Constructing Discrete Unbounded Distributions with Gaussian-Copula Dependence and Given Rank Correlation

Athanassios N. Avramidis
School of Mathematics, University of Southampton
Highfield, Southampton, SO17 1BJ, UNITED KINGDOM
aa1w07@soton.ac.uk

August 19, 2011

Abstract

A random vector $\mathbf{X}$ with given univariate marginals can be obtained by first applying the normal distribution function to each coordinate of a vector $\mathbf{Z}$ of correlated standard normals to produce a vector $\mathbf{U}$ of correlated uniforms over $(0, 1)$ and then transforming each coordinate of $\mathbf{U}$ by the relevant inverse marginal. One approach to fitting requires, separately for each pair of coordinates of $\mathbf{X}$, the rank correlation, $r(\rho)$, where ρ is the correlation of the corresponding coordinates of $\mathbf{Z}$, to equal some target r^*. When the marginals are discrete and unbounded (that is, with infinite support), infinite sums arise in the mean and variance (marginal) terms and in the mean-product term of $r(\rho)$. We approximate these sums by truncation and we develop lower and upper bounds on the truncation errors. An approximation of $r(\rho)$ is obtained and enables solving the correlation-matching problem to any desired accuracy. The absolute-error bound is asymptotically a sum of weighted squared tail probabilities as truncation points tend to infinity. We determine truncation points that roughly minimize the number of summands remaining after truncation of the mean product (which is costlier to compute) subject to given limits on the error bounds. An analogous program is undertaken when one of the two marginals is continuous. Numerical examples show that rank correlations can be matched more efficiently than was previously possible, and the gain can be large when tails are heavy.

Keywords: statistics; multivariate distribution; unbounded discrete distribution; correlation; Gaussian copula.

1 Introduction

A multivariate distribution may be specified via marginal univariate distributions and with dependence between marginals induced via a Gaussian (normal) copula. This is also known as the

NORmal To Anything (NORTA) approach (Cario and Nelson, 1996, 1997). More precisely, let F_k, $k = 1, \ldots, d$ be univariate (cumulative) distribution functions, write $\mathcal{N}_{\mathbf{R}}$ for the multivariate normal distribution with mean the zero vector and $d \times d$ correlation matrix $\mathbf{R}$, and construct $\mathbf{X}$ as

$$\begin{aligned} &\mathbf{Z} = (Z_1, \ldots Z_d) \sim \mathcal{N}_{\mathbf{R}} \\ &\mathbf{X} = (X_1, \ldots, X_d) = \left(F_1^{-1}[\Phi(Z_1)], \ldots, F_d^{-1}([\Phi(Z_d)])\right), \end{aligned} \tag{1}$$

where Φ is the standard normal distribution function (with mean 0 and variance 1) and $F_k^{-1}(u) = \inf\{x : F_k(x) \geq u\}$ for $0 < u < 1$ is the inverse of F_k. By construction, the k-th marginal of $\mathbf{X}$ is F_k. Relative to other multivariate approaches, this model may be appealing by its separating the marginals from the dependence, which is contained in $\mathbf{R}$. The choice of Gaussian copula, while restrictive, facilitates fitting the model and sampling from it.

Consider the case $d = 2$. The construction reduces to selecting the scalar correlation $\rho = \operatorname{Corr}(Z_1, Z_2)$. One approach to specifying ρ is to require that the *rank correlation* between X_1 and X_2, $r(\rho) = r(\rho; F_1, F_2) = \operatorname{Corr}(F_1(X_1), F_2(X_2))$, equals (matches) a target value r^*, which may be the sample rank correlation computed from data (observations of $\mathbf{X}$), or determined otherwise. This leads to the *rank-correlation matching* problem of solving

$$r(\rho; F_1, F_2) = r^*. \tag{2}$$

The problem is nontrivial only when a non-continuous marginal is involved: if F_1 and F_2 are both continuous (that is, absolutely continuous with respect to Lebesgue measure), then $r(\rho) = \operatorname{Corr}(\Phi(Z_1), \Phi(Z_2)) = (6/\pi) \arcsin(\rho/2)$ (Kruskal, 1958). Alternatively, one may seek ρ so that the product-moment correlation between X_1 and X_2 matches a target.

The problem in dimension $d = 2$ is central to Gaussian-copula-based constructions of random vectors in dimension $d > 2$ and the VARTA class of stationary multivariate time series (Biller and Nelson, 2003). In these constructions, a correlation-matching problem is solved for certain pairs of coordinates. In the random-vector construction, a positive semi-definite matrix $\mathbf{R}$ is computed from the solutions of all coordinate pairs (Ghosh and Henderson, 2003).

We are mainly interested in rank-correlation matching problems with discrete and unbounded marginals. In emphasizing rank correlation, we are guided by arguments that it is a more appropriate measure of dependence than product-moment correlation (Embrechts et al., 2002), but also because of technical difficulties that arise with the latter. We consider the *discrete* problem, where each marginal is discrete; and the *mixed* problem, where one marginal is discrete and the other one is continuous. A discrete and unbounded X_1 gives rise to an infinite sum in the mean $\mathbb{E}[F_1(X_1)]$ and the variance $\operatorname{Var}[F_1(X_1)]$ and contributes its own infinite sum in the mean product $\mathbb{E}[F_1(X_1)F_2(X_2)]$; if X_2 is also discrete and unbounded, then the mean product leads to nested infinite sums with a term corresponding to each pairing of marginal support points. How does one approximate $r(\rho)$? One heuristic would be to truncate the sums. This is the approach in Avramidis et al. (2009) and Channouf and L'Ecuyer (2009) for discrete and mixed problems, respectively. Each infinite tail to the right is truncated at the quantile x_p associated to a tail probability p (that is, the quantile of order $1-p$); they set $p = 10^{-6}$. Such a heuristic may become impractical if truncation point(s) are large. For example, for a discrete Pareto marginal with tail index $\alpha > 1$, x_p is asymptotically a constant times $p^{1/(1-\alpha)}$ as $p \to 0$, which can easily become very large. Apart from this, truncation means a solution ρ involves an error, $r(\rho) - r^*$. Avramidis et al. (2009, Proposition 6) bound in principle the truncation error in the discrete problem; but the bound is not computable exactly, as a numerical-integration error of unknown size is involved. In the product-moment-correlation setting for Poisson marginals, Shin and Pasupathy (2010, Proposition 11) bound the error of truncating a doubly infinite sum representing $\mathbb{E}[X_1 X_2]$.

Our first contribution is to express $\mathbb{E}[X_1 X_2]$ and $\mathbb{E}[F_1(X_1)F_2(X_2)]$ and their derivatives with respect to ρ without imposing restrictions on the marginals. A corollary is that $r(\cdot)$ and the product-moment correlation are strictly increasing functions on $(-1, 1)$, so the associated correlation-matching problem has a unique solution, provided the target is within the feasible range. This development generalizes results obtained previously by imposing various restrictions on the marginals (Cario and Nelson, 1996; Avramidis et al., 2009).

Our main contribution is to approximate the mean, variance, and mean-product terms of the rank correlation by truncating the infinite sums and to bound these truncation errors. The sums representing the mean and variance terms are truncated to the right only. In the discrete problem, the nested infinite sums in the mean product are truncated to the left and to the right each. With $\tilde{r}(\rho)$ being the approximated rank-correlation function, we bound the error $r(\rho) - r^*$, where ρ is the (unique) solution to the approximating problem, $\tilde{r}(\rho) = r^*$. The error bound decomposes into a component due to rightward truncation that is asymptotically a weighted sum of the squared tail probabilities at the truncation points and a component due to leftward truncation that can be made zero by not truncating to the left. Then, we choose truncation points iteratively, first by increasing the two rightward truncation points and then by decreasing the two leftward truncation points, stopping in each case when the error-bound component falls below a given tolerance. A similar development is given for the mixed problem.

We compare the solution work associated to heuristic truncation at the quantile x_p, where $p = 10^{-6}$, to our truncation with tolerance $\delta = 10^{-3}$. With Poisson and negative binomial marginals, there is a notable gain in efficiency, meaning work reduction relative to the heuristic. With discrete Pareto marginals the gain is much larger, and it appears to be related to the heaviness of tails.

The remainder of the paper is organized as follows. In Section 2 we develop expressions for the joint expectations and their derivatives. The discrete problem is studied in Section 3; preliminary results have appeared in Avramidis (2009). The mixed problem is studied in Section 4. Numerical examples appear in Section 5.

2 General Marginals

We assume throughout the paper that the marginals are non-degenerate. The rank correlation between X_1 and X_2 as in (1) is

$$r(\rho) = \text{Corr}(F_1(X_1), F_2(X_2)) = \frac{g(\rho) - \mu_1\mu_2}{\sigma_1\sigma_2}, \tag{3}$$

where $\mu_k = \mathbb{E}[F_k(X_k)]$; $\sigma_k^2 = \text{Var}[F_k(X_k)]$; and

$$g(\rho) = \mathbb{E}[F_1(X_1)F_2(X_2)] = \int_0^1 \int_0^1 \mathbb{P}(h_1(Z_1) > x, h_2(Z_2) > y)\mathrm{d}x\mathrm{d}y \tag{4}$$

where $h_k = F_k \circ F_k^{-1} \circ \Phi$ (the composite function). The last equality comes from the well-known identity

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{P}(X > x, Y > y)\mathrm{d}x\mathrm{d}y, \tag{5}$$

for any random variables X and Y, provided the expectation is finite (e.g. Lehmann, 1966, Lemma 2).

We now represent g and its derivative with respect to ρ as integrals involving the bivariate normal density. We need a (generalized) inverse of the functions h_k. To this end, let F be a (cumulative) distribution function (c.d.f. in short), let D_F be the set of discontinuity points of F, and put $G(F) = \cup_{x \in D_F}(F(x-), F(x))$; this is the set of u for which there exists no v such that $F(v) = u$, due to discontinuity of F. The inverse of F is $F^{-1}(u) = \inf\{x : F(x) \geq u\}$. We now define the inverse of $F \circ F^{-1}$ as

$$(F \circ F^{-1})^{-1}(v) = \begin{cases} F(F^{-1}(v)-) & v \in G(F) \\ v & \text{otherwise} \end{cases} \tag{6}$$

for $v \in (0, 1)$, where $F(F^{-1}(v)-)$ is the left limit of F at $F^{-1}(v)$. A special case of (6) that we need later has a discrete F with cumulative probabilities $0 = f_0 < f_1 < \ldots$; then $(F \circ F^{-1})^{-1}(v) = f_{i-1}$ whenever $v \in (f_{i-1}, f_i)$. Now define $h_k^{-1} = \Phi^{-1} \circ (F_k \circ F_k^{-1})^{-1}$, where Φ^{-1} is the inverse of Φ. For any F, one may verify that

$$F \circ F^{-1}(u) > v \iff u > (F \circ F^{-1})^{-1}(v) \tag{7}$$

and thus (4) gives

$$g(\rho) = \int_0^1 \int_0^1 \mathbb{P}(Z_1 > h_1^{-1}(x), Z_2 > h_2^{-1}(y))\mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^1 \bar{\Phi}_\rho(h_1^{-1}(x), h_2^{-1}(y))\mathrm{d}x\mathrm{d}y, \tag{8}$$

where $\bar{\Phi}_\rho(x, y) = \int_x^\infty \int_y^\infty \phi_\rho(z, w)\mathrm{d}z\mathrm{d}w$, where $\phi_\rho(x, y)$ is the density at (x, y) of the bivariate standard normal distribution with correlation ρ. Differentiation of (8) gives

$$\frac{\mathrm{d}}{\mathrm{d}\rho} g(\rho) = \int_0^1 \int_0^1 \frac{d}{\mathrm{d}\rho} \bar{\Phi}_\rho(h_1^{-1}(x), h_2^{-1}(y))\mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^1 \phi_\rho(h_1^{-1}(x), h_2^{-1}(y))\mathrm{d}x\mathrm{d}y \tag{9}$$

for $\rho \in (-1,1)$. The derivative can pass inside the integral by an argument as in the proof of Theorem 9.42 in Rudin (1976) on noting that $\phi_\rho(h_1^{-1}(x), h_2^{-1}(y))$ has a bounded gradient with respect to (ρ, x, y) almost everywhere on $(-1,1) \times \mathbb{R}^2$. We then use that $(\mathrm{d}/\mathrm{d}\rho)\bar{\Phi}_\rho(x,y) = \phi_\rho(x,y)$ (e.g. Avramidis et al., 2009, eq. (13)).

An analogous development for the product-moment correlation follows: $r_{\mathrm{L}}(\rho) = \mathrm{Corr}(X_1, X_2) = (g_{\mathrm{L}}(\rho) - \mu_1\mu_2)/(\sigma_1\sigma_2)$, where $\mu_k = \mathbb{E}[X_k]$; $\sigma_k^2 = \mathrm{Var}(X_k)$; and, using (5),

$$\begin{aligned} g_{\mathrm{L}}(\rho) = \mathbb{E}[X_1X_2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{P}(F_1^{-1}(\Phi(Z_1)) > x, F_2^{-1}(\Phi(Z_2)) > y)\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbb{P}(Z_1 > \Phi^{-1}(F_1(x)), Z_2 > \Phi^{-1}(F_2(y)))\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{\Phi}_\rho(\Phi^{-1}(F_1(x)), \Phi^{-1}(F_2(y)))\mathrm{d}x\mathrm{d}y. \end{aligned} \tag{10}$$

We used above the equivalence $F^{-1}(u) > x \iff u > F(x)$, valid for any c.d.f. F and $u \in (0,1)$ (Asmusssen and Glynn, 2007, Proposition 2.2(a), page 38). Differentiation of (10) gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\rho} g_{\mathrm{L}}(\rho) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}\rho}\bar{\Phi}_\rho((\Phi^{-1}(F_1(x)), \Phi^{-1}(F_2(y))\mathrm{d}x\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_\rho((\Phi^{-1}(F_1(x)), \Phi^{-1}(F_2(y))\mathrm{d}x\mathrm{d}y \end{aligned} \tag{11}$$

for $\rho \in (-1,1)$. The derivative can pass inside the integral because $\phi_\rho(\Phi^{-1}(F_1(x)), \Phi^{-1}(F_2(x)))$ has a bounded gradient with respect to (ρ, x, y) almost everywhere on $(-1,1) \times \mathbb{R}^2$. We have shown the following.

Proposition 1 *Let F_1 and F_2 be c.d.f.'s of non-degenerate distributions. Put $X_k = F_k^{-1}(\Phi(Z_k))$ for $k = 1,2$, where (Z_1, Z_2) is bivariate normal with standard-normal marginals and correlation ρ. For $\rho \in (-1,1)$, the mean product $g(\rho)$ in (4) has derivative (9), and the mean product $g_{\mathrm{L}}(\rho)$ in (10) has derivative (11).*

Each of (9) and (11) is positive because the integrand is positive on a set of positive Lebesgue measure and non-negative everywhere.

Corollary 2 *The functions g and g_{L} are strictly increasing. For $r^* \in (r(-1), r(1))$, the equation $r(\rho) = r^*$ has a unique solution. For $r^* \in (r_{\mathrm{L}}(-1), r_{\mathrm{L}}(1))$, the equation $r_{\mathrm{L}}(\rho) = r^*$ has a unique solution.*

These results strengthen the known result that product-moment correlation is non-decreasing (Cario and Nelson, 1996, Theorem 1) and generalize analogous results of Avramidis et al. (2009) for discrete marginals.

3 The Discrete Problem

3.1 Preliminaries

The *discrete rank-correlation-matching problem* refers to solving $r(\rho) = r^*$ where the marginals F_1 and F_2 are both discrete. For simplicity, we assume the marginals have an infinite tail to the right only. We enumerate the support points after putting them in increasing order as $\{0, 1, 2, \ldots\}$. For the k-th marginal, $p_{k,i}$ denotes the probability mass at i; we put $f_{k,i} = \sum_{j=0}^{i} p_{k,j}$ and $f_{k,-1} = 0$. As stated following (6), $(F_k \circ F_k^{-1})^{-1}(v) = f_{k,i-1}$ for all $v \in (f_{k,i-1}, f_{k,i})$. Then (8) gives

$$g(\rho) = \sum_{i=0}^{\infty} p_{1,i} \sum_{j=0}^{\infty} p_{2,j} \bar{\Phi}_\rho(z_{1,i-1}, z_{2,j-1}), \tag{12}$$

where $z_{k,i} = \Phi^{-1}(f_{k,i})$ and $z_{k,-1} = -\infty$. This is equation (10) of Avramidis et al. (2009), seen here to be a special case of (8).

3.2 Approximation of the Mean and the Variance

The task is to approximate the means and variances in (3). To lighten notation, we work with a single marginal and later apply the forthcoming results to each marginal. Denote p_i the probability mass at i and $f_i = \sum_{j=0}^{i} p_j$ the cumulative probability at i. We will approximate the mean $\mu = \sum_{i=0}^{\infty} f_i p_i$ and the variance $\sigma^2 = \sum_{i=0}^{\infty} f_i^2 p_i - \mu^2$. Note that $\mu < 1$ and $\sigma^2 > 0$, by non-degeneracy. The approximation is via the corresponding exact moments of the finite-support random variable, X_n, obtained by shifting to the point $n+1$ the probability mass of all the points to its right, so that the resulting mass at $n+1$ is the tail probability $t_n = 1 - f_n = \sum_{i>n} p_i$. With F_n denoting the c.d.f. of X_n, the approximate mean is

$$\tilde{\mu}_n = \mathbb{E}[F_n(X_n)] = \sum_{i=0}^{n} f_i p_i + 1 - f_n$$

and the approximate variance is

$$\tilde{\sigma}_n^2 = \text{Var}[F_n(X_n)] = \tilde{\mu}_n^{(2)} - \tilde{\mu}_n^2, \tag{13}$$

where $\tilde{\mu}_n^{(2)} = \mathbb{E}[F_n^2(X_n)] = \sum_{i=0}^{n} f_i^2 p_i + 1 - f_n$.

We now derive sequences that bound μ and σ^2 from below and above and that converge to these targets in each case. We define $x^+ = \max(x, 0)$.

Lemma 3 *(i) (Sequences bounding μ below and above and converging to it.) Define $\underline{\mu} = \underline{\mu}_n = (\tilde{\mu}_n - t_n t_{n+1})^+$. We have*

$$\underline{\mu}_n \le \mu \le \tilde{\mu}_n \quad \textit{for all } n, \tag{14}$$

and $\tilde{\mu}_n \downarrow \mu$ and $\underline{\mu}_n \to \mu$ as $n \to \infty$.

(ii) (Sequences bounding σ^2 below and above and converging to it.) Define

$$\underline{\sigma}_n^2 = \tilde{\sigma}_n^2 - 2(1-\underline{\mu}_n)t_n t_{n+1} \quad \text{and} \quad \bar{\sigma}_n^2 = \begin{cases} \tilde{\sigma}_n^2 - l_n & n < n^+ \\ \tilde{\sigma}_n^2 & n \geq n^+ \end{cases}$$

where $l_n = (1 + f_n - \tilde{\mu}_{n-1} - \tilde{\mu}_n)t_n t_{n+1}$ and $n^+ = \min\{n : 1 + f_n - \tilde{\mu}_{n-1} - \tilde{\mu}_n > 0\} < \infty$. We have

$$\underline{\sigma}_n^2 \leq \sigma^2 \leq \bar{\sigma}_n^2 \quad \text{for all } n, \tag{15}$$

and $\{\tilde{\sigma}_n^2\}_{n=n^+}^\infty \downarrow \sigma^2$ and $\underline{\sigma}_n^2 \to \sigma^2$ as $n \to \infty$.

Proof. Part (i). Write $\tilde{\mu}_n - \mu = \sum_{i>n}(\tilde{\mu}_{i-1} - \tilde{\mu}_i)$ and note that

$$\tilde{\mu}_{i-1} - \tilde{\mu}_i = \sum_{k \leq i-1} f_k p_k + 1 - f_{i-1} - \sum_{k \leq i} f_k p_k - (1 - f_i) = p_i(1 - f_i) > 0. \tag{16}$$

Thus

$$0 < \tilde{\mu}_n - \mu = \sum_{i>n}(\tilde{\mu}_{i-1} - \tilde{\mu}_i) = \sum_{i>n} p_i(1 - f_i) \leq (1 - f_{n+1})(1 - f_n) = t_{n+1}t_n. \tag{17}$$

The assertion $\lim \underline{\mu}_n = \mu$ follows from $\lim_{n\to\infty} t_n t_{n+1} = 0$.

Part (ii). Write

$$\tilde{\sigma}_n^2 - \sigma^2 = \sum_{i>n}(\tilde{\sigma}_{i-1}^2 - \tilde{\sigma}_i^2) \tag{18}$$

and observe that the i-th summand above is

$$\begin{aligned} \tilde{\sigma}_{i-1}^2 - \tilde{\sigma}_i^2 &= \tilde{\mu}_{i-1}^{(2)} - \tilde{\mu}_i^{(2)} - (\tilde{\mu}_{i-1}^2 - \tilde{\mu}_i^2) \\ &= \sum_{k \leq i-1} f_k^2 p_k + 1 - f_{i-1} - \sum_{k \leq i} f_k^2 p_k - (1 - f_i) - (\tilde{\mu}_{i-1} - \tilde{\mu}_i)(\tilde{\mu}_{i-1} + \tilde{\mu}_i) \\ &= p_i(1 - f_i^2) - p_i(1 - f_i)(\tilde{\mu}_{i-1} + \tilde{\mu}_i) \\ &= p_i(1 - f_i)(1 + f_i - \tilde{\mu}_{i-1} - \tilde{\mu}_i) \end{aligned} \tag{19}$$

by using (16) in the third step. We claim that

$$2(1-\underline{\mu}_n)t_n t_{n+1} \geq \sum_{i>n} p_i(1-f_i)(1+f_i-\tilde{\mu}_{i-1}-\tilde{\mu}_i) \geq \begin{cases} (1 + f_n - \tilde{\mu}_{n-1} - \tilde{\mu}_n)t_n t_{n+1}, & n < n^+ \\ 0, & n \geq n^+ \end{cases}. \tag{20}$$

The quantity in the middle is $\tilde{\sigma}_n^2 - \sigma^2$, and a simple rearrangement proves (15). It remains to prove (20). Note that the sequence $\{1 + f_i - \tilde{\mu}_{i-1} - \tilde{\mu}_i\}_{i=1}^\infty$ is monotonically increasing to $2(1-\mu)$ (since $\{f_i\}_{i=0}^\infty \uparrow 1$ and $\{\tilde{\mu}_i\}_{i=0}^\infty \downarrow \mu$), so

$$2(1-\mu)\sum_{i>n} p_i(1-f_i) \geq \sum_{i>n}(1+f_i-\tilde{\mu}_{i-1}-\tilde{\mu}_i)p_i(1-f_i) \geq (1+f_n-\tilde{\mu}_{n-1}-\tilde{\mu}_n)\sum_{i>n} p_i(1-f_i). \tag{21}$$

In the above, we may substitute looser bounds, as follows. The upper bound (left side) is positive, so we may substitute for $\sum_{i>n} p_i(1-f_i)$ and $1-\mu$ the respective upper bounds $t_n t_{n+1}$ and $1-\underline{\mu}_n$.

The lower bound (right side) is negative (positive) when $n < n^+$ ($n \geq n^+$) respectively; in the negative case, we may substitute for $\sum_{i>n} p_i(1-f_i)$ the upper bound $t_n t_{n+1}$; in the positive case, we may substitute zero. These substitutions give (20), and this completes the proof of (15). The assertion $\{\tilde{\sigma}_n^2\}_{n=n^+}^{\infty} \downarrow \sigma^2$ holds on noting that the sequence $\{1+f_i-\tilde{\mu}_{i-1}-\tilde{\mu}_i\}_{i=0}^{\infty}$ is monotonically increasing and its n^+-th term is positive, so each summand in (18) is positive for $n \geq n^+$. The assertion $\lim \underline{\sigma}_n^2 = \sigma^2$ follows from $\lim_{n\to\infty} t_n t_{n+1} = 0$. □

In view of $\lim \underline{\sigma}_n^2 = \sigma^2 > 0$, we may define for n large enough the real number $\underline{\sigma}_n = \sqrt{\underline{\sigma}_n^2}$.

3.3 Approximation of the Mean Product

For a vector $\mathbf{n} = (l_1, r_1, l_2, r_2)$, define the approximation $\tilde{g}_{\mathbf{n}}(\rho)$ of $g(\rho)$ as the right side of (12) truncated so the range of i is restricted to $l_1 \leq i \leq r_1$ and the range of j is restricted to $l_2 \leq j \leq r_2$. For $k \in \{1,2\}$, put $f_k = f_{k,l_k-1}$ and $t_k = t_{k,r_k} = 1 - f_{k,r_k}$. To allow vacuous leftward truncation where $l_k = 0$, we set $f_{k,-1} = 0$.

Lemma 4 *We have*

$$0 \leq g(\rho) - \tilde{g}_{\mathbf{n}}(\rho) \leq f_1 + t_1^2 + f_2 + t_2^2 \quad \textit{for all } \rho. \tag{22}$$

Proof. By the non-negativity of each summand in (12), we have, for any ρ,

$$\begin{aligned} 0 \leq g(\rho) - \tilde{g}_{\mathbf{n}}(\rho) \quad \leq \quad & \sum_{i<l_1} p_{1,i} \sum_{j=0}^{\infty} p_{2,j} \bar{\Phi}_\rho(z_{1,i-1}, z_{2,j-1}) + \sum_{i>r_1} p_{1,i} \sum_{j=0}^{\infty} p_{2,j} \bar{\Phi}_\rho(z_{1,i-1}, z_{2,j-1}) \\ & + \sum_{j<l_2} p_{2,j} \sum_{i=0}^{\infty} p_{1,i} \bar{\Phi}_\rho(z_{1,i-1}, z_{2,j-1}) + \sum_{j>r_2} p_{2,j} \sum_{i=0}^{\infty} p_{1,i} \bar{\Phi}_\rho(z_{1,i-1}, z_{2,j-1}). \end{aligned} \tag{23}$$

Since $\bar{\Phi}_\rho(x,y)$ is non-decreasing in ρ, we have

$$\bar{\Phi}_\rho(x,y) \leq \bar{\Phi}_1(x,y) = \bar{\Phi}(\max(x,y)) = \min(\bar{\Phi}(x), \bar{\Phi}(y)) \quad \text{for all } \rho, \tag{24}$$

where $\bar{\Phi} = 1 - \Phi$ is the standard univariate normal complementary c.d.f.. Using this, an upper bound for the first of the four terms on the right in (23) is

$$\sum_{i<l_1} p_{1,i} \bar{\Phi}(z_{1,i-1}) \sum_{j=0}^{\infty} p_{2,j} = \sum_{i<l_1} p_{1,i} t_{1,i-1} \leq \sum_{i<l_1} p_{1,i} = f_{1,l_1-1} \tag{25}$$

upon noting that $\bar{\Phi}(z_{1,i-1}) = t_{1,i-1}$ and $\sum_{j=0}^{\infty} p_{2,j} = 1$; and an upper bound for the second term on the right of (23) is

$$\sum_{i>r_1} p_{1,i} \bar{\Phi}(z_{1,i-1}) \sum_{j=0}^{\infty} p_{2,j} = \sum_{i>r_1} p_{1,i} t_{1,i-1} \leq t_{1,r_1} t_{1,r_1+1}. \tag{26}$$

The bounds (25) and (26) and their analogs for the third and fourth term in (23) give (22). □

3.4 Approximation of the Rank Correlation

For $k \in \{1, 2\}$, and for the purpose of approximating μ_k and σ_k, we truncate marginal k to the right of r_k, as described in Section 3.2. We put $f_k = f_{k,l_k-1}$, $t_k = t_{k,r_k}$, $\tilde{\mu}_k = \tilde{\mu}_{k,r_k}$, $\underline{\mu}_k = \underline{\mu}_{k,r_k}$, $\tilde{\sigma}_k = \tilde{\sigma}_{k,r_k}$, $\underline{\sigma}_k = \underline{\sigma}_{k,r_k}$, and $\bar{\sigma}_k = \bar{\sigma}_{k,r_k}$. The function $\tilde{r}_{\mathbf{n}}(\rho) = (\tilde{g}_{\mathbf{n}}(\rho) - \tilde{\mu}_1\tilde{\mu}_2)/(\tilde{\sigma}_1\tilde{\sigma}_2)$, where the vector $\mathbf{n} = (l_1, r_1, l_2, r_2)$ gives the truncation of all relevant sums (note that the l_k apply to $\tilde{g}_{\mathbf{n}}$ and not to $\tilde{\mu}_k$ or $\tilde{\sigma}_k$) will approximate $r(\rho)$. With ρ being the unique root of $\tilde{r}_{\mathbf{n}}(\rho) = r^*$, which is assumed to exist, we will bound the error $r(\rho) - r^*$. To see the uniqueness, observe that $\tilde{g}_{\mathbf{n}}$ is the g in (12) corresponding to the finite support that results when for each $k \in \{1, 2\}$ we shift to the point r_k the probability mass of the points to its right and we shift to the point l_k the probability mass of the points to its left. By Corollary 2, $\tilde{g}_{\mathbf{n}}$ is strictly increasing in ρ, and this proves the uniqueness. Our main result is as follows.

Proposition 5 *Let ρ be the unique solution to $\tilde{r}_{\mathbf{n}}(\rho) = r^*$, assuming it exists. Provided that $\underline{\sigma}_1^2$ and $\underline{\sigma}_2^2$ are positive, we have*

$$\zeta_{\mathbf{n}} \le r(\rho) - r^* \le \eta_{\mathbf{n}} + \theta_{\mathbf{n}} \quad \textit{for all } n, \tag{27}$$

where

$$\zeta_{\mathbf{n}} = \begin{cases} r^*\left(\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\bar{\sigma}_1\bar{\sigma}_2} - 1\right), & r^* > 0 \\ r^*\left(\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\underline{\sigma}_1\underline{\sigma}_2} - 1\right), & r^* < 0, \end{cases}, \qquad \eta_{\mathbf{n}} = \frac{f_1 + f_2}{\underline{\sigma}_1\underline{\sigma}_2},$$

and

$$\theta_{\mathbf{n}} = \begin{cases} \frac{t_1^2 + t_2^2 + \tilde{\mu}_1\tilde{\mu}_2 - \underline{\mu}_1\underline{\mu}_2}{\underline{\sigma}_1\underline{\sigma}_2} + r^*\left(\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\underline{\sigma}_1\underline{\sigma}_2} - 1\right), & r^* > 0 \\ \frac{t_1^2 + t_2^2 + \tilde{\mu}_1\tilde{\mu}_2 - \underline{\mu}_1\underline{\mu}_2}{\underline{\sigma}_1\underline{\sigma}_2} + r^*\left(\frac{\tilde{\sigma}_1\tilde{\sigma}_2}{\bar{\sigma}_1\bar{\sigma}_2} - 1\right), & r^* < 0. \end{cases}$$

Proof. Putting $\tilde{h}_{\mathbf{n}}(y) = \tilde{g}_{\mathbf{n}}(y) - \tilde{\mu}_1\tilde{\mu}_2 - r^*\tilde{\sigma}_1\tilde{\sigma}_2$, we have $\tilde{h}_{\mathbf{n}}(\rho) = 0$ and

$$r(\rho) - r^* = \frac{g(\rho) - \tilde{g}_{\mathbf{n}}(\rho) + \tilde{h}_{\mathbf{n}}(\rho) + \tilde{\mu}_1\tilde{\mu}_2 - \mu_1\mu_2 + r^*(\tilde{\sigma}_1\tilde{\sigma}_2 - \sigma_1\sigma_2)}{\sigma_1\sigma_2}. \tag{28}$$

Now (27) follows from the bounds on $g(\rho) - \tilde{g}_{\mathbf{n}}(\rho)$ in (22); the bounds on μ_k as in (14); and the bounds on σ_k^2 as in (15). □

Note that $\zeta_{\mathbf{n}} \le 0$ and $\eta_{\mathbf{n}}, \theta_{\mathbf{n}} \ge 0$. We now derive asymptotic relations about the error bounds under the assumption $l_1 = l_2 = 0$ and $r_1, r_2 \to \infty$. First we note: (a) for all r_k large enough, we have $\underline{\mu}_k = \tilde{\mu}_k - t_{k,r_k}t_{k,r_k+1}$, which gives $\tilde{\mu}_1\tilde{\mu}_2 - \underline{\mu}_1\underline{\mu}_2 = \mu_2 t_1^2 + \mu_1 t_2^2 + o(t_1^2 + t_2^2)$, where $o(\cdot)$ has the usual meaning; (b) from $\underline{\sigma}_k = \tilde{\sigma}_k\sqrt{1 - 2(1 - \underline{\mu}_k)t_{k,r_k}t_{k,r_k+1}/\tilde{\sigma}_k^2}$ and the Taylor expansion $\sqrt{1-x} = 1 - x/2 + o(x)$ as $x \downarrow 0$, we obtain $\underline{\sigma}_k = \tilde{\sigma}_k[1 - (1 - \mu_k)t_k^2/\sigma_k^2] + o(t_k^2)$; then a simple calculation gives $\tilde{\sigma}_1\tilde{\sigma}_2/\underline{\sigma}_1\underline{\sigma}_2 - 1 = a_1 t_1^2 + a_2 t_2^2 + o(t_1^2 + t_2^2)$, where $a_k = (1 - \mu_k)/\sigma_k^2$; and (c) $\tilde{\sigma}_1\tilde{\sigma}_2/(\bar{\sigma}_1\bar{\sigma}_2) = 1$, provided $r_k \ge n_k^+$, the n_k^+ being as in Lemma 3. From these observations, we obtain: (i) when $r^* < 0$, we have $\zeta_{\mathbf{n}} = r^*(a_1 t_1^2 + a_2 t_2^2) + o(t_1^2 + t_2^2)$; when $r^* > 0$, we have $\zeta_{\mathbf{n}} = 0$, provided $r_k \geq n_k^+$; and (ii) $\theta_{\mathbf{n}} = b_1 t_1^2 + b_2 t_2^2 + o(t_1^2 + t_2^2)$, where we define $b_k = (1+\mu_{3-k})/(\sigma_1\sigma_2) + r^* a_k$ for $r^* > 0$ and $b_k = (1+\mu_{3-k})/(\sigma_1\sigma_2)$ for $r^* < 0$.

Remark 6 The bounds on $g(\rho) - \tilde{g}_{\mathbf{n}}(\rho)$ in (22), the bounds on μ_k as in (14), and the bounds on σ_k^2 as in (15), give immediately

$$\left.\begin{array}{ll} \dfrac{\tilde{g}_{\mathbf{n}}(\rho) - \tilde{\mu}_1\tilde{\mu}_2}{\bar{\sigma}_1\bar{\sigma}_2} \leq \dfrac{g(\rho) - \mu_1\mu_2}{\sigma_1\sigma_2} \leq \dfrac{\tilde{g}_{\mathbf{n}}(\rho) + f_1 + t_1^2 + f_2 + t_2^2 - \underline{\mu}_1\underline{\mu}_2}{\underline{\sigma}_1\underline{\sigma}_2}, & \rho > 0 \\ \dfrac{\tilde{g}_{\mathbf{n}}(\rho) - \tilde{\mu}_1\tilde{\mu}_2}{\underline{\sigma}_1\underline{\sigma}_2} \leq \dfrac{g(\rho) - \mu_1\mu_2}{\sigma_1\sigma_2} \leq \dfrac{\tilde{g}_{\mathbf{n}}(\rho) + f_1 + t_1^2 + f_2 + t_2^2 - \underline{\mu}_1\underline{\mu}_2}{\bar{\sigma}_1\bar{\sigma}_2}, & \rho < 0 \end{array}\right\} \quad (29)$$

The distance between the lower and upper bounds above converges to zero when $l_1 = l_2 = 0$ and $r_1, r_2 \to \infty$ (this follows from the asymptotics (a) to (c) following Proposition 5). Thus, (29) may be used to compute $r(\rho)$ for any ρ, including the extreme correlations $r(-1)$ and $r(1)$, to any desired accuracy.

3.5 Truncation Algorithm

The work to compute the root of $\tilde{r}_{\mathbf{n}}(\rho) = r^*$ can be expected to be *roughly* linear in $w = (r_1 - l_1 + 1)(r_2 - l_2 + 1)$. This is because w bivariate normal integrals are involved in evaluating $\tilde{g}_{\mathbf{n}}(\rho)$ at any candidate; if derivatives are to be used (Avramidis et al., 2009), then w derivatives, one for each term of (12), are involved at any candidate; and empirical results in Avramidis et al. (2009) are consistent with our claim. Then, accuracy and efficiency considerations suggest that $\mathbf{n}$ be chosen to minimize w subject to the error bounds in (27) being within given limits.

Rather than solving such a minimization problem exactly, we proceed as follows. First we reduce the quantity $\max(-\zeta_{\mathbf{n}}, \theta_{\mathbf{n}})$ —called the *rightward* error bound, as it only depends on the r_k—as follows: we initialize r_1 and r_2 as the smallest support point, 0, and iteratively increase r_1 or r_2 by one, choosing for simplicity the one that corresponds to the larger tail probability to the right, t_{k,r_k}, until the rightward error bound is no larger than δ_{r}, where $\delta_{\mathrm{r}} > 0$ is a specified tolerance. Having determined r_1 and r_2, we then reduce the quantity $\eta_{\mathbf{n}}$ —called the *leftward* error bound because the r_k have been fixed—as follows: we initialize l_k as the r_k ($k = 1, 2$) determined in phase one, and iteratively decrease l_1 or l_2 by one, choosing the one that corresponds to the larger probability to the left, f_{k,l_k-1}, until the leftward error bound is no larger than δ_{l}, where $\delta_{\mathrm{l}} \geq 0$ is a specified tolerance. The output is a truncation $\mathbf{n} = (l_1, r_1, l_2, r_2)$ and the numbers $\zeta_{\mathbf{n}}$, $\eta_{\mathbf{n}}$, and $\theta_{\mathbf{n}}$. The solution to $r_{\mathbf{n}}(\rho) = r^*$ (to be computed elsewhere) satisfies (27) and in particular $-\delta_{\mathrm{r}} \leq r(\rho) - r^* \leq \delta_{\mathrm{r}} + \delta_{\mathrm{l}}$. The algorithm is stated in more detail as Algorithm 1. By convention, the input $\delta_{\mathrm{l}} = 0$ means that no truncation to the left is intended, that is, $l_1 = l_2 = 0$.

We briefly examine how Algorithm 1 with $l_1 = l_2 = 0$ and some $\delta_{\mathrm{r}} > 0$ relates to the heuristic truncation discussed in the introduction. Suppose $r^* > 0$. In the limit as $\delta_{\mathrm{r}} \to 0$, the only requirement that remains active is $\theta_{\mathbf{n}} \leq \delta_{\mathrm{r}}$, and we have seen that $\theta_{\mathbf{n}} \sim b_1 t_1^2 + b_2 t_2^2$ (where $x_n \sim y_n$

Algorithm 1: Truncate

Input: Probability masses $\{p_{k,i}\}_{i=0}^{\infty}$ for $k = 1, 2$; target r^*; tolerances $\delta_{\mathrm{r}} > 0$ and $\delta_{\mathrm{l}} \geq 0$.

Output: Vector $\mathbf{n} = (l_1, r_1, l_2, r_2)$; error-bound components $\zeta_{\mathbf{n}}$, $\eta_{\mathbf{n}}$, and $\theta_{\mathbf{n}}$.

$r_1 \leftarrow 0$; $r_2 \leftarrow 0$; $\theta_{\mathbf{n}} \leftarrow \infty$; $\zeta_{\mathbf{n}} \leftarrow -\infty$

`/* Phase 1: rightward truncation */`

while $\max(-\zeta_{\mathbf{n}}, \theta_{\mathbf{n}}) > \delta_{\mathrm{r}}$ **do**
    if $t_{1,r_1} > t_{2,r_2}$ **then**
        $r_1 \leftarrow r_1 + 1$
        Update f_{1,r_1}, t_{1,r_1}, $\tilde{\mu}_{1,r_1}$, $\underline{\mu}_{1,r_1}$, $\tilde{\sigma}^2_{1,r_1}$, $\underline{\sigma}^2_{1,r_1}$ and $\bar{\sigma}^2_{1,r_1}$
        if $\underline{\sigma}^2_{1,r_1} \leq 0$ **then**
            continue while
        end
    else
        $r_2 \leftarrow r_2 + 1$
        Update f_{2,r_2}, t_{2,r_2}, $\tilde{\mu}_{2,r_2}$, $\underline{\mu}_{2,r_2}$, $\tilde{\sigma}^2_{2,r_2}$, $\underline{\sigma}^2_{2,r_2}$ and $\bar{\sigma}^2_{2,r_2}$
        if $\underline{\sigma}^2_{2,r_2} \leq 0$ **then**
            continue while
        end
    end
    Update $\zeta_{\mathbf{n}}$ and $\theta_{\mathbf{n}}$
end

$l_1 \leftarrow r_1$ `/* Phase 2: leftward truncation */`
$\epsilon \leftarrow f_{1,l_1} - p_{1,l_1}$
$l_2 \leftarrow r_2$
$\epsilon' \leftarrow f_{2,l_2} - p_{2,l_2}$
while $\epsilon + \epsilon' > \underline{\sigma}_{1,r_1}\underline{\sigma}_{2,r_2}\delta_{\mathrm{l}}$ **do**
    if $\epsilon > \epsilon'$ **then**
        $l_1 \leftarrow l_1 - 1$
        $\epsilon \leftarrow \epsilon - p_{1,l_1}$
    else
        $l_2 \leftarrow l_2 - 1$
        $\epsilon' \leftarrow \epsilon' - p_{2,l_2}$
    end
end
$\eta_{\mathbf{n}} \leftarrow (\epsilon + \epsilon')/(\underline{\sigma}_{1,r_1}\underline{\sigma}_{2,r_2})$

means $x_n/y_n \to 1$), with b_k defined in point (ii) following Proposition 5. Our algorithm imposes $t_1 \sim t_2$, so $t_1 \sim t_2 \sim \sqrt{\delta_r/(b_1+b_2)}$. That is, the tail probability at which we truncate is asymptotically proportional to $\sqrt{\delta_r}$. Then, r_k is the quantile of F_k corresponding to this tail probability, so it will scale according to the marginal. The case $r^* < 0$ gives similar behavior.

4 The Mixed Problem

4.1 Preliminaries

The *mixed correlation-matching problem* refers to solving $r(\rho) = r^*$ where F_1 is discrete and F_2 is continuous. This indexing involves no loss of generality. The discrete support points are $0, 1, 2, \ldots$; p_i is the probability mass at i; and $f_i = \sum_{j=0}^{i} p_j$. The continuity of F_2 means that $F_2(X_2)$ is uniformly distributed on (0,1), so its mean is $\mu_2 = 1/2$ and its variance is $\sigma_2^2 = 1/12$.

The general expression (8) of g would lead to bivariate normal integrals. We work with an alternative expression developed in Channouf and L'Ecuyer (2009) that appears easier to compute; we repeat the details for completeness:

$$
\begin{aligned}
g(\rho) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(F_1 \circ F_1^{-1} \circ \Phi(t)\right)\left(F_2 \circ F_2^{-1} \circ \Phi(s)\right)\phi_\rho(s,t)\mathrm{d}s\mathrm{d}t \\
&= (1-\rho^2)^{-1/2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(F_1 \circ F_1^{-1} \circ \Phi(t)\right)\Phi(s)\phi(s)\phi\left(\frac{t-\rho s}{\sqrt{1-\rho^2}}\right)\mathrm{d}s\mathrm{d}t \\
&= (1-\rho^2)^{-1/2} \int_{-\infty}^{\infty} \Phi(s)\phi(s)\left[\sum_{i=0}^{\infty}\int_{z_{i-1}}^{z_i} f_i \phi\left(\frac{t-\rho s}{\sqrt{1-\rho^2}}\right)\mathrm{d}t\right]\mathrm{d}s
\end{aligned}
$$

where $z_i = \Phi^{-1}(f_i)$ and $z_{-1} = -\infty$. The change of variable $(t-\rho s)/\sqrt{1-\rho^2} = y$ gives

$$
(1-\rho^2)^{-1/2}\int_{z_{i-1}}^{z_i} \phi\left(\frac{t-\rho s}{\sqrt{1-\rho^2}}\right)\mathrm{d}t = \Phi\left(\frac{z_i-\rho s}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{z_{i-1}-\rho s}{\sqrt{1-\rho^2}}\right)
$$

and the change of variable $\Phi(s) = u$ gives

$$
g(\rho) = \sum_{i=0}^{\infty} f_i \int_0^1 u[\Phi(i,u) - \Phi(i-1,u)]\mathrm{d}u, \tag{30}
$$

where $\Phi(i,u) = \Phi(i,u,\rho) = \Phi\left((z_i - \rho\Phi^{-1}(u))/\sqrt{1-\rho^2}\right)$. This is equation (9) in Channouf and L'Ecuyer (2009), except that the support there is unbounded in both directions. Note that $g(\cdot)$ does not depend on the continuous marginal; and likewise for $r(\cdot)$.

4.2 Approximation of the Mean Product and the Rank Correlation

We will develop an approximation of (30) and associated error bounds. Based on this and the usual approximations of μ_1 and σ_1, we will develop an approximation of $r(\rho)$ and error bounds in analogy to the discrete problem.

Put $\bar{\Phi}(i,u) = 1 - \Phi(i,u)$ and $f_{-1} = 0$. Rewrite (30) as $g(\rho) = \int_0^1 I(u,\rho)\mathrm{d}u$, where

$$\begin{aligned} I(u,\rho) &= u\sum_{i=0}^{\infty} f_i[\Phi(i,u) - \Phi(i-1,u)] = u\sum_{i=0}^{\infty} f_i[\bar{\Phi}(i-1,u) - \bar{\Phi}(i,u)] \\ &= u\sum_{i=0}^{\infty} \bar{\Phi}(i-1,u)(f_i - f_{i-1}) = u\sum_{i=0}^{\infty} \bar{\Phi}(i-1,u)p_i. \end{aligned} \tag{31}$$

For an integer n, we truncate the sum expression of the integrand I to obtain

$$I_n(u,\rho) = u\sum_{i=0}^{n} \bar{\Phi}(i-1,u,\rho)p_i, \tag{32}$$

and we approximate $g(\rho)$ by

$$\tilde{g}_n(\rho) = \int_0^1 I_n(u,\rho)\mathrm{d}u. \tag{33}$$

We do not consider truncation to the left for simplicity and because our numerical evidence suggests that the mixed problem is less demanding computationally than the discrete one. To bound the error, observe that

$$I(u,\rho) - I_n(u,\rho) = u\sum_{i>n} \bar{\Phi}(i-1,u,\rho)p_i \geq 0.$$

Integrating this over u, we obtain lower and upper bounds on the error:

$$0 \leq g(\rho) - \tilde{g}_n(\rho) = \int_0^1 u\sum_{i>n} \bar{\Phi}(i-1,u,\rho)p_i\mathrm{d}u \leq t_n \int_0^1 u\bar{\Phi}(n,u,\rho)\mathrm{d}u \leq \frac{t_n}{2}. \tag{34}$$

Computing the integral upper bound above (second from the right) would require numerical integration. For simplicity, we will forego this and use instead the looser upper bound on the right.

To approximate $r(\rho)$ and bound the error, let t_n, $\tilde{\mu}_n$, $\underline{\mu}_n$, $\tilde{\sigma}_n$, $\underline{\sigma}_n$, and $\bar{\sigma}_n$ be as in Section 3.2, referring to the discrete marginal, where n is the truncation point and is chosen same to that in (32). The function $\tilde{r}_n(\rho) = (\tilde{g}_n(\rho) - \tilde{\mu}_n/2)/(\tilde{\sigma}_n/\sqrt{12})$ is an approximation of $r(\rho)$. With ρ being the unique root of $\tilde{r}_n(\rho) = r^*$, which is assumed to exist, we will bound the error $r(\rho) - r^*$. To see the uniqueness, observe that $\tilde{g}_n$ is the g in (30) that results when we shift to the point n the probability mass of the points to its right. By Corollary 2, $\tilde{g}_n$ is strictly increasing in ρ, and this proves the uniqueness. Our main result is as follows.

Proposition 7 *Let ρ be the unique solution to $\tilde{r}_n(\rho) = r^*$, assuming it exists. Provided that $\underline{\sigma}_n^2$ is positive, we have*

$$\zeta_n \leq r(\rho) - r^* \leq \theta_n \quad \textit{for all } n, \tag{35}$$

where

$$\zeta_n = \begin{cases} r^*\left(\frac{\tilde{\sigma}_n}{\bar{\sigma}_n} - 1\right), & r^* > 0 \\ r^*\left(\frac{\tilde{\sigma}_n}{\underline{\sigma}_n} - 1\right), & r^* < 0 \end{cases}$$

and

$$\theta_n = \begin{cases} \frac{\sqrt{12}(t_n+\tilde{\mu}_n-\underline{\mu}_n)}{2\underline{\sigma}_n} + r^* \left(\frac{\tilde{\sigma}_n}{\underline{\sigma}_n} - 1\right), & r^* > 0 \\ \frac{\sqrt{12}(t_n+\tilde{\mu}_n-\underline{\mu}_n)}{2\underline{\sigma}_n} + r^* \left(\frac{\tilde{\sigma}_n}{\bar{\sigma}_n} - 1\right), & r^* < 0. \end{cases}$$

Proof. Putting $\tilde{h}_n(y) = \tilde{g}_n(y) - \tilde{\mu}_n/2 - r^*\tilde{\sigma}_n/\sqrt{12}$, we have $\tilde{h}_n(\rho) = 0$. Equation (28) holds, where $\tilde{\mu}_2 = \mu_2 = 1/2$ and $\tilde{\sigma}_2 = \sigma_2 = 1/\sqrt{12}$ refer to the continuous marginal. We then use the bounds on $g(\rho) - \tilde{g}_n(\rho)$ from (34); the bounds on μ_n as in (14); and the bounds on σ_n^2 as in (15). □

Note that $\zeta_n \leq 0$ and $\theta_n > 0$. We can easily see the asymptotics of the error bounds in (35) as $n \to \infty$, which will show that the error converges to zero. The quantity ζ_n behaves according to point (i) following Proposition 5, modified to eliminate the tail corresponding to the continuous marginal; in particular, $\zeta_n = O(t_n^2)$. A simple calculation gives $\theta_n = [\sqrt{12}/(2\sigma^2)]t_n + o(t_n)$; note the order is t_n, in contrast to the t_n^2 obtained elsewhere.

Remark 8 The bounds in (34), (14), and (15) imply lower and upper bounds on $r(\rho)$ analogous to (29). The distance between these bounds converges to zero as $n \to \infty$. This enables the computation of $r(\rho)$, for any $\rho \in [-1, 1]$, to any desired accuracy.

The work to compute the root of $\tilde{r}_n(\rho) = r^*$ can be expected to be *roughly* linear in n. This is because n univariate-normal tail probabilities are involved in evaluating (and integrating numerically) $I_n(u, \rho)$ at any candidate ρ; and empirical results in Channouf and L'Ecuyer (2009) and in this paper are consistent with this claim. Then, accuracy and efficiency considerations suggest that n be minimized subject to the error bounds in (35) being within given limits. Rather than solving such a minimization problem exactly, we initialize n as the smallest support point and iteratively increase it by one until $\max(-\zeta_n, \theta_n)$ is at most a specified tolerance δ; for any $\delta > 0$, clearly there exists a finite n satisfying this.

5 Numerical Results

We solve a set of test problems where marginals belong to one of three families: discrete Pareto, Poisson, and negative binomial. For these problems, solutions to equations $\tilde{r}_{\mathbf{n}}(\rho) = r^*$ associated to two different truncations, that is, different $\mathbf{n}$, are computed, as detailed later. We are particularly interested in the work of obtaining these solutions. Computations are done in MATLAB, and CPU times are measured by the `cputime` function. We do not claim these times are competitive; for example, in solving a few problems from Avramidis et al. (2009) with identical truncation and root-finder, our CPU times are larger by a factor of about one thousand. The large timing gap seems to be primarily due to the computer language (these authors use Java). We compute $\bar{\Phi}_\rho(x, y)$, the standard bivariate-normal c.d.f. at $(-x, -y)$, via MATLAB's function `mvncdf` to tolerance 10^{-9}; this method cites Drezner and Wesolowsky (1989), so we think it is reasonably efficient.

Discrete and mixed problems appear in Sections 5.1 and 5.2, respectively.

5.1 Discrete Problems

For the Pareto and Poisson families, four values r^* are chosen between $0.999\tilde{r}_{\mathbf{n_0}}(-0.9999)$ and $0.999\tilde{r}_{\mathbf{n_0}}(0.9999)$, that is, close to the minimal and maximal rank correlation, respectively, inclusively of these and in equal distance.

The *benchmark* uses the truncation vector $\mathbf{n_0} = (l_{1,0}, r_{1,0}, l_{2,0}, r_{2,0})$, where $l_{k,0}$ is the leftmost support point and $r_{k,0}$ is the quantile of order $1-p$, where $p = 10^{-6}$. This p value is also the choice of Avramidis et al. (2009) and Channouf and L'Ecuyer (2009). We compare this against truncation via Algorithm 1 with tolerances specified later. The root-finding problem is solved by a hybrid of the Newton-Raphson method and bisection, identical to Press et al. (1992, routine `rtsafe`, pp. 366–367) and to method NI3 in Avramidis et al. (2009, Section 3.1.4), to which we refer for analytical derivatives of $\tilde{g}_{\mathbf{n}}$.

The first family that we consider is the (discrete) Pareto. The discrete Pareto(α) distribution, defined for $\alpha > 1$, has support the positive integers and probability mass $f(k) = k^{-\alpha}/\zeta(\alpha)$, where $\zeta(\alpha) = \sum_{k=1}^{\infty} k^{-\alpha}$ is Riemann's *zeta function*. It is also known as "zeta" or "Zipf" distribution and is henceforth called "Pareto". It has been used in a wide range of contexts, including the modeling of service times in queues (Parulekar and Makowski, 1997; Suárez-González et al., 2002); firm sizes (Axtell, 2001); and counts of a person's sexual partners (Deuchert and Brody, 2007). Maximum-likelihood estimation is studied in Seal (1952). To explain that the benchmark can become impractical in more generality than our experiments will demonstrate, we give the asymptotic of the quantile associated to right-tail probability p, $x_p = \min\{x : \mathbb{P}(X > x) \le p\}$, as $p \to 0$. For $X \sim$ Pareto(α), we have $\mathbb{P}(X > x) \sim [(\alpha-1)\zeta(\alpha)]^{-1} x^{-\alpha+1}$, as $x \to \infty$, where $a_x \sim b_x$ means $a_x/b_x \to 1$. (This follows from $\int_{x+1}^{\infty} y^{-\alpha} \mathrm{d}y \le \sum_{k>x} k^{-\alpha} \le \int_{x+1}^{\infty} (y-1)^{-\alpha} \mathrm{d}y$.) Setting $p = \mathbb{P}(X > x_p)$ and making p the independent variable shows that $x_p \sim \big((\alpha-1)\zeta(\alpha)p\big)^{1/(1-\alpha)}$ as $p \to 0$; and rounding the asymptotic to the nearest positive integer gives an approximation, $\hat{x}_p$, of x_p. To report a few quantiles and their approximation, we compute $\zeta(2) \doteq 1.644934$ and $\zeta(1.5) \doteq 2.612375$ via MATLAB's `zeta` function. For $\alpha = 2$ and for $p = 10^{-2}$, 10^{-4} and 10^{-6}, the pairs $(x_p, \hat{x}_p)$ are identical, equal to 61, 6079, and 607927, respectively. For $\alpha = 1.5$ and for $p = 10^{-2}$ and 10^{-4}, the pairs are again identical, equal to 5861 and 58,612,310, respectively. We see that depending on how small α is, x_p can be very large—even when p is not very small.

Table 1 gives results for the Pareto family, where we only truncate to the right with $\delta_{\mathrm{r}} = 10^{-3}$ ($\delta_{\mathrm{l}} = 0$). Each of the six panels specifies a pair of marginals and the benchmark number $w_0 = (r_{1,0} - l_{1,0} + 1)(r_{2,0} - l_{2,0} + 1)$. Each row within a panel corresponds to the problem instance with target r^*; we report the (approximate) solution ρ; our method's number $w = (r_1 - l_1 + 1)(r_2 - l_2 + 1)$; our CPU time; the error estimate $\tilde{r}_{\mathbf{n_0}}(\rho) - r^*$ (where "3e-04" means 3×10^{-4}); and the ratio of the benchmark's CPU time to our CPU time (CPU ratio). Heavier tails (smaller α) are associated with more work for our method, as evidenced by larger w and CPU, and much larger work for the benchmark, as evidenced by larger w_0/w and larger CPU ratio. It is noteworthy that in the first row of each panel, the target is $r^* = 0.999 r_{\mathbf{n_0}}(-0.9999)$ and the (approximate) solution ρ is far from -0.9999; this happens because the function $r_{\mathbf{n_0}}$ increases very slowly between -0.9999 and that ρ. We also note here that we solved the same problems with a second root finder, MATLAB's `fzero`, and observed that CPU ratios were roughly linear in w_0/w, as seen above.

Table 1: Discrete problem with Pareto(α_1) and Pareto(α_2) marginals. $\delta_\mathrm{r} = 10^{-3}$, $\delta_\mathrm{l} = 0$.

	r^*	ρ	w	CPU (sec)	$\tilde{r}_{\mathbf{n_0}}(\rho) - r^*$	CPU ratio
$\alpha_1 = 5$, $\alpha_2 = 5$	-0.0368	-0.5160	49	0.20	3e-04	15.6
$w_0 = 484$	0.3044	0.6541	49	0.19	3e-04	11.5
	0.6455	0.9157	49	0.33	3e-04	10.9
	0.9867	0.9999	49	0.58	3e-04	10.1
$\alpha_1 = 5$, $\alpha_2 = 4$	-0.0547	-0.5677	72	0.34	4e-04	35.6
$w_0 = 1496$	0.2001	0.4849	72	0.25	4e-04	27.3
	0.4550	0.7892	72	0.34	5e-04	27.2
	0.7099	0.9923	72	0.52	5e-04	24.0
$\alpha_1 = 5$, $\alpha_2 = 3$	-0.0846	-0.6420	190	0.96	5e-04	112.6
$w_0 = 14190$	0.1311	0.3341	190	0.76	5e-04	87.5
	0.3468	0.6875	190	0.76	5e-04	90.1
	0.5625	0.9999	190	1.20	6e-04	81.2
$\alpha_1 = 4$, $\alpha_2 = 4$	-0.0815	-0.6277	100	0.55	5e-04	72.4
$w_0 = 4624$	0.2752	0.5436	100	0.42	5e-04	56.9
	0.6319	0.8777	100	0.61	5e-04	55.9
	0.9887	0.9999	100	1.23	5e-04	57.4
$\alpha_1 = 4$, $\alpha_2 = 3$	-0.1259	-0.7134	261	1.66	4e-04	200.8
$w_0 = 43860$	0.1659	0.3426	261	1.14	4e-04	191.9
	0.4576	0.7269	261	1.11	5e-04	179.9
	0.7494	0.9987	261	1.96	5e-04	173.6
$\alpha_1 = 3$, $\alpha_2 = 3$	-0.1945	-0.7882	529	3.51	5e-04	950.6
$w_0 = 416025$	0.2008	0.3475	529	2.43	5e-04	809.7
	0.5960	0.7933	552	2.86	5e-04	839.3
	0.9913	0.9999	552	7.40	5e-04	816.7

The second set of examples has Poisson marginals. We initially ran Algorithm 1 in two ways that ensure the absolute error is at most 10^{-3}: truncation to the right only with $\delta_\mathrm{r} = 10^{-3}$ ($\delta_\mathrm{l} = 0$); and truncation to the left and right (two-sided truncation) with $\delta_\mathrm{l} = \delta_\mathrm{r} = 0.5 \times 10^{-3}$. We found that for larger scale parameters (means), solution with two-sided truncation took less work (smaller w and CPU time); for smaller means, there was no clear winner, but this is much less important because the work is much smaller. Table 2 contains results for the two-sided truncation, in the existing format. Larger means (the k-th mean is λ_k) are associated with more work for our method, as indicated by w and CPU, and even more work for the benchmark, as indicated by w_0/w and the CPU ratio. The gain reflected by these ratios is much smaller in comparison to the discrete Pareto examples.

Table 2: Discrete problem with Poisson(λ_1) and Poisson(λ_2) marginals. $\delta_{\mathrm{r}} = \delta_{\mathrm{l}} = 0.5 \times 10^{-3}$.

	r^*	ρ	w	CPU (sec)	$\tilde{r}_{\mathbf{n_0}}(\rho) - r^*$	CPU ratio
$\lambda_1 = 1$	-0.8501	-0.9898	30	0.19	9e-05	3.9
$\lambda_2 = 1$	-0.2359	-0.2922	30	0.12	1e-04	3.2
$w_0 = 100$	0.3783	0.4635	30	0.09	2e-04	3.7
	0.9925	0.9999	30	0.38	3e-04	3.1
$\lambda_1 = 1$	-0.9248	-0.9963	100	0.69	7e-05	3.2
$\lambda_2 = 10$	-0.3075	-0.3505	100	0.36	1e-04	2.7
$w_0 = 290$	0.3099	0.3539	100	0.31	3e-04	3.2
	0.9272	0.9981	100	0.68	4e-04	2.9
$\lambda_1 = 1$	-0.9352	-0.9987	330	2.11	3e-04	5.9
$\lambda_2 = 100$	-0.3116	-0.3532	330	1.10	4e-04	5.1
$w_0 = 1520$	0.3121	0.3550	330	1.05	6e-04	5.3
	0.9358	0.9997	330	2.73	6e-04	4.6
$\lambda_1 = 10$	-0.9818	-0.9985	400	2.92	1e-05	2.2
$\lambda_2 = 10$	-0.3222	-0.3394	400	1.47	9e-05	2.2
$w_0 = 841$	0.3374	0.3549	400	1.44	2e-04	2.2
	0.9970	0.9998	400	4.80	3e-04	2.4
$\lambda_1 = 10$	-0.9906	-0.9987	1300	10.27	2e-04	3.3
$\lambda_2 = 100$	-0.3294	-0.3450	1300	5.25	3e-04	3.1
$w_0 = 4408$	0.3317	0.3478	1300	5.08	5e-04	3.4
	0.9928	0.9993	1320	11.56	5e-04	2.9
$\lambda_1 = 100$	-0.9972	-0.9988	4422	35.80	2e-04	5.1
$\lambda_2 = 100$	-0.3320	-0.3460	4422	18.03	3e-04	5.1
$w_0 = 23104$	0.3332	0.3479	4422	17.48	4e-04	5.3
	0.9984	0.9996	4489	42.18	4e-04	5.1

In the third set of examples, the marginals come from the negative binomial family with (shape, scale) parameter (p, s), whose probability mass function is $f(k) = \binom{s+k-1}{k} p^s (1-p)^k$. Here, the marginals and the values r^* are chosen to match those in Avramidis et al. (2009). Table 3 contains results for the two-sided truncation, which for larger means was more efficient than one-sided truncation. Patterns similar to the Poisson examples are seen: larger means (the k-th mean is $s_k(1-p_k)/p_k$)) are associated with more work for our method, as indicated by w and CPU, and even more work for the benchmark, as indicated by w_0/w and the CPU ratio. The gain reflected by these ratios is much smaller in comparison to the discrete Pareto examples.

Table 3: Discrete problem with negative binomial(p_1, s_1) and negative binomial(p_2, s_2) marginals. $\delta_{\mathrm{r}} = \delta_{\mathrm{l}} = 0.5 \times 10^{-3}$.

	r^*	ρ	w	CPU (sec)	$\tilde{r}_{\mathbf{n_0}}(\rho) - r^*$	CPU ratio
$p_1 = 0.386,\ s_1 = 1.568$	-0.50	-0.5340	182	0.75	9e-05	4.2
$p_2 = 0.621,\ s_2 = 6.021$	0.05	0.0544	182	0.67	2e-04	4.3
$w_0 = 768$	0.43	0.4618	182	0.63	3e-04	4.6
	0.90	0.9338	195	0.84	2e-04	4.1
	0.96	0.9905	195	1.34	2e-04	4.3
$p_1 = 0.386,\ s_1 = 15.68$	-0.50	-0.5181	2352	10.06	3e-04	2.6
$p_2 = 0.621,\ s_2 = 60.21$	0.05	0.0528	2401	10.08	4e-04	2.6
$w_0 = 6560$	0.43	0.4474	2401	9.75	4e-04	2.7
	0.90	0.9096	2401	11.54	5e-04	2.7
	0.98	0.9836	2401	15.44	5e-04	2.7
$p_1 = 0.386,\ s_1 = 156.8$	-0.50	-0.5174	26726	107.42	3e-04	6.9
$p_2 = 0.621,\ s_2 = 602.1$	0.05	0.0528	26892	108.88	4e-04	6.8
$w_0 = 189912$	0.43	0.4470	27054	107.94	5e-04	6.9
	0.90	0.9085	27216	128.65	5e-04	6.9
	0.98	0.9824	27216	172.02	5e-04	6.9

5.2 Mixed Problems

We compare two alternative truncation points: (i) a *benchmark* n_0, set as the quantile of order $1-10^{-6}$; and (ii) the smallest n such that the error bound $\max(-\zeta_n, \theta_n)$ is no larger than $\delta = 10^{-3}$. The values r^* are chosen via near-extremes $0.999\tilde{r}_{n_0}(\pm 0.9999)$, as before. The respective equations, $\tilde{r}_{n_0}(\rho) = r^*$ and $r_n(\rho) = r^*$, are solved with MATLAB's `fzero`, described as "a combination of bisection, secant, and inverse quadratic interpolation methods". The integral in (33) is evaluated via MATLAB's `quadgk` function, described as "adaptive quadrature based on a Gauss-Kronrod pair (15th- and 7th-order formulas)", with error tolerance 10^{-12}.

Table 4 gives results for mixed problems whose two marginals are a discrete Pareto and any continuous marginal. Heavier tails (smaller α) are associated with more work for our method, as indicated by n and CPU, and even more work for the benchmark, as indicated by n_0/n and the CPU ratio. Compared with the discrete problem with the same marginals, the mixed problem is solved much faster; this is in agreement with findings in Channouf and L'Ecuyer (2009).

6 Conclusion

We contributed to the mathematical underpinning of how a random vector $\mathbf{X}$ of the form (1) can be constructed by controlling, separately for each pair of coordinates of $\mathbf{X}$, the rank correlations or product-moment correlations by setting them to target values. For arbitrary univariate distribution functions F_1 and F_2, we gave expressions for $\mathbb{E}[F_1(X_1)F_2(X_2)]$ and $\mathbb{E}[X_1X_2]$ and their

Table 4: Mixed problem with a Pareto(α) discrete marginal. $\delta = 10^{-3}$.

	r^*	ρ	n	CPU (sec)	$\tilde{r}_{n_0}(\rho) - r^*$	CPU ratio
$\alpha = 5$	-0.3204	-0.9970	16	0.02	-4e-10	1.3
$n_0 = 22$	-0.1068	-0.2606	16	0.02	-3e-10	1.2
	0.1068	0.2606	16	0.03	6e-09	1.1
	0.3205	0.9971	16	0.03	-1e-08	1.2
$\alpha = 4$	-0.4580	-0.9981	32	0.03	-1e-09	1.5
$n_0 = 68$	-0.1527	-0.2884	32	0.03	-3e-07	1.6
	0.1527	0.2884	32	0.04	3e-07	1.6
	0.4580	0.9981	32	0.05	3e-08	1.6
$\alpha = 3$	-0.6465	-0.9987	126	0.11	-3e-08	4.5
$n_0 = 645$	-0.2155	-0.3194	126	0.11	2e-08	4.8
	0.2155	0.3194	126	0.18	3e-08	5.1
	0.6465	0.9987	126	0.18	-1e-07	5.0
$\alpha = 2.2$	-0.8254	-0.9990	2055	2.34	1e-07	32.7
$n_0 = 61597$	-0.2751	-0.3416	2055	2.39	3e-08	32.7
	0.2751	0.3416	2055	4.76	9e-08	33.2
	0.8254	0.9990	2055	4.70	1e-07	33.9

derivatives with respect to ρ and showed that both the rank correlation $r(\rho)$ and the product-moment correlation are strictly increasing on $(-1, 1)$. This extended results obtained previously by imposing restrictions on the marginals. Assuming one marginal is discrete and unbounded and the other one is discrete and unbounded or continuous, we approximated $\mathbb{E}[F_k(X_k)]$, $\mathrm{Var}[F_k(X_k)]$, and $\mathbb{E}[F_1(X_1)F_2(X_2)]$ by truncating the relevant infinite sums, and we bounded the error in each case. The resulting approximation $\tilde{r}(\rho)$ of $r(\rho)$ is such that equations of the form $r(\rho) = r^*$ can be solved to any desired accuracy. As truncation points tend to infinity, the error bound is asymptotically a weighted sum of the squared tail probabilities at these points.

In addition to ensuring accuracy, our approach to rank-correlation matching was found to generally require less work than a heuristic whose truncation points are the respective quantiles x_p associated to a small tail probability p. Such a heuristic may be impractical if the quantiles grow too large as $p \to 0$, and this is prone to happen when tails are heavy. The improvement (work reduction) over the heuristic was very large with Pareto marginals and modest with Poisson and negative binomial marginals, suggesting it is closely related to the heaviness of tails. For marginals with large scale parameter, splitting our "truncation error budget" across the two directions (left and right) resulted in less work compared with allocating all the error budget to the right.

Some ideas for future inquiry are now proposed. Problem instances with large scale parameters can require substantial work, even with our approach. More efficient solution of such problems is an open research problem. Another line of inquiry could be to see if our approach can be extended to the product-moment correlation for general discrete and unbounded marginals. A difficulty in this program is that the summands in the corresponding infinite sums do not seem to obey a simple bound such as the bound "1" we used for the cumulative probabilities.

Acknowledgments

The author thanks Professor Pierre L'Ecuyer for his comments on an early draft of the paper.

References

Asmusssen, S., P.W. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.

Avramidis, A. N. 2009. Fitting discrete multivariate distributions with unbounded marginals and normal-copula dependence. M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, R. G. Ingalls, eds., *Proceedings of the 2009 Winter Simulation Conference*. Institute of Electrical and Electronics Engineers, Inc., Piscataway, New Jersey, 452–458.

Avramidis, A. N., N. Channouf, P. L'Ecuyer. 2009. Efficient correlation matching for fitting discrete multivariate distributions with arbitrary marginals and normal-copula dependence. *INFORMS Journal on Computing* **21** 88–106.

Axtell, R. L. 2001. Zipf distribution of U.S. firm sizes. *Science* **293** 1818–1820.

Biller, B., B.L. Nelson. 2003. Modeling and generating multivariate time-series input processes using a vector autoregressive technique. *ACM Transactions on Modeling and Computer Simulation* **13** 211–237.

Cario, M. C., B. L. Nelson. 1996. Autoregressive to anything: Time series input processes for simulation. *Operations Research Letters* **19** 51–58.

Cario, M. C., B. L. Nelson. 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical Report, Department of Industrial Engineering and Management Science, Northwestern University.

Channouf, N., P. L'Ecuyer. 2009. Fitting a normal copula for a multivariate distribution with both discrete and continuous marginals. *Proceedings of the 2009 Winter Simulation Conference*. IEEE Press, 352–358.

Deuchert, E., S. Brody. 2007. Plausible and implausible parameters for mathematical modeling of nominal heterosexual HIV transmission. *Annals of Epidemiology* **17** 237–244.

Drezner, Z., G. O. Wesolowsky. 1989. On the computation of the bivariate normal integral. *Journal of Statistical Computation and Simulation* **35** 101–107.

Embrechts, P., A. McNeil, D. Straumann. 2002. Correlation and dependence in risk management: properties and pitfalls. M. A. H. Dempster, ed., *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, 176–223.

Ghosh, S., S.G. Henderson. 2003. Behavior of the NORTA method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation* **13** 276–294.

Kruskal, W. 1958. Ordinal measures of association. *Journal of the Americal Statistical Association* **53** 814–861.

Lehmann, E. L. 1966. Some concepts of dependence. *Annals of Mathematical Statistics* **37** 1137–1153.

Parulekar, Minothi, Armand M. Makowski. 1997. Tail probabilities for M/G/∞ input processes (I): Preliminary asymptotics. *Queueing Systems* **27** 271–296.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, B. P. Flannery. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. 2nd ed. Combridge University Press, New York.

Rudin, W. 1976. *Principles of Mathematical Analysis*. 3rd ed. McGraw-Hill, New York.

Seal, H. L. 1952. The maximum likelihood fitting of the discrete Pareto law. *Journal of the Institute of Actuaries* **78** 115–121.

Shin, K., R. Pasupathy. 2010. A method for fast generation of Poisson random vectors. *INFORMS Journal on Computing* **22** 81–92.

Suárez-González, A., J. C. López-Ardao, C. López-García, M. Fernández-Veiga, R. F. R. Rubio, M. E. S. Vieira. 2002. A new heavy-tailed discrete distribution for LRD M/G/∞ sample generation. *Perform. Eval.* **47** 197–219.