) - \bar{z}), \tag{5}$$

where $\bar{z} = N^{-1}\sum_{\boldsymbol{\alpha} \in \Gamma} z(\boldsymbol{\alpha})$ and $\boldsymbol{M} = \boldsymbol{I}_N - N^{-1}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'$, with $\boldsymbol{\iota}_N$ denoting the $N \times 1$ vector of all ones. Generalizations to the linear regression case are also straightforward, but will not be considered in this paper.

Another popular statistic associated to the spatial design matrices is the sample correlation coefficient

$$r_h = \frac{\widehat{c}_h}{\widehat{c}_0}, \tag{6}$$

[1] We define h-neighborhood with respect to the *squared* Euclidean distance, rather than the Euclidean distance, only for notational convenience. Note that squared Euclidean distances between the elements of a grid Γ are integer.

where, from (5), $\widehat{c}_0$ is the sample variance $N^{-1}\boldsymbol{z}'\boldsymbol{M}\boldsymbol{z}$.

In the context of testing for spatial autocorrelation, r_h is referred to as the *Moran statistic* (Moran, 1950). Similarly, the ratio

$$g_h = \frac{\widehat{\gamma}_h}{\widehat{c}_0} \tag{7}$$

can be used as a test statistic for spatial autocorrelation (Geary, 1954), or as a normalized semivariogram estimator (e.g., Gorsich *et al.* 2002, p. 161).[2] It is worth noting that, when $d = 1$, r_h reduces to the standard serial correlation coefficient at time lag $\sqrt{h}$ (e.g., Anderson, 1971, p. 299), and g_h reduces to (half) the von Neumann ratio at time lag $\sqrt{h}$ (von Neumann, 1941).

Spatial design matrices are also relevant for modeling purposes. For example, the $\boldsymbol{A}_h$'s can be used as weights matrices in *conditional* or *simultaneous autoregressions* on Γ (e.g., Cressie 1993, Ch. 6.3), and the $\boldsymbol{L}_h$'s can be used as precision matrices of *intrinsic autoregressions* on Γ (e.g., Rue and Held, 2005, Ch. 3). In this context, the results in the present paper are useful to construct an approximate likelihood in a way similar to what is done Kent and Mardia (1996), but this will not be our emphasis here.

Throughout the paper, we assume that the random vector $\boldsymbol{z}$ is Gaussian, although, as we mention in Section 5, various generalizations are possible. Several authors, especially in geostatistics, have been concerned with the cumulants of the quadratic forms associated to spatial design matrices under Gaussianity. Such cumulants are of direct interest in the context of estimation (for instance, as we shall see below, variances and covariances of $\widehat{\gamma}_h$ for different values of h are required for generalized least squares fitting of the variogram) and are useful to derive approximations to the densities of the various statistics defined above (for instance, by matching the cumulants to those of some family of distribution, or by saddlepoint approximation). Cressie (1985) considers the case when $\boldsymbol{z}$ is a Gaussian intrinsically stationary process, and studies the covariance structure of the non-isotropic versions of the $\widehat{\gamma}_h$'s. Genton (1998) deals with the isotropic case for general d, when observations are independent and when only "non-diagonal directions" are considered.[3] Gorsich et al. (2002) provides generalizations, and study $\mathrm{var}(\widehat{\gamma}_h)$ by simulation, under second-order stationarity and isotropy. Hillier and Martellosio (2006), henceforth HM, derives a complete structural representation of the matrices $\boldsymbol{A}_h$ and $\boldsymbol{L}_h$, and studies generating functions for the computation of the cumulants of the associated quadratic forms.

The purpose of this paper is to approximate the matrices $\boldsymbol{A}_h$ and $\boldsymbol{L}_h$ with matrices having a "more convenient" structure. Part of the difficulty in working with the matrices $\boldsymbol{A}_h$ and $\boldsymbol{L}_h$ is that, in dimension d higher than 1, their eigenvalues are generally not known in closed form. This is problematic, because in many important cases the cumulants of the statistics mentioned above are functions of the eigenvalues. Also,

[2] Geary (1954) used the unbiased sample variance $[N/(N-1)]\widehat{c}_0$, rather than $\widehat{c}_0$, as a normalization factor in (7). Under the assumption $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \boldsymbol{I}_N)$, such a normalization has the advantage of making $\mathrm{E}(g_h)$ independent of N.

[3] Restricting attention to non-diagonal directions amounts to setting $(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 1$ if $\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2 = h$ and $\boldsymbol{\alpha} - \boldsymbol{\beta}$ contains $d-1$ zeros, $(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0$ otherwise.

since all those statistics are based on quadratic forms, it follows that under the specific assumption of a spherically symmetric distribution, *all* their properties depend only on the eigenvalues of the spatial design matrices. It should be noted that, in principle, the computation of the cumulants is possible even without knowing the eigenvalues, using the generating functions given in HM. However, the required computation becomes prohibitive when N is large. This is true not only under general assumptions on the underlying spatial process (such as second-order stationarity and isotropy), but, in the case of cumulants of order higher than two, even for i.i.d. data; see Section 3 of HM. In the present contribution, we overcome these problems by approximating $\boldsymbol{A}_h$ and $\boldsymbol{L}_h$ with matrices $\widetilde{\boldsymbol{A}}_h$ and $\widetilde{\boldsymbol{L}}_h$ whose eigenvalues are available in closed form. Because they are constructed on the basis of circulant matrices, the matrices $\widetilde{\boldsymbol{A}}_h$ and $\widetilde{\boldsymbol{L}}_h$ will be named *circular* spatial design matrices.

The rest of the paper is organized as follows. In Section 2 we briefly review the basic structure of spatial design matrices, and we introduce the matrices that will be used as building blocks for our approximation. In Section 3 we construct the approximation, and we study some of its properties. In Section 4 we investigate the use of the circular spatial design matrices in the context of three applications: the study of the properties of the sample autocovariance, a test of significance of the Moran correlogram, and generalized least squares (GLS) estimation of the variogram. Some illustrative numerical results are included. Section 5 concludes. The appendices contain some additional technical material and all the proofs.

2 Preliminaries

Let $\boldsymbol{F}_r^{(n)}$, for $r = 0, 1, ..., n-1$, denote the $n \times n$ matrices with (i, j) entry equal to 1 if $|i - j| = r$, 0 otherwise. For example,

$$\boldsymbol{F}_1^{(4)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \ \boldsymbol{F}_2^{(4)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

The $\boldsymbol{F}_r^{(n)}$'s are the off-diagonal spatial design matrices $\boldsymbol{A}_h$ in dimension $d = 1$ (with $r = \sqrt{h}$). When $d > 1$, our circular approximation rests on a representation of the spatial design matrices in terms of sums of Kronecker products ($\otimes$) of the matrices $\boldsymbol{F}_r^{(n)}$. Namely, the extension of Proposition 1 in HM to the case in which $n_1, n_2, ..., n_d$ are not necessarily the same yield

$$\boldsymbol{A}_h = \sum_{\boldsymbol{\alpha} \in \Gamma_h} \boldsymbol{F}_{\boldsymbol{\alpha}}^{\otimes}, \tag{8}$$

where

$$\Gamma_h = \{\boldsymbol{\alpha} \in \Gamma : \|\boldsymbol{\alpha}\|^2 = h\},$$

and

$$\boldsymbol{F}_{\boldsymbol{\alpha}}^{\otimes} = \boldsymbol{F}_{\alpha(1)}^{(n_1)} \otimes \boldsymbol{F}_{\alpha(2)}^{(n_2)} \otimes \ldots \otimes \boldsymbol{F}_{\alpha(d)}^{(n_d)} = \bigotimes_{i=1}^{d} \boldsymbol{F}_{\alpha(i)}^{(n_i)}. \tag{9}$$

The eigenvalues of the matrices $\boldsymbol{F}_r^{(n)}$, $r = 0, \ldots, n-1$, are easily derived in closed form (e.g., Biggs, 1993). Hence, by (9), the eigenvalues of any $\boldsymbol{F}_{\boldsymbol{\alpha}}^{\otimes}$ are also known. However, this is generally of no help in obtaining the eigenvalues of the $\boldsymbol{A}_h$'s when $d > 1$, because the summands in (8) are typically not pairwise commutative; see Section 2.3 of HM for details. For many purposes, it would be useful to approximate $\boldsymbol{A}_h$ with a matrix having a simple eigenstructure. We shall show that this can be achieved by replacing the matrices $\boldsymbol{F}_r^{(n)}$ with their circular counterparts $\widetilde{\boldsymbol{F}}_r^{(n)}$. These are circulant matrices with (i,j)-th entry

$$(\widetilde{\boldsymbol{F}}_r^{(n)})_{i,j} = \begin{cases} 1 & \text{if } |i-j| = r \text{ or } \ |i-j| = n-r \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

For example,

$$\widetilde{\boldsymbol{F}}_1^{(4)} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}; \ \widetilde{\boldsymbol{F}}_2^{(4)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Observe that $\widetilde{\boldsymbol{F}}_r^{(n)} = \boldsymbol{F}_r^{(n)}$ if and only if $r = 0$ or $n/2$. Also, note that $\widetilde{\boldsymbol{F}}_r^{(n)} = \widetilde{\boldsymbol{F}}_{n-r}^{(n)}$, and hence only $\lfloor n/2 \rfloor + 1$ of the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$, $r = 0, 1, \ldots, n-1$, are distinct ($\lfloor \cdot \rfloor$ denotes the floor function).

The matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$, $r = 0, \ldots, \lfloor n/2 \rfloor$, are used to define the so-called *circular serial correlation coefficients* (e.g., Anderson, 1971, Section 6.5), which are well-known to have simpler statistical properties than the corresponding statistics based on the matrices $\boldsymbol{F}_r^{(n)}$. Most of the analytical advantages of replacing the $\boldsymbol{F}_r^{(n)}$'s with the $\widetilde{\boldsymbol{F}}_r^{(n)}$'s are a consequence of the fact that the latter matrices span an algebra—known as *Bose-Mesner* algebra—that admits a basis of symmetric and pairwise orthogonal idempotents, and hence is commutative and closed under multiplication and generalized inversion (see, e.g., Bannai and Ito, 1984).[4]

[4] A simple way to verify that the span of the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$ is a Bose-Mesner algebra is to exploit the well-known result that the *distance matrices* of a *distance-regular* graph span a Bose-Mesner algebra (e.g., Biggs, 1993, p. 159-161). Indeed, the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$, $r = 0, \ldots, \lfloor n/2 \rfloor$, can be interpreted as the distance matrices of the graph with vertex set $\{0, 1, \ldots, n-1\}$ and edges the pairs $(i, i+1)$, $i = 0, 1, \ldots, n-2$, and the pair $(0, n-1)$. Such a graph is easily seen to be distance-regular.

3 The Circular Spatial Design Matrices

3.1 Definition and Properties

For each $h = 1, 2, ..., \sum_{i=1}^{d}(n_i - 1)^2$, we define the circular off-diagonal spatial design matrices by analogy with expression (8):

$$\widetilde{\boldsymbol{A}}_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h} \widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes}, \tag{11}$$

where

$$\widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes} = \bigotimes_{i=1}^{d} \widetilde{\boldsymbol{F}}_{\alpha(i)}^{(n_i)}.$$

We can regard $\widetilde{\boldsymbol{A}}_h$ either as an approximation to the "true" matrix $\boldsymbol{A}_h$, or as an alternative spatial design matrix in its own right. Letting $\widetilde{\boldsymbol{D}}_h$ be the diagonal matrix with $(\widetilde{\boldsymbol{D}}_h)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}\in\Gamma}(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, for each $\boldsymbol{\alpha} \in \Gamma$, it is natural to consider also the circular full spatial design matrix

$$\widetilde{\boldsymbol{L}}_h = \widetilde{\boldsymbol{D}}_h - \widetilde{\boldsymbol{A}}_h.$$

In order to describe the structure of $\widetilde{\boldsymbol{A}}_h$, we need some new notation. Let Δ be the collection of all (proper and improper) subsets of $\{1, ..., d\}$. For each $D \in \Delta$, we define the sequence $\boldsymbol{\pi}_D = (\pi_D(1), ..., \pi_D(d))$ with components

$$\pi_D(i) = \begin{cases} 0 & \text{for } i \in D \\ n_i & \text{otherwise.} \end{cases}$$

For each $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \Gamma \times \Gamma$, we also define the sequence $\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \Gamma$ with components $\epsilon_{\boldsymbol{\alpha},\boldsymbol{\beta}}(i) = |\alpha(i) - \beta(i)|$, $i = 1, ..., d$.

Theorem 1 *For each $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \Gamma \times \Gamma$,*

$$(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{cases} 1 & \text{if } \exists\ D \in \Delta \text{ such that } \left\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_D\right\|^2 = h \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Geometrically, the 2^d sequences $\boldsymbol{\pi}_D$ are the corners of the grid $\Gamma(n_1 + 1, ..., n_d + 1)$. Thus, Theorem 1 asserts that $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 1$ if and only if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are at squared distance h on the toroidal grid $\widetilde{\Gamma}$ obtained by joining the "opposite sides" of the grid Γ.[5] Replacing a lattice by its toroidal counterpart has often proved useful in the statistical literature to approximate properties of spatial processes; see, for instance, Besag and Moran (1975), Martin (1986), and Kent and Mardia (1996).

Some immediate consequences of Theorem 1 are given in the following corollary.

[5] In graph theoretic terminology, $\widetilde{\boldsymbol{A}}_h$ and $\widetilde{\boldsymbol{L}}_h$ are, respectively, the adjacency matrix and the Laplacian matrix of the graph having $\widetilde{\Gamma}$ as vertex set, and edges the pairs $(\boldsymbol{\alpha},\boldsymbol{\beta})$ such that $\left\|\boldsymbol{\alpha} - \boldsymbol{\beta}\right\|^2 = h$ (see, e.g., Biggs 1993).

Corollary 1 *$\widetilde{\boldsymbol{A}}_h = \boldsymbol{A}_h$ if and only if $\alpha(i)$ equals 0 or $n_i/2$, for each $\boldsymbol{\alpha} \in \Gamma_h$ and each $i = 1, ..., d$. For each $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \Gamma \times \Gamma$, $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ if $\epsilon_{\boldsymbol{\alpha},\boldsymbol{\beta}}(i) < n_i - \sqrt{h}$ for all $i = 1, ..., d$.*

The first part of Corollary 1 asserts that, except for very restrictive cases, $\widetilde{\boldsymbol{A}}_h \neq \boldsymbol{A}_h$. However, when h is small relative to each n_i, many of the entries of $\boldsymbol{A}_h$ and $\widetilde{\boldsymbol{A}}_h$ agree. Indeed, according to the second part of the Corollary 1, the proportion of entries $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ that are different from $(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is non-increasing in each n_i, and, for a fixed N, non-decreasing in d.[6] This suggests that the approximation should work particularly well when all n_i's are large and d is small, which is precisely the case when an approximation is most needed.

Note that the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$ have constant row sum, equal to 1 if $r = 0$ or $r = n/2$, to 2 otherwise. An important consequence of this is that $\widetilde{\boldsymbol{A}}_h$, in contrast to $\boldsymbol{A}_h$, also has constant row-sum, to be denote by $\widetilde{m}_h$. To see this, let $\delta_{r,s}$ be the Kronecker delta ($\delta_{r,s} = 1$ if $r = s$, $\delta_{r,s} = 0$ otherwise), and write

$$\begin{aligned}
\widetilde{\boldsymbol{A}}_h \boldsymbol{\iota}_N &= \sum_{\boldsymbol{\alpha} \in \Gamma_h} \widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes} \bigotimes_{i=1}^{d} \boldsymbol{\iota}_{n_i} = \sum_{\boldsymbol{\alpha} \in \Gamma_h} \bigotimes_{i=1}^{d} \widetilde{\boldsymbol{F}}_{\alpha(i)}^{(n_i)} \boldsymbol{\iota}_{n_i} \\
&= \sum_{\boldsymbol{\alpha} \in \Gamma_h} \prod_{i=1}^{d} \left(2 - \delta_{\alpha(i),0} - \delta_{\alpha(i),\frac{n_i}{2}} \right) \bigotimes_{i=1}^{d} \boldsymbol{\iota}_{n_i} = \widetilde{m}_h \boldsymbol{\iota}_N. \qquad (13)
\end{aligned}$$

Observe that $\widetilde{m}_h$ is an eigenvalue of $\widetilde{\boldsymbol{A}}_h$, associated to the eigenvector $\boldsymbol{\iota}_N$. Also, note that when $n_1 = ... = n_d = n$, expression (13) for $\widetilde{m}_h$ yields

$$\widetilde{m}_h = \sum_{i=0}^{d} 2^{d-i} f_i, \qquad (14)$$

where f_i is the number of sequences in Γ_h with exactly i elements equal to 0 or $n/2$. In particular, since $\sum_{i=0}^{d} f_i = |\Gamma_h|$, it follows that if no $\boldsymbol{\alpha} \in \Gamma_h$ contains either 0 or $n/2$, then (14) simplifies to $\widetilde{m}_h = 2^d |\Gamma_h|$. We denote the sum $N\widetilde{m}_h$ of all elements in $\widetilde{\boldsymbol{A}}_h$ by $2\widetilde{N}_h$, so that $\widetilde{N}_h$ is the number of (unordered) pairs of h-neighbors on $\widetilde{\Gamma}$. The corresponding quantity on Γ is N_h, defined in the Introduction.

Remark 1 The computation of $\widetilde{m}_h$ from equations (13) or (14) requires previous computation of the elements of Γ_h (see Section 3.2). Alternatively, following the approach described in HM, it is easily checked that $\widetilde{m}_h$ can be efficiently computed as the coefficient of t^h in the formal expansion of the generating function

$$\prod_{i=1}^{d} \left\{ 1 + \sum_{r=1}^{n_i - 1} \left(2 - \delta_{r,n_i/2} \right) t^{r^2} \right\}. \qquad (15)$$

[6] A more formal justification of the approximation is provided by the fact that, as $n_1, n_2, .., , n_d \to \infty$, $\widetilde{\boldsymbol{A}}_h$ and $\boldsymbol{A}_h$ are asymptotically equivalent sequences of matrices, in the sense of, e.g., Gray (2006), Section 2.3.

The matrices $\widetilde{\boldsymbol{A}}_h$ do not retain all properties of the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$. In particular, the $\widetilde{\boldsymbol{A}}_h$'s are not necessarily circulant matrices, and do not necessarily span an algebra that is closed under multiplication and generalized inversion. One crucial property that the $\widetilde{\boldsymbol{A}}_h$'s do inherit from the $\widetilde{\boldsymbol{F}}_r^{(n)}$'s is that of pairwise commutativity.

Theorem 2 *For all h and k, the matrices $\widetilde{\boldsymbol{A}}_h$ and $\widetilde{\boldsymbol{A}}_k$ commute.*

Recall that if two symmetric matrices commute, they are simultaneously diagonalizable (e.g., Horn and Johnson, 1985). Because of this, Theorem 2 has two important implications: (i) the eigenvalues of any linear combination of matrices $\widetilde{\boldsymbol{A}}_h$ are the same linear combination of the eigenvalues of the $\widetilde{\boldsymbol{A}}_h$'s; (ii) all positive or negative powers of all linear combinations of matrices $\widetilde{\boldsymbol{A}}_h$ share the same eigenvectors. As an illustration of (ii), consider a spatial process $\boldsymbol{z}$ with $\mathrm{E}(\boldsymbol{z}) = \boldsymbol{X\beta}$, where $\boldsymbol{X}$ is a fixed full-rank $n \times k$ matrix and $\boldsymbol{\beta} \in \mathbb{R}^k$ is an unknown parameter, and with variance matrix equal to some integer power q of a linear combination of the $\widetilde{\boldsymbol{A}}_j$', i.e., $\mathrm{var}(\boldsymbol{z}) = \widetilde{\boldsymbol{\Sigma}}^q$, with

$$\widetilde{\boldsymbol{\Sigma}} = \sum_{j \in J} c(j) \widetilde{\boldsymbol{A}}_j, \tag{16}$$

where J is a set of nonnegative integers including 0, and the coefficients $c(j)$ are such that $\widetilde{\boldsymbol{\Sigma}}^q$ exists (if $q < 0$) and is positive definite. For example, $q = 1$ corresponds to *second-order stationary and isotropic processes* on $\widetilde{\Gamma}$ (cf. Section 4.3 below); $q = -1$ and $q = -2$ correspond to, respectively, *conditional autoregressions* and *simultaneous autoregressions* constructed using the $\widetilde{\boldsymbol{A}}_h$'s as weights matrices (see, e.g., Cressie 1993, Ch. 6). By the Gauss-Markov theorem, $\widehat{\boldsymbol{\beta}}_{GLS} = (\boldsymbol{X}\widetilde{\boldsymbol{\Sigma}}^{-q}\boldsymbol{X})^{-1}\boldsymbol{X}'\widetilde{\boldsymbol{\Sigma}}^{-q}\boldsymbol{z}$ is the best linear unbiased estimator. Assume that the column space of $\boldsymbol{X}$ is spanned by k eigenvectors of the $\widetilde{\boldsymbol{A}}_h$'s. By implication (ii), it follows that $\widehat{\boldsymbol{\beta}}_{GLS} = \widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{z}$, i.e., the OLS estimator is best linear unbiased. Since $\boldsymbol{\iota}_N$ is an eigenvector of any $\widetilde{\boldsymbol{A}}_h$, a particular case of this result is that, if $\boldsymbol{z}$ has constant unknown mean $\mu \in \mathbb{R}$, then the sample mean $\boldsymbol{\iota}_N'\boldsymbol{z}/N$ is the best linear unbiased estimator of μ.

3.2 Eigenvalues

We now derive all eigenvalues of $\widetilde{\boldsymbol{A}}_h$. First observe that

$$\widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes}\widetilde{\boldsymbol{F}}_{\boldsymbol{\beta}}^{\otimes} = \bigotimes_{i=1}^{d}\left(\widetilde{\boldsymbol{F}}_{\alpha(i)}^{(n_i)}\widetilde{\boldsymbol{F}}_{\beta(i)}^{(n_i)}\right) = \bigotimes_{i=1}^{d}\left(\widetilde{\boldsymbol{F}}_{\beta(i)}^{(n_i)}\widetilde{\boldsymbol{F}}_{\alpha(i)}^{(n_i)}\right) = \widetilde{\boldsymbol{F}}_{\boldsymbol{\beta}}^{\otimes}\widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes}. \tag{17}$$

That is, like the $\widetilde{\boldsymbol{F}}_r$, the matrices $\widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes}$ are pairwise commutative, and hence admit a set of common eigenvectors (see, Horn and Johnson, 1985, 51-53). Thus, the eigenvalues of the matrices $\widetilde{\boldsymbol{A}}_h$, which are sums of the matrices $\widetilde{\boldsymbol{F}}_{\boldsymbol{\alpha}}^{\otimes}$'s, are easily obtained from those of the summands. Specifically, letting the known eigenvalues of $\widetilde{\boldsymbol{F}}_r^{(n)}$ be denoted by $\lambda_{r,0}^{(n)}, \lambda_{r,1}^{(n)}, ..., \lambda_{r,n-1}^{(n)}$, we have:[7]

[7] The eigenvalues of the matrices $\widetilde{\boldsymbol{F}}_r^{(n)}$ are given, e.g., in Theorem 6.5.3 of Anderson (1971). If n is odd and $r \neq 0$, they are $2\cos(2\pi ri/n)$, $i = 1, ..., (n-1)/2$, each with multiplicity two, and 2; if

Theorem 3 *The eigenvalues of the matrix* $\widetilde{\boldsymbol{A}}_h$ *are*

$$\lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta}) = \sum_{\boldsymbol{\alpha}\in\Gamma_h}\left(\prod_{i=1}^{d}\lambda^{(n_i)}_{\alpha(i),\beta(i)}\right),\ \boldsymbol{\beta}\in\Gamma. \tag{18}$$

Since $\widetilde{\boldsymbol{D}}_h = \widetilde{m}_h\boldsymbol{I}_N$, it follows immediately that the eigenvalues of $\widetilde{\boldsymbol{L}}_h$ are $\widetilde{m}_h - \lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta})$, $\boldsymbol{\beta}\in\Gamma$. To obtain the $\lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta})$'s, one needs to know Γ_h, which, depending on Γ and on h, may be a complicated set. Fortunately, again following the approach described in HM, the set Γ_h can be obtained from a suitable generating function. Letting $\bar{h} = \sum_{i=1}^{d}(n_i-1)^2$, it is easy to verify that $\boldsymbol{\alpha}\in\Gamma_h$ if and only if the term $\prod_{i=1}^{d}x_{\alpha(i)}$ appears as a term in the coefficient of t^h in the formal expansion of the generating function

$$\prod_{i=1}^{d}\left\{\sum_{r=0}^{n_i-1}(x_r t^{r^2})\right\} = \sum_{h=0}^{\bar{h}}\left\{t^h\sum_{\boldsymbol{\alpha}\in\Gamma_h}\left(\prod_{i=1}^{d}x_{\alpha(i)}\right)\right\}.$$

The members of the set Γ_h can therefore be obtained simply by expanding the generating function using symbolic computer package, for any Γ and any h.

Remark 2 Theorem 3 holds not only for the matrix $\widetilde{\boldsymbol{A}}_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}$, but, more generally, for any matrix $\sum_{\boldsymbol{\alpha}\in U}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}$, with $U\subset\Gamma$. From the proof of the theorem, it is clear that the eigenvalues of $\sum_{\boldsymbol{\alpha}\in U}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}$ are $\lambda^{\otimes}_{U}(\boldsymbol{\beta})$, $\boldsymbol{\beta}\in\Gamma$. This is useful for extensions of the theory in this paper to cases when neighborhood on Γ is defined according to a metric different from the Euclidean one.

4 Applications

In this section we study three applications of the above theoretical results, in increasing order of complexity. Many more applications seem possible, but are left for future investigation. Section 4.1 analyzes the sample autocovariance of a spatial process, in the simple case when the mean is zero and data are identically and independently distributed. Section 4.2 considers assessing the significance of a Moran correlogram. Section 4.3 is concerned with GLS estimation of the variogram of an isotropic and second order stationary process.

4.1 The Sample Autocovariance

As a first application of the circular spatial design matrices, we consider the circular sample autocovariance

$$\widetilde{b}_h = \frac{1}{2\widetilde{N}_h}\boldsymbol{z}'\widetilde{\boldsymbol{A}}_h\boldsymbol{z}.$$

---

n is even and $r\neq 0, n/2$, they are $2\cos(2\pi ri/n)$, $i=1,...,(n-2)/2$, each with multiplicity two, 2, and $2(-1)^r$. Finally, $\widetilde{\boldsymbol{F}}^{(n)}_{n/2}$ has eigenvalues 1 and -1, each with multiplicity $n/2$, and $\widetilde{\boldsymbol{F}}^{(n)}_0 = \boldsymbol{I}_n$ and hence has the unique eigenvalue 1.

in approximating the sample autocovariance $\widehat{b}_h$, given in (4). As we mentioned already above, the main reason why it is convenient to approximate $\widehat{b}_h$ with $\widetilde{b}_h$ is that the cumulants of the latter are easier to obtain. To see this, we need to state the following standard lemma, which is proved for instance in Kendall and Stuart (1969), Ch. 15.

Lemma 1 *Let $\boldsymbol{y}$ be an $N \times 1$ random vector, and $\boldsymbol{R}, \boldsymbol{R}_1$ and $\boldsymbol{R}_2$ $N \times N$ symmetric matrices. If $\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{\tau}, \boldsymbol{\Omega})$, then the p-th cumulant of $\boldsymbol{y}'\boldsymbol{R}\boldsymbol{y}$ is*

$$\kappa_p(\boldsymbol{y}'\boldsymbol{R}\boldsymbol{y}) = 2^{p-1}(p-1)! \left(\mathrm{tr}\left[(\boldsymbol{R}\boldsymbol{\Omega})^p\right] + p\boldsymbol{\tau}'\boldsymbol{R}(\boldsymbol{\Omega}\boldsymbol{R})^{p-1}\boldsymbol{\tau}\right),\ p = 1, 2, ... \tag{19}$$

and

$$\mathrm{cov}(\boldsymbol{y}'\boldsymbol{R}_1\boldsymbol{y}, \boldsymbol{y}'\boldsymbol{R}_1\boldsymbol{y}) = 2\,\mathrm{tr}\,(\boldsymbol{R}_1\boldsymbol{\Omega}\boldsymbol{R}_2\boldsymbol{\Omega}) + 4\boldsymbol{\tau}'\boldsymbol{R}_1\boldsymbol{\Omega}\boldsymbol{R}_2\boldsymbol{\tau}. \tag{20}$$

Using Theorem 2, one can see immediately that if $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \widetilde{\boldsymbol{\Sigma}}^q)$, then $\kappa_p(\boldsymbol{z}'\widetilde{\boldsymbol{A}}_h\boldsymbol{z})$ and $\mathrm{cov}(\boldsymbol{z}'\widetilde{\boldsymbol{A}}_h\boldsymbol{z}, \boldsymbol{z}'\widetilde{\boldsymbol{A}}_k\boldsymbol{z})$ depend only on the (known) eigenvalues of the $\widetilde{\boldsymbol{A}}_h$'s. In Section 4.3, we shall also use the fact that, again by Theorem 2, if $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \widetilde{\boldsymbol{\Sigma}}^q)$, then $\kappa_p(\boldsymbol{z}'\widetilde{\boldsymbol{L}}_h\boldsymbol{z})$ and $\mathrm{cov}(\boldsymbol{z}'\widetilde{\boldsymbol{L}}_h\boldsymbol{z}, \boldsymbol{z}'\widetilde{\boldsymbol{L}}_k\boldsymbol{z})$ depend only on the (known) eigenvalues of the $\widetilde{\boldsymbol{L}}_h$'s.

Having explained the main advantage related to the use of $\widetilde{b}_h$, we now move to analyze the performance of $\widetilde{b}_h$ in approximating $\widehat{b}_h$. We only consider the case $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$ (which can of course be seen as a particular case of $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \widetilde{\boldsymbol{\Sigma}}^q)$), although extensions are certainly possible (the case $\mathrm{var}(\boldsymbol{z}) = \widetilde{\boldsymbol{\Sigma}}$ can be dealt with as in Section 4.3).

Ideally, one would like to compare the densities of $\widetilde{b}_h$ and $\widehat{b}_h$, but this is outside the scope of the present paper. In the rest of this subsection, we confine ourselves to the first few cumulants of $\widehat{b}_h$.

Proposition 1 *When $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, and for $h, k > 0$ such that $N_h, N_k > 0$,*

(i) $\mathrm{E}(\widehat{b}_h) = 0$, $\mathrm{var}(\widehat{b}_h) = 1/N_h$;

(ii) $\widehat{b}_h$ *and* $\widehat{b}_k$, $h \neq k$, *are uncorrelated.*

(iii) if $d = 1, 2$ or h is odd, the density of $\widehat{b}_h$ is symmetric about zero.

To derive corresponding results for $\widetilde{b}_h$ we need the following condition, which is stated for a fixed squared distance h.

Condition A *For any $\boldsymbol{\alpha} \in \Gamma_h$ and any $i = 1, ..., d$, $\alpha\,(i) < n_i/2$.*

In applications, one is usually concerned only with values of h satisfying Condition A. Indeed, when Condition A is not satisfied, the number N_h of h-neighbors may be too small for inferential purposes.

Proposition 2 *When $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, and for $h, k > 0$ such that $N_h, N_k > 0$,*

(i) $\mathrm{E}(\widetilde{b}_h) = 0$, $\mathrm{var}(\widetilde{b}_h) = 1/\widetilde{N}_h$;

(ii) $\widetilde{b}_h$ *and* $\widetilde{b}_k$, $h \neq k$, *are uncorrelated if Condition A holds for both squared distances h and k.*

The extension of Proposition 1(iii) to $\widetilde{b}_h$ is more complicated and is given in Appendix A. The main result there is that, although $\widetilde{b}_h$ does not need to be symmetric about zero when $\widehat{b}_h$ is, this is not a problem in practice from the point of view of approximating $\widehat{b}_h$ with $\widetilde{b}_h$, because the low-order cumulants of $\widetilde{b}_h$ are zero if and only if those of $\widehat{b}_h$ are.

Regarding part (ii) of Proposition 2, it is worth noting that if Condition A does not hold for h or k, then $\widetilde{b}_h$ and $\widetilde{b}_k$ are generally correlated. The magnitude of the correlation can be obtained from the eigenvalues of $\widetilde{\boldsymbol{A}}_h$ and $\widetilde{\boldsymbol{A}}_k$, because,

$$\operatorname{cov}(\widetilde{b}_h, \widetilde{b}_k) = 2\frac{1}{4\widetilde{N}_h\widetilde{N}_k}\operatorname{tr}(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k) = \frac{1}{2\widetilde{N}_h\widetilde{N}_k}\sum_{\boldsymbol{\beta}\in\Gamma}\{\lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta})\,\lambda^{\otimes}_{\Gamma_k}(\boldsymbol{\beta})\},$$

where the first equality follows from Lemma 1 and the second from Theorem 2.

To give an indication of the accuracy of the circular approximation, in Tables 1 and 2 we report some values of the ratios

$$\eta_{p,h} = \frac{\kappa_p(\widehat{b}_h)}{\kappa_p(\widetilde{b}_h)},$$

where $\kappa_p(y)$ denotes the p-th cumulant of a random variable y. Table 1 displays values of $\eta_{2,h}$ and $\eta_{4,h}$ for square grids in 2 and 4 dimensions (values of $\eta_{3,h}$ are not reported, because $\kappa_3(\widehat{b}_h) = \kappa_3(\widetilde{b}_h) = 0$ for most values of n, d, and h). The values of $\eta_{2,h}$ have been computed using formulae derived in Appendix B, whereas those of $\eta_{4,h}$ have been obtained by deriving $\kappa_4(\widehat{b}_1)$ from the generating functions given in HM and $\kappa_4(\widetilde{b}_1)$ from the known eigenvalues of $\widetilde{\boldsymbol{A}}_h$. It is also of interest to look at the performance of the approximation of non-square grid. Table 2 displays values of $\eta_{2,h}$ for 2-dimensional grids with $n_1/n_2 = 1, 4, 16, 25$.

		$\eta_{2,h}$					$\eta_{4,h}$				
		n					n				
	h	20	40	60	80	100	20	40	60	80	100
$d=2$	1	1.053	1.026	1.017	1.013	1.010	1.114	1.055	1.036	1.027	1.021
	2	1.108	1.052	1.034	1.025	1.020	1.267	1.125	1.081	1.060	1.048
	5	1.170	1.080	1.052	1.039	1.031	1.416	1.188	1.121	1.089	1.071
	10	1.238	1.109	1.070	1.052	1.041	1.590	1.256	1.163	1.120	1.094
	100	1.866	1.422	1.253	1.181	1.141	2.265	2.083	1.631	1.439	1.336
$d=3$	1	1.053	1.026	1.017	1.013	1.010	1.112	1.054	1.035	1.026	1.021
	2	1.108	1.052	1.034	1.025	1.020	1.251	1.117	1.077	1.057	1.045
	5	1.170	1.080	1.052	1.039	1.031	1.391	1.177	1.114	1.084	1.067
	10	1.238	1.109	1.070	1.052	1.041	1.572	1.247	1.157	1.115	1.091
	100	2.064	1.441	1.265	1.189	1.147	2.834	2.134	1.643	1.446	1.340

Table 1: Some values of $\eta_{2,h}$ and $\eta_{4,h}$ on 2- and 3-dimensional grids.

The numerical results suggest that our approximation is generally very satisfactory, but it deteriorates when N is small, or h is large. Also, the approximation works better

		N					
n_1/n_2	h	400	1600	3600	6400	10^4	10^6
1	1	1.053	1.026	1.017	1.013	1.010	1.001
	2	1.108	1.052	1.034	1.026	1.020	1.002
	5	1.170	1.080	1.052	1.039	1.030	1.003
	10	1.238	1.109	1.070	1.052	1.041	1.004
	100	1.866	1.422	1.253	1.181	1.141	1.013
4	1	1.067	1.032	1.021	1.016	1.013	1.001
	2	1.140	1.066	1.043	1.032	1.026	1.003
	5	1.223	1.102	1.066	1.049	1.039	1.004
	10	1.320	1.140	1.090	1.066	1.052	1.005
	100	2.890	1.469	1.342	1.239	1.184	1.016
16	1	1.119	1.056	1.037	1.027	1.022	1.002
	2	1.266	1.118	1.076	1.056	1.044	1.004
	5	1.457	1.188	1.118	1.086	1.068	1.006
	10	1.717	1.267	1.164	1.118	1.093	1.009
	100	1.143	2.401	1.785	1.371	1.362	1.028
25	1	1.149	1.070	1.054	1.034	1.027	1.004
	2	1.347	1.149	1.095	1.069	1.055	1.007
	5	1.220	1.240	1.149	1.108	1.084	1.011
	10	2.051	1.348	1.208	1.149	1.116	1.014
	100	1.111	2.098	1.812	1.404	1.360	1.046

Table 2: Some values of $\eta_{2,h}$ on 2-dimensional grids.

on square grids than on rectangular ones. Finally, note that, while it is straightforward to show that $\kappa_p(\boldsymbol{z}'\widetilde{\boldsymbol{A}}_h\boldsymbol{z}) \geq \kappa_p(\boldsymbol{z}'\boldsymbol{A}_h\boldsymbol{z})$ (in particular, $\widetilde{N}_h \geq N_h$ for $p = 2$), a simple relationship does not necessarily hold for the cumulants of the normalized quadratic forms $\widetilde{b}_h$ and $\widehat{b}_h$, although the results in Tables 1 and 2 suggest that $\kappa_p(\widetilde{b}_h)$ underestimates $\kappa_p(\widehat{b}_h)$ in most cases of interest.

4.2 Moran Correlogram

Our second application concerns testing for the joint significance of a Moran correlogram. For an isotropic process on Γ, a Moran correlogram is a plot of the Moran statistic r_h, defined in expression (6), against h. Note that, when $d = 1$, this reduces to the usual time-series correlogram. A popular test for assessing the joint significance of a Moran correlogram has been proposed by Oden (1984). Suppose that r_h is computed for a set H of nonzero squared distances h such that $N_h > 0$, and let $\boldsymbol{r}$ be the $|H| \times 1$ vector $(r_h, h \in H)'$. Then, Oden's test rejects the null hypothesis $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \boldsymbol{I}_N)$ for large values of the statistic

$$O = (\boldsymbol{r} - \mathrm{E}(\boldsymbol{r}))'(\mathrm{var}(\boldsymbol{r}))^{-1}(\boldsymbol{r} - \mathrm{E}(\boldsymbol{r})), \tag{21}$$

where the two moments are evaluated under the null hypothesis.[8] Under the null, $\boldsymbol{r}$ is asymptotically (as $N \to \infty$) multivariate normal; see Kelejian and Prucha (2001). Thus, asymptotic critical values for Oden's test can be obtained from a χ^2 distribution with $|H|$ degrees of freedom.

For an $h > 0$ such that N_h (and hence $\widetilde{N}_h$) is positive, we define the circular counterpart of c_h in (5) as

$$\widetilde{c}_h = \frac{1}{\widetilde{N}_h}\boldsymbol{z}'\boldsymbol{M}\widetilde{\boldsymbol{A}}_h\boldsymbol{M}\boldsymbol{z},$$

and the sample correlation coefficient, or circular Moran statistic, as

$$\widetilde{r}_h = \frac{\widetilde{c}_h}{\widetilde{c}_0} = \frac{1}{\widetilde{m}_h}\frac{\boldsymbol{z}\boldsymbol{M}\widetilde{\boldsymbol{A}}_h\boldsymbol{M}\boldsymbol{z}}{\boldsymbol{z}'\boldsymbol{M}\boldsymbol{z}}.$$

Proposition 3 *Assume $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \boldsymbol{I}_N)$. For any $h, k > 0$ such that $N_h, N_k > 0$,*

$$\begin{aligned}\mathrm{E}(r_h) = \mathrm{E}(\widetilde{r}_h) &= -\frac{1}{N-1},\\ \mathrm{cov}(r_h, r_k) &= \frac{1}{N^2-1}\left(\frac{N^2}{N_h}1_{h=k} - \frac{N_{h,k}}{N_h N_k}N + 2\frac{N-2}{N-1}\right),\end{aligned} \tag{22}$$

where $N_{h,k} = \boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{A}_k\boldsymbol{\iota}_N$, and, provided that Condition A holds for both h and k,

$$\mathrm{cov}(\widetilde{r}_h, \widetilde{r}_k) = \frac{2N}{N^2-1}\left(\frac{1}{\widetilde{m}_h}1_{h=k} - \frac{1}{N-1}\right). \tag{23}$$

Remark 3 Proposition 3 generalizes results available in the time-series literature ($d = 1$) to higher dimension. When $d = 1$, Condition A is equivalent to $\sqrt{h}, \sqrt{k} < n/2$, where $n = n_1$. Under this condition, expression (32) in Appendix C yields $N_{h,k} = 4(n-k)-2h$. It is then easily checked that expressions (4.4) and (4.5) in Dufour and Roy (1985) are particular cases of (22) and (23), respectively.[9] Also, when $d = h = 1$ Proposition 3 gives $\mathrm{var}(r_1) = (n-2)^2/(n-1)^3$ and $\mathrm{var}(\widetilde{r}_1) = n(n-3)(n+1)^{-1}(n-1)^{-2}$, which correspond to the results derived by Moran (1948) for the first-order serial correlation coefficient and its circular counterpart.

Computation of the term $N_{h,k}$ defined in Proposition 3 is discussed in Appendix C. Exact computation of Oden's statistic, by means of expression (22), may be cumbersome when N or $|H|$ are large. We propose to replace the test statistic (21) with

$$\widetilde{O} = (\widetilde{\boldsymbol{r}} - \mathrm{E}(\widetilde{\boldsymbol{r}}))'(\mathrm{var}(\widetilde{\boldsymbol{r}}))^{-1}(\widetilde{\boldsymbol{r}} - \mathrm{E}(\widetilde{\boldsymbol{r}})), \tag{24}$$

[8] Of course, under the null hypothesis $\mathrm{var}(\boldsymbol{z})$ may be known only up to a parameter σ^2, which is here taken to be 1 without loss of generality.

[9] The reader should notice that r_h in Dufour and Roy (1984) is not the same as our r_h, because of a different normalization of the sample covariances, and because our h denotes a squared distance.

where, with obvious notation, $\widetilde{\boldsymbol{r}}$ is the $|H| \times 1$ vector $(\widetilde{r}_h, h \in H)')$. As for $\boldsymbol{r}$, it is straightforward to check that $\widetilde{\boldsymbol{r}}$ is asymptotically normal, and hence that $\widetilde{O}$ is asymptotically distributed as a χ^2 distribution with $|H|$ degrees of freedom. A substantial reduction in computational effort comes from the fact that expression (23) does not depend on h and k, when $h \neq k$.

Monte Carlo simulation shows that the test based on $\widetilde{O}$ provides a satisfactory approximation to Oden's test, both under the null and under typical alternative hypotheses. This is true as long as N is not too small (say, smaller than 1000). As a representative example, in Table 3 we report the rejection rates obtained for a first-order conditional autoregressive (CAR(1)) process $\boldsymbol{z} \sim \mathrm{N}\left(\boldsymbol{\iota}_N, (\boldsymbol{I}_N - \rho \boldsymbol{A}_1)^{-1}\right)$, with $N = 2^{12} = 4096$. The process is simulated on a 2-dimensional grid with $n_1 = n_2 = 2^6$, and on a 3-dimensional one with $n_1 = n_2 = n_3 = 2^4$. We consider four values of $\rho/\rho_{\max}$, where $\rho_{\max}$ is the right boundary of the set of admissible values of ρ: 0 (corresponding to the null hypothesis), $0.1, 0.2, 0.3$. The set H is taken to be the set of squared distances $h = 1, ..., 10$ such that $N_h > 0$ (that is, $H = \{1, 2, 4, 5, 8, 9, 10\}$ when $d = 2$, and $H = \{1, 2, 3, 4, 5, 6, 8, 9, 10\}$ when $d = 3$). Two values of the nominal size α of the test are considered, 0.01 and 0.05. The number of repetitions is 200,000. From the numerical results it appears that size is essentially unaffected by the approximation, and at the same time the power is not seriously compromised.

	$\alpha = 0.01$				$\alpha = 0.05$			
	$d = 2$		$d = 3$		$d = 2$		$d = 3$	
$\rho/\rho_{\max}$	O	$\widetilde{O}$	O	$\widetilde{O}$	O	$\widetilde{O}$	O	$\widetilde{O}$
0	0.010	0.010	0.011	0.011	0.050	0.050	0.051	0.051
0.1	0.139	0.137	0.068	0.063	0.322	0.317	0.192	0.182
0.2	0.825	0.817	0.504	0.464	0.936	0.932	0.725	0.692
0.3	0.999	0.999	0.955	0.938	1	1	0.988	0.983

Table 3: Monte Carlo rejection rates for Oden's test and its circular approximation.

4.3 GLS Estimation of the Variogram

In this section we assume that the spatial process $\{z(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in \Gamma\}$ is second-order stationary and isotropic, i.e., that $\mathrm{E}(z(\boldsymbol{\alpha}))$ does not depend on $\boldsymbol{\alpha}$ and $\mathrm{cov}(z(\boldsymbol{\alpha}), z(\boldsymbol{\beta}))$ depends on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ only through $\|\boldsymbol{\alpha} - \boldsymbol{\beta}\|$. The covariance matrix of a such process can be represented as

$$\boldsymbol{\Sigma} = \sum_{j \in J} c(j) \boldsymbol{A}_j,$$

where $c(j)$ denotes the covariance between variables observed at j-neighbors, and J is a set of nonnegative integers including 0. We assume that the $c(j)$'s, for $j \in J$, are nonzero and such that $\boldsymbol{\Sigma}$ is positive-definite. Note that the semivariogram of a second-order stationary and isotropic process is $\gamma(h) = c(0) - c(h)$.

Given a parametric semivariogram model $\gamma(h;\boldsymbol{\theta})$, a popular way of estimating the parameter vector $\boldsymbol{\theta}$ is by a least squares procedure; see, e.g., Cressie (1993). Suppose the estimator $\widehat{\gamma}_h$ in (3) (or some other nonparametric estimator of $\gamma(h)$) is computed for a set H of distances h such that $N_h > 0$. Let $\widehat{\boldsymbol{\gamma}}$ and $\boldsymbol{\gamma}(\boldsymbol{\theta})$ be $|H| \times 1$ vectors with entries, respectively, $\widehat{\gamma}_h$ and $\gamma(h;\boldsymbol{\theta})$, $h \in H$. Then, a least squares estimator of $\boldsymbol{\theta}$ is found by minimizing

$$(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\boldsymbol{\theta}))' \boldsymbol{W}(\boldsymbol{\theta}) (\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}(\boldsymbol{\theta})), \tag{25}$$

for some $|H| \times |H|$ positive definite weighting matrix $\boldsymbol{W}(\boldsymbol{\theta})$. Under weak conditions, the most efficient estimator in this class is the GLS estimator, say $\widehat{\boldsymbol{\theta}}_{GLS}$, obtained by taking

$$\boldsymbol{W}(\boldsymbol{\theta}) = (\mathrm{var}(\widehat{\boldsymbol{\gamma}}))^{-1} \tag{26}$$

(e.g., Cressie, 1993). Since, for a general spatial process, $\mathrm{var}(\widehat{\boldsymbol{\gamma}})$ is non-diagonal, it follows that a diagonal $\boldsymbol{W}(\boldsymbol{\theta})$ (as in ordinary or weighted least squares) may lead to a substantial loss of efficiency. On the other hand, the computation of $\widehat{\boldsymbol{\theta}}_{GLS}$ may be prohibitive, because it requires deriving $\mathrm{var}(\widehat{\boldsymbol{\gamma}})$, inverting it, and minimizing (25) (see, again, Cressie, 1993). For a Gaussian second-order stationary and isotropic process on Γ, $\mathrm{var}(\widehat{\boldsymbol{\gamma}})$ can be obtained exactly by Lemma 7 in HM,[10] but it is clear the required computation is still cumbersome if N or $|J|$ is large. In what follows we show that our circular approximation can dramatically reduce the computational effort associated to $\widehat{\boldsymbol{\theta}}_{GLS}$, and we study its performance numerically.

In order to approximate $\mathrm{var}(\widehat{\boldsymbol{\gamma}})$, we define a "circular process" $\boldsymbol{z}$ with $\mathrm{var}(\boldsymbol{z}) = \sum_{j\in J} c(j)\widetilde{\boldsymbol{A}}_j$, which has already been denoted by $\widetilde{\boldsymbol{\Sigma}}$ in (16). The following proposition establishes that, under a suitable condition, the circular variogram estimator

$$\widetilde{\gamma}_h = \frac{1}{2\widetilde{N}_h}\boldsymbol{z}'\widetilde{\boldsymbol{L}}_h\boldsymbol{z}$$

is unbiased for $\gamma(h)$. For a general proof of the unbiasedeness of $\widehat{\gamma}_h$ see, e.g., Cressie (1993).

Proposition 4 *Assume that $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \widetilde{\boldsymbol{\Sigma}})$. Then, provided that $N_h > 0$ and that Condition A holds for the squared distance h and for all the squared distances $j \in J$, $\widetilde{\gamma}_h$ is unbiased.*

Letting $\widetilde{\boldsymbol{\gamma}}$ be the vector $(\widetilde{\gamma}_h, h \in H)'$, the circular approximation $\widetilde{\boldsymbol{\theta}}_{GLS}$ to $\widehat{\boldsymbol{\theta}}_{GLS}$ is then obtained by replacing $\mathrm{var}(\widehat{\boldsymbol{\gamma}})$ in (26) with $\mathrm{var}(\widetilde{\boldsymbol{\gamma}})$, computed under the assumption $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \widetilde{\boldsymbol{\Sigma}})$. The next result shows that $\mathrm{var}(\widetilde{\boldsymbol{\gamma}})$ can be conveniently obtained from the known eigenvalues of the matrices $\widetilde{\boldsymbol{A}}_h$.

Proposition 5 *Assume that $\boldsymbol{z} \sim \mathrm{N}(\mu\boldsymbol{\iota}_N, \widetilde{\boldsymbol{\Sigma}})$. For $h, k \in H$,*

$$\mathrm{cov}(\widetilde{\gamma}_h, \widetilde{\gamma}_k) = \frac{1}{2\widetilde{N}_h\widetilde{N}_k}\sum_{j,l\in J}\left\{c(j)c(l)\sum_{\boldsymbol{\beta}\in\Gamma}\left[\lambda^{\otimes}_{\Gamma_j}(\boldsymbol{\beta})\lambda^{\otimes}_{\Gamma_l}(\boldsymbol{\beta})\left(\widetilde{m}_h - \lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta})\right)\left(\widetilde{m}_k - \lambda^{\otimes}_{\Gamma_k}(\boldsymbol{\beta})\right)\right]\right\}.$$

[10]Although it is not pointed out there, expression (37) in HM also holds when $\mathrm{E}(\boldsymbol{z}) = \mu\boldsymbol{\iota}_N$, where $\mu \in \mathbb{R}$ is an unknown parameter.

Since $\boldsymbol{A}_h$ and $\widetilde{\boldsymbol{A}}_h$ are asymptotically equivalent (see footnote 6), it seems clear that $\widetilde{\boldsymbol{\theta}}_{GLS}$ should be asymptotically efficient (in the sense of Lahiri et al, 2002). Rather than providing a formal proof of this property, in the following we report numerical results illustrating the finite sample behavior of $\widetilde{\boldsymbol{\theta}}_{GLS}$.

Consider the spherical variogram

$$\gamma(h;\theta) = \begin{cases} 0 & \text{if } h = 0, \\ \frac{1}{2}\left[3\left(\frac{h}{\theta}\right)^{\frac{1}{2}} - \left(\frac{h}{\theta}\right)^{\frac{3}{2}}\right] & \text{if } 0 \leq h \leq \theta, \\ 1 & h > \theta. \end{cases} \tag{27}$$

For simplicity, model (27) depends only on one parameter, θ, representing the square of the range. The sill and the nugget have been fixed to 1 and 0, respectively. Table 4 displays Monte Carlo results concerning estimation of θ given observation from a Gaussian spatial process on a 2-dimensional grid with $\theta = 20$ in model (27). We consider two choices of H, namely, $H = \{1 \leq h \leq h_{\max}\}$, with $h_{\max} = 10, 20$. The number of repetitions is 5000. It can be seen that the performance of $\widetilde{\theta}_{GLS}$ is very similar to that of $\widehat{\theta}_{GLS}$, even when the grid is not square. In all the 16 cases considered, the reduction in computational time associated to using $\widetilde{\theta}_{GLS}$ rather than $\widehat{\theta}_{GLS}$ is larger than 99% (average over the repetitions). To show the advantages of GLS estimation, Table 4 also displays results concerning the OLS and the WLS estimators of θ, which are obtained by setting $\boldsymbol{W}(\boldsymbol{\theta}) = \boldsymbol{I}_N$ and $\boldsymbol{W}(\boldsymbol{\theta}) = \mathrm{diag}\left((\mathrm{var}(\widehat{\boldsymbol{\gamma}}))^{-1}\right)$ in (26), respectively. Both the OLS and the WLS estimators are very inefficient.

n_1, n_2	$h_{\max}$	$\widehat{\theta}_{GLS}$	$\widetilde{\theta}_{GLS}$	$\widehat{\theta}_{OLS}$	$\widehat{\theta}_{WLS}$
$50, 50$	10	19.944	19.919	20.244	20.134
		(1.065)	(1.064)	(2.935)	(2.042)
	20	19.887	19.847	20.616	20.181
		(0.938)	(0.952)	(3.715)	(2.159)
$25, 100$	10	19.928	19.900	20.101	20.039
		(1.117)	(1.124)	(2.954)	(2.038)
	20	19.810	19.753	20.578	20.112
		(0.947)	(0.961)	(3.930)	(2.243)

Table 4: Monte Carlo results concerning the estimation of θ in the variogram model (27), standard errors in parentheses.

5 Discussion

The (suitably normalized) quadratic forms associated to spatial design matrices play a central role in various inferential procedures in the context of isotropic spatial processes defined on uniform grids. In many cases of interest, the cumulants of such statistics are a function of the eigenvalues of the spatial design matrices (e.g., for second-order stationary data, and for spatial autoregressions based on the Euclidean distance). Unfortunately, the eigenvalues are not known analytically in general. This is not a problem in principle, because the cumulants can be obtained exactly by the procedures outlined in Hillier and Martellosio (2006). However, the required computation may become prohibitive as the sample size increases. The present paper has proposed an approximation to the cumulants, based on circular counterparts of the spatial design matrices. Depending on the particular application, the reduction in computational time can be dramatic, because the eigenvalues of the circular spatial design matrices are known in closed form. The approximation performs well especially when it is most needed, that is, when the sample size is large. Our numerical results indicate that, for a fixed sample size, the approximation works best if the sides of the grid are of similar magnitude, and deteriorates as the grid becomes more rectangular.

The following two extensions are left for future research. First, while we have confined attention to Gaussian processes, various generalizations are possible. For instance, the methods that we have used admit simple modifications in the case of elliptically contoured and skew-symmetric distributions (see Genton, 1999, Genton et al., 2001, and Arellano-Valle and Genton, 2009).[11] Second, in some applications other norms may be more appropriate than the Euclidean one to define neighborhood. Extensions of the framework proposed in this paper to any other L_p-norm are theoretically simple.

Appendix A Symmetry of the Circular Sample Autocovariance

Proposition 1 asserts that the sample autocovariance $\widehat{b}_h$ is symmetric about zero if $d = 1, 2$ or h is odd. Of course, it would be desirable if the circular sample autocovariance $\widetilde{b}_h$ inherited the same symmetry, at least approximately. In this appendix we check whether this is the case.

It is convenient to start from the 1-dimensional case. Let

$$\widetilde{b}_{r,1} = \frac{1}{2n_1} \boldsymbol{z}' \widetilde{\boldsymbol{F}}_r^{(n_1)} \boldsymbol{z}$$

denote the circular sample autocovariance $\widetilde{b}_h$ when $d = 1$, with $r = \sqrt{h}$. Observe that $\widetilde{b}_{r,1} = \widetilde{b}_{n_1-r,1}$, for any $r = \lfloor n_1/2 \rfloor, ..., n_1 - 1$, and that $\widetilde{b}_{n_1/2,1} = \widehat{b}_{n_1/2,1} = \boldsymbol{z}' \boldsymbol{F}_{n_1/2}^{(n_1)} \boldsymbol{z}$, for any even n_1. Hence, we only need to consider the autocovariances $\widetilde{b}_{r,1}$, $r = 1, 2, ..., \lfloor n_1/2 \rfloor - 1$.

Proposition A.1 *Assume $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_{n_1})$. For $r = 1, 2, ..., \lfloor n_1/2 \rfloor - 1$, the density of $\widetilde{b}_{r,1}$ is symmetric about zero if and only if either n_1 is even and r is odd, or n_1 is an even multiple of r.*

[11] It is also worth mentioning that the distribution of a ratio of quadratic forms such as the Moran statistic r_h is the same for any spherically symmetric distribution; see, e.g., Dufour and Roy (1985).

By comparing Proposition A.1 with part (iii) of Proposition 1, it emerges that the density of $\widetilde{b}_h$ may be nonsymmetric when that of $\widehat{b}_h$ is symmetric. Recall that the the density of a statistic is symmetric about zero if and only if all its odd cumulants vanish. Then, by Proposition A.1, $\kappa_p(\widehat{b}_h) = 0$ does not imply $\kappa_p(\widetilde{b}_h) = 0$ for all odd p (at least when $d = 1$). It is therefore of interest to establish the circumstances under which $\kappa_p(\widetilde{b}_h) = 0$ is necessary and sufficient for $\kappa_p(\widehat{b}_h) = 0$. To this purpose, we need the following extension of Condition A.

Condition B *For any $\boldsymbol{\alpha} \in \Gamma_h$ and any $i = 1, ..., d$, $\alpha(i) < n_i/p$.*

For a general dimension d, we have the following result.

Proposition A.2 *Assume $\boldsymbol{z} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_N)$. If $\kappa_p(\widetilde{b}_h) = 0$ then $\kappa_p(\widehat{b}_h) = 0$, for any $p, h = 1, 2, ...$ The converse does not hold in general, but it does hold if $p = 1$, p is even, or under Condition B.*

According to Proposition A.2, an odd cumulant of $\widetilde{b}_h$ may be nonzero when the corresponding cumulant of $\widehat{b}_h$ is 0 only if p is large (specifically, if $p > n_{\min}/h$, where $n_{\min} = \min\{n_1, ..., n_d\}$). We can therefore conclude that, although $\widetilde{b}_h$ does not need to be symmetric about zero when $\widehat{b}_h$ is, in practice this does not represent a serious problem from the point of view of approximating $\widehat{b}_h$ with $\widetilde{b}_h$.

Remark A.1 For any h and Γ and when $\boldsymbol{z} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_N)$, whether the density of $\widetilde{b}_h$ is symmetric or not can be established by looking at the eigenvalues of $\widetilde{\boldsymbol{A}}_h$. Indeed, since symmetry requires all odd cumulants to be zero, it follows that $\widetilde{b}_h$ has symmetric density if and only if the spectrum of $\widetilde{\boldsymbol{A}}_h$ is symmetric about zero (in the sense that if λ is an eigenvalue of $\widetilde{\boldsymbol{A}}_h$, then $-\lambda$ is an eigenvalue too, with same multiplicity).

Appendix B Formulae for $\eta_{2,h}$

This appendix gives formulae for computing the ratio $\eta_{2,h} = \mathrm{var}(\widehat{b}_h)/\,\mathrm{var}(\widetilde{b}_h)$, considered at the end of Section 4.1.

Theorem B.1 *Let $\zeta_{\boldsymbol{\alpha}}$ be the number of zeros in a sequence $\boldsymbol{\alpha} \in \Gamma_h$. Then,*

$$\eta_{2,h} = \frac{N \sum_{\boldsymbol{\alpha}\in\Gamma_h} 2^{-\zeta_{\boldsymbol{\alpha}}}}{\sum_{\boldsymbol{\alpha}\in\Gamma_h} \left\{ 2^{-\zeta_{\boldsymbol{\alpha}}} \prod_{i=1}^{d} (n_i - \alpha(i)) \right\}}. \tag{28}$$

In some cases (28) simplifies considerably. Three such cases are considered in the three corollaries below. Before stating the first corollary, some notions concerning the structure of the set Γ_h are needed. Let $\sigma\boldsymbol{\alpha}$ denote the action of a permutation $\sigma \in S_d$ on an element $\boldsymbol{\alpha} \in \Gamma$, S_d being the symmetric group on d objects. Observe that for some triplet $h \in \mathbb{N}$, $\sigma \in S_d$, $\boldsymbol{\alpha} \in \Gamma_h$, it holds that $\sigma\boldsymbol{\alpha}$ is in Γ_h if it is in Γ. Now, a necessary and sufficient condition—to be denoted by $\mathcal{C}$—for $\sigma\boldsymbol{\alpha}$ to be in Γ for any $\sigma \in S_d$ and any $\boldsymbol{\alpha} \in \Gamma_h$ is that $\alpha(i) < n_{\min}$, for any $\boldsymbol{\alpha} \in \Gamma_h$ and any $i = 1, ..., d$, where $n_{\min} = \min\{n_1, ..., n_d\}$. Note that $\mathcal{C}$ is trivially satisfied when $n_1 = n_2 = ... = n_d$, but not necessarily otherwise. Under $\mathcal{C}$, Γ_h is the union of one or more orbits in Γ under the action of S_d. A set of orbit representatives is provided by the set of *non-decreasing* sequences $\omega \in \Gamma(n_{\min}, ..., n_{\min})$, with $\omega(1) \leq \omega(2) \leq ... \leq \omega(d)$. For simplicity, we denote such a set by Ω, without explicit reference to the dependence on d and $n_{\min}$. The set $\Omega_h = \Gamma_h \cap \Omega$ plays a central role in determining the structure of spatial design matrices (see HM, Theorem 3). In particular, $|\Omega_h|$ equals the number of orbits (of the action of S_d on Γ) in Γ_h. As an example, consider $h = 25$ on a 2-dimensional grid. Under $\mathcal{C}$, i.e., when $n_1, n_2 > 5$, $\Omega_h = \{(0,5),(3,4)\}$ and Γ_h is made of two orbits. More generally, when $\mathcal{C}$ does not necessarily hold, Γ_h is the union of one or more subsets of orbits in Γ under the action of S_d. For example, when $n_1 = 5$ and $n_2 > 5$, $\Gamma_h = \{(0,5),(3,4),(4,3)\}$.

The first corollary of Theorem B.1 is concerned with the case when the sides of the grid are of equal length, and S_d acts transitively on Γ_h, i.e., Ω_h has a single element.

Corollary B.1 *If $n_1 = ... = n_d = n$ and $\Omega_h = \{\boldsymbol{\omega}_h\}$, then*

$$\eta_{2,h} = \frac{n^{d-\zeta_{\boldsymbol{\omega}_h}}}{\prod_{i=1;\omega_h(i)\neq 0}^{d} (n - \omega_h(i))}.$$

Consider now $h = 1, 2, 3$. These are the only distances such that the action of S_d on Γ_h is transitive in any dimension d.[12]

Corollary B.2 *If $n_1 = ... = n_d = n$ and $h = 1, 2, 3$, then*

$$\eta_{2,h} = \left(\frac{n}{n-1}\right)^h.$$

Another case in which $\eta_{2,h}$ takes a simple form is when only "non-diagonal directions" are considered, i.e., when Γ_h has—or is restricted to have—only sequences lying on the d main axes of Γ (see footnote 3 and Gorsich et al., 2002).

Corollary B.3 *If only non-diagonal directions are considered, then*

$$\eta_{2,h} = \frac{d}{d - \sqrt{h}\sum_{i=1}^{d} \frac{1}{n_i}}. \tag{29}$$

Appendix C Evaluation of $N_{h,k}$

This appendix discusses computation of the term $N_{h,k}$ required for (22). From the definition $N_{h,k} = \boldsymbol{\iota}_N' \boldsymbol{A}_h \boldsymbol{A}_k \boldsymbol{\iota}_N$, we have

$$N_{h,k} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\Gamma} (\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\gamma}} (\boldsymbol{A}_k)_{\boldsymbol{\gamma},\boldsymbol{\beta}} \tag{30}$$

[12] Recall that we are assuming $N_h > 0$ and that Condition A holds. When $h = 1, 2, 3$, such assumptions require $d \geq h$ and $n > 2$.

(that is, $N_{h,k}$ is the number of triangles $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$ on Γ such that $\|\boldsymbol{\alpha}-\boldsymbol{\gamma}\|^2 = h$ and $\|\boldsymbol{\beta}-\boldsymbol{\gamma}\|^2 = k$). It follows immediately that

$$N_{h,k} = \operatorname{tr}(\boldsymbol{D}_h\boldsymbol{D}_k). \tag{31}$$

For $r = 0, ..., n-1$, let $\boldsymbol{M}_r^{(n)}$ be the diagonal matrix with (i,i) entry the i-th row sum of $\boldsymbol{F}_r^{(n)}$, and let $\boldsymbol{C}_{n_i}(t) = \sum_{r=0}^{n-1} t^{r^2}\boldsymbol{M}_r^{(n_i)}$. Also, using Wilf (1994) notation, let $[t^h]$ denote the operator extracting the coefficient of t^h from the expansion in powers of t of the function which follow. Then, by a trivial extension of equation (31) in HM, (31) yields

$$\begin{aligned} N_{h,k} &= [t^h][s^k]\prod_{i=1}^{d} \operatorname{tr}\left(\boldsymbol{C}_{n_i}(t)\boldsymbol{C}_{n_i}(s)\right) \\ &= [t^h][s^k]\prod_{i=1}^{d}\sum_{p_1,p_2=0}^{n_i-1} t^{p_1^2}s^{p_2^2}\operatorname{tr}(\boldsymbol{M}_{p_1}^{(n_i)}\boldsymbol{M}_{p_2}^{(n_i)}). \end{aligned} \tag{32}$$

The implementation of expression (32) in a symbolic package is straightforward, because all terms $\operatorname{tr}(\boldsymbol{M}_{p_1}^{(n_i)}\boldsymbol{M}_{p_2}^{(n_i)})$ are simple functions of p_1 and p_2; see equation (29) in HM. Nevertheless, the computational effort can be substantial when the grid is large, or when $N_{h,k}$ is required for many values of h and k. These are precisely the cases when it is convenient to approximate expression (22) with its circular counterpart (23).

Appendix D Proofs

We first give an auxiliary lemma, whose proof can be found in, e.g., Anderson (1971), p. 304, and then we prove the results given in the main text of the paper and in Appendices A and B.

Lemma 2 *Assume that $\boldsymbol{y} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$. For any idempotent $N\times N$ matrix $\boldsymbol{Q}$, and any symmetric $N \times N$ matrices $\boldsymbol{R}_i$, $i = 1, ..., n$,*

$$\mathrm{E}\left[\frac{\prod_{i=1}^n \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{R}_i\boldsymbol{Q}\boldsymbol{y}}{(\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y})^n}\right] = \frac{\mathrm{E}\left(\prod_{i=1}^n \boldsymbol{y}'\boldsymbol{Q}\boldsymbol{R}_i\boldsymbol{Q}\boldsymbol{y}\right)}{\mathrm{E}\left[(\boldsymbol{y}'\boldsymbol{Q}\boldsymbol{y})^n\right]}. \tag{33}$$

Proof of Theorem 1 Note that $(\widetilde{\boldsymbol{F}}_{\boldsymbol{\gamma}}^{\otimes})_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 1$ only if $|\alpha(i) - \beta(i)|$ equals either $\gamma(i)$ or $n - \gamma(i)$ for each $i = 1, ..., d$. Hence, $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 1$ if and only if $\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \{\boldsymbol{\pi}_D - \boldsymbol{\gamma},\ \boldsymbol{\gamma} \in \Gamma_h,\ D \in \Delta\}$. That is, there must exist $D \in \Delta$ such that $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_D\|^2 = h$. ∎

Proof of Corollary 1 We first prove the necessity and sufficiency of the condition for $\widetilde{\boldsymbol{A}}_h = \boldsymbol{A}_h$. If there exists an $\boldsymbol{\alpha} \in \Gamma_h$ with a component $\alpha(i)$ different from 0 or $n_i/2$, then, by Theorem 1, there exists at least one pair $(\boldsymbol{\alpha},\boldsymbol{\beta})$ such that $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} \neq (\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. This establishes the necessity of the condition. Next, observe that the condition in the corollary implies that there is no pair $(\boldsymbol{\alpha},\boldsymbol{\beta}) \in \Gamma\times\Gamma$ such that $\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|^2 \neq h$ and $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_D\|^2 = h$ for a $D \in \Delta$ other that $\{0, ..., 0\}$. The sufficiency of the condition follows, again by Theorem 1. We now move to the second part of the corollary. Suppose that $\epsilon_{\alpha,\beta}(i) < n_i - \sqrt{h}$, $i = 1, ..., d$. Then $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_D\|^2 = \sum_{i=1}^{n}(\epsilon_{\boldsymbol{\alpha},\boldsymbol{\beta}}(i) - \pi_D(i))^2$ is greater than h for all $D \in \Delta$ other than $D = \{0, ..., 0\}$. Suppose also that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not h-neighbors, i.e., $(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 0$. Then $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_D\|^2$ cannot be equal to h when $D = \{0, ..., 0\}$. It follows that $(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is also zero. The proof is completed, because if $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are h-neighbors, then $(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = (\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}} = 1$. ∎

Proof of Theorem 2 Using (17), we obtain

$$\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k = \sum_{\boldsymbol{\alpha}\in\Gamma_h}\sum_{\boldsymbol{\beta}\in\Gamma_k}\left(\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\beta}}\right) = \sum_{\boldsymbol{\beta}\in\Gamma_k}\sum_{\boldsymbol{\alpha}\in\Gamma_h}\left(\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\beta}}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}\right) = \widetilde{\boldsymbol{A}}_k\widetilde{\boldsymbol{A}}_h,$$

which is the desired result. ∎

Proof of Theorem 3 Let $x_1, ..., x_n$ be the common eigenvectors of each $\widetilde{\boldsymbol{F}}_r$. Then, for each $\boldsymbol{\beta} \in \Gamma$,

$$\begin{aligned}\widetilde{\boldsymbol{A}}_h\bigotimes_{i=1}^{d}x_{\beta(i)} &= \sum_{\boldsymbol{\alpha}\in\Gamma_h}\widetilde{\boldsymbol{F}}^{\otimes}_{\boldsymbol{\alpha}}\bigotimes_{i=1}^{d}x_{\beta(i)} = \sum_{\boldsymbol{\alpha}\in\Gamma_h}\bigotimes_{i=1}^{d}\widetilde{\boldsymbol{F}}_{\alpha(i)}x_{\beta(i)}\\ &= \left(\sum_{\boldsymbol{\alpha}\in\Gamma_h}\prod_{i=1}^{d}\lambda_{\alpha(i),\beta(i)}\right)\bigotimes_{i=1}^{d}x_{\beta(i)},\end{aligned}$$

so that the $\lambda^{\otimes}_{\Gamma_h}(\boldsymbol{\beta})$, $\boldsymbol{\beta} \in \Gamma$, are eigenvalues of $\widetilde{\boldsymbol{A}}_h$. ∎

Proof of Proposition 1 Parts (i) and (ii) follow immediately from Lemma 3 of HM. As for part (iii), the symmetry of the density of $\widehat{b}_h$ when $d = 1$ or when h is odd is a consequence of Lemmata 2 and 4 in HM. It only remains to prove that the density of $\widehat{b}_h$ is symmetric, for any $h \geq 1$ and when $d = 2$. A necessary and sufficient condition for symmetry of $\widehat{b}_h$ is that all its odd-cumulants are zero, or equivalently, $\mathrm{tr}(\boldsymbol{A}_h^p) = 0$ for any odd p. For an arbitrary $\boldsymbol{\alpha} \in \Gamma$,

$$(\boldsymbol{A}_h^p)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_{p-1}\in\Gamma}\{(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}_1}(\boldsymbol{A}_h)_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2}...(\boldsymbol{A}_h)_{\boldsymbol{\beta}_{p-2},\boldsymbol{\beta}_{p-1}}(\boldsymbol{A}_h)_{\boldsymbol{\beta}_{p-1},\boldsymbol{\alpha}}\}. \tag{34}$$

This is nonzero if and only if there is at least one $(p-1)$-tuple $(\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p-1}) \in \Gamma$ such that

$$\|\boldsymbol{\alpha} - \boldsymbol{\beta}_1\|^2 = \|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\|^2 =, ..., = \left\|\boldsymbol{\beta}_{p-2} - \boldsymbol{\beta}_{p-1}\right\|^2 = \left\|\boldsymbol{\beta}_{p-1} - \boldsymbol{\alpha}\right\|^2 = h. \tag{35}$$

Hence, $\mathrm{tr}(\boldsymbol{A}_h^p) \neq 0$ only if there is a sequence $(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p-1}, \boldsymbol{\alpha})$ of elements of Γ such that all consecutive elements of the sequence are at squared Euclidean distance h. We call a such sequence a cycle, more precisely an odd-cycle when p is odd, and we refer to h as the step of the cycle. To establish that $\mathrm{tr}(\boldsymbol{A}_h^p) = 0$, for $d = 2$, any $h \geq 1$, and any odd p, we then need to show that a 2-dimensional grid does not contain any odd-cycle. When h is odd, this is guaranteed by Lemma 4 of HM. For the case when h is even, suppose that one or more odd-cycles of even step exist and let h^* denote the minimum step of such cycles. Define Γ_1 and Γ_2 as the subsets of the 2-dimensional square grid $\Gamma(n,n)$ such that a sequence $\boldsymbol{\alpha} \in \Gamma(n,n)$ belongs to Γ_1 if $\alpha(1)+\alpha(2)$ is even, to Γ_2 if $\alpha(1)+\alpha(2)$ is odd. Observe that a cycle of even step belongs to either Γ_1 or Γ_2. Thus, an odd-cycle of step h^* on $\Gamma(n,n)$ is also an odd-cycle on either Γ_1 or Γ_2. But, after an obvious rescaling, Γ_1 and Γ_2 are themselves square uniform grids, leading to the contradiction that h^* cannot be the minimum even step of an odd-cycle on a 2-dimensional uniform grid. This completes the proof. ∎

Proof of Proposition 2 *(i)* For any $h > 0$ such that $N_h > 0$, and hence $\widetilde{N}_h > 0$, $\mathrm{E}(\widetilde{b}_h) = \mathrm{tr}(\widetilde{\boldsymbol{A}}_h)/(2\widetilde{N}_h)$. This is zero because no $\boldsymbol{\alpha} \in \Gamma(n_1, ..., n_d)$ can be an h-neighbor of itself when $h > 0$, and therefore $\mathrm{tr}(\widetilde{\boldsymbol{A}}_h) = \sum_{\boldsymbol{\alpha}\in\Gamma}(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = 0$. The variance is

$$\mathrm{var}(\widetilde{b}_h) = \frac{1}{(2\widetilde{N}_h)^2} 2\,\mathrm{tr}(\widetilde{\boldsymbol{A}}_h^2) = \frac{2}{\widetilde{N}_h^2}\widetilde{m}_h N = \frac{1}{\widetilde{N}_h}.$$

(ii) When $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, $\mathrm{cov}(\widetilde{b}_h, \widetilde{b}_k) = 2\,\mathrm{tr}(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k)$. Now, for any $\boldsymbol{\alpha} \in \Gamma$, $(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}\in\Gamma}(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\widetilde{\boldsymbol{A}}_k)_{\boldsymbol{\beta},\boldsymbol{\alpha}}$, and hence $\mathrm{cov}(\widetilde{b}_h, \widetilde{b}_k) = 0$ unless there exists at least one pair $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \Gamma \times \Gamma$ and two (different) subsets D_1 and D_2 of Δ such that $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_{D_1}\|^2 = h$ and $\|\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} - \boldsymbol{\pi}_{D_2}\|^2 = k$. That is, $\mathrm{cov}(\widetilde{b}_h, \widetilde{b}_k) \neq 0$ if and only if there exists a sequence $\boldsymbol{\epsilon}_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \Gamma$ at distance h from one vertex of $\Gamma\left(n_1+1, ..., n_d+1\right)$ and at distance k from another vertex of $\Gamma\left(n_1+1, ..., n_d+1\right)$. It follows that a necessary condition for $\mathrm{cov}(\widetilde{b}_h, \widetilde{b}_k) \neq 0$ is that there exist $\boldsymbol{\alpha} \in \Gamma_h$ and $\boldsymbol{\beta} \in \Gamma_k$ such that, for at least one $i = 1, ..., d$, $\alpha\,(i) + \beta\,(i) = n_i$. As a consequence, $\mathrm{cov}(\widetilde{b}_h, \widetilde{b}_k) = 0$ if no $\boldsymbol{\alpha} \in \Gamma_h \cup \Gamma_k$ contains an element $\alpha\,(i) \geq n_i/2$, which completes the proof. ∎

Proof of Proposition 3 We start from the non-circular case. Let $u_h = \boldsymbol{z}'\boldsymbol{M}\boldsymbol{A}_h\boldsymbol{M}\boldsymbol{z}$. From (19), for any $h > 0$,

$$\begin{aligned}\mathrm{E}(u_h) &= \mathrm{tr}(\boldsymbol{M}\boldsymbol{A}_h\boldsymbol{M}) = \mathrm{tr}\left[\left(\boldsymbol{I}_N - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\right)\boldsymbol{A}_h\right] \\ &= -\frac{1}{N}\,\mathrm{tr}(\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{\iota}_N) = -\frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{\iota}_N = -2\frac{N_h}{N} \qquad (36)\end{aligned}$$

and hence, $\mathrm{E}(\widehat{c}_h) = \mathrm{E}(u_h)/2N_h = -1/N$. Observe that $\mathrm{E}(\widehat{c}_0) = \mathrm{tr}(\boldsymbol{z}'\boldsymbol{M}\boldsymbol{z})/N = (N-1)/N$. It follows from Lemma 2 that, for any $h > 0$ such that $N_h > 0$,

$$\mathrm{E}(r_h) = \frac{\mathrm{E}(\widehat{c}_h)}{\mathrm{E}(\widehat{c}_0)} = -\frac{1}{N-1}.$$

Next, using again Lemma 2, write

$$\mathrm{E}\left(r_h r_k\right) = \frac{N^2}{4N_hN_k}\frac{\mathrm{E}\left(u_hu_k\right)}{\mathrm{E}\left(u_0^2\right)} = \frac{N^2}{4N_hN_k}\frac{\mathrm{cov}\left(u_h,u_k\right) + \mathrm{E}\left(u_h\right)\mathrm{E}(u_k)}{\mathrm{E}\left(u_0^2\right)}. \qquad (37)$$

Observe that

$$\begin{aligned}\mathrm{E}\left(u_0^2\right) &= \mathrm{var}(\boldsymbol{z}'\boldsymbol{M}\boldsymbol{z}) + \left[\mathrm{E}(\boldsymbol{z}'\boldsymbol{M}\boldsymbol{z})\right]^2 \\ &= 2(N-1) + (N-1)^2 = N^2 - 1,\end{aligned}$$

and, from (20) and for $h, k > 0$,

$$\begin{aligned}\operatorname{cov}(u_h, u_k) &= 2\operatorname{tr}(\boldsymbol{M}\boldsymbol{A}_h\boldsymbol{M}\boldsymbol{M}\boldsymbol{A}_k\boldsymbol{M}) = 2\operatorname{tr}(\boldsymbol{M}\boldsymbol{A}_h\boldsymbol{M}\boldsymbol{A}_k)\\ &= 2\operatorname{tr}\left(\boldsymbol{A}_h\boldsymbol{A}_k - \frac{1}{N}\boldsymbol{A}_h\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\boldsymbol{A}_k - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{A}_k + \frac{1}{N^2}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\boldsymbol{A}_k\right)\\ &= 2\left[\operatorname{tr}(\boldsymbol{A}_h\boldsymbol{A}_k) - \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{A}_k\boldsymbol{A}_h\boldsymbol{\iota}_N - \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{A}_k\boldsymbol{\iota}_N + \frac{1}{N^2}\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\boldsymbol{A}_k\boldsymbol{\iota}_N\right].\end{aligned}$$

If $h = k$, then $\operatorname{tr}(\boldsymbol{A}_h\boldsymbol{A}_k) = \operatorname{tr}(\boldsymbol{A}_h^2) = 2N_h$. Conversely, if $h \neq k$, then $\operatorname{tr}(\boldsymbol{A}_h\boldsymbol{A}_k) = \sum_{\boldsymbol{\alpha}\in\Gamma}\sum_{\boldsymbol{\beta}\in\Gamma}(\boldsymbol{A}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{A}_k)_{\boldsymbol{\beta},\boldsymbol{\alpha}} = 0$, because no sequences $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be both h- and k-neighbors. Also, note that $\boldsymbol{\iota}_N'\boldsymbol{A}_k\boldsymbol{A}_h\boldsymbol{\iota}_N = \boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{A}_k\boldsymbol{\iota}_N = N_{h,k}$, and that $\boldsymbol{\iota}_N'\boldsymbol{A}_h\boldsymbol{\iota}_N = 2N_h$. It follows that

$$\operatorname{cov}(u_h, u_k) = 4\left(N_h 1_{h=k} - \frac{1}{N}N_{h,k} + 2\frac{1}{N^2}N_hN_k\right).$$

Thus, from (37),

$$\mathrm{E}\left(r_hr_k\right) = \frac{N^2}{\left(N^2-1\right)N_hN_k}\left(N_h 1_{h=k} - \frac{1}{N}N_{h,k} + 3\frac{1}{N^2}N_hN_k\right).$$

From the above calculations we obtain

$$\begin{aligned}\operatorname{cov}(r_h, r_k) &= \mathrm{E}(r_hr_k) - \mathrm{E}(r_h)\,\mathrm{E}(r_k)\\ &= \frac{N^2}{\left(N^2-1\right)N_h}1_{h=k} - \frac{N}{N^2-1}\frac{N_{h,k}}{N_hN_k} + \frac{3}{N^2-1} - \frac{1}{(N-1)^2}\\ &= \frac{1}{N^2-1}\left(\frac{N^2}{N_h}1_{h=k} - \frac{N_{h,k}}{N_hN_k}N + 2\frac{N-2}{N-1}\right).\end{aligned}$$

Turning to the circular case, let $\widetilde{u}_h = \boldsymbol{z}'\boldsymbol{M}\widetilde{\boldsymbol{A}}_h\boldsymbol{M}\boldsymbol{z}$. By obvious modification of (36), we find $\mathrm{E}(\widetilde{u}_h) = -2\widetilde{N}_h/N$. Thus, for any $h > 0$ such that N_h, and hence $\widetilde{N}_h$, is positive, $\mathrm{E}(\widetilde{r}_h) = \mathrm{E}(r_h)$. To obtain $\operatorname{cov}(\widetilde{r}_h, \widetilde{r}_k)$, we follow the same steps as for $\operatorname{cov}(r_h, r_k)$. In particular, for $h, k > 0$,

$$\begin{aligned}\operatorname{cov}(\widetilde{u}_h, \widetilde{u}_k) &= 2\operatorname{tr}\left(\boldsymbol{M}\widetilde{\boldsymbol{A}}_h\boldsymbol{M}\boldsymbol{M}\widetilde{\boldsymbol{A}}_k\boldsymbol{M}\right) = 2\operatorname{tr}\left[\left(\boldsymbol{I}_N - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'\right)\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k\right]\\ &= 2\operatorname{tr}\left(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k - \frac{1}{N}\boldsymbol{\iota}_N'\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k\boldsymbol{\iota}_N\right) = 2\left[\operatorname{tr}(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k) - 4\frac{1}{N^2}\widetilde{N}_h\widetilde{N}_k\right]\\ &= 2\widetilde{m}_h(N1_{h=k} - \widetilde{m}_k),\end{aligned}$$

where the first line of the display uses commutativity of $\boldsymbol{M}$ and $\widetilde{\boldsymbol{A}}_h$, and the last line follows by observing that, when Condition A is satisfied for both h and k, $\operatorname{tr}(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k) = \operatorname{tr}(\widetilde{\boldsymbol{A}}_h^2) = 2\widetilde{N}_h$ if $h = k$, $\operatorname{tr}(\widetilde{\boldsymbol{A}}_h\widetilde{\boldsymbol{A}}_k) = 0$ otherwise. Given $\operatorname{cov}(\widetilde{u}_h, \widetilde{u}_k)$, we can compute $\mathrm{E}\left(\widetilde{r}_h\widetilde{r}_k\right)$ by the obvious analog of (37), which gives

$$\mathrm{E}\left(\widetilde{r}_h\widetilde{r}_k\right) = \frac{1}{N^2-1}\left(\frac{2N}{\widetilde{m}_h}1_{h=k} - \frac{1}{N^2-1}\right).$$

Thus,

$$\begin{aligned}
\operatorname{cov}(\widetilde{r}_h, \widetilde{r}_k) &= \mathrm{E}\left(\widetilde{r}_h \widetilde{r}_k\right) - \mathrm{E}\left(\widetilde{r}_h\right) \mathrm{E}\left(\widetilde{r}_k\right) \\
&= \frac{1}{N^2-1}\left(\frac{2N}{\widetilde{m}_h} 1_{h=k} - \frac{1}{N^2-1}\right) - \frac{1}{(N-1)^2} \\
&= \frac{2N}{N^2-1}\left(\frac{1}{\widetilde{m}_h} 1_{h=k} - \frac{1}{N-1}\right),
\end{aligned}$$

which completes the proof of the proposition. ∎

Proof of Proposition 4 For any $h \geq 1$ such that N_h, and hence $\widetilde{N}_h$, is positive, equation (19) gives

$$\mathrm{E}(\widetilde{\gamma}_h) = \frac{1}{2\widetilde{N}_h} \operatorname{tr}(\widetilde{\boldsymbol{L}}_h \widetilde{\boldsymbol{\Sigma}}) = \frac{1}{2\widetilde{N}_h} \operatorname{tr}\left[\left(\widetilde{m}_h \boldsymbol{I}_N - \widetilde{\boldsymbol{A}}_h\right)\left(\sigma^2 \boldsymbol{I}_N + \sum_{j \in J\setminus\{0\}} c(j) \widetilde{\boldsymbol{A}}_j\right)\right].$$

Since, for any $h > 0$, $\operatorname{tr}(\widetilde{\boldsymbol{A}}_h) = 0$ it follows that

$$\mathrm{E}(\widetilde{\gamma}_h) = \frac{1}{2\widetilde{N}_h}\left\{\sigma^2 \widetilde{m}_h N - \sum_{j \in J\setminus\{0\}} c(j) \operatorname{tr}(\widetilde{\boldsymbol{A}}_h \widetilde{\boldsymbol{A}}_j)\right\}.$$

Recall now that $\widetilde{m}_h N = 2\widetilde{N}_h$, and note that, when h and j satisfy Condition A, $\operatorname{tr}(\widetilde{\boldsymbol{A}}_h \widetilde{\boldsymbol{A}}_j) = \operatorname{tr}(\widetilde{\boldsymbol{A}}_h^2) = 2\widetilde{N}_h$ if $h = j$, and $\operatorname{tr}(\widetilde{\boldsymbol{A}}_h \widetilde{\boldsymbol{A}}_j) = 0$ otherwise. It follows that $\mathrm{E}(\widetilde{\gamma}_h) = c(0) - c(h) 1_{h\in J} = \gamma_h$. ∎

Proof of Proposition 5 When $\boldsymbol{z} \sim \mathrm{N}(\mu \boldsymbol{\iota}_N, \widetilde{\boldsymbol{\Sigma}})$, Lemma 1 yields

$$\begin{aligned}
\operatorname{cov}(\widetilde{\gamma}_h, \widetilde{\gamma}_k) &= \frac{1}{2\widetilde{N}_h 2\widetilde{N}_k} 2 \operatorname{tr}(\widetilde{\boldsymbol{L}}_h \widetilde{\boldsymbol{\Sigma}} \widetilde{\boldsymbol{L}}_k \widetilde{\boldsymbol{\Sigma}}) \\
&= \frac{1}{2\widetilde{N}_h \widetilde{N}_k} \sum_{j,l \in J} c(j) c(l) \operatorname{tr}(\widetilde{\boldsymbol{L}}_h \widetilde{\boldsymbol{A}}_j \widetilde{\boldsymbol{L}}_k \widetilde{\boldsymbol{A}}_l).
\end{aligned}$$

The proposition follows by using the expression $\widetilde{\boldsymbol{L}}_h = \widetilde{m}_h \boldsymbol{I}_N - \widetilde{\boldsymbol{A}}_h$, Theorem 2, and the fact that the eigenvalue $\lambda^{\otimes}_{\Gamma_j}(\boldsymbol{\beta})$ of $\widetilde{\boldsymbol{A}}_j$ and the eigenvalue $\lambda^{\otimes}_{\Gamma_l}(\boldsymbol{\beta})$ of $\widetilde{\boldsymbol{A}}_l$ are associated to the same eigenvector, for any fixed $\boldsymbol{\beta} \in \Gamma$. ∎

Proof of Proposition A.1 The density of $\widetilde{b}_{r,1}$ is symmetric about zero if and only if all its odd cumulants vanish, that is, if and only if

$$\operatorname{tr}\left[\left(\widetilde{\boldsymbol{F}}_r^{(n_1)}\right)^p\right] = 0$$

for all odd p. Note that the diagonal entries of $\left(\widetilde{\boldsymbol{F}}_r^{(n_1)}\right)^p$ are all the same, with the (i,i)-th diagonal entry being

$$\left(\left(\widetilde{\boldsymbol{F}}_r^{(n_1)}\right)^p\right)_{i,i} = \sum_{l_1,\ldots,l_{p-1}=1,\ldots,n_1} (\widetilde{\boldsymbol{F}}_r^{(n_1)})_{i,l_1} (\widetilde{\boldsymbol{F}}_r^{(n_1)})_{l_1,l_2} \cdots (\widetilde{\boldsymbol{F}}_r^{(n_1)})_{l_{p-2},l_{p-1}} (\widetilde{\boldsymbol{F}}_r^{(n_1)})_{l_{p-1},i},$$

which, by expression (10), is nonzero if and only if there exists one $(p-1)$-tuple $(l_1, ..., l_{p-1})$ such that each of the absolute values $|i - l_1|, |l_1 - l_2|, ..., |l_{p-2}, l_{p-1}|, |l_{p-1}, i|$ equals r or $n_1 - r$. This requires that $kn_1 = rp$ for some $k = 1, 2, ...$ Hence, the density of $\widetilde{b}_{r,1}$ is symmetric about zero if and only if there is no odd p such that $kn_1 = rp$. Such a condition is satisfied if either n_1 is even and r is odd, or n_1 is an even multiple of r, and is not satisfied in all other possible cases. ∎

Proof of Proposition A.2 By Lemma 1, when $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_N)$, $\kappa_p(\widetilde{b}_h)$ is equal to $(2N_h)^{-1}2^{p-1}(p-1)!\,\mathrm{tr}(\boldsymbol{A}_h^p)$ and $\kappa_p(\widehat{b}_h)$ to $(2\widetilde{N}_h)^{-1}2^{p-1}(p-1)!\,\mathrm{tr}(\widetilde{\boldsymbol{A}}_h^p)$. The first part of the proposition follows trivially from noting that $\widetilde{\boldsymbol{A}}_h$ and $\boldsymbol{A}_h$ are nonnegative matrices, and $\widetilde{\boldsymbol{A}}_h = \boldsymbol{A}_h + \boldsymbol{R}_h$, for some nonnegative matrix $\boldsymbol{R}_h$. The case $p = 1$ is also straightforward, because $\mathrm{E}(\widetilde{b}_h) = \mathrm{E}(\widehat{b}_h) = 0$. Next, observe that

$$(\widetilde{\boldsymbol{A}}_h^p)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_{p-1}\in\Gamma} \{(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\alpha},\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2}...(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\beta}_{p-2},\boldsymbol{\beta}_{p-1}}(\widetilde{\boldsymbol{A}}_h)_{\boldsymbol{\beta}_{p-1},\boldsymbol{\alpha}}\}. \tag{38}$$

Thus, in view of Theorem 1, $(\widetilde{\boldsymbol{A}}_h^p)_{\boldsymbol{\alpha},\boldsymbol{\alpha}} \neq 0$ if and only if there is at least one $(p-1)$-tuple $\left(\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p-1}\right)$ such that

$$\exists D_j \in \Delta : \left\|\boldsymbol{\epsilon}_j - \boldsymbol{\pi}_{D_j}\right\|^2 = h,\ j = 1, ..., p, \tag{39}$$

where the sequences $\boldsymbol{\epsilon}_j \in \Gamma$ are defined by $\epsilon_j(i) = |\beta_{j-1}(i) - \beta_j(i)|$, $i = 1, ..., d$, with $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_p = \boldsymbol{\alpha}$. Call a distance h feasible if Γ contains at least one pair of h-neighbors. For even p, $\kappa_p(\widetilde{b}_h) = 0$ is necessary and sufficient for $\kappa_p(\widehat{b}_h) = 0$ because: (i) if h is not feasible then both $(\boldsymbol{A}_h^p)_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$ and $(\widetilde{\boldsymbol{A}}_h^p)_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$ are zero, for any $\boldsymbol{\alpha} \in \Gamma$, by definition; (ii) if h is feasible then both $\kappa_p(\widetilde{b}_h)$ and $\kappa_p(\widehat{b}_h)$ are positive, because for any pair of h-neighbors $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, the $(p-1)$-tuple $(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, ..., \boldsymbol{\alpha}, \boldsymbol{\beta})$ satisfies both expression (39) and expression (35) in the proof of Lemma 1. To prove the part of the Lemma relative to Condition B, we establish that, under Condition B, $\kappa_p(\widetilde{b}_h) \neq 0$ implies $\kappa_p(\widehat{b}_h) \neq 0$. First, observe that $\widetilde{\boldsymbol{A}}_h^p(\boldsymbol{\alpha}, \boldsymbol{\alpha})$ does not depend on $\boldsymbol{\alpha}$, as is easily seen by considering formula (11) plus the fact that the product of any two matrices $\boldsymbol{F}_r^{(n)}$ has constant diagonal. If $\kappa_p(\widetilde{b}_h) \neq 0$, then for each $\boldsymbol{\alpha} \in \Gamma$, there is at least one $(p-1)$-tuple $\left(\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p-1}\right)$ satisfying (39). This clearly implies that, under Condition B, it is always possible to find a p-tuple $\left(\boldsymbol{\alpha}, \boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p-1}\right)$ such that (39) is satisfied with $D_j = \{0, ..., 0\}$, for $j = 1, ..., p$. But this in turn implies that the same p-tuple satisfies (35), and hence that $\kappa_p(\widehat{b}_h) \neq 0$, as was to be shown. To complete the proof, we need to show that when $p \geq 3$ and without Condition B, $\kappa_p(\widehat{b}_h)$ may be zero even if $\kappa_p(\widetilde{b}_h)$ is not. An example suffices. When $d = 2$ and $h = 20$, $\kappa_3(\widehat{b}_{20}) = 0$, since there is no equilateral triangle with vertices on a 2-dimensional planar lattice (e.g., Beeson, 1992), but $\kappa_3(\widetilde{b}_{20}) > 0$ because, for instance, the sequences $\boldsymbol{\beta}_1 = (4, 2)$ and $\boldsymbol{\beta}_2 = (2, 6)$ satisfy (39) when $\boldsymbol{\alpha} = (0, 0)$. ∎

Proof of Theorem B.1 Recall that $\eta_{2,h} = \widetilde{N}_h/N_h = N\widetilde{m}_h/(2N_h)$. Under Condition A, from the generating function in Remark 1 we have $\widetilde{m}_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h} 2^{d-\zeta_{\boldsymbol{\alpha}}}$. The desired result follows by applying equation (25) in HM, which yields

$$2N_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h} \prod_{i=1}^{d} \left\{ \left(2 - \delta_{\alpha(i),0}\right)(n_i - \alpha(i)) \right\} = \sum_{\boldsymbol{\alpha}\in\Gamma_h} \left\{ 2^{d-\zeta_{\boldsymbol{\alpha}}} \prod_{i=1}^{d} (n_i - \alpha(i)) \right\}.$$

Under Condition A, the generating function in Remark 1 yields $\widetilde{m}_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h} 2^{d-\zeta_{\boldsymbol{\alpha}}}$. From equation (25) in HM, we obtain

$$2N_h = \sum_{\boldsymbol{\alpha}\in\Gamma_h} \prod_{i=1}^{d} \left\{ \left(2 - \delta_{\alpha(i),0}\right) \left(n_i - \alpha(i)\right) \right\}$$
$$= \sum_{\boldsymbol{\alpha}\in\Gamma_h} \left\{ 2^{d-\zeta_{\boldsymbol{\alpha}}} \prod_{i=1}^{d} (n_i - \alpha(i)) \right\}.$$

Expression (28) follows on recalling that $\eta_{2,h} = \widetilde{N}_h/N_h = N\widetilde{m}_h/(2N_h)$. ∎

Proof of Corollary B.1 For any $\boldsymbol{\omega} \in \Omega_h$, let $v(\boldsymbol{\omega}) = \prod_{j=0}^{n-1} \zeta_{\boldsymbol{\omega}}(j)!$, where $\zeta_{\boldsymbol{\omega}}(j)$ denotes the multiplicity of j in $\boldsymbol{\omega}$ (so that $\zeta_{\boldsymbol{\omega}}(0) = \zeta_{\boldsymbol{\omega}}$). Observe that the numerator in (28) is $n^d \sum_{\boldsymbol{\omega}\in\Omega_h} \{d!2^{-\zeta_{\boldsymbol{\omega}}}/\nu(\boldsymbol{\omega})\}$, and the denominator is $\sum_{\boldsymbol{\omega}\in\Omega_h} \{d!2^{-\zeta_{\boldsymbol{\omega}}} \prod_{i=1}^{d} (n - \omega(i))/\nu(\boldsymbol{\omega})\}$. The corollary follows straightforwardly. ∎

Proof of Corollary B.2 When $h = 1, 2, 3$ $|\Omega_h| = 1$ for any dimension $d \geq h$, because any decomposition in d squares of h consists of h ones and $d - h$ zeros. The result follows by applying Corollary B.1. ∎

Proof of Corollary B.3 If only non-diagonal directions are considered, $\zeta_{\boldsymbol{\alpha}} = d - 1$, for any $\boldsymbol{\alpha} \in \Gamma_h$. The corollary then follows from expression (28), because under the restriction of only diagonal directions the only nonzero element of any $\boldsymbol{\alpha} \in \Gamma_h$ is $\sqrt{h}$. ∎

References

Anderson, T.W., 1971. The Statistical Analysis of Time Series. Wiley, New York.

Arellano-Valle, R.B., Genton, M.G. 2010. An invariance property of quadratic forms in random vectors with a selection distribution, with application to sample variogram and covariogram estimators, Annals of the Institute of Statistical Mathematics, forthcoming.

Bannai, E., Ito, T., 1984. Algebraic Combinatorics I: Association Schemes. Benjamin/Cummings, London.

Beeson, M.J., 1992. Triangles with vertices on lattice points, American Mathematical Monthly 99, 243-252.

Besag, J.E., Moran, P.A.P., 1975. On the estimation and testing of spatial interaction in Gaussian lattice processes, Biometrika 62, 555-562.

Biggs, N.L., 1993. Algebraic Graph Theory, second edition. Cambridge University Press, Cambridge.

Cressie, N., 1985. Fitting variogram models by weighted least squares, Mathematical. Geology 17, 563-586.

Cressie, N., 1993. Statistics for Spatial Data, revised edition. Wiley, New York.

Dufour, J.-M., Roy, R., 1985. Some robust exact results on sample autocorrelations and tests of randomness, Journal of Econometrics 29, 257-273.

Genton, M.G., 1998. Variogram fitting by generalized least squares using an explicit formula for the covariance structure, Mathematical Geology 30, 323-345.

Genton, M.G., 1999. The correlation structure of the sample autocovariance function for a particular class of time series with elliptically contoured distribution, Statistics and Probability Letters 41, 131-137.

Genton, M.G., He, L., Liu, X., 2001. Moments of skew-normal random vectors and their quadratic forms, Statistics and Probability Letters 51, 319-325.

Geary, R.C., 1954. The contiguity ratio and statistical mapping, The Incorporated Statistician 5, 115-145.

Gorsich, D.J., Genton, M.G., Strang, G., 2002. Eigenstructures of Spatial Design Matrices, Journal of Multivariate Analysis 80, 138-165.

Gray, R.M., 2006. Toeplitz and circulant matrices: a review, Foundations and Trends in Communications and Information Theory 3, 155-239.

Hillier, G.H., Martellosio, F., 2006. Spatial design matrices and associated quadratic forms: structure and properties, Journal of Multivariate Analysis 97, 1-18.

Horn, R., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, Cambridge..

Lahiri, S.N., Lee, Y., Cressie, N., 2001. On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters, Journal of Statistical Planning and Inference 103, 65-85.

Kelejian, H.H., Prucha, I. R., 2001. On the asymptotic distribution of the Moran I test statistic with applications, Journal of Econometrics 104, 219–257.

Kendall, M.G., Stuart, A., 1969. The Advanced Theory of Statistics, Vol. 1, Distribution Theory, Griffin and Co., London.

Kent, J.T., Mardia, K.V., 1996. Spectral and circulant approximations to the likelihood for stationary Gaussian random fields, Journal of Statistical Planning and Inference 50, 379-394.

Martin, R.J., 1986. A note on the asymptotic eigenvalues and eigenvectors of the dispersion matrix of a second-order stationary process on a d-dimensional lattice, Journal of Applied Probability 23, 529-535.

Moran, P.A.P., 1948. Some theorems on time series II. The significance of the serial correlation coefficient, Biometrika 64, 523-529.

Moran, P.A.P., 1950. Notes on continuos stochastic phenomena, Biometrika 37, 17-23.

Oden, N.L., 1984. Assessing the significance of a spatial correlogram, Geographical Analysis 16, 1-16.

Rue, H., Held, L., 2005. Gaussian Markov Random Fields. Chapman & Hall, London.

von Neumann, J., 1941. Distribution of the ratio of the mean square successive difference to the variance, Annals of Mathematical Statistics 12, 367-395.

Wilf, H.S., 1994. generatingfunctionology, second edition, Academic Press Inc.