Research Article

Jan-Frederik Mai*

About the exact simulation of bivariate (reciprocal) Archimax copulas

https://doi.org/10.1515/demo-2022-0102
received November 23, 2021; accepted March 8, 2022

Abstract: We provide an exact simulation algorithm for bivariate Archimax copulas, including instances with negative association. In contrast to existing simulation approaches, the feasibility of our algorithm is directly linked to the availability of an exact simulation algorithm for the probability measure described by the derivative of the parameterizing Pickands dependence function. We demonstrate that this hypothesis is satisfied in many cases of interest and, in particular, it is satisfied for piece-wise constant Pickands dependence functions, which can approximate the general case to a given level of desired accuracy. Finally, the algorithm can be leveraged to an exact simulation algorithm for bivariate copulas associated with max-infinitely divisible random vectors whose exponent measure has norm-symmetric survival function, so-called reciprocal Archimax copulas.

Keywords: Archimax copula, extreme-value copula, Archimedean copula, max-infinitely divisible, simulation algorithm

MSC 2020: 65C05, 60E05, 62H05

1 Introduction

We recall that a bivariate copula is the restriction to $[0,1]^2$ of a joint distribution function of a random vector $(X, Y)$ with the property that both $X$ and $Y$ are uniformly distributed on the interval $[0,1]$, i.e., $F_X(x) = \mathbb{P}(X \leq x) = x$ and $F_Y(y) = \mathbb{P}(Y \leq y) = y$ for $0 \leq x, y \leq 1$. There exists a lot of research on copulas, since the distribution function of an arbitrary bivariate random vector $(X, Y)$ can always be written as $C(F_X(x), F_Y(y))$ for some copula $C$ due to Sklar’s theorem, so that copulas are an analytic tool to separate dependence and marginals. Similarly, the survival function of $(X, Y)$ can always be written as $\hat{C}(1 - F_X(x), 1 - F_Y(y))$ for some copula $\hat{C}$, called survival copula. If $F_X$, $F_Y$ are continuous, the copula and survival copula of $(X, Y)$ are even unique, see [5,17] for popular textbooks on the topic.

The present article deals with the specific family of bivariate Archimax copulas that have been introduced in [1] with the intention to create a large parametric family of copulas with desirable properties in the context of extreme-value theory. For detailed background on dependence properties we also refer to [6], and inference for Archimax copulas is investigated in [3]. Bivariate Archimax copulas are parameterized in terms of a pair $(\varphi, \ell)$ of a continuous and convex function $\varphi : [0, \infty) \to [0,1]$ with $\varphi(0) = 1$ and $\lim_{x\to\infty} \varphi(x) = 0$, and a norm $\ell$ on $\mathbb{R}^2$ with normalization $\ell(1,0) = \ell(0,1) = 1$ that is orthant–monotonic, which means that $\ell(x,y) = \ell(|x|,|y|)$. Intuitively, this means that its unit circle is axis-symmetric so that its unit simplex $\{ (x,y) \in [0,1]^2 : \ell(x,y) = 1 \}$ lies between the ones corresponding to $\|\cdot\|_1$ and $\|\cdot\|_\infty$. It is well-known from [1,2] that there exists a random vector $(X, Y)$ on the positive orthant $(0,\infty)^2$ whose survival

* Corresponding author: Jan-Frederik Mai, XAIA Investment GmbH, Sonnenstr. 19, 80331 München, Germany, e-mail: jan-frederik.mai@xaia.com

Open Access. © 2022 Jan-Frederik Mai, published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.
function is given by $P(X > x, Y > y) = \varphi \circ \ell(x, y)$, $x, y \geq 0$. The unique survival copula of $(X, Y)$ is called Archimax copula, given by

$$C_{\varphi, \ell}(x, y) = \varphi \circ \ell[\varphi^{-1}(x), \varphi^{-1}(y)], \quad 0 \leq x, y \leq 1.$$  

(1)

In Section 5, we will further provide an exact simulation algorithm for reciprocal Archimax copulas, which have the algebraic form

$$C_{\varphi, \ell}(x, y) = \frac{x y}{C_{\varphi, \ell}(x, y)}, \quad 0 \leq x, y \leq 1,$$  

(2)

where $\ell$ is the same as above, and $\psi : [0, \infty) \to [0, \infty)$ is a convex function with $\lim_{x \to \infty} \psi(x) = 0$ (just like $\varphi$) and $\lim_{x \to \infty} \varphi(x) = \infty$ (unlike $\varphi$). The nomenclature originates from the algebraic reminiscence with Archimax copulas in (1) and was introduced in [7] for the special case $\ell = \| . \|_1$. These arise as copulas of bivariate max-infinitely divisible random vectors whose exponent measure has $\ell$-symmetric survival function.

The function $\varphi$ is called Archimedean generator, since the subfamily of copulas of the form $C_{\varphi, \ell} = \| . \|_1$ coincides with Archimedean copulas, see [16] for background on the matter. Similarly, $\psi$ will be called a generator for the reciprocal Archimax copula, since the special case $C_{\varphi, \ell} = \| . \|_1$ coincides with bivariate reciprocal Archimedean copulas introduced in [7]. The norm $\ell$ is called a stable tail dependence function, and the function $A_{\ell}(x) = \ell(x, 1-x)$ for $x \in [0, 1]$ is called Pickands dependence function. (Note that $\ell$ is fully determined by $A_{\ell}$ via $\ell(x, y) = (x + y)A_{\ell}(\frac{x}{x+y})$.) The nomenclature originates from the field of multivariate extreme-value theory, since the subfamily of copulas of the form $C_{\varphi(x) = \exp(-x), \ell}$ coincides with bivariate extreme-value copulas, see [8] for background on the matter.

Regarding the strength of association between the components $X$ and $Y$ there are two different regimes. If the function $\varphi$ happens to be the Laplace transform of a positive random variable $M$, we have non-negative association, while we can have negative association in the complementary case, when $\varphi$ is a so-called strict Archimedean generator (not a Laplace transform). The case of positive association is significantly better understood, because in this case the survival function $\varphi \circ \ell(x, y) = E[\exp(-M \ell(x, y))]$ equals that of a scale mixture of a bivariate min-stable exponential distribution, which is a well-investigated concept of positive association, see [9, Chapter 6, p. 169 ff]. In particular, for random vectors with survival function $\exp(-mt(x, y))$, or equivalently for extreme-value copulas $C_{\varphi(x) = \exp(-x), \ell}$ there exist exact simulation algorithms due to [4], which are feasible for many $\ell$. Our intention is to provide an algorithm that is equally general with respect to the choice of $\ell$, but is not restricted to Laplace transforms $\varphi(x) = E[\exp(-Mx)]$, but works also for strict generators $\varphi$, in particular for the important special case $\varphi(x) = (1-x)^{+}$.

We recap the existing literature that forms the basis for our derivation. On the one hand, there is the following decomposition result due to [2].

**Theorem 1.1.** (Charpentier et al. [2]) Let $\ell$ be an orthonormal–monotonic norm on $\mathbb{R}^2$ subject to the normalization $\ell(1, 0) = \ell(0, 1) = 1$, and let $\varphi : [0, \infty) \to [0, 1]$ be a continuous and convex function with $\varphi(0) = 1$ and $\lim_{x \to \infty} \varphi(x) = 0$. Then there exists a random vector $(X, Y)$ with survival function $P(X > x, Y > y) = (1 - \ell(x, y))$, and an independent positive random variable $R_{\varphi}$ with the property $E[1 - (1 - x/R_{\varphi})] = \varphi(x)$, whose distribution is uniquely determined by the function $\varphi$. Furthermore, the random vector $(X, Y) = R_{\varphi}(X_\ell, Y)$ has survival function $P(X > x, Y > y) = \varphi \circ \ell(x, y)$, $x, y \geq 0$, and survival copula $C_{\varphi, \ell}$. Thus, the random vector $(\varphi(R_{\varphi}), \varphi(R_{\varphi}))$ has distribution function $C_{\varphi, \ell}$.

In words, Theorem 1.1 states that every Archimax copula arises as scale mixture of the special Archimax copula with generator $\varphi(x) = (1-x)^{+}$, which equals the Archimax copula with minimal association for a fixed $\ell$. When simulating $(X, Y)$ with survival copula $C_{\varphi, \ell}$, this allows us to perform two independent subsequent simulation steps: first simulation of $R_{\varphi}$ and second simulation of $(X_\ell, Y)$ with survival copula given in terms of $\ell$ by

$$C_{\ell}(x, y) = C_{\varphi(x) = (1-x)^{+}, \ell}(x, y) = (1 - \ell(1-x, 1-y))^{+}.$$
On the other hand, in the original reference [1] the authors propose an exact simulation algorithm, which we are going to refine. According to Theorem 1.1, we may focus on the simulation of \((X_t, Y_t)\), which corresponds to the special case \(\varphi(x) = (1 - x)_+\), and in which the method of [1] can be formulated as in the following theorem.

**Theorem 1.2.** (Capéraà et al. [1]) Assume that the second derivative \(A''_t\) of \(A_t\) exists and is continuous everywhere on \((0, 1)\). Let \(U, V\) and \(Z_t\) be three independent random variables, where \(U, V \sim \mathcal{U}[0, 1]\) and

\[
P(Z_t \leq z) = z + z(1 - z) \frac{A'_t(z)}{A_t(z)}, \quad z \in [0, 1].
\]

Further denote the density of \(Z_t\) by \(f_{Z_t}\) (it exists by the assumption of existing \(A''_t\)) and define the threshold function

\[
p_t(z) = \frac{z(1 - z) A''_t(z)}{f_{Z_t}(z)A_t(z)}, \quad z \in (0, 1).
\]

Under these assumptions, the random vector \((X_t, Y_t)\) from Theorem 1.1 admits the stochastic representation

\[
(X_t, Y_t) \sim \begin{cases} 
\left( \frac{Z_t}{A_t(Z_t)}, \frac{1 - Z_t}{A_t(Z_t)} \right), & \text{if } U > p_t(Z_t) \\
\mathcal{V}\left( \frac{Z_t}{A_t(Z_t)}, \frac{1 - Z_t}{A_t(Z_t)} \right), & \text{if } U \leq p_t(Z_t).
\end{cases}
\]

Our proposed algorithm below provides concrete implementation advice for simulation from \(C_t\). For the simulation of \(C_t\) we use an alternative approach than Theorem 1.2, in particular we do not require \(A''_t\) to exist, but we do make use of the distributional statement (3) in our derivation. The main idea is to use the well-known representation

\[
A_t(x) = 2\mathbb{E}[\max(Qx, (1 - Q)(1 - x))] = 1 + \int_0^x 1 - 2\mathbb{P}(Q \leq 1 - q)\,dq,
\]

which defines a one-to-one relationship between the parameterizing norm \(\ell\) and the distribution of a random variable \(Q \in [0, 1]\) with \(2\mathbb{E}[Q] = 1\). This representation of \(\ell\) in terms of \(Q\) is called Pickands representation in extreme-value theory. The last equality in (4) shows that the right-hand version of the derivative \(A'_t\), which is well-defined everywhere by convexity of \(A_t\) and which we simply denote by \(A'_t\) throughout, essentially determines the distribution function of \(Q\) and, if existing, the second derivative \(A''_t\) equals the density of \(1 - Q\). Whereas the analytical treatment of extreme-value copulas in terms of \(A_t\) is traditional, the article [18] provides some convincing arguments as to why a treatment in terms of the cdf of \(Q\) (i.e., essentially \(A'_t\)) can sometimes be more convenient. The appearance of \(A'_t\) and \(A''_t\) in (3) above thus hints at the possibility to reformulate the simulation algorithm in terms of \(Q\). In fact, we manage to provide a simulation algorithm for \(C_t\) that is exact and feasible whenever the following pre-requisites are met:

**Assumption 1.** (Exact evaluation of \(\ell\)) We can evaluate \(\ell(y, x)\) exactly and efficiently for arbitrary \(x, y \geq 0\).

**Assumption 2.** (Exact evaluation of \(A'_t\)) We can evaluate \(A'_t(x)\) exactly and efficiently for arbitrary \(x \in (0, 1)\).

**Assumption 3.** (Exact simulation of \(Q\)) The random variable \(Q\) can be simulated exactly and efficiently.

Assumptions 1 and 2 are clearly mild, in particular as they are minimal requirements in order to have hope for simulation according to (3) or the inverse Rosenblatt transformation method via partial derivatives to work at all. The decisive question the thoughtful reader should raise at this point is: “Why is it useful to re-formulate the known model (3) in terms of \(Q\)?” To this end, we point out that due to the aforementioned Pickands representation the random variable \(Q\) plays a central role in extreme-value theory, in particular...
the law of \( Q \) is feasible very often, as we will demonstrate below in terms of a comprehensive survey of models. For instance, [4] provides two simulation algorithms for extreme-value copulas precisely under Assumption 3. While the algorithms of [4] are restricted to extreme-value copulas but treat the general case of \( d \)-dimensional extreme-value copulas also for \( d > 2 \), the algorithm we develop is restricted to the bivariate case only but can be leveraged to an algorithm for arbitrary bivariate Archimax copulas. The \( d \)-dimensional generalization remains an interesting open problem, to date only solved for the special case \( \ell = \| \cdot \|_p \), see [15].

The remainder of the article is organized as follows. Section 2 derives our simulation algorithm, while Section 3 demonstrates it by providing examples. Section 4 discusses some interesting aspects of the particular case when \( \ell \) is symmetric, and Section 5 treats the situation of bivariate reciprocal Archimax copulas, in particular indicating that our simulation algorithm for bivariate Archimax copulas immediately implies an exact simulation algorithm for bivariate reciprocal Archimax copulas as well. Section 6 finally concludes.

## 2 Exact simulation from \( C_\ell \) in terms of \( Q \)

We derive an alternative stochastic representation for the random vector \((X_t, Y_t)\) in Theorem 1.1, which serves as convenient basis for simulation under our Assumptions 1–3. One problem with the stochastic representation of Theorem 1.2 is that it requires \( A_\ell^n \) to exist, which is restrictive. Our method will resolve this problem. The other problem is that the probability law of \( Z_\ell \) is rather complicated to simulate in many cases. Our proposed method replaces the requirement to simulate \( Z_\ell \) with the requirement to simulate \( Q \) from the Pickands representation (4), and we demonstrate that this is feasible in many cases.

For the sake of clarity, we recall that the survival function of \((X_t, Y_t)\) is given by \((1 - \ell(x, y)) = C_\ell(1 - x, 1 - y)\), and the distribution function of \((X_t, Y_t)\) equals

\[
\hat{C}_\ell(x, y) = x + y - 1 + C_\ell(1 - x, 1 - y),
\]

which is itself a copula (the survival copula of \( C_\ell \)) and \((1 - X_t, 1 - Y_t) \sim C_\ell\). Our proposed method to simulate \((X_t, Y_t)\) requires the exact simulation of an absolutely continuous random variable on \((0, 1)\) whose density \( h_0 \) depends on a realization of \( Q \) from the Pickands representation (4). As a first auxiliary step, the following lemma states the explicit form of this density. It furthermore provides the basis for an exact and efficient simulation from \( h_0 \) via the so-called rejection–acceptance sampling method, see [12, p. 234–235] for background. The idea of rejection–acceptance sampling for a (possibly complicated) density \( h \) is to find an easy-to-simulate density \( g \), called comparison density, and a finite constant \( c > 0 \) such that \( h \leq cg \). Given this, one may iteratively simulate \( X_g \sim g \) and independent \( U \sim \mathcal{U}[0, 1] \) until finally \( U \leq h(X_g) / (cg(X_g)) \), and then return the resulting sample \( X_g \) satisfying this condition as a sample from \( h \).

**Lemma 2.1.** (Auxiliary density and associated comparison density) Let \( \ell \) be an orthant–monotonic norm on \( \mathbb{R}^2 \) which is normalized to satisfy \( \ell(0, 1) = \ell(1, 0) = 1 \), and fix \( q \in [0, 1] \). With the positive constant

\[
c_q = \begin{cases} 
1 - \ell(\frac{1}{q}, \frac{1}{1-q}), & \text{if } q \in (0, 1) \\
1, & \text{if } q \in \{0, 1\}
\end{cases}
\]

the function

\[
h_q(x) = \frac{1}{c_q} \begin{cases} 
1 - \frac{1}{A_\ell(x)} \left(1 - 2x - \frac{x(1-x)A_\ell'(x)}{A_\ell(x)}\right), & \text{if } x < 1 - q \\
1 + \frac{1}{A_\ell(x)} \left(1 - 2x - \frac{x(1-x)A_\ell'(x)}{A_\ell(x)}\right), & \text{if } x \geq 1 - q
\end{cases}, \quad x \in (0, 1),
\]
defines a probability density on \((0, 1)\). Furthermore, we have that \(c_q h_q \leq d_q g_q\), where \(g_q\) is a piece-wise polynomial probability density on \((0, 1)\) that is given by

\[
g_q(x) = \frac{1}{d_q} [1_{x<1-q} (1 + 2x) + 1_{x \geq 1-q} (3 - 2x)], \quad x \in (0, 1),
\]

with normalizing constant \(d_q = 2(1 - q(1 - q)) > 0\).

**Proof.** We observe that the function

\[
H_q(x) = \frac{1}{c_q} \begin{cases} 
    x - \frac{x(1-x)}{A_q(x)}, & \text{if } x < 1 - q \\
    x + \frac{x(1-x)}{A_q(x)} - \frac{2(1-q)q}{A_q(1-q)}, & \text{if } x \geq 1 - q,
\end{cases}
\]

is a continuous distribution function on \([0, 1]\) for arbitrary \(q \in [0, 1]\) with \(H_q(0) = 0\) and \(H_q(1) = 1\). To verify this, we need to check that \(x \pm x(1-x)/A_q(x)\) is non-decreasing, which follows from the identity

\[
\frac{\partial}{\partial x} \left( x \pm \frac{x(1-x)}{A_q(x)} \right) = 1 \pm \frac{P(Z_q > x) - x}{A_q(x)},
\]

and the bounds \(P(Z_q > x) \in [0, 1]\) as well as \(A_q(x) \geq \max(x, 1-x)\). We further observe easily with similar estimates, using again the similarity of the definition of \(h_q(x)\) and \(P(Z_q \leq x)\), that the density \(h_q = H_q'\) of \(H_q\) is bounded from above by \(d_q g_q / c_q\) for \(g_q\) given in (6).

A random variable \(X_{v_q}\) with density \(g_q\) can be simulated via the inversion method, see [12, p. 234–235] for background, since its cumulative distribution function

\[
G_q(x) = P(X_{v_q} \leq x) = \int_0^x g_q(y) dy, \quad x \in [0, 1],
\]

can be inverted in closed form according to

\[
G_q^{-1}(v) = \begin{cases} 
    -\frac{1}{2} + \frac{1}{2} \sqrt{1 + 8(1-q(1-q))v}, & \text{if } v < \frac{(1-q)(2-q)}{2(1-q(1-q))} \\
    \frac{3}{2} - \frac{1}{2} \sqrt{9 - 8((1-q(1-q))v + q(1-q))}, & \text{else}
\end{cases}
\]

for \(v \in (0, 1)\). Consequently, Lemma 2.1 implies that a random variable \(X_{v_q}\) with density \(h_q\) can be simulated exactly via the rejection–acceptance method under Assumptions 1 and 2. Algorithm 1 summarizes the resulting simulation method for \(X_{h_q} \sim h_q\) based on this idea.

---

**Algorithm 1.** Simulation of \(X_{h_q} \sim h_q\) via rejection–acceptance

1: \hspace{1em} \textbf{procedure} \texttt{SIMULATEH}(q)
2: \hspace{2em} \(U \leftarrow 1\) and \(T \leftarrow 0\)
3: \hspace{2em} \textbf{while} \(U > T\) \textbf{do}
4: \hspace{3em} \textbf{if} \(V < \frac{(1-q)^2-q}{2(1-q(1-q))}\) \textbf{then}
5: \hspace{4em} \(X_{v_q} \leftarrow -\frac{1}{2} + \frac{1}{2} \sqrt{1 + 8(1-q(1-q))V}\)
6: \hspace{3em} \text{\textbf{else}}
7: \hspace{4em} \(X_{v_q} \leftarrow \frac{3}{2} - \frac{1}{2} \sqrt{9 - 8((1-q(1-q))V + q(1-q))}\)


Based on the availability of an efficient simulation of $X_{h_q}$ via Algorithm 1, the following theorem is the main result of the present article.

**Theorem 2.2.** (Exact simulation of $C_t$) Let $t$ be an orthant–monotonic norm on $\mathbb{R}^2$ subject to the normalization $t(0, 1) = t(1, 0) = 1$. Let $V \sim \mathcal{U}[0, 1]$ and $Q$ from the Pickands representation (4) of $A_t$ be independent random variables. Given $Q$ (and independent of $V$), let $X_{h_q}$ be a random variable with density $h_Q$, simulated according to Algorithm 1. The random vector $(X_t, Y_t)$ with survival function $C_t(1 - x, 1 - y)$ and distribution function $\hat{C}_t$ has the stochastic representation

$$(X_t, Y_t) \sim \begin{cases} 
\left( \frac{X_{h_q}}{A_t(X_{h_q})}, \frac{1 - X_{h_q}}{A_t(X_{h_q})} \right), & \text{if } \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) \geq 1 \\
\left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right), & \text{if } \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) < 1.
\end{cases} \tag{7}$$

**Proof.** It is sufficient to prove the claim for absolutely continuous $Q$. Why? An arbitrary $Q$ (possibly with atoms) with $\mathbb{E}[Q] = 1/2$ may be approximated arbitrarily well in distribution with a sequence $Q_n$ of absolutely continuous random variables satisfying $\mathbb{E}[Q_n] = 1/2$. These correspond to a sequence of stable tail dependence functions $\ell_n$ that converge point-wise, and this implies that the respective distributional statement also holds in the limit $t$ associated with $Q$. Note that this logic relies on the fact that convergence in distribution $Q_n \to Q$ stands in one-to-one relation with the point-wise convergence $\ell_n \to \ell$, resp. $C_n \to C_t$.

Under the assumption of absolute continuity of $Q$, $\ell$ is differentiable and it is clear that $\hat{C}_t$ has a singular component on the $\ell$-simplex $\{(x, y) \in [0, 1]^2 : \ell(x, y) = 1\}$, has zero mass outside the simplex and has a density inside the simplex. Consequently, we may compute this density of $\hat{C}_t$ inside the simplex by taking the partial derivative with respect to both components. To this end, we denote by $f_t(q_1, q_2)$ the density of the random vector $(V/2Q, V/2(1 - Q))$ for arbitrary but fixed $V \in (0, 1)$, and obtain for $x, y$ with $\ell(x, y) < 1$

$$\frac{\partial^2}{\partial x \partial y} \hat{C}_t(x, y) = -\frac{\partial^2}{\partial y \partial x} \ell(x, y) = \frac{\partial^2}{\partial x \partial y} \int_0^1 \mathbb{P}(\max\{2Qx, 2(1 - Q)y\} \leq v) dv$$

$$= \int_0^1 \frac{\partial^2}{\partial x \partial y} \mathbb{P}\left(\frac{V}{2Q} > x, \frac{V}{2(1 - Q)} > y\right) dv = \int_0^1 f_t(x, y) dv.$$

Denoting by $p_{\hat{C}_t}$ the mass that $\hat{C}_t$ allocates on the $\ell$-simplex, by $\hat{C}_t^s$ its singular and by $\hat{C}_t^c$ its continuous part, we have the decomposition $\hat{C}_t = (1 - p_{\hat{C}_t})\hat{C}_t^s + p_{\hat{C}_t}\hat{C}_t^c$. We obtain for $x, y \in [0, 1]$ from the above expression...
Furthermore, we obtain the probability $1 - p_{C_t}$ by plugging in $x = y = 1$, which yields

$$1 - p_{C_t} = \mathbb{P} \left( V < \min \left\{ 2Q, 2(1 - Q), \frac{1}{\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)} \right\} \right) = \mathbb{P} \left( V < \frac{1}{\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)} \right)$$

$$= \mathbb{P} \left( \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) < 1 \right) = \mathbb{E} \left[ 1 - \frac{1}{\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)} \right].$$

Note that the last equality is stated only for later reference, and implicitly these computations rely on the fact that $1 \geq \min\{2Q, 2(1 - Q)\} \geq 1/\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)$ almost surely. This already proves the first part of our claimed statement.

For the second part, from Theorem 1.2 we obtain that

$$\hat{\zeta}_t^s (x, y) = \mathbb{P} \left( \frac{Z_t}{\ell (Z_t, 1 - Z_t)} \leq x, \frac{1 - Z_t}{\ell (Z_t, 1 - Z_t)} \leq y \mid U > p_t(Z_t) \right).$$

Using the well-known identity $A_t(z) = 1 - \mathbb{P}(Q \leq 1 - z)$ from (4) and the definition of $p_t$ and $\mathbb{P}(Z_t \leq z)$ in Theorem 1.2, we find

$$\mathbb{E}[1 - p_t(Z_t)] = \mathbb{E} \left[ 1 - \frac{1}{\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)} \right] = \mathbb{P} \left( Q \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) > 1 \right) = p_{C_t},$$

as well as

$$p_{C_t} \mathbb{P}(Z_t \leq z \mid U > p_t(Z_t)) = \mathbb{E} \left[ x + \frac{z(1 - z)}{A_t(z)} \begin{cases} \frac{1}{\ell \left( \frac{1}{2(1 - z)}, 1 \right)}, & \text{if } z \leq 1 - Q \\ \frac{1}{\ell \left( \frac{1}{2Q}, \frac{1}{2(1 - Q)} \right)}, & \text{if } z > 1 - Q \end{cases} \right].$$

In the last expression we recognize the distribution function $H_Q$ from the proof of Lemma 2.1, so that this finally implies

$$\mathbb{P}(Z_t \leq z \mid U > p_t(Z_t)) = \mathbb{E} \left[ \frac{c_t}{p_{C_t}} H_Q(z) \right] = \mathbb{P} \left( X_{h_Q} \leq z \mid \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) > 1 \right),$$

which establishes the claimed statement, when comparing with (3).
As mentioned in Section 1, the simulation of an arbitrary bivariate Archimax copula \( C_{\varphi, \psi} \) requires simulation from \( C_{\psi} \), which we have now established, and, independently, simulation of a random radius \( R_\varphi > 0 \), whose probability distribution is determined uniquely by the Williamson-2-transform \( \varphi(x) = E[(1 - x/R_\varphi),] \). We first note that the distribution function of \( R_\varphi \) is given by the Williamson-inversion formula

\[
P(R_\varphi \leq x) = 1 - \varphi(x) + x\varphi'(x), \tag{8}
\]

where \( \varphi' \) is meant to be the right-hand version of the (by convexity almost everywhere existing) derivative of \( \varphi \). Regarding examples for feasible \( R_\varphi \), we refer to [16]. Generally speaking, rejection-acceptance sampling of \( R_\varphi \) based on (8) is feasible whenever \( \varphi'' \) exists and is analytically feasible and if the then existing density

\[
f_{R_\varphi}(x) = x\varphi''(x)
\]

can be bounded from above by a scalar multiple of an easy-to-simulate density. In contrast, we end this section by providing a simulation algorithm for the special case

\[
\varphi(x) = \begin{cases} 1 - \int_{(0,\infty)} (1 - e^{-sx}) \nu(ds) & \text{, } x \geq 0, \end{cases}
\tag{9}
\]

where \( (1 - e^{-s})\nu(ds) \) is a probability measure on \((0, \infty)\), under the following assumption:

**Assumption 4.** (Requirement to simulate \( R_\varphi \) with generator (9)) The probability measure \( \rho_s(ds) = (1 - e^{-s})\nu(ds) \) can be simulated exactly.

A generator of the form (9) always implies \( R_\varphi \leq 1 \). There are examples for generators \( \varphi \) of that form which are not analytically tractable on first glimpse, in particular for which it is not straightforward to apply direct acceptance-rejection sampling based on the density \( f_{R_\varphi} \), an example is provided below.

The function

\[
\Psi_s(x) = \int_{(0,\infty)} (1 - e^{-sx}) \nu(ds), \quad x \geq 0,
\]

is a so-called Bernstein function satisfying \( \Psi_s(1) = 1 \), in particular is concave, so that \( \varphi(x) = (1 - \Psi_s(x)) \), equals the Williamson-2-transform of some random variable taking values in \((0, 1)\), i.e., \( \varphi \) is a proper Archimedean generator. Every \( R_\varphi \) that arises in this way can be simulated under Assumption 4. Why?

We observe that

\[
P(R_\varphi \leq x) = \Psi_s(x) - x\Psi'_s(x) = \int_{(0,\infty)} \frac{1 - e^{-sx}(1 + xS)}{1 - e^{-s}} \rho_s(ds).
\]

For fixed \( s \in (0, \infty) \), the function \( G_s : [0, 1] \rightarrow [0, 1] \) is the distribution function of a random variable \( R_s \) on \([0, 1] \). Furthermore, simulation of \( R_s \sim G_s \) is efficient and exact via acceptance-rejection, since \( R_s \) takes values in \((0, 1)\), has an atom at 1 with probability \( \exp(-s)/(1 - \exp(-s)) \) and a density \( G_s' \) on \((0, 1)\) that is bounded from above by \( x \mapsto xs^2/(1 - \exp(-s)) \). Summing up, in order to simulate \( R_\varphi \), the following steps need to be carried out:

- Simulate \( S \sim \rho_s \), using Assumption 4.
- Simulate \( B \sim \text{Bernoulli}(\exp(-S), S/(1 - \exp(-S))) \).
- If \( B = 1 \), return \( R_\varphi = 1 \).
- Otherwise, repeatedly simulate \( U, V \sim \mathcal{U}[0, 1] \) independent, until finally \( U \leq e^{-\sqrt{V}S} \), and then return \( R_\varphi = \sqrt{V} \).
There exist some examples for closed-form analytic functions $\Psi$ for which $\Psi''$ (hence the resulting density $f_{R\phi}$) is nasty to deal with, but for which Assumption 4 is still satisfied. For instance, we may extract the following example from [13, Example 3.3]. With parameters $\alpha \in [0, 1)$ and $\theta \geq 0$ we consider the Bernstein function

$$
\Psi(x) = x \frac{\Gamma(x + \theta)\Gamma(2 + \theta - \alpha)}{\Gamma(1 + \theta)\Gamma(x + \theta + 1 - \alpha)}, \quad x \geq 0,
$$

which corresponds to the measure

$$
v(ds) = \frac{\Gamma(2 + \theta - \alpha)}{\Gamma(\theta + 1)\Gamma(1 - \alpha)} \frac{e^{\theta s}}{(1 - e^{-s})^\alpha + \theta(1 - e^{-s})}.\]

The associated random variable $S$ with probability distribution $\rho$ has density given by $(1 - \alpha)g_{\theta,2-\alpha} + \alpha g_{\theta,1,-\alpha}$, where $g_{\theta,a}$ denotes the density of $-\log(B_{a,b})$ and $B_{a,b}$ has Beta distribution with density proportional to $x^{-a}(1 - x)^{b-1}$. The simulation of $S$ is thus straightforward as a mixture of logarithms of Beta distributions, see [12, p. 243] for advice on how to simulate $B_{a,b}$. Apparently, computing the second derivative of $\Psi$ and developing a rejection–acceptance sampling method for $R\phi$ in this example appears to be rather inconvenient, whereas the presented method to simulate $R\phi$ and its associated Archimax copulas only requires to evaluate $\Psi$ (in order to transform margins to uniforms), which is easy.

### 3 Examples for feasible $\ell$

We review some popular families of stable tail dependence functions $\ell$, for which the simulation of the associated $Q$ is viable, thereby demonstrating that our algorithm is feasible in many interesting cases.

#### 3.1 $Q$ corresponding to conditionally iid extreme-value copulas

Many (but not all, see [14]) parametric families of symmetric stable tail dependence functions can be considered jointly under one common umbrella, by observing that they are of the particular form

$$
\ell(x, y) = \ell_F(x, y) = \int_0^\infty 1 - F\left(\frac{u}{x}\right)F\left(\frac{u}{y}\right)du,
$$

where $F : [0, \infty) \to [0, 1]$ equals the distribution function of a non-negative random variable $X_F$ with unit mean. In probabilistic terms, this means that

$$
\ell_F(x, y) = \mathbb{E}[\max(x X_{F,1}, y X_{F,2})], \quad x, y \geq 0,
$$

where $X_{F,1}$, $X_{F,2}$ are iid random variables. These particular stable tail dependence functions are studied in [10], where it is shown that a random vector $(X, Y)$ with survival function $\exp(-\ell_F(x, y))$ is conditionally iid. The probability distribution of $Q = Q_F$ in this case can be simulated by

$$
Q = 1_{[0,\infty)} \frac{Y_F}{Y_F + X_{F,2}} + 1_{[0,1]} \frac{X_{F,1}}{X_{F,1} + Y_F},
$$

where $X_{F,1}, X_{F,2} \sim F$, $B \sim \text{Bernoulli}(1/2)$ and $Y \sim xdf(x)$ are independent random variables. For many popular models, the distribution $F$ as well as its so-called size bias $Y_F$ can be simulated efficiently. We provide some examples below. Note in particular that for both of the following two examples, the Hüsler-Reiss example and the Galambos example, simulation algorithms based on partial derivatives require numerical root search algorithms, while the presented algorithm is exact.
Example 3.1. (Hüsler-Reiss family) With a parameter \( \theta > 0 \) we consider the norm

\[
\ell(x, y) = y \Phi \left[ \theta + \frac{1}{2\theta} \log \left( \frac{y}{x} \right) \right] + x \Phi \left[ \theta + \frac{1}{2\theta} \log \left( \frac{1 - x}{1 - y} \right) \right], \quad x, y \geq 0,
\]

where \( \Phi \) denotes the cdf of a standard normally distributed random variable. This is of the form \( \ell_F \) with \( F(x) = \Phi \left( \log(x) + \theta^2/2 \right)/\theta \) a log-normal distribution. The random variables \( X_F \sim F \) and \( Y_F \sim x dF(x) \) can be simulated according to

\[
X_F \sim e^{\theta + \sqrt{2} \theta W}, \quad Y_F \sim e^{\theta + \sqrt{2} \theta W} W, \quad W \sim \Phi.
\]

The required derivative of \( A'_1 \) is available in closed form, given by

\[
A'_1(x) = \Phi \left[ \theta + \frac{1}{2\theta} \log \left( \frac{x}{1 - x} \right) \right] - \Phi \left[ \theta + \frac{1}{2\theta} \log \left( \frac{1 - x}{x} \right) \right] + \phi \left[ \theta + \frac{1}{2\theta} \log \left( \frac{x}{1 - x} \right) \right] \frac{1 - x}{2\theta x},
\]

where \( \phi(x) = \Phi'(x) = \exp(-x^2/2) / \sqrt{2\pi} \) equals the density of \( W \).

Example 3.2. (Galambos family) With a parameter \( \theta > 0 \) we consider the norm

\[
\ell(x, y) = x + y - (x^\theta + y^\theta)^{1/\theta}, \quad x, y \geq 0.
\]

This is of the form \( \ell_F \) with \( F(x) = 1 - \exp(-\Gamma(1 + 1/\theta)x^\beta) \) a Weibull distribution. (Note that in [10, Example 2] there is a typo, mistaking \( \theta \) for \( 1/\theta \).) The random variable \( X_F \sim F \) can be simulated exactly via the inversion method and \( Y_F \sim x dF(x) \) according to

\[
Y_F \sim \frac{\Gamma^{1/\theta}}{\Gamma(1 + 1/\theta)}, \quad G \sim \Gamma(1 + 1/\theta, 1),
\]

where \( \Gamma(\beta, \eta) \) denotes the Gamma distribution with density proportional to \( x^{\beta-1} \exp(-x/\eta) \). The required derivative of \( A'_1 \) is available in closed form, given by

\[
A'_1(x) = (x^\theta + (1 - x)^\theta)^{1/\theta - 1}((1 - x)^{-\theta - 1} - x^{-\theta - 1}).
\]

Example 3.3. (Cuadras-Augé family) With a parameter \( \theta \in (0, 1] \) we consider the norm

\[
\ell(x, y) = \max\{x, y\} + (1 - \theta) \min\{x, y\} = \theta \|\|x, y\|\|_{\infty} + (1 - \theta) \|\|x, y\|\|_1.
\]

This is of the form \( \ell_F \) with \( F(x) = 1 - \theta + \theta \|\|0, 1\|\|_1\) \( F \), meaning that \( P(X_F = 0) = 1 - \theta = 1 - P(X_F = 1/\theta) \). Furthermore, the associated size bias \( Y_F = 1/\theta \) equals a constant in this case. The required derivative \( A'_1 \) is piece-wise constant, taking the value \( -\theta \) on \( (0, 1/2) \) and the value \( \theta \) on \( [1/2, 1) \). The case of general piece-wise constant \( A'_1 \) is further discussed in Section 3.2.

Example 3.4. (Gumbel family \( \ell = \|\| \)) When \( \ell(x, y) = \|(x, y)\|_p = (x^p + y^p)^{1/p} \) for some \( p > 1 \), then \( \ell = \ell_F \) with associated Fréchet distribution function \( F(x) = \exp(-\Gamma(1 - 1/p)x^\beta) \). In particular, \( X_F \sim F \) may be simulated exactly via the inversion method, and the size bias \( Y_F \) may be simulated according to

\[
Y_F \sim \frac{\Gamma^{-1/p}}{\Gamma(1 - 1/p)}, \quad G \sim \Gamma(1 - 1/p, 1),
\]

where \( \Gamma(\beta, \eta) \) denotes the Gamma distribution with density proportional to \( x^{\beta-1} \exp(-x/\eta) \). The required derivative \( A'_1 \) is known in closed form, to wit

\[
A'_1(x) = (x^p + (1 - x)^p)^{1/p - 1}((1 - x)^{p-1} - x^{p-1}).
\]
In [13,11] it is further explained how to simulate $Q$ associated with convex mixtures of stable tail dependence functions of the type $\xi_F$, and how this can be used to construct interesting and tractable new families. The following is an example in this spirit.

**Example 3.5.** (Holmgren-Liouville family) Let $S$ be some random variable taking values in $(0,1)$. Then the function

$$
\Psi(x) = \frac{1 - \mathbb{E}[(1-S)^x]}{\mathbb{E}[S]}, \quad x \geq 0,
$$

is a so-called Bernstein function. Assume further that $S$ is either absolutely continuous or discrete. In the absolutely continuous case we denote by $f_S$ its density, and in the discrete case we denote by $f_S$ the function $f_S(z) = \mathbb{P}(S = z)$, thus treating both the absolutely continuous case and the discrete case jointly with our notation. Similarly, we consider a random variable $Z$ with associated function

$$
f_z(z) = \frac{1}{\mathbb{E}[1 + S/\log(1-S)]} \left(\frac{1}{z} + \frac{1}{\log(1-z)}\right) f_S(z), \quad z \in (0,1].
$$

This function is bounded from above by

$$
f_z(z) \leq \frac{2f_S(z)}{\mathbb{E}[S]}, \quad z \in (0,1].
$$

Consequently, we can simulate $Z$ exactly via acceptance-rejection, whenever we can simulate $S$, see [12, p. 235–237]. Now we consider

$$
\ell(x,y) = \frac{2\Psi(2)xy}{x+y} - \left(\frac{2xy}{x+y} - \max[x,y]\right) \left(1 - \frac{\min[x,y]}{\max[x,y]}\right),
$$

which is an orthant–monotonic norm with $\ell(1,0) = \ell(0,1) = 1$, hence a stable tail dependence function. This norm arises as a convex probability mixture of norms of the form $\ell_F$ with $F(x) = \min[\exp(x-1),1]$, when randomizing the power $z$, as demonstrated in [13, Example 3.6]. The associated random variable $Q$ can be simulated as follows:

- Simulate $Z$ according to the aforementioned rejection–acceptance method (requires simulation of $S$).
- Simulate $Y_F$ from the continuous density

$$
f_y(x) = \frac{Z e^{-z}}{Z + e^{-z} - 1} x e^{2x}, \quad x \in (0,1),
$$

which can be done via acceptance-rejection with comparison density $g(x) = 2x$.
- Simulate $X_{F,1}, X_{F,2}$ iid from the distribution function $F(x) = \min[\exp(2x - Z), 1]$, which is straightforward by the inversion method.
- Apply formula (10) in order to simulate $Q$ from the simulated $X_{F,1}, X_{F,2}, Y_F$ and a Bernoulli variable $B$.

In order to make this example feasible, the choice of $S$ requires availability of $\Psi$ in closed form, as well as an exact simulation method for $S$. Examples that work certainly include finite-valued $S$, as well as $S$ with beta distribution whose density is proportional to $x \mapsto x^a (1-x)^{b-1}$ for parameters $a, b > 0$. In the latter case, we observe that

$$
\Psi(x) = \frac{a + b}{b} \left(1 - \frac{\beta(a, b + x)}{\beta(a, b)}\right),
$$

where $\beta(\cdot, \cdot)$ denotes the beta function. This example originates from a random vector $(X, Y)$ with survival function $\exp(-\ell(x,y))$ that is defined by $(X, Y) = (L^{-1}_\epsilon, L^{-1}_\nu)$, where $\epsilon, \nu$ are iid standard exponential random variables and, independently, $L = \{L_t\}_{t \geq 0}$ is a so-called Holmgren-Liouville subordinator. The latter is a non-decreasing stochastic process with long-range dependence properties. In particular, $(X, Y)$ are iid conditioned on $L$. For the sake of completeness, we point out that $A_t$ is explicitly given by
\[ A_t'(x) = 2\Psi(2(1-2x)) - \left[(1 - 4 \max[x, 1-x])\Psi\left(1 - \frac{\min[x, 1-x]}{\max[x, 1-x]}\right)\right. \]
\[ + \left. \frac{1 - 2 \max[x, 1-x]}{\max[x, 1-x]}(x - 1)\Psi\left(1 - \frac{\min[x, 1-x]}{\max[x, 1-x]}\right)\right]. \]

In the particular special case with \( S \) having a beta distribution, the derivative \( \Psi' \) of \( \Psi \) in (11) is given in terms of the Digamma function \( D_t(x) = \Gamma'(x)/\Gamma(x) \), which is pre-implemented in most software packages just like the Gamma function \( \Gamma(.) \) and Beta function \( \beta(.) \), to wit
\[
\Psi'(x) = \frac{a + b \beta(a, b + x)}{b \beta(a, b)} (D_t(a + b + x) - D_t(b + x)).
\]

### 3.2 Finite-valued \( Q \), resp. piece-wise constant \( A_t \)

The Pickands dependence function \( A_t \) is piece-wise linear if and only if the associated random variable \( Q \) is discrete. We may approximate an arbitrary \( A_t \) (from above due to convexity) by the function that takes the identical values on the grid \( i/n \) for \( n \geq 1 \) and \( i = 0, \ldots, n \) and which is piece-wise linear in between these grid points. The latter approximation corresponds to some finite-valued \( Q_n \). Since the corresponding approximation \( A_{t,n} \) converges point-wise to \( A_t \) by construction, \( Q_n \) converges in distribution to \( Q \), and also the associated copulas \( C_{t,n} \) approximate \( C_t \) point-wise. From this perspective, the case of finite-valued \( Q \) is of utmost importance.

Regarding the implementation, we recommend to apply the strategy proposed in [18], that is to work on the level of \( A_t' \). To this end, we consider a Pickands dependence function \( A_{t,n} \) that is piece-wise constant between given values \( v_0 = 0 < v_1 < \ldots < v_n = 1 \), with \( n \geq 1 \) arbitrary but fixed. According to the methodology in [18], we parameterize this function in terms of the vectors \( v = (v_0, \ldots, v_n) \) and \( a = (a_1, \ldots, a_n) \) subject to the conditions
\[ -1 \leq a_1 \leq a_2 \leq \ldots \leq a_n \leq 1, \quad \sum_{i=1}^{n} a_i(v_i - v_{i-1}) = 0. \]

Intuitively, we define \( A_{t,n} \) piece-wise constant with \( A_{t,n}'(x) = a_i \) for \( v_{i-1} \leq x < v_i, \ i = 1, \ldots, n \). With \( i(x) = \max\{i \in \{0, \ldots, n\} : v_i \leq x\} \), \( A_{t,n} \) is given by
\[
A_{t,n}(x) = 1 + (x - v_{i(x)})a_{i(x)+1} + \sum_{j=1}^{i(x)} a_j(v_j - v_{j-1}).
\]

With the convenient notation \( a_0 = -1 \), the associated random variable \( Q_n \) has distribution function \( P(Q_n \leq q) = (1 - a_{(1-q)})/2 \). The probability mass function is given by
\[
P(Q_n = 0) = \frac{1 - a_n}{2}, \quad P(Q_n = 1) = \frac{1 + a_1}{2}, \quad P(Q_n = 1 - v_i) = \frac{a_{i+1} - a_i}{2}, \quad i = 1, \ldots, n - 1.
\]

An arbitrary \( A_t \) may now be approximated by a piece-wise constant one of the above form \( A_{t,n} \) with \( v = (0, 1/n, \ldots, (n-1)/n, 1) \) by computing the parameters \( a_1, \ldots, a_n \) iteratively from the relation
\[
A_t \left( \frac{i}{n} \right) = 1 + \frac{1}{n} \sum_{j=1}^{i} a_j, \quad i = 1, \ldots, n,
\]
meaning that
\[ a_i = n \left[ A_t \left( \frac{i}{n} \right) - 1 \right], \quad a_i = n \left[ A_t \left( \frac{i}{n} \right) - 1 - \frac{1}{n} \sum_{j=1}^{i-1} a_j \right], \quad i \geq 2. \]
Since exact simulation of the finite-valued $Q_n$ is possible, like the evaluation of $A_{n,n}$ and $A'_{n,n}$, our algorithm is feasible. Figure 1 visualizes points simulated from $\hat{C}_{t,n}$ when $t$ is from the Hüsler-Reiss family in Example 3.1.

4 The specific case of symmetric norms

From the proof of Theorem 2.2 we recall the decomposition of the survival copula $\hat{C}_t$ in (5) of the Archimax copula $C_t$ into the parts on the $t$-simplex and inside the $t$-simplex. To wit, the distribution function $\hat{C}_t$ of the random vector $(X_t, Y_t)$ in Theorem 2.2 admits the decomposition

$$\hat{C}_t = (1 - p_{C_t})\hat{C}_t^c + p_{C_t}\hat{C}_t^s,$$

where $p_{C_t} = P(t(X_t, Y_t) = 1)$ denotes the probability mass on the $t$-simplex, $\hat{C}_t^s$ denotes the cdf of a singular distribution concentrated on the $t$-simplex and $\hat{C}_t^c$ denotes the cdf of a probability distribution concentrated inside the $t$-simplex. Note that $\hat{C}_t^c$ is only absolutely continuous, if the associated random variable $Q$ from the Pickands representation (4) of $t$ is absolutely continuous, such as assumed in the proof of Theorem 2.2. In the general case, it has singular parts concentrated on straight lines corresponding to atoms of $Q$, as can be seen in Figure 1.

In the special case of symmetric $t$, that is if $t(x, y) = t(y, x)$ for arbitrary $x, y$, there is one phenomenon that we find interesting enough to discuss in the present context. The distribution function $\hat{C}_t^s$ equals the distribution function of a counter-monotonic random vector, meaning that

$$\hat{C}_t^s(x, y) = (G_{1,t}(x) + G_{2,t}(y) - 1), \quad x, y \in [0, 1],$$

with marginal distribution functions $G_{1,t}$ and $G_{2,t}$, respectively. A stochastic representation for $(X_{t}, Y_{t}) \sim \hat{C}_t^s$ is hence well-known to be given by $(G_{1,t}^{-1}(W), G_{2,t}^{-1}(1 - W))$ for $W \sim \mathcal{U}[0, 1]$. Since we know that this random vector is concentrated on the $t$-unit simplex, the norm $t$ determines the implicit relationship between $G_{1,t}$ and $G_{2,t}$ via the equation

$$t(G_{1,t}^{-1}(w), G_{2,t}^{-1}(1 - w)) = 1 \quad \text{for all } w \in [0, 1],$$

so that $G_{1,t}$ determines $G_{2,t}$ and vice versa. Now if $t$ is symmetric, it follows from exchangeability of $\hat{C}_t^s$ that $G_{1,t} = G_{2,t} = G_t$. The distribution function $G_t$ is then uniquely and implicitly determined by $t$ via the equation

$$t(G_t^{-1}(w), G_t^{-1}(1 - w)) = 1 \quad \text{for all } w \in [0, 1].$$

(12)
We furthermore see that
\[ \ell(x, y) = 1 \iff \exists w \in [0, 1] \colon (x, y) = \left( G^{-1}(w), G^{-1}(1-w) \right) \implies G_t(x) + G_t(y) = 1, \]
which provides an alternative way to describe the unit simplex of \( \ell \) in terms of the distribution function \( G_t \).

Summarizing, we obtain a correspondence between symmetric \( \ell \) and convex and continuous distribution functions \( G_t \) on \([0, 1]\).

(An arbitrary distribution function on \([0, 1]\) that is convex is always continuous on \([0, 1]\), but in general may have a jump at 1, which we exclude by saying that \( G_t \) is continuous on all of \([0, 1]\).)

**Lemma 4.1.** (Correspondence between symmetric \( \ell \) and \( G_t \)) For a symmetric stable tail dependence function \( \ell \) the distribution function \( G_t \) is continuous and convex on \([0, 1]\). Conversely, any continuous and convex distribution function on \([0, 1]\) implicitly defines a symmetric stable tail dependence function \( \ell = \ell_G \), which is uniquely determined by its unit simplex \( \{(x, y) \in [0, 1]^2 : G(x) + G(y) = 1\} \).

**Proof.** Continuity of \( G_t \) is immediate, since a jump would imply a non-empty interval \((a, b) \subset (0, 1)\) on which \( G_t^{-1} \) was constant, contradicting the fact that \( w \mapsto (G_t^{-1}(w), G_t^{-1}(1-w)) \) is the closed curve that determines the \( \ell \)-unit simplex. To verify convexity, assume the converse, that is assume \( G_t \) is not convex. Then we necessarily find two distinct \( x, y \in [0, 1] \) and \( a \in (0, 1) \) such that

\[ G_t(ax + (1-a)y) > aG_t(x) + (1-a)G_t(y). \]

By continuity of \( G_t \), we can always choose \( x \) and \( y \) such that the last inequality holds for all \( a \in (0, 1) \) jointly. In particular, for fixed \( a \in (0, 1) \) we also have this inequality for \( 1-a \), that is

\[ G_t(ay + (1-a)x) > aG_t(y) + (1-a)G_t(x). \]

But this implies

\[ G_t(x) + G_t(y) = aG_t(x) + (1-a)G_t(y) + aG_t(y) + (1-a)G_t(x) < G_t(ax + (1-a)y) + G_t(ay + (1-a)x). \]

This implies that

\[ \ell(x, y) < \ell(ax + (1-a)y, ay + (1-a)x), \]

or equivalently in terms of the convex function \( A_t \) that

\[ A_t \left( \frac{x}{x+y} \right) < A_t \left( \frac{ax + (1-a)y}{x+y} \right) \leq aA_t \left( \frac{x}{x+y} \right) + (1-a)A_t \left( \frac{y}{x+y} \right), \]

for all \( a \in (0, 1) \). Reversing the roles of \( x \) and \( y \) in this derivation, we hence obtain that

\[ \max \left\{ A_t \left( \frac{x}{x+y} \right), A_t \left( \frac{y}{x+y} \right) \right\} < aA_t \left( \frac{x}{x+y} \right) + (1-a)A_t \left( \frac{y}{x+y} \right), \]

for all \( a \in (0, 1) \), contradicting convexity of \( A_t \).

Conversely, let \( G \) be a convex and continuous distribution function on \([0, 1]\). Since norms in \( \mathbb{R}^2 \) are uniquely determined by their unit balls, which are in one-to-one correspondence with closed convex sets containing the origin in their interior, all we need to prove is that the closed set

\[ B_G = \{(x, y) \in \mathbb{R}^2 : G(x) + G(y) \leq 1\} \]

is convex, which follows immediately from the convexity of \( G \), as can readily be verified by the readers themselves.

As a consequence, we can improve our simulation algorithm based on the stochastic representation (7) for symmetric \( \ell \) under the following (strong) assumption that is sometimes satisfied:

**Assumption 5.** (Strong additional assumption for symmetric \( \ell \)) We can evaluate exactly the unique implicit function \( G_t^{-1} \) that satisfies (12).

Having the possibility to evaluate \( G_t^{-1} \) explicitly, the stochastic representation (7) in the exchangeable case of symmetric \( \ell \) can be replaced with a simpler expression.
Lemma 4.2. (Simpler sampling algorithm for symmetric \( \ell \)) We consider the same setting as in Theorem 2.2 with the additional assumption that \( \ell(x, y) = \ell(y, x) \) for all \( x, y \geq 0 \), and with Assumption 5 satisfied. We further assume that \( W \sim \mathcal{U}(0, 1) \) is independent of \( V, Q \), which are the same as in Theorem 2.2. Then a random vector \((X, Y) \sim \tilde{C}_\ell\) admits the stochastic representation

\[
(X, Y) \sim \begin{cases} 
(G_\ell^{-1}(W), G_\ell^{-1}(1 - W)), & \text{if } \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) \geq 1 \\
\left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right), & \text{if } \ell \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) < 1.
\end{cases}
\]

Proof. The statement follows from the text preceding this lemma. \( \square \)

We provide two examples for situations in which the implicit definition of \( G_\ell \) is simple to make explicit, hence Assumption 5 is met.

Example 4.3. (The case \( \ell = \| \cdot \|_p \)) Obviously, \( G_\ell^{-1}(w) = w^{1/p} \) does the job. Note that this is well in line with known results about \( \| \cdot \|_p \)-symmetric survival functions in [15].

Example 4.4. (The case of \( Q \) with just two atoms) If we consider for \( q \in [0, 1/2] \) the norm

\[
\ell(x, y) = \max\{qx, (1 - q)y\} + \max\{(1 - q)x, qy\},
\]

it is not difficult to verify that \( G_\ell^{-1}(w) = (q + (1 - 2q)w)/(1 - q) \) does the job. Note that these stable tail dependence functions form the extremal boundary of the simplex of all symmetric bivariate stable tail dependence functions, see [14]. The probability mass of \( \tilde{C}_\ell \) in this case is fully concentrated on the boundary of the triangle with edges \((0, 0), (q/(1 - q), 1) \) and \((0, q/(1 - q))\).

As a final remark, we point out that our simulation algorithm can be utilized to simulate a random variable \( \tilde{S}_G \sim \tilde{C}_\ell \) under Assumptions 1–3 on the norm \( \ell \) only, that is, even without having \( G_\ell \) or even \( G_\ell^{-1} \) in explicit form. To this end, we perform the following steps, again with \( Q \) denoting the random variable associated with \( \ell \) according to the Pickands representation (4):

- Until \( \left( \frac{V}{2Q}, \frac{V}{2(1 - Q)} \right) \geq 1 \), repeatedly simulate independent instances of \( V \) and \( Q \).
- With \( Q \) from the previous step, run Algorithm 1 in order to simulate \( X_{hQ} \).
- Return \( S = X_{hQ}/A_\ell(X_{hQ}) \sim G_\ell \).

In probabilistic terms, \( S \sim G_\ell \) has the same distribution as \( X_{hQ}/A_\ell(X_{hQ}) \), where \( X_{hQ} \sim hQ \) and \( Q \) has distribution given by

\[
P(Q \in dq) = \frac{cQ}{pC_\ell}P(Q \in dq).
\]

5 Bivariate reciprocal Archimax copulas

The presented stochastic model for Archimax copulas can be leveraged to obtain an exact simulation algorithm for bivariate random vectors that are max-infinitely divisible and whose exponent measure has \( \ell \)-symmetric survival function. Their associated copulas take the form (2), so that it is plausible to call them reciprocal Archimax copulas, generalizing the terminology of [7] from the special case \( \ell = \| \cdot \|_p \) to general \( \ell \). We briefly explain this, the logic being identical to the logic outlined in [15, Section 5] for the special case \( \ell = \| \cdot \|_p \). Concretely, a measure \( \mu \) on \( E = [0, \infty)^2 \setminus \{(0, 0)\} \) with

\[
\mu(E \setminus [0, x] \times [0, y]) < \infty \quad \forall x, y > 0, \quad \lim_{x, y \to \infty} \mu(E \setminus [0, x] \times [0, y]) = 0,
\]

about the exact simulation of bivariate (reciprocal) Archimax copulas — 43
is called an exponent measure. We say that \( \mu \) has \( \ell \)-symmetric survival function, if the survival function of \( \mu \) depends on its argument \((x, y)\) only through its norm \((x, y)\), i.e., \( \mu((x, y), (y, x)) = \psi(\ell(x, y)) \) for some function \( \psi \) in one variable. Feasible functions are well-known to comprise such \( \psi : (0, \infty) \rightarrow [0, \infty] \) that are convex with \( \lim_{x \rightarrow 0} \psi(x) = \infty \) and \( \lim_{x \rightarrow \infty} \psi(x) = 0 \), and this is the case if and only if there exists a non-finite Radon measure \( \nu_\psi \) on \((0, \infty)\) with \( \{x\} \mapsto \nu_\psi(x) \) and \( \int_{\mathbb{R}^+} \psi(x) \nu_\psi(dx) = 0 \).

Given a pair \((\ell, \psi)\) of a stable tail dependence function and such convex \( \psi \), defining an exponent measure \( \mu \) with \( \ell \)-norm symmetric survival function, the random vector \((\ell \psi X, \ell \psi Y)\), on \([0, \infty) \times [0, \infty)\), with distribution function

\[
\frac{\psi(x)}{\psi(x) + \psi(y)} = \ell \psi X, \ell \psi Y, 0, 0,
\]

is said to be max-infinitely divisible with exponent measure \( \mu \). Note that \( \ell \psi X \) and \( \ell \psi Y \) both have the identical marginal distribution function \( \{x\} \mapsto \psi(x) \exp \frac{-x}{G_\psi^{-1}(x)} \), so their associated copula is computed to be

\[
C_{\psi, \ell}(x, y) = xy \exp[\psi[\ell(\psi^{-1}[-\log(x)], \psi^{-1}[-\log(y)])] = \frac{x y}{C_{-\psi, \ell}(x, y)}.
\]

An exact simulation algorithm for \( C_{\psi, \ell} \) requires exact simulation of \((\ell \psi X, \ell \psi Y)\), which is feasible under Assumptions 1–3, as well as the following assumption on \( \psi \), respectively, its representing measure \( \nu_\psi \):

**Assumption 6.** (Additional requirement for reciprocal Archimax copulas) The function \( \psi \) and the inverse \( G_{\psi}^{-1} \) of the survival function \( G_{\psi}(x) = \nu_\psi((0, \infty]) \) can be evaluated exactly.

Algorithm 2 provides the pseudo-code for simulation of a sample \((U, V) \sim C_{\psi, \ell}\) under Assumptions 1, 2, 3, and 6, the proof idea being an immediate generalization of [15, Section 5] from \( \| \cdot \|_p \) to general \( \ell \) that the readers can verify themselves. Note that the occurring while-loop certainly ends, since \( \eta = G_{\psi}^{-1}(T) \) decreases to zero due to the non-finiteness of \( \nu_\psi \). In Algorithm 2, the sub-routine \( \text{SimulateExp} \) returns an exponential random variable with unit mean.

Figure 2 depicts scatter plots for two examples of copulas of the form \( C_{\psi, \ell} \). In both plots, the function \( \psi = \psi_a \) with a parameter \( a > 0 \) is chosen to be

\[
\psi_a(x) = a \left[ 1 - \frac{1}{x} \left( 1 - \frac{x \left( \frac{1}{2} + 1 \right) + 1}{2} \right) \right], \quad x > 0,
\]

Algorithm 2. Simulation of \((U, V) \sim C_{\psi, \ell}\)

1: **procedure** \( \text{Simulate} \) \( C_{\psi, \ell}(\psi) \)
2: \((X_0, Y_0) \leftarrow (0, 0)\)
3: \(T \leftarrow \text{SimulateExp}\)
4: \(\eta \leftarrow G_{\psi}^{-1}(T)\)
5: **while** \(\eta > \min\{X_0, Y_0\}\)
6: \(\text{Simulate} (X_0, Y_0) \sim \tilde{C}_{\ell} \) according to (7)
7: \(X_{\ell} \leftarrow \max\{X_0, \eta X_{\ell}\}\)
8: \(Y_{\ell} \leftarrow \max\{Y_0, \eta Y_{\ell}\}\)
9: \(T \leftarrow T + \text{SimulateExp}\)
10: \(\eta \leftarrow G_{\psi}^{-1}(T)\)
11: **end while**
12: **return** \((U, V) = (e^{-\psi(X_{\ell})}, e^{-\psi(Y_{\ell})})\)
13: **end procedure**
corresponding to a discrete measure \( \nu_\psi \) whose required inverse of the survival function equals \( G_\psi^{-1}(x) = \frac{1}{x/a} \), see [15, Example 3].

Finally, we collect some immediate properties of \( C_{x, \psi, t} \). Whereas traditional concordance measures like Spearman’s \( \text{Rho} \) or Kendall’s \( \text{Tau} \) appear to be difficult to compute in convenient form for \( C_{x, \psi, t} \), the following properties are straightforward to observe:

- **Support of \( C_{x, \psi, t} \):** The left and right end points of the support of \( Q \) have an influence on the support of \( C_{x, \psi, t} \). Assume that \( Q \) has support inside the interval \([b_Q, u_Q] \) for either \( b_Q > 0 \) and/or \( u_Q < 1 \). It follows that \( \ell = \frac{x}{b_Q} - \psi^{-1}[-\log(y)] \) on \([b_Q, u_Q] \) and/or \( \ell = \frac{x}{u_Q} - \psi^{-1}[-\log(y)] \) on \([u_Q, 1] \). A simple computation shows that this implies

  \[
  C_{x, \psi, t}(x, y) = \begin{cases}
  x, & \text{if } x \leq \exp\left(-\psi\left(\frac{b_Q - \psi^{-1}[-\log(y)]}{1 - b_Q}\right)\right), \\
  y, & \text{if } y \leq \exp\left(-\psi\left(\frac{1 - u_Q - \psi^{-1}[-\log(x)]}{u_Q}\right)\right).
  \end{cases}
  \]

  so that the copula assigns no mass to these two areas, since the second derivative vanishes. Consequently, if \((x, y)\) is in the support of \( C_{x, \psi, t} \), then necessarily

  \[
  \exp\left(-\psi\left(\frac{1 - u_Q - \psi^{-1}[-\log(x)]}{u_Q}\right)\right) \leq y \leq \exp\left(-\psi\left(\frac{1 - b_Q - \psi^{-1}[-\log(x)]}{b_Q}\right)\right).
  \]

  Note that the lower (upper) bound becomes trivially equal to 0 (equal to 1) if \( u_Q = 1 \) (\( b_Q = 0 \)). The left plot in Figure 2 depicts an example with support truly inside the unit square \([0, 1]^2 \), since \( b_Q = 0.45 = 1 - u_Q \) in this case.

- **Tail dependence:** Just like in case of bivariate Archimax copulas, the coefficients of upper- and lower-tail dependence depend on regular variation properties of the generator \( \psi \) as well as on the value \( \ell(1, 1) \in [1, 2] \), and both lower- and upper-tail dependence can be present or not. To this end, we recall that a real function \( f : \mathbb{R} \rightarrow \mathbb{R} \) is called regularly varying at some point \( x \in (-\infty, \infty] \) with index \( -\rho \) for \( \rho \in (0, \infty) \) if \( \lim_{t \rightarrow x} f(tx)/f(t) = x^{-\rho} \) for each \( x \in \mathbb{R} \), this property being denoted \( f \in \mathcal{R}^\rho_{-\rho} \). With this terminology, we immediately obtain

  \[
  e^\theta \in \mathcal{R}^\rho_{-\rho} \quad \text{for } \rho \in (0, \infty) \quad \Rightarrow \quad \lim_{x \rightarrow 0} \frac{C_{x, \psi, t}(x, x)}{x} = \ell(1, 1)^\rho.
  \]

Figure 2: Scatter plot of 25,000 samples from \( C_{x, \psi, t} \) with \( \psi \) given in (13) for \( a = 1.5 \) and two different \( \ell \) (left: \( \ell \) from Example 4.4 with \( q = 0.45 \), right: \( \ell \) piece-wise constant corresponding to some (asymmetric) \( Q \) with six atoms and \( P(Q = 0), P(Q = 1) > 0 \), i.e., full support).
The limit is called lower-tail dependence coefficient and measures the degree of accumulation of mass in the lower-left corner of $[0, 1]^2$. A concrete example is given by the generator $\psi(x) = -\log(1 - \exp(-x^\theta))$ for some $\theta \geq 1$, for which we observe $\exp(-\psi) \in R_{\geq 1}^\infty$.

Similarly, if $\psi'$ exists on an interval $(T, \infty)$ for some $T > 0$, we obtain

$$\psi' \in R_{\geq 1}^\infty \quad \text{for } p \in (1, \infty) = \lim_{x \to 1} \frac{C_{\psi'}(x, x) - 2x + 1}{1 - x} = \ell(1, 1)^{-p}.$$ 

The limit is called upper-tail dependence coefficient and measures the degree of accumulation of mass in the upper-right corner of $[0, 1]^2$. A concrete example is given by the generator $\psi(x) = x^{-\theta}$ with $\theta > 0$, for which we observe $\psi' \in R_{\geq 1}^\infty$.

6 Conclusion and outlook

We have derived a novel simulation algorithm for bivariate Archimax copulas, demonstrating its feasibility by many examples. Furthermore, the algorithm was used to obtain an exact simulation algorithm also for the family of bivariate reciprocal Archimax copulas. A generalization of the algorithm to the multi-dimensional case $d \geq 3$ remains an interesting open problem, which to date is only solved for the particular case $\ell = ||.||_p$, see [15].

Acknowledgments: The author gratefully acknowledges helpful comments by two anonymous referees on an earlier version of this paper.

Conflict of interest: The author states no conflict of interest.

References


