

# Behavior of the NORTA Method for Correlated Random Vector Generation as the Dimension Increases

SOUMYADIP GHOSH and SHANE G. HENDERSON

Cornell University

---

The NORTA method is a fast general-purpose method for generating samples of a random vector with given marginal distributions and given correlation matrix. It is known that there exist marginal distributions and correlation matrices that the NORTA method cannot match, even though a random vector with the prescribed qualities exists. We investigate this problem as the dimension of the random vector increases. Simulation results show that the problem rapidly becomes acute, in the sense that NORTA fails to work with an increasingly large proportion of correlation matrices. Simulation results also show that if one is willing to settle for a correlation matrix that is “close” to the desired one, then NORTA performs well with increasing dimension. As part of our analysis, we develop a method for sampling correlation matrices *uniformly* (in a certain precise sense) from the set of all such matrices. This procedure can be used more generally for sampling uniformly from the space of all symmetric positive definite matrices with diagonal elements fixed at positive values.

Categories and Subject Descriptors: G.3 [Probability and Statistics]—*Random Number Generation*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: NORTA method, semidefinite programming, sampling random matrices, onion method

---

## 1. INTRODUCTION

In many settings, there is a need to capture dependence between random variables that serve as primitive inputs to stochastic models. In a manufacturing setting, for example, the processing times of a single job at different stations may be correlated due to characteristics of the job such as size. In financial engineering, Das et al. [2001] claim that the risk profile of credit portfolios can be

---

The work of S. Ghosh was partially supported by an IBM Ph.D. Fellowship Award.

This research was partially supported by National Science Foundation (NSF) grant number DMI-0230528.

Authors' address: School of Operations Research and Industrial Engineering, Cornell University, 230 Rhodes Hall, Ithaca, NY 14853; email: sdghosh@orie.cornell.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 1515 Broadway, New York, NY 10036 USA, fax: +1 (212) 869-0481, or [permissions@acm.org](mailto:permissions@acm.org).

© 2003 ACM 1049-3301/03/0700-0001 \$5.00

2 • S. Ghosh and S. G. Henderson

understated if correlation is ignored. Further applications have recently been reported in cost analysis [Lurie and Goldberg 1998], decision and risk analysis [Clemen and Reilly 1999], and generation of test problems [Hill and Reilly 2000]. If simulation is used with such models, then we need methods for efficiently generating samples of correlated random variables. We examine the case where the correlated primitive inputs of a model are finite in number and hence can be characterized jointly as a random vector.

We focus on the NORmal To Anything method of random vector generation described by Cario and Nelson [1997]. The NORTA method belongs to a family of methods that address the problem of generating samples of a finite-dimensional random vector with a given set of distributions for the individual components (their *marginal* distributions) and a given measure of dependence between them, which in the NORTA case is usually the product-moment or the rank correlation matrix.

*Definition 1.1.* The *product-moment* correlation matrix for a random vector  $X = (X_1, \dots, X_d)$  is the correlation matrix (i.e., a symmetric positive semidefinite matrix with unit diagonal elements)  $\Sigma_X = (\Sigma_X(i, j) : 1 \leq i, j \leq d)$  where

$$\Sigma_X(i, j) = \frac{\text{Cov}(X_i, X_j)}{(\text{Var}X_i \text{Var}X_j)^{1/2}}.$$

The *rank* correlation matrix is of the same form except that now

$$\Sigma_X(i, j) = \frac{\text{Cov}(F_i(X_i), F_j(X_j))}{(\text{Var}F_i(X_i) \text{Var}F_j(X_j))^{1/2}},$$

where  $F_i$  and  $F_j$  are the distribution functions of  $X_i$  and  $X_j$  respectively.

*Remark 1.2.* The product-moment correlation is a measure of *linear* correlation between random variables. Rank correlation is often preferred to product-moment correlation as a measure of dependence for two reasons. First, it is always defined, even if the random variables involved have infinite variance. Second, it is invariant with respect to strictly increasing transformations of the random variables involved.

The philosophy of specifying marginals and correlations to model dependent random variables is clearly an approximate one, since the joint distribution is not completely specified by this information. One hopes (but is not guaranteed) to capture the essence of the dependence between the random variables. This approach has the advantage that it is easy to specify the marginal distributions and correlation matrix. An alternative is to fully specify the joint distribution. The primary difficulty in this case is that a tremendous amount of information is typically required to specify (and fit) such a joint distribution. Furthermore, special methods must be devised to generate random vectors with the given joint distribution, and this can be a practically insurmountable problem for a model of even moderate complexity [Law and Kelton 2000, p. 479]. Another alternative is to settle for a parametric family of distributions where the marginal distributions come from a restricted class [Devroye 1986; Johnson 1987].

Another argument in support of modelling random vectors using marginals and correlations relates to the use of diffusion approximations for modelling stochastic systems. In many cases, the limiting diffusions depend only on the first two moments of the input distributions. Therefore, there is some insensitivity in performance measures computed from these models to the exact form of the input distributions. In general, then, if a form of this insensitivity is present in a model, then the approach discussed here for modelling random vectors is quite reasonable.

The NORTA method produces a random vector with the desired properties via a component-wise transformation of a multivariate normal random vector, and capitalizes on the fact that multivariate normal random vectors are easily generated; see, for example, Law and Kelton [2000, p. 480]. Cario and Nelson [1997] traced the roots of the method back to Mardia [1970] who studied bivariate distributions, and to Li and Hammond [1975] who concentrated on the case where all of the marginals have densities (with respect to Lebesgue measure). Iman and Conover [1982] implemented the same transformation procedure to induce a given rank correlation in the output. Their method is only approximate, in that the output has only very approximately the desired rank correlation.

The NORTA method is a very efficient and easy-to-implement sampling method, and has seen use in a variety of contexts. Clemen and Reilly [1999] use the NORTA procedure to induce a desired rank correlation in the context of decision and risk analysis. Lurie and Goldberg [1998] implement a variant of the NORTA method for use in cost analysis. Henderson et al. [2000] adapt the NORTA method to generate samples of dependent quasi-random vectors. The NORTA method is also routinely used in portfolio models in industry.

So the NORTA procedure is often the method of choice for generating random vectors with prescribed marginals and correlation matrix. But can the NORTA procedure be used in all circumstances that require sampling such random vectors? To pose this question more precisely, we need a definition.

*Definition 1.3.* A product-moment (rank) correlation matrix  $\Sigma$  is *feasible* for a given set of marginal distributions  $F_1, \dots, F_d$  if there exists a random vector  $X$  with marginal distributions  $F_1, \dots, F_d$  and product-moment (rank) correlation matrix  $\Sigma$ .

In Ghosh and Henderson [2002a], we examined whether the NORTA procedure can match *any* feasible correlation matrix for a given set of marginals. For 2-dimensional random vectors, the NORTA method can match any feasible correlation matrix. This follows immediately from the characterizations in Whitt [1976]. However, this does not hold for dimensions 3 and greater. Both Li and Hammond [1975] and Lurie and Goldberg [1998] postulate examples of 3-dimensional random vectors where the NORTA procedure might fail, but do not establish that the counterexamples exist, that is, that the example correlation matrices are feasible.

A necessary, but not always sufficient, condition for feasibility is that the given matrix is a *correlation* matrix, that is, a symmetric, positive semidefinite matrix with unit diagonal elements. To see why this is not sufficient, consider a random vector that consists of an exponential random variable with mean 1 and

4 • S. Ghosh and S. G. Henderson

a uniform  $(0,1]$  random variable. The maximum (minimum) product-moment correlation that can be induced between the two random variables is  $\pm\sqrt{3}/2$ , which is strictly less (greater) than  $+1$  ( $-1$ ). (Note that if correlations of  $1$  or  $-1$  were achievable, then one random variable would be a linear function of the other, which is clearly impossible given the distributions specified.) In general, then, the range of correlations that can be achieved between two random variables is a subinterval of  $[-1,1]$ . Hence, for the case of the exponential-uniform pair, any matrix

$$\begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix},$$

where  $\sqrt{3}/2 < |r| \leq 1$ , defines a  $2 \times 2$  correlation matrix, but not a feasible correlation matrix. Another good example can be found in Schmeiser and Lal [1982], where they show that the range of correlation that can be achieved between two gamma variables is often a subset of  $[-1, 1]$  if their parameters differ. Thus, determining the feasibility of a given matrix for a specified set of marginals is not trivial.

In Ghosh and Henderson [2001, 2002a] a computational procedure is developed based on chessboard distributions to determine whether a given correlation matrix is feasible for the marginal distributions or not. Using this procedure, one can rigorously establish that counterexamples like those postulated by Li and Hammond [1975] and Lurie and Goldberg [1998] do exist.

Let us call feasible correlation matrices that cannot be matched using the NORTA method *NORTA defective* matrices. Based on the numerical results obtained in Ghosh and Henderson [2002a], we conjectured that as the dimension grew, more and more feasible correlation matrices would be NORTA defective. This is the first aspect of the problem that we investigate in this article. We estimate, for each dimension, the probability that the NORTA procedure fails to work for a rank correlation matrix chosen *uniformly* from the set of all correlation matrices.

We should note here that NORTA might fail for two reasons. First, a given correlation matrix might not be feasible so that the random vector as specified does not exist, in which case no generation method can achieve the given correlation matrix. Second, a given correlation matrix might be feasible, but might be NORTA defective. In general, there is no clear characterization of the set of feasible correlation matrices for a given set of marginal distributions. Therefore, it is unreasonable to expect a practitioner to specify a *feasible* correlation matrix, but we can reasonably expect the practitioner to specify a correlation matrix. The probability we estimate is over all correlation matrices, and not necessarily *feasible* correlation matrices.

Kurowicka and Cooke [2001] report simulation results on a similar problem, and find that the probability the NORTA procedure fails to work grows rapidly with dimension. They provide these results to help motivate their copula-vine method for modeling random vectors. The probability distribution they used is not uniform over the set of all correlation matrices, so one possible explanation for their results prior to our study is that their probability distribution gives excessive weight to matrices for which NORTA fails. However, our

results using uniform distributions confirm their finding, suggesting that perhaps the NORTA procedure is unlikely to be effective in high-dimensional problems.

Now, suppose that we are willing to accept a random vector with the prescribed marginals, and a correlation matrix that is, at least approximately, the desired correlation matrix. In Ghosh and Henderson [2002a], we describe a semidefinite programming (SDP) approach that can assist in this regard.

The augmented NORTA method proposed in Ghosh and Henderson [2002a] works in exactly the same manner as the original method unless a NORTA defective matrix is encountered. For such a matrix, a semidefinite program is solved, and the results are then used to modify the inputs given to the NORTA generation step in the hope that the generated random vector has a correlation matrix that is “close” to the desired one (it has the same marginal distributions). The numerical results in Ghosh and Henderson [2002a] indicate that this holds for the 3-dimensional case. In this article, we examine higher dimensions, exploring how the augmented NORTA method performs as the dimension increases. Specifically, we estimate the average distance (measured in an appropriate sense) between the correlation matrix given by the augmented NORTA procedure and the desired (possibly NORTA defective) matrix.

The results indicate that NORTA can typically get very close to a target correlation matrix, even in very high dimensions. So in high dimensions, while NORTA is (very) unlikely to be able to exactly match a desired correlation matrix, it is typically able to match a correlation matrix that is very close to the desired one.

An important part of our analysis is the development of a method for sampling uniformly from the set of all correlation matrices of a given dimension. As discussed in the next section, this set can be viewed as a closed, bounded and full-dimensional subset of a real vector space, and we sample uniformly over this bounded set. We choose to call our sampling method the *onion* method for reasons that will be clear once the working of the method is explained. The method is easily generalized to generate from the set of symmetric positive definite matrices with arbitrary (fixed) positive diagonal entries. Thus, a possible use of an appropriately modified version of the sampling scheme might be to study the performance of algorithms that operate on such matrices. In related work, Marsaglia and Olkin [1984] survey methods for sampling random correlation matrices, but none of the methods they mention samples *uniformly* over the set of all correlation matrices (of fixed dimension). Theorems 3.1 and 3.2 of Edelman [1989] show how the distribution of a symmetric positive definite matrix can be expressed as a function of the distributions of the matrices of its eigenvalue decomposition, although we do not use this result in our analysis. The onion method can also be used to sample from bounded nonuniform densities on sets of the kind mentioned above, via standard sampling frameworks like the acceptance-rejection method.

The onion method is simple to implement since it uses nothing more than standard tools from the simulation input modelling toolkit, and sample generation is very fast. Indeed, the most complex and computationally demanding part of the method involves sampling from univariate beta distributions, which

is a very well-studied problem with many efficient algorithms available (see Law and Kelton [2000, p. 467]).

This article, which is an outgrowth of Ghosh and Henderson [2002b], is organized as follows: The next section reviews the NORTA procedure and indicates why some matrices may be NORTA defective. Section 3 discusses our simulation framework for estimating the performance of NORTA as the dimension increases. Section 4 describes the random matrix sampling procedure. In Section 5, we return to the NORTA method, briefly describing the SDP augmentation proposed in Ghosh and Henderson [2002a], and studying how this augmented method performs in higher dimensions.

## 2. THE NORTA PROCEDURE

Suppose that we wish to generate independent and identically distributed replicates of a random vector  $X = (X_1, \dots, X_d)$  with prescribed marginal distributions

$$F_i(\cdot) = P(X_i \leq \cdot), \quad i = 1, \dots, d,$$

and product-moment or rank correlation matrix

$$\Sigma_X = (\Sigma_X(i, j), \quad 1 \leq i, j \leq d).$$

The NORTA method generates independent and identically distributed replicates of  $X$  through the following procedure.

- (1) Generate an  $\mathbb{R}^d$ -valued joint-normal random vector  $Z = (Z_1, \dots, Z_d)$  with mean vector 0 and (product-moment) covariance matrix  $\Sigma_Z = (\Sigma_Z(i, j) : 1 \leq i, j \leq d)$ , where  $\Sigma_Z(i, i) = 1$  for  $i = 1, \dots, d$ . (Since the variances of each component have been chosen to be 1,  $\Sigma_Z$  also represents the product-moment correlation matrix of  $Z$ .)
- (2) Compute the vector  $X = (X_1, \dots, X_d)$  via

$$X_i = F_i^{-1}(\Phi(Z_i)), \quad (1)$$

for  $i = 1, \dots, d$ , where  $\Phi$  is the distribution function of a standard normal random variable, and

$$F_i^{-1}(u) = \inf\{x : F_i(x) \geq u\}. \quad (2)$$

The vector  $X$  generated by this procedure will have the prescribed marginal distributions. To see this, note that each  $Z_i$  has a standard normal distribution, so that  $\Phi(Z_i)$  is uniformly distributed on  $(0, 1)$ , and so  $F_i^{-1}(\Phi(Z_i))$  will have the required marginal distribution.

The covariance matrix  $\Sigma_Z$  should be chosen, in a preprocessing phase, so that it induces the prescribed correlation matrix  $\Sigma_X$  on  $X$ . However, there is no general closed-form expression that gives  $\Sigma_Z$  in terms of  $\Sigma_X$ . Indeed, determining the right  $\Sigma_Z$  is perhaps the most difficult step in implementing the NORTA method.

Each component of  $\Sigma_X$  is known to depend only on the corresponding component of  $\Sigma_Z$ . As in Cario and Nelson [1997], we can define  $c_{ij}(z) = \Sigma_X(i, j)$  to represent the correlation between  $X_i$  and  $X_j$  as a function of the correlation  $z$

between  $Z_i$  and  $Z_j$ , when  $X_i$  and  $X_j$  are generated as in (1). Cario and Nelson [1997] show that under certain very mild conditions  $c_{ij}(\cdot)$  is a non-decreasing, continuous function. This result helps us perform an efficient numerical search for a value  $\Lambda_Z(i, j)$  that solves

$$c_{ij}(\Lambda_Z(i, j)) = \Sigma_X(i, j). \quad (3)$$

Hence, a numerical estimate  $\Lambda_Z$  of  $\Sigma_Z$  can be determined by solving a number of one-dimensional root-finding problems. Unless stated otherwise, we assume that a solution exists for (3) for all  $i$  and  $j$ . This assumption is not unreasonable, since if  $\Sigma_X$  is feasible, then (3) must have a solution for all  $i$  and  $j$ . Henderson et al. [2000] also show that under slightly stronger assumptions,  $\Lambda_Z(i, j)$  in (3) is uniquely determined by  $\Sigma_X(i, j)$ .

The matrix  $\Lambda_Z$  is constructed in a way that does not necessarily ensure that it is positive semidefinite. It might indeed turn out to be indefinite, in which case it cannot be a valid covariance matrix for a joint normal distribution, and NORTA will fail. Li and Hammond [1975] postulated the following example to demonstrate this possibility. Suppose that  $X = (X_1, X_2, X_3)$  is a random vector with uniform  $(0, 1]$  marginals, and correlation matrix

$$\Sigma_X = \begin{pmatrix} 1 & -0.4 & 0.2 \\ -0.4 & 1 & 0.8 \\ 0.2 & 0.8 & 1 \end{pmatrix}.$$

For the special case of uniform marginals, (3) can be solved analytically (e.g., Kruskal [1958]) to give

$$\Lambda_Z(i, j) = 2 \sin \left( \frac{\pi}{6} \Sigma_X(i, j) \right). \quad (4)$$

The unique solution  $\Lambda_Z$  for the given  $\Sigma_X$  turns out to be indefinite.

This example is valid only if such a uniform random vector  $X$  exists. Li and Hammond [1975] did not show this. Joe [1997, p. 137] asserts that for 3-dimensional random vectors with uniform marginals, as in the Li and Hammond example, any correlation matrix is feasible. Hence, the example exists. Furthermore, Ghosh and Henderson [2002a] develop a computational procedure that can determine, for *almost* any (in a Lebesgue measure sense) given correlation matrix, whether it is feasible for a given set of marginal distributions (assumed to be continuous and have bounded support) or not. Applying this procedure to the Li and Hammond example gives a construction of the random vector. Ghosh and Henderson [2002a] generate a number of such feasible matrices for three-dimensional uniform random vectors that are NORTA defective. The numerical results suggest a structure to the failure of NORTA. To explain this observation more carefully, we need some notation.

With an abuse of notation, we can view a  $d \times d$  correlation matrix as an element of a  $d(d-1)/2$  dimensional vector space, since there are  $d(d-1)/2$  independent elements above the diagonal, the matrix is symmetric, and the diagonal elements are equal to 1. Let  $\Omega$  denote the set of correlation matrices, that is, symmetric positive semidefinite matrices with unit diagonal elements. We view this set as a subset of the real vector space  $\mathbb{R}^{d(d-1)/2}$ . This set has been

established to be nonempty, convex, closed and full-dimensional; see Wolkowicz et al. [2000].

Returning to the discussion above, Ghosh and Henderson [2002a] found that in 3 dimensions, NORTA defective matrices tended to occur near the boundary of  $\Omega$ . Moreover, the indefinite correlation matrices  $\Lambda_Z$  determined for the joint normal distribution from (3) seemed to lie close to (but outside of) the set of symmetric positive semidefinite matrices. So NORTA defective matrices tended to occur near the boundary, and they were never too distant from a NORTA feasible matrix.

### 3. NORTA IN HIGHER DIMENSIONS

NORTA appears to fail most often when the correlation matrix is close to the boundary of the set  $\Omega$  for 3-dimensional random vectors. Now, in a sense that can be made precise, “most” points in certain sets in high dimensions lie close to the boundary. For example, consider the interior of the unit hypercube  $[-\frac{1}{2}, \frac{1}{2}]^d$  in  $\mathbb{R}^d$  represented by the hypercube  $[-\frac{1-\epsilon}{2}, \frac{1-\epsilon}{2}]^d$ , where  $\epsilon \in (0, 1)$ . The ratio of the volume of the interior to that of the whole set is  $(1 - \epsilon)^d$ , which decreases rapidly to 0 as  $d$  increases.

This suggests that matrices within the set  $\Omega$  may be increasingly likely to fail with NORTA as the dimension of the problem increases, so that the problem that a user faces in using NORTA may become increasingly acute as the dimension increases.

Let us consider this possibility in the context of generating samples of a uniform random vector, that is, a random vector with uniform  $(0, 1]$  marginal distributions. This case has special significance to the NORTA method because, by construction, the method has to generate a uniform random vector  $(\Phi(Z_1), \dots, \Phi(Z_d))$  as an intermediary step. Furthermore, the rank correlation matrix of a NORTA-generated vector with continuous marginal distributions coincides with the product-moment correlation matrix for the intermediate uniform random vector.

This special case also has an advantage in that the function  $c_{ij}$  is explicitly known; see (4). Hence, a correlation matrix for the uniform random vector can be easily tested for NORTA feasibility. One simply computes the (symmetric) matrix  $\Lambda_Z$  as given by (4) and checks whether it is positive semidefinite or not. Note that if  $\Lambda_Z$  is positive semidefinite, then a joint normal random vector  $Z$  with this correlation matrix exists, and this vector transforms through NORTA to a random vector with the desired correlation matrix. On the other hand, a correlation matrix must be positive semidefinite, so if  $\Lambda_Z$  is indefinite, then it is not a correlation matrix and NORTA will fail.

Thus, the problem of estimating the probability that NORTA fails reduces to the following algorithm.

- (1) Let  $n \geq 1$  be given.
- (2) Let  $\Sigma_X(1), \dots, \Sigma_X(n)$  be an independent and identically distributed sample chosen uniformly from

$$\Omega = \{\Sigma : \Sigma = \Sigma^T, \Sigma \succeq 0, \Sigma_{jj} = 1 \quad j = 1, \dots, d\}. \quad (5)$$

- (3) For each  $i = 1, \dots, n$  let  $\Lambda_Z(i)$  be obtained from  $\Sigma_X(i)$  using the component-wise relation (4).



- (4) Estimate the probability that NORTA fails by the proportion of matrices in  $\{\Lambda_Z(i) : i = 1, \dots, n\}$  that are not positive semidefinite.

(The matrix inequality  $A \succeq 0$  signifies a constraint that the matrix  $A$  be positive semidefinite.)

Note that, in estimating the probability of NORTA ineffectiveness, we have had to choose a probability distribution on  $\Omega$ . The uniform distribution (with respect to Lebesgue measure) is a natural choice and is the one we prefer to work with. Kurowicka and Cooke [2001] give estimates for this probability using a different distribution on  $\Omega$ .

A straightforward approach (and one that we adopted) to estimate the probability of NORTA ineffectiveness is to combine three well-known methods in simulation estimation: acceptance-rejection, importance sampling and ratio estimation. We used importance sampling and acceptance-rejection on the hypercube  $H = [-1, 1]^{\frac{d(d-1)}{2}}$  ( $\Omega$  is a strict subset of  $H$ ) to choose correlation vectors from  $\Omega$ . We then used ratio estimation (see, e.g., Henderson [2001]) to estimate the probability that NORTA fails. The probability we aim to estimate can be rewritten as

$$P(\Lambda_Z \not\geq 0 | \Sigma_X \succeq 0) = \frac{P(\Lambda_Z \not\geq 0, \Sigma_X \succeq 0)}{P(\Sigma_X \succeq 0)},$$

where  $\Sigma_X$  is uniformly distributed on  $H$  and  $\Lambda_Z$  is computed from (4). An estimator of this probability is therefore of the form

$$\frac{\sum_{i=1}^n \left[ I(\Sigma_X(i) \succeq 0, \Lambda_Z(i) \not\geq 0) \frac{2^{-d(d-1)/2}}{\phi(\Sigma_X(i))} \right]}{\sum_{i=1}^n \left[ I(\Sigma_X(i) \succeq 0) \frac{2^{-d(d-1)/2}}{\phi(\Sigma_X(i))} \right]}, \quad (6)$$

where  $I$  represents the indicator function that equals 1 if its argument is true, and 0 otherwise, and the matrices  $\Sigma_X(i)$  are chosen independently with density  $\phi$  from the hypercube  $H$ . We chose the density  $\phi$  in a heuristic fashion.

This method of estimation works well in lower dimensions but we found that it became excessively slow as the dimension increased. Indeed, it took more than two days to generate on the order of a thousand samples of positive-definite matrices even for a dimension as low as  $d = 12$ . With a better choice of  $\phi$ , the algorithm would presumably be much faster, but it is not clear how to choose  $\phi$ . A better sampling technique is needed.

#### 4. THE ONION METHOD

Our goal is to construct a method that samples *exactly*, and very quickly, from the uniform distribution on the set  $\Omega_d$  as defined in (5), when viewed as a subset of  $\mathbb{R}^{d(d-1)/2}$ . We use the suffix  $d$  to emphasize the dependence on the dimension  $d$ . We thus have to construct a procedure that samples uniformly from the convex, closed, compact and full-dimensional set  $\Omega_d$ , that is, generate samples from the density

$$f(\Sigma) \propto 1, \quad \forall \Sigma \in \Omega_d, \quad (7)$$

where  $f$  is a function of the  $d(d-1)/2$  upper-diagonal elements of  $\Sigma$ .

For the random matrix  $\Sigma$ , let  $\Sigma_{\mathbf{k}}$  represent its  $k \times k$  dimensional principal leading minor (i.e., the upper-left  $k \times k$  submatrix of  $\Sigma$ ), and  $f_{\mathbf{k}}$  represent the marginal density of  $\Sigma_{\mathbf{k}}$  when  $\Sigma$  has the joint density (7). Let  $\mathbf{q}$  be the vector such that

$$\Sigma_{\mathbf{k}} = \begin{bmatrix} \Sigma_{\mathbf{k}-1} & \mathbf{q} \\ \mathbf{q}^t & 1 \end{bmatrix}.$$

We call  $\mathbf{q}$  the *completion* of  $\Sigma_{\mathbf{k}-1}$  in  $\Sigma_{\mathbf{k}}$ .

The onion method is iterative in that it starts with the one-dimensional matrix 1 and then “grows out” the matrix to the dimension desired by successively adding an extra row (and the corresponding mirrored column) chosen from an appropriate distribution. This successive layering approach is the inspiration behind its name. Marsaglia and Olkin [1984] use a similar matrix-growing approach in their algorithm to sample correlation matrices with a given set of eigenvalues, but they apply it to transform diagonal elements of arbitrary positive definite matrices to 1 in order to form correlation matrices from them. Ouellette [1981] points out some other uses of the layering approach, notable among them being the numerical method proposed by Guttman [1946] to compute inverses of large non-singular matrices.

To be more precise, the onion method is as follows:

- (1) Let  $\Sigma_1$  be the  $1 \times 1$  matrix 1.
- (2) For  $k = 2, \dots, d$ 
  - (a) Let  $\mathbf{q}$  be a column vector in  $\mathbb{R}^{k-1}$  sampled, independently of all else, from density  $\varphi_k(\cdot; \Sigma_{\mathbf{k}-1})$  say.
  - (b) Set  $\Sigma_{\mathbf{k}} = \begin{bmatrix} \Sigma_{\mathbf{k}-1} & \mathbf{q} \\ \mathbf{q}^t & 1 \end{bmatrix}$ .
  - (c) Next  $k$ .

The densities  $\varphi_k$ , which determine the  $k$ th layer, are conditional densities that depend on the partial matrix  $\Sigma_{\mathbf{k}-1}$  constructed thus far. We now state the key result that motivates the iterative sampling scheme, and in particular provides the form of the  $\varphi_k$ s.

**PROPOSITION 4.1.** *Let  $f_{\mathbf{k}}$  be the marginal density of  $\Sigma_{\mathbf{k}}$  when  $\Sigma$  is distributed as in (7). Then*

$$f_{\mathbf{k}}(\Sigma_{\mathbf{k}}) \propto (\det(\Sigma_{\mathbf{k}}))^{\frac{d-k}{2}} \quad \forall \Sigma_{\mathbf{k}} \in \Omega_k, \quad \forall 2 \leq k \leq d.$$

The marginal density  $f_{\mathbf{k}}$  represents the *joint* marginal density of the components  $\Sigma_{\mathbf{k}-1}$  and  $\mathbf{q}$  of  $\Sigma_{\mathbf{k}}$ , and Proposition 4.1 expresses  $f_{\mathbf{k}}$  as a function of  $\Sigma_{\mathbf{k}-1}$  and  $\mathbf{q}$  (through  $\Sigma_{\mathbf{k}}$ ). The density  $\varphi_k$  of the completion  $\mathbf{q}$  can then be obtained from  $f_{\mathbf{k}}$  by conditioning on a fixed  $\Sigma_{\mathbf{k}-1}$ , for each  $k = 2, \dots, d$ . The key to the generation scheme is the fact that the expression obtained for  $\varphi_k$  by this conditioning argument can be unravelled in terms of  $\Sigma_{\mathbf{k}-1}$  and  $\mathbf{q}$  in a way that allows for easy generation of  $\mathbf{q}$  for a fixed  $\Sigma_{\mathbf{k}-1}$ .

We describe an efficient scheme to sample  $\mathbf{q}$  from  $\varphi_k$  after we prove Proposition 4.1. We need two preliminary results for the proof.

LEMMA 4.2. *If  $m \geq 0$  and  $A$  is some symmetric p.d. matrix in  $\Omega_d$ , then*

$$C \triangleq \int_{\mathbb{R}^d} I(x^t A x \leq 1)(1 - x^t A x)^m dx = L(m, d) \cdot \det(A)^{-\frac{1}{2}},$$

where  $0 < L(m, d) < \infty$  does not depend on  $A$ .

PROOF. Since  $A$  is symmetric and positive definite, it has a unique upper triangular Cholesky factor  $A^{1/2}$  say, so that  $A = (A^{1/2})^t A^{1/2}$ . Applying the linear change of variables  $w = A^{1/2}x$  gives

$$\begin{aligned} C &= |\det(A)^{-\frac{1}{2}}| \cdot \int_{\mathbb{R}^d} I(w^t w \leq 1)(1 - w^t w)^m dw \\ &= \det(A)^{-\frac{1}{2}} \cdot L(m, d). \end{aligned}$$

The function  $g(w) = I(w^t w \leq 1)(1 - w^t w)^m$  is nonnegative, bounded and nonzero only over the compact region that forms the unit ball in  $\mathbb{R}^d$ . Hence,  $0 \leq L(m, d) < \infty$ . Since  $g(0) = 1$  and  $g$  is continuous about 0,  $L(m, d) > 0$ .  $\square$

For the second result that we use in the proof of Proposition 4.1, first note that any positive definite symmetric  $d \times d$  matrix  $A$  can be written as a product of two matrices, akin to a first step in an  $LU$  factorization of  $A$ , as

$$A = \begin{bmatrix} B & b \\ b^t & 1 \end{bmatrix} = \begin{bmatrix} B & 0 \\ b^t & 1 \end{bmatrix} \begin{bmatrix} I & B^{-1}b \\ 0 & 1 - b^t B^{-1}b \end{bmatrix}. \quad (8)$$

The quantity  $(1 - b^t B^{-1}b)$  is called the *Schur complement* of  $B$  in  $A$ . Ouellette [1981] is a useful source of literature on Schur complements. In particular, Ouellette [1981] points to the result obtained in Guttman [1946] that  $\text{rank}(A) = \text{rank}(B) + I(1 - b^t B^{-1}b > 0)$ . This immediately gives us the following lemma.

LEMMA 4.3. *A necessary and sufficient condition for  $A$  to be positive definite is that  $B$  be positive definite, and  $1 - b^t B^{-1}b > 0$ .*

Ouellette [1981] also describes a result that follows from (8) and was first shown by Frobenius [1968]:

$$\det(A) = \det(B)(1 - b^t B^{-1}b). \quad (9)$$

We are now ready to prove Proposition 4.1.

PROOF OF PROPOSITION 4.1. We use induction on  $k$  from  $d$  to 2 to complete the proof. The result is immediate for  $k = d$ , since the density (7) in this case is the density we are aiming for in the first place. This establishes the base case.

Let  $\Psi_k = \{q \in \mathbb{R}^k | q^t (\Sigma_k)^{-1} q \leq 1\}$ . Then, by Lemma 4.3,  $\Psi_k$  represents the set of all completion vectors  $\mathbf{q}$  of  $\Sigma_k$  in  $\Sigma_{k+1}$ . For any general  $k$ , assuming that the induction hypothesis holds for  $k + 1$ , we get

$$f_{\mathbf{k}}(\Sigma_{\mathbf{k}}) = \int_{\Psi_k} f_{\mathbf{k}+1}(\Sigma_{\mathbf{k}+1}) dq \quad (10)$$

$$\propto \int_{\Psi_k} \det(\Sigma_{\mathbf{k}+1})^{\frac{d-k-1}{2}} dq \quad (11)$$

12 • S. Ghosh and S. G. Henderson

$$= \int_{\Psi_k} \det(\Sigma_{\mathbf{k}})^{\frac{d-k-1}{2}} (1 - q^t \Sigma_{\mathbf{k}}^{-1} q)^{\frac{d-k-1}{2}} dq \quad (12)$$

$$\propto \det(\Sigma_{\mathbf{k}})^{\frac{d-k-1}{2}} \cdot \det(\Sigma_{\mathbf{k}})^{\frac{1}{2}} = \det(\Sigma_{\mathbf{k}})^{\frac{d-k}{2}}. \quad (13)$$

The first step (10) above expresses the marginal density of  $\Sigma_{\mathbf{k}}$  as the function that results from integrating out the  $(k + 1)$ st column, the completion vector  $\mathbf{q}$ , from the marginal density of  $\Sigma_{\mathbf{k}+1}$  over the set  $\Psi_k$ . The inductive hypothesis gives (11). The equality (12) uses (9), and (13) follows from Lemma 4.2.

Thus, the induction hypothesis holds for  $k$ , and hence is true for all  $k$  from  $d$  to 2.  $\square$

We now determine the densities  $\varphi_k$  used in the iterative generation procedure from the marginal densities of Proposition 4.1. As mentioned before, the densities  $f_{\mathbf{k}}$  represent the joint densities of  $\Sigma_{\mathbf{k}-1}$  and its corresponding completion vector  $\mathbf{q}$  in  $\Sigma_{\mathbf{k}}$ . Hence, if  $\Sigma_{\mathbf{k}-1}$  were fixed at  $A$  say, we would have that for a  $q \in \Psi_{k-1}$ ,

$$\begin{aligned} \varphi_k(q) &= f_{\mathbf{k}|\{\Sigma_{\mathbf{k}-1}=A\}}(\Sigma_{\mathbf{k}}) \propto \det \left( \begin{bmatrix} A & q \\ q^t & 1 \end{bmatrix} \right)^{\frac{d-k}{2}} \\ &= \det(A)^{\frac{d-k}{2}} \cdot (1 - q^t A^{-1} q)^{\frac{d-k}{2}}. \end{aligned}$$

Therefore, given  $\Sigma_{\mathbf{k}-1}$ , the conditional density for its completion vector  $\mathbf{q}$  is

$$\varphi_k(q) \propto (1 - q^t \Sigma_{\mathbf{k}-1}^{-1} q)^{\frac{d-k}{2}} \forall q \in \Psi_{k-1}. \quad (14)$$

Next comes the question of generating from densities of the form (14). For this, we employ a sequence of variable transformations. First, we apply the linear transformation  $w = \Sigma_{\mathbf{k}-1}^{-1/2} q$ , where  $\Sigma_{\mathbf{k}-1}^{-1/2}$  represents the upper triangular Cholesky factor of  $\Sigma_{\mathbf{k}-1}^{-1}$ , to get that

$$\varphi_k(q) dq \propto \tilde{\varphi}_k(w) dw,$$

where  $\tilde{\varphi}_k(w) \propto (1 - w^t w)^{(d-k)/2}$ , and  $w \in \mathbb{B}^{k-1}$ , the unit ball in  $\mathbb{R}^{k-1}$  (the constant Jacobian term that arises out of the transformation is included in the proportionality constant). Hence, to sample  $\mathbf{q}$  from  $\varphi_k$ , we could equivalently generate a  $\mathbf{w}$  from  $\tilde{\varphi}_k$  and set  $\mathbf{q}$  to be the appropriate linear transformation of  $\mathbf{w}$ .

Now,  $\tilde{\varphi}_k$  is radially symmetric, as is the set  $\mathbb{B}^{k-1}$ . Thus, if we apply a polar transformation  $w = (r, \theta)$ , where  $r$  is the  $L_2$ -norm of  $w$  and  $\theta = (\theta_1, \dots, \theta_{k-2})$  represents the angles of the polar transformation (see Kendall [1961, p. 15] for a treatment of polar transformations in higher dimensions), then

$$\begin{aligned} \tilde{\varphi}_k(w) dw &\propto (1 - r^2)^{\frac{d-k}{2}} J(r, \theta) dr d\theta_1 \cdots d\theta_{k-2} \\ &= (1 - r^2)^{\frac{d-k}{2}} r^{k-2} dr (\cos \theta_1)^{k-3} (\cos \theta_2)^{k-4} \cdots \cos \theta_{k-3} d\theta_1 \cdots d\theta_{k-2} \\ &\propto h(r) dr \end{aligned}$$

where  $J(r, \theta)$  represents the Jacobian term of the variable transformation and expands out as given in the second equation, and  $h(r) = (1 - r^2)^{\frac{d-k}{2}} r^{k-2}$ . The

second equation implies that the distribution of  $r$  is independent of the distributions of the angles  $\theta_i$ . Moreover, the radial symmetry of the integrand also gives us that  $\tilde{\varphi}_k(w)$  affects only the distribution of  $r$ , and the angles need to be sampled such that a point is chosen uniformly on the surface of the unit hyper-ball  $\mathbb{B}^{k-1}$ .

This suggests that we can sample a  $\mathbf{w}$  from  $\tilde{\varphi}_k$  by instead first sampling a radius from a normalized version of  $h$  and then multiplying by a point chosen *uniformly* over the surface of the unit ball  $\mathbb{B}^{k-1}$ . Such a point can be generated by normalizing a joint-normal independent random vector (i.e., one with the identity matrix as its correlation matrix) to have unit norm. The radius has to be sampled from  $h$ , but note that under yet another change of variable  $y = r^2$ , we have that

$$h(r) dr \propto y^{\alpha_1-1}(1-y)^{\alpha_2-1} dy,$$

which (after normalization) is simply a univariate beta density function with parameters  $\alpha_1$  and  $\alpha_2$ . For our case, the parameters  $\alpha_1$  and  $\alpha_2$  turn out to be  $(k-1)/2$  and  $(d-k)/2$  respectively. Law and Kelton [2000, p. 467] points to extensive literature on generating from beta distributions.

To recap, at the  $k$ th stage of the iterative generating procedure, to generate a realization  $q$  of  $\mathbf{q}$  from  $\varphi_k$  given the matrix  $\Sigma_{\mathbf{k}-1}$  already constructed, we do the following:

- Sample  $y$  from a beta distribution with  $\alpha_1 = (k-1)/2$  and  $\alpha_2 = (d-k)/2$ ,
- Set  $r = \sqrt{y}$ ,
- Sample a unit vector  $\theta$  uniformly from the surface of  $\mathbb{B}^{k-1}$ ,
- Set  $w = r\theta$ , and finally
- Set  $q = \Sigma_{\mathbf{k}-1}^{\frac{1}{2}}w$ .

This completes the description of the onion method.

This exact sampling method is very competitive when estimating statistical properties of the set  $\Omega_d$  when compared to methods like the one described in Section 3. First, since sampling from  $\varphi_i$  can be reduced to the problem of sampling from a univariate beta distribution and a joint-normal independent random vector, the method scales very well with dimension. In our study we were able to generate samples consisting of many thousands of matrices up to dimension  $d = 25$  in a matter of hours. Second, this method does not involve a ratio-estimation step, which means that the estimation is more straightforward to implement. Third, for a given sample size, we also found the results to be more accurate, in the sense that confidence-interval widths are smaller for this sampling method.

As noted before, this method can be generalized and applied very easily to generate uniformly from sets of symmetric positive-definite matrices with any arbitrary (fixed) positive diagonal elements. One simply has to modify the method by substituting the diagonal values of 1 assumed in this section with the corresponding positive values at the appropriate places (the definition of  $\Psi_k$  in the proof of Proposition 4.1 is one such place). The constants of the variate

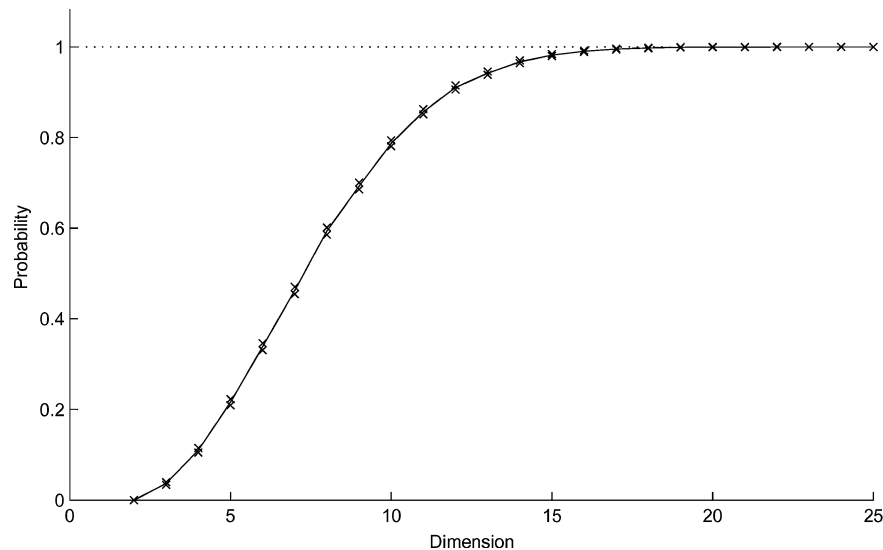


Fig. 1. Probability of NORTA ineffectiveness, based on sampling 15,000 matrices uniformly from  $\Omega$  in each dimension. Also shown are 95% confidence intervals.

generation method would be affected accordingly (for instance, the beta variate generation would not be over  $(0, 1]$ ).

One can also use this method to sample from any bounded non-uniform density  $f$  defined on a set of symmetric positive definite matrices of the kind mentioned above. One simply uses the acceptance-rejection framework of random variate generation to do this, namely by first generating a point  $s$  uniformly from the set and then checking whether  $f_{\max} * U \leq f(s)$  (where  $U$  is an independent uniform random variable, and  $f_{\max} = \max_x f(x)$ ) in order to accept  $s$  as a sample.

The beta distribution used to sample the polar variable  $r$  above can be replaced with any distribution over the positive real line. Thus the onion method can also sample from any member of the family of distributions on the set of symmetric p.s.d. random matrices that are radially symmetric under an affine transformation (the first transformation in the sequence above).

## 5. FIXING NORTA

We used the exact sampling approach of Section 4 to estimate the probability that NORTA fails for various dimensions. Our results are given in Figure 1, where the probability is plotted against dimension. The plot establishes that the problem rapidly becomes acute as the dimension increases.

We stress here that this increasing rate of failure of correlation matrices with the NORTA procedure observed in Figure 1 could be the combined effect of two causes. First, the chosen correlation matrix might not be feasible, in which case this matrix fails not only with NORTA but with any method. Second, the correlation matrix might be feasible but NORTA defective. One needs a characterization of feasible correlation matrices to be able to differentiate

between these two effects. In the absence of any such characterization, the set of correlation matrices remains the natural set for a user to choose from, and so the probability we calculate is important from the practical standpoint.

It would thus seem that NORTA can only be successful in low-dimensional problems, but this is in fact not the case. Recall that in Ghosh and Henderson [2002a] the indefinite matrices  $\Lambda_Z$  were observed to lie very close to the set of feasible correlation matrices for joint normal random vectors (i.e., the set of positive semidefinite matrices with ones on the diagonal). This led to the suggestion in Ghosh and Henderson [2002a] that the setup stage of NORTA be augmented with an SDP that is used, if  $\Lambda_Z$  turns out indefinite, to find a matrix  $\Sigma_Z$  that is “close” to  $\Lambda_Z$  and is positive semidefinite. The matrix  $\Sigma_Z$  is then used within the NORTA method.

Why is this approach reasonable? In Theorem 2 of Cario and Nelson [1997], it is shown that, under a certain moment condition, the output correlation matrix is a continuous function of the input covariance matrix  $\Sigma_Z$  used in the NORTA procedure. So if  $\Sigma_Z$  is “close” to  $\Lambda_Z$ , then we can expect the correlation matrix of the NORTA generated random vectors to be close to the desired matrix  $\Sigma_X$ . The moment condition always holds when we are attempting to match rank correlations, and we can expect it to hold almost invariably when matching product-moment correlations. Therefore, it is eminently reasonable to try to minimize some measure of distance  $r(\Lambda_Z, \Sigma_Z)$  between  $\Lambda_Z$  and  $\Sigma_Z$ .

The SDP falls under the broad class of matrix completion problems; see Alfakih and Wolkowicz [2000], or Johnson [1990]. For this case, given  $\Lambda_Z$  as data, we wish to choose a symmetric matrix  $\Sigma_Z$  to

$$\begin{aligned} & \text{minimize} && r(\Sigma_Z, \Lambda_Z) \\ & \text{subject to} && \Sigma_Z \geq 0, \quad \Sigma_Z(i, i) = 1. \end{aligned} \quad (15)$$

The metric  $r(\cdot, \cdot)$  can be chosen as desired. In particular, choosing either the  $L_1$  metric

$$r(A, B) = \sum_{i>j} |A_{ij} - B_{ij}|$$

or the  $L_\infty$  metric

$$r(A, B) = \max_{i>j} |A_{ij} - B_{ij}|$$

makes the minimization problem an SDP-constrained problem with a linear objective function. This is a convex optimization problem and therefore any local minimum is, in fact, a global minimum. Efficient algorithms, and public domain codes implementing them, are available for solving semidefinite programs of this type to optimality; see Wolkowicz et al. [2000].

Numerical studies conducted in Ghosh and Henderson [2002a] indicate that in 3 dimensions this SDP augmentation yields NORTA generated random vectors with correlation matrices that are close to the desired ones. One might then ask whether this remains the case as the dimension increases.

To study this, we use a setting identical to that used in Section 3, and choose our measure of performance to be the expected  $L_1$  distance that we have

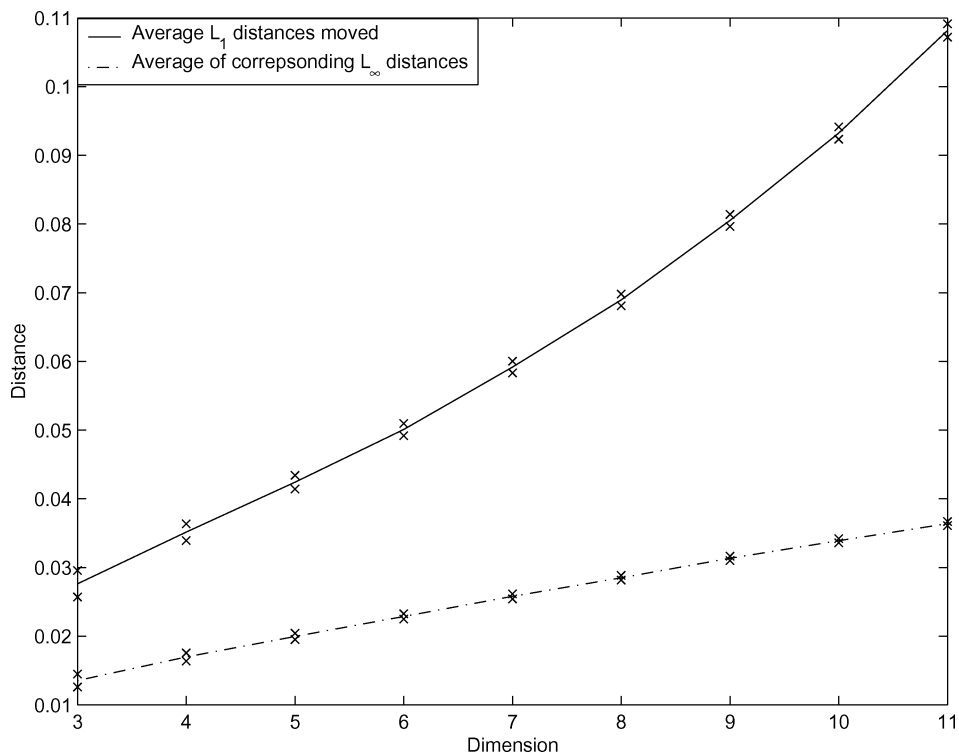


Fig. 2. Performance of the SDP-augmented NORTA in higher dimensions. 15,000 matrices were generated uniformly from  $\Omega$  and the semidefinite program, with  $r$  taken as the  $L_1$  distance, solved for the cases where NORTA was found to fail. The solid line gives the expected  $L_1$  distance with 95% confidence intervals as marked, with the average taken only over matrices that fail to work with NORTA. The dotted line gives the corresponding expected distance as measured in the  $L_\infty$  metric.

to move from the desired correlation matrix to reach a NORTA feasible one. This means that the minimization problem (15) is solved with  $r(\cdot, \cdot)$  as the  $L_1$  metric.

Figure 2 plots the results. We see that the expected  $L_1$  distance increases as the dimension  $d$  increases at what might be perceived as a linear rate, although a super-linear rate seems more likely. If the rate of increase is indeed linear, then, since there are  $d(d-1)/2$  matrix entries above the diagonal, the *average* change per entry is (eventually) decreasing with dimension. Of course, it is possible that a small number of entries change by a large amount. The  $L_\infty$  distance is also shown, and we see that indeed, at least one entry is changed by an increasing amount as the dimension increases.

It might be preferable from a modelling standpoint to instead minimize the  $L_\infty$  distance, so that one tries to minimize the maximum deviation between the achievable and target correlations. The results in this case are shown in Figure 3. We see that the expected  $L_\infty$  distance appears to remain constant at around 0.005 or even decrease with dimension. The corresponding  $L_1$  distance seems to grow at a superlinear rate.



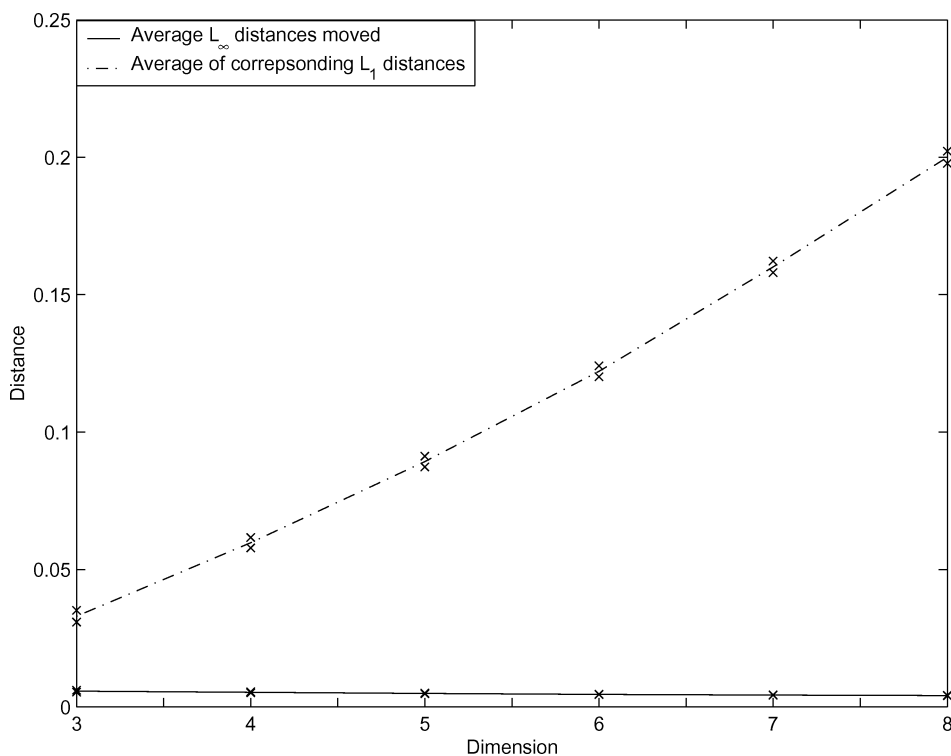


Fig. 3. The SDP-augmented NORTA in higher dimensions. For each dimension, 15,000 matrices were generated uniformly from  $\Omega$  and the semidefinite program, with  $r$  taken as the  $L_\infty$  distance, solved for NORTA defective matrices. The solid line gives the expected  $L_\infty$  distance with 95% confidence intervals as marked, with the average taken only over matrices that fail with NORTA. The dotted line gives the corresponding expected distance as measured in the  $L_1$  metric.

Table I. Maximum Deviations in Correlations ( $L_\infty$  Distances)

$d$	Taking $r$ as	
	$L_1$	$L_\infty$
3	0.040	0.017
4	0.045	0.013
5	0.055	0.015
6	0.069	0.013
7	0.072	0.012
8	0.079	0.012
9	0.077	0.011
10	0.091	0.009
11	0.099	0.010

So, on average, the SDP appears to perform well, but what about worst-case performance? Table I shows the maximum deviation observed in correlations (the  $L_\infty$  distance) for a sample of 1000 matrices. The NORTA-defective matrices were put through both the SDP-augmented methods, and the maximum deviation in correlations observed is recorded in the table. For the second column,

the SDP was solved with  $r$  as the  $L_1$  metric, while the third column takes  $r$  to be the  $L_\infty$  metric. The trends observed in the expected values of the  $L_\infty$  distances in Figures 2 and 3 are reflected in the maximum values of the  $L_\infty$  distances in Table I. The maximum deviation for the case where the SDP is solved for the  $L_1$  metric is seen to grow at a linear rate, while in the  $L_\infty$  case the maximum remains constant at around 0.015 or even decreases with dimension. This suggests that the corrective step involved in the SDP-augmented NORTA method appears to do well even in worst cases.

While the total absolute change in correlations seems to grow in either case, Figure 3 suggests that one could attempt a hybrid of the  $L_1$  and  $L_\infty$  approaches by, for instance, minimizing the  $L_1$  distance subject to an upper bound on the  $L_\infty$  distance, and thus avoid changing any single component of the correlation matrix by too large a value, while keeping the total change within reasonable limits.

We remark here that the SDP framework used here in searching for a “close” positive definite matrix in the SDP problem (15) allows us a certain degree of control on how the search is performed. For instance, we can restrict the change in certain components by adding additional constraints on the SDP.

In summary, the SDP-augmented NORTA problem performs well even in higher dimensions. It generates random vectors with correlation matrices that are close to the desired ones, while keeping changes to the individual correlations within reasonable limits.

We conclude that despite the feasibility problem, the NORTA method *is* a viable method even in high-dimensional problems.

Computational results show that the SDP problem in the SDP-augmented method is solved within a reasonable amount of time for dimensions less than 10. However, we find that the SDP problems, especially those where the  $L_\infty$  metric is used as  $r$ , become increasingly harder to solve as the dimension increases. In the instances that  $L_\infty$  SDP problems were solved for  $d = 10$  in this test, the publicly available implementations of the solvers took, on an average, around 10 minutes to terminate! So we are presently investigating alternative computational methods to SDP, as well as alternatives to the SDP formulation presented above.

## REFERENCES

- ALFAKIH, A. AND WOLKOWICZ, H. 2000. Matrix completion problems. In *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds. Kluwer, Boston, Mass. 533–545.
- CARIO, M. C. AND NELSON, B. L. 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Tech. Rep., Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Ill.
- CLEMEN, R. T. AND REILLY, T. 1999. Correlations and copulas for decision and risk analysis. *Manage. Sci.* 45, 208–224.
- DAS, S. R., FONG, G., AND GENG, G. 2001. Impact of correlated default risk on credit portfolios. *J. Fixed Inc.* 11, 3, 9–19.
- DEVROYE, L. 1986. *Non-uniform Random Variate Generation*. Springer-Verlag, New York.
- EDELMAN, A. 1989. Eigenvalues and condition numbers of random matrices. Ph.D. dissertation, Dep. Mathematics, MIT, Cambridge, Mass.
- FROBENIUS, F. G. 1968. *Gesammelte Abhandlungen*. Vol. 3. Springer, Berlin, Germany.

- GHOSH, S. AND HENDERSON, S. G. 2001. Chessboard distributions. In *Proceedings of the 2001 Winter Simulation Conference*, B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, Eds. IEEE, Piscataway, N.J., 385–393.
- GHOSH, S. AND HENDERSON, S. G. 2002a. Chessboard distributions and random vectors with specified marginals and covariance matrix. *Oper. Res.* 50, 5, 820–834.
- GHOSH, S. AND HENDERSON, S. G. 2002b. Properties of the NORTA method in higher dimensions. In *Proceedings of the 2002 Winter Simulation Conference*, E. Yücesan, C. H. Chen, J. L. Snowdon, and J. M. Charnes, Eds. IEEE, Piscataway, N.J., 263–269.
- GUTTMAN, L. 1946. Enlargement methods for computing the inverse matrix. *Ann. Math. Stat.* 17, 336–343.
- HENDERSON, S. G. 2001. Mathematics for simulation. In *Proceedings of the 2001 Winter Simulation Conference*, B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, Eds. IEEE, Piscataway, N.J., 83–94.
- HENDERSON, S. G., CHIERA, B. A., AND COOKE, R. M. 2000. Generating “dependent” quasi-random numbers. In *Proceedings of the 2000 Winter Simulation Conference*, J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, Eds. IEEE, Piscataway, N.J., 527–536.
- HILL, R. R. AND REILLY, C. H. 2000. The effects of coefficient correlation structure in two-dimensional knapsack problems on solution procedure performance. *Manage. Sci.* 46, 302–317.
- IMAN, R. AND CONOVER, W. 1982. A distribution-free approach to inducing rank correlation among input variables. *Commun. Stat. Simulat. Comput.* 11, 311–334.
- JOE, H. 1997. *Multivariate Models and Dependence Concepts*. Chapman & Hall, London, England.
- JOHNSON, C. R. 1990. Matrix completion problems: A survey. *Proc. Symp. Appl. Math.* 40, 171–198.
- JOHNSON, M. E. 1987. *Multivariate Statistical Simulation*. Wiley, New York.
- KENDALL, M. G. 1961. *A Course in the Geometry of n Dimensions*. Charles Griffin and Co., London, England.
- KRUSKAL, W. 1958. Ordinal measures of association. *J. Amer. Stat. Assoc.* 53, 814–861.
- KUROWICKA, D. AND COOKE, R. M. 2001. Conditional, partial and rank correlation for the elliptical copula; Dependence modelling in uncertainty analysis. Tech. Rep., Delft University of Technology, Mekelweg 4, 2628CD Delft, Netherlands.
- LAW, A. M. AND KELTON, W. D. 2000. *Simulation Modeling and Analysis*, 3rd ed. McGraw-Hill, New York.
- LI, S. T. AND HAMMOND, J. L. 1975. Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients. *IEEE Trans. Syst. Man, and Cybernet.* 5, 557–561.
- LURIE, P. M. AND GOLDBERG, M. S. 1998. An approximate method for sampling correlated random variables from partially-specified distributions. *Manage. Sci.* 44, 203–218.
- MARDIA, K. V. 1970. A translation family of bivariate distributions and Fréchet’s bounds. *Sankhya A*32, 119–122.
- MARSAGLIA, G. AND OLKIN, I. 1984. Generating correlation matrices. *SIAM J. Sci. Stat. Comput.* 5, 470–475.
- OUELLETTE, D. V. 1981. Schur complements and statistics. *Lin. Alg. Appl.* 36, 187–295.
- SCHMEISER, B. AND LAL, R. 1982. Bivariate gamma random vectors. *Oper. Res.* 30, 355–374.
- WHITT, W. 1976. Bivariate distributions with given marginals. *Ann. Stat.* 4, 1280–1289.
- WOLKOWICZ, H., SAIGAL, R., AND VANDENBERGHE, L., Eds. 2000. *Handbook of Semidefinite Programming: Theory, Algorithms and Applications*. Kluwer, Boston, Mass.

Received December 2002; revised April 2003 and May 2003; accepted June 2003