A Generator of Bivariate Distributions: Properties, Estimation, and Applications

Manuel Franco 1,*,‡, Juana-María Vivo 1,‡ and Debasis Kundu 2,§

1 Department of Statistics and Operations Research, University of Murcia, CEIR Campus Mare Nostrum, IMIB-Arrixaca, 30100 Murcia, Spain; jmvivomo@um.es
2 Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur 208016, India; kundu@iitk.ac.in
* Correspondence: mfranco@um.es

Received: 3 September 2020; Accepted: 8 October 2020; Published: 14 October 2020

Abstract: In 2020, El-Morshedy et al. introduced a bivariate extension of the Burr type X generator (BBX-G) of distributions, and Muhammed presented a bivariate generalized inverted Kumaraswamy (BGIK) distribution. In this paper, we propose a more flexible generator of bivariate distributions based on the maximization process from an arbitrary three-dimensional baseline distribution vector, which is of interest for maintenance and stress models, and expands the BBX-G and BGIK distributions, among others. This proposed generator allows one to generate new bivariate distributions by combining non-identically distributed baseline components. The bivariate distributions belonging to the proposed family have a singular part due to the latent component which makes them suitable for modeling two-dimensional data sets with ties. Several distributional and stochastic properties are studied for such bivariate models, as well as for its marginals, conditional distributions, and order statistics. Furthermore, we analyze its copula representation and some related association measures. The EM algorithm is proposed to compute the maximum likelihood estimations of the unknown parameters, which is illustrated by using two particular distributions of this bivariate family for modeling two real data sets.

Keywords: bivariate distribution generator; copula; reversed hazard gradient; maximum likelihood estimation; EM algorithm; multivariate distribution generator

MSC: 60E05; 62H05; 62H10

1. Introduction

Gumbel [1], Freund [2], and Marshall and Olkin [3] in their pioneering papers developed bivariate exponential distributions. Since then, an extensive amount of work has been done on these models and their different generalizations, which have played a crucial role in the construction of multivariate distributions and modeling in a wide variety of applications, such as physic, economy, biology, health, engineering, computer science, etc. Several continuous bivariate distributions can be found in Balakrishnan and Lai [4], and some generalizations and multivariate extensions have been studied by Franco and Vivo [5], Kundu and Gupta [6], Franco et al. [7], Gupta et al. [8], Kundu et al. [9], among others, and recently by Muhammed [10], Franco et al. [11], and El-Morsheedy et al. [12], also see the references cited therein.

Kundu and Gupta [13] introduced a bivariate generalized exponential (BGE) distribution by using the trivariate reduction technique with generalized exponential (GE) random variables, which is based on the maximization process between components with a latent random variable, suitable for modeling of some stress and maintenance models. This procedure has also been applied in the
literature to generate other bivariate distributions, for example, the bivariate generalized linear failure rate (BGLFR) given by Sarhan et al. [14], the bivariate log-exponentiated Kumaraswamy (BlogEK) introduced by Elsherpieny et al. [15], the bivariate exponentiated modified Weibull extension (BEMWE) given by El-Gohary et al. [16], the bivariate inverse Weibull (BIW) studied by Muhammed [17] and Kundu and Gupta [18], the bivariate Dagum (BD) provided by Muhammed [19], the bivariate generalized Rayleigh (BGR) depicted by Sarhan [20], the bivariate Gumbel-G (BGu-G) presented by Eliawi and El-Morshedy [21], the bivariate generalized inverted Kumaraswamy (BGIK) given by Muhammed [10], and the bivariate Burr typeX-G (BBX-G) proposed by El-Morshedy et al. [12]. Some associated inferential issues have been discussed in these articles, and all of them are based on considering the same kind of baseline components. In each of these bivariate models, the baseline components belong to the proportional reversed hazard rate (PRH) family with a certain underlying distribution (Gupta et al. [22] and Di Crescenzo [23]). It is worth mentioning that Kundu and Gupta [24] extended the BGE model by using components within the PRH family, called a bivariate proportional reversed hazard rate (BPRH) family, and a multivariate extension of the BPRH model was studied by Kundu et al. [9].

The main aim of this paper is to provide a more flexible generator of bivariate distributions based on the maximization process from an arbitrary three-dimensional baseline continuous distribution vector, i.e., not necessarily identical continuous distributions. Hence, this proposed generator allows researchers and practitioners to generate new bivariate distributions even by combining non-identically distributed baseline components, which may be interpreted as a stress model or as a maintenance model. We refer to the bivariate models from this generator as the generalized bivariate distribution (GBD) family, which contains as special cases the aforementioned bivariate distributions. Note that a two-dimensional random variable \((X_1, X_2)\), belonging to the GBD family, has dependent components due to a latent factor, and its joint cumulative distribution function (cdf) is not absolutely continuous, i.e., the joint cdf is a mixture of an absolutely continuous part and a singular part due to the positive probability of the event \(X_1 = X_2\), whereas the line \(x_1 = x_2\) has two-dimensional Lebesgue measure zero. In general, the maximum likelihood estimation (MLE) of the unknown parameters a GBD model cannot be obtained in closed form, and we propose using an EM algorithm to compute the MLEs of such parameters.

The rest of the paper is organized as follows. The construction of the GBD family is given in Section 2, and we obtain its decomposition in absolutely continuous and singular parts and its joint probability density function (pdf). In Section 3, several special bivariate models are presented. The cdf and pdf of the marginals and conditional distributions are derived in Section 4, as well as for its order statistics. Some dependence and two-dimensional ageing properties for the GBD family, and stochastic properties of their marginals and order statistics are studied in Section 5, as well as its copula representation and some related association measures. The EM algorithm is proposed in Section 6, which is applied in Section 7, for illustrative purposes, to find the MLEs of particular models of the GBD family in the analysis of two real data sets. Finally, the multivariate extension is discussed in Section 8, as well as the concluding remarks. Some of the proofs are relegated to Appendix A for a fluent presentation of the results, and some technical details of the applications can be found in Appendix B.

2. The GBD Family

In this section, we define the generalized bivariate distribution family as a generator system from any three-dimensional baseline continuous distribution, and then we provide its joint cdf, decomposition, and joint pdf.

Let \(U_1, U_2,\) and \(U_3\) be mutually independent random variables with any continuous distribution functions \(F_{U_1}, F_{U_2}\) and \(F_{U_3}\), respectively. Let \(X_1 = \max(U_1, U_3)\) and \(X_2 = \max(U_2, U_3)\). Then, the random vector \((X_1, X_2)\) is said to be a GBD model with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\).
**Theorem 1.** Let \((X_1, X_2)\) be a GBD model with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\), then its joint cdf is given by

\[
F(x_1, x_2) = F_{U_1}(x_1)F_{U_2}(x_2)F_{U_3}(z),
\]

where \(z = \min(x_1, x_2)\), for all \(x_1, x_2 \in \mathbb{R}\).

**Proof.** It is immediate since

\[
F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = P(\max(U_1, U_3) \leq x_1, \max(U_2, U_3) \leq x_2)
\]

\[
= P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq \min(x_1, x_2)) = F_{U_1}(x_1)F_{U_2}(x_2)F_{U_3}(\min(x_1, x_2)).
\]

\[\square\]

For instance, a stress model may lead to the GBD family, as in Kundu and Gupta [13]. Suppose a two-component system where each component is subject to an individual independent stress, say \(U_1\) and \(U_2\), respectively. The system has an overall stress \(U_3\) which has been equally transmitted to both the components, independent of their individual stresses. Then, the observed stress for each component is the maximum of both, individual and overall stresses, i.e., \(X_1 = \max(U_1, U_3)\) and \(X_2 = \max(U_2, U_3)\), and \((X_1, X_2)\) is a GBD model.

Analogously, a GBD model is also plausible for a maintenance model. Suppose a system has two components, and it is assumed that each component has been maintained independently and there is also an overall maintenance. Due to component maintenance, the lifetime of the individual component is increased by a random time, say \(U_1\) and \(U_2\) respectively, and, because of the overall maintenance, the lifetime of each component is increased by another random time \(U_3\). Then, the increased lifetime of each component is the maximum of both, individual and overall maintenances, \(X_1 = \max(U_1, U_3)\) and \(X_2 = \max(U_2, U_3)\), respectively.

As mentioned before, a bivariate model belonging to the GBD family does not have an absolutely continuous cdf. Let us see now the decomposition of a GBD model as a mixture of bivariate absolutely continuous and singular cdfs, the proof is provided in Appendix A.

**Theorem 2.** Let \((X_1, X_2)\) be a GBD model with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\). Then,

\[
F(x_1, x_2) = \alpha F_s(x_1, x_2) + (1 - \alpha)F_{ac}(x_1, x_2)
\]

where

\[
F_s(x_1, x_2) = \frac{1}{\alpha} \int_{-\infty}^{z} F_{U_1}(u)F_{U_2}(u)dF_{U_3}(u)
\]

and

\[
F_{ac}(x_1, x_2) = \frac{1}{1 - \alpha} \left( F_{U_1}(x_1)F_{U_2}(x_2)F_{U_3}(z) - \int_{-\infty}^{z} F_{U_1}(u)F_{U_2}(u)dF_{U_3}(u) \right)
\]

with \(z = \min(x_1, x_2)\), are the singular and absolutely continuous parts, respectively, and

\[
\alpha = \int_{-\infty}^{\infty} F_{U_1}(u)F_{U_2}(u)dF_{U_3}(u).
\]

In addition, due to the singular part \(F_s\) in (2), the GBD family does not have a pdf with respect to the two-dimensional Lebesgue measure even when the distribution functions \(F_{U_1}, F_{U_2},\) and \(F_{U_3}\) are absolutely continuous. However, it is possible to construct a joint pdf for \((X_1, X_2)\) through a mixture between a pdf with respect to the two-dimensional Lebesgue measure and a pdf with respect to the one-dimensional Lebesgue measure (the proof is provided in Appendix A).
Theorem 3. If \((X_1, X_2)\) is a GBD model with joint cdf given by (1), then the joint pdf with respect to \(\mu\), the measure associated with \(F\), is

\[
f(x_1, x_2) = \begin{cases} 
    f_1(x_1, x_2), & \text{if } x_1 < x_2 \\
    f_2(x_1, x_2), & \text{if } x_1 > x_2 \\
    f_0(x), & \text{if } x_1 = x_2 = x,
\end{cases}
\]

where

\[
f_i(x_1, x_2) = f_{U_i}(x_i) \left( f_{U_i}(x_i) F_{U_3}(x_i) + F_{U_i}(x_i) f_{U_3}(x_i) \right), \text{ with } i \neq j \in \{1, 2\},
\]

and

\[
f_0(x) = f_{U_0}(x) F_{U_1}(x) F_{U_2}(x),
\]

when the pdf \(f_{U_i}\) of \(U_i\) exists, \(i = 1, 2, 3\).

3. Special Cases

In this section, we derive new bivariate models from Theorem 1, taking into account particular baseline distribution vectors \(\left(F_{U_1}, F_{U_2}, F_{U_3}\right)\).

Note that, if the baseline components \(U_i\)s belong to the same distribution family, say \(F_{U_i}\), then the proposed generator provides novel extended bivariate versions of that distribution \(F_{U_i}\). Furthermore, under certain restrictions on the underlying parameters of each \(U_i\), bivariate distributions given in the literature are obtained. From now on, it is assumed that all parameters of each \(F_{U_i}\) are positive unless otherwise mentioned.

**Extended bivariate generalized exponential model.** A random variable \(U\) follows a GE distribution, \(U \sim GE(\theta, \lambda)\) (see Gupta and Kundu [25]), if its cdf is given by

\[
F_{GE}(u; \theta, \lambda) = \left(1 - e^{-\lambda u}\right)^\theta, \text{ for } u > 0.
\]

If \(U_i \sim GE(\theta_i, \lambda_i)\) \(i = 1, 2, 3\), then the GBD model with the GE baseline distribution vector is an extended BGE model with \(\theta = (\theta_1, \theta_2, \theta_3)\) and \(\lambda = (\lambda_1, \lambda_2, \lambda_3)\) parameter vectors, denoted as \((X_1, X_2) \sim EBGE(\theta, \lambda)\), and its joint cdf is

\[
F_{EBGE}(x_1, x_2) = F_{GE}(x_1; \theta_1, \lambda_1) F_{GE}(x_2; \theta_2, \lambda_2) F_{GE}(z; \theta_3, \lambda_3), \text{ for } x_1 > 0, x_2 > 0,
\]

where \(z = \min(x_1, x_2)\).

As a particular case, if \(\lambda = \lambda_i, i = 1, 2, 3\), \((X_1, X_2) \sim BGE(\theta, \lambda)\) given by Kundu and Gupta [13].

**Extended bivariate proportional reversed hazard rate model.** If \(U_i \sim PRH(\theta_i)\) with base distribution \(F_B\) \(i = 1, 2, 3\), i.e., its cdf can be expressed as \(F_{U_i} = F_{B_i}^\theta\) (see Gupta et al. [22] and Di Crescenzo [23]), then the GBD model with PRH baseline distribution vector provides an extended BPRH model, \((X_1, X_2) \sim EBPRH(\theta, \lambda)\), with \(\theta = (\theta_1, \theta_2, \theta_3)\) parameter vector of the PRH components and \(\lambda = (\lambda_1, \lambda_2, \lambda_3)\) parameter vector of the underlying distributions \(F_B\)’s. From (1), its joint cdf is given by

\[
F_{EBPRH}(x_1, x_2) = F_{B_1}^{\theta_1}(x_1; \lambda_1) F_{B_2}^{\theta_2}(x_2; \lambda_2) F_{B_3}^{\theta_3}(z; \lambda_3), \text{ for } x_1 > 0, x_2 > 0,
\]

where \(z = \min(x_1, x_2)\).

In particular, if the PRH components have the same base distribution, \(F_B = F_B\) \(i = 1, 2, 3\), then \((X_1, X_2) \sim BPRH(\theta, \lambda)\) with baseline distribution \(F_B(\cdot; \lambda)\) introduced by Kundu and Gupta [24].

**Extended bivariate generalized linear failure rate model.** It is said that a random variable \(U\) follows a GLFR distribution, \(U \sim GLFR(\theta, \lambda, \gamma)\) (see Sarhan and Kundu [26]), if its cdf is given by

\[
F_{GLFR}(u; \theta, \lambda, \gamma) = \left(1 - \exp\left(-\lambda u - \gamma \frac{u^2}{2}\right)\right)^\theta, \text{ for } u > 0.
\]
If $U_i \sim GLFR(\theta_i, \lambda_i, \gamma_i)$ $i = 1, 2, 3$, then the GBD model with GLFRs baseline distribution vector is an extended BGLFR model, $(X_1, X_2) \sim EBGLFR(\theta, \lambda, \gamma)$, with parameters $\theta = (\theta_1, \theta_2, \theta_3)$, $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, having joint cdf

$$F_{EBGLFR}(x_1, x_2) = F_{GLFR}(x_1; \theta_1, \lambda_1, \gamma_1)F_{GLFR}(x_2; \theta_2, \lambda_2, \gamma_2)F_{GLFR}(z; \theta_3, \lambda_3, \gamma_3),$$

for $x_1 > 0, x_2 > 0$,

where $z = \min(x_1, x_2)$.

When $\lambda_i = \lambda$ and $\gamma_i = \gamma$, $i = 1, 2, 3$, it is obtained that $(X_1, X_2) \sim BGLFR(\theta, \lambda, \gamma)$ given by Sarhan et al. [14].

**Extended bivariate log-exponentiated Kumaraswamy model.** Let $U$ be a random variable with logEK distribution, $U \sim logEK(\theta, \lambda, \gamma)$ (see Lemonte et al. [27]), then its cdf

$$F_{logEK}(u; \theta, \lambda, \gamma) = \left(1 - \left(1 - e^{-u}\right)^{\lambda}\right)^{\gamma}, \text{ for } u > 0.$$

If $U_i \sim logEK(\theta_i, \lambda_i, \gamma_i)$ $i = 1, 2, 3$, then the GBD model with logEKS baseline distribution vector is an extended BlogEK model, $(X_1, X_2) \sim EBlogEK(\theta, \lambda, \gamma)$ with parameters $\theta = (\theta_1, \theta_2, \theta_3)$, $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, and its joint cdf is given by

$$F_{EBlogEK}(x_1, x_2) = F_{logEK}(x_1; \theta_1, \lambda_1, \gamma_1)F_{logEK}(x_2; \theta_2, \lambda_2, \gamma_2)F_{logEK}(z; \theta_3, \lambda_3, \gamma_3),$$

for $x_1 > 0, x_2 > 0$,

where $z = \min(x_1, x_2)$.

Clearly, it can be seen that $(X_1, X_2) \sim BlogEK(\theta, \lambda, \gamma)$ given by Elsheapieny et al. [15], when $\lambda = \gamma$, $i = 1, 2, 3$.

**Extended bivariate exponentiated modified Weibull extension model.** A random variable $U$ follows an EMWE distribution, $U \sim EMWE(\theta, a, \beta, \lambda)$ (see Sarhan and Apaloo [28]), if its cdf can be expressed as

$$F_{EMWE}(u; \theta, a, \beta, \lambda) = \left(1 - \exp \left( a \lambda \left(1 - e^{(u/a)\beta}\right) \right) \right)^{\theta}, \text{ for } u > 0.$$

If $U_i \sim EMWE(\theta_i, a_i, \beta_i, \lambda_i)$ $i = 1, 2, 3$, then the GBD model with EMWEs baseline distribution vector is an extended BEMWE model, $(X_1, X_2) \sim EBEMWE(\theta, a, \beta, \lambda)$ with $\theta = (\theta_1, \theta_2, \theta_3)$ and $a = (a_1, a_2, a_3)$, $\beta = (\beta_1, \beta_2, \beta_3)$, and $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ parameter vectors, and its joint cdf is given by

$$F_{EBEMWE}(x_1, x_2) = F_{EMWE}(x_1; \theta_1, a_1, \beta_1, \lambda_1)F_{EMWE}(x_2; \theta_2, a_2, \beta_2, \lambda_2)F_{EMWE}(z; \theta_3, a_3, \beta_3, \lambda_3),$$

for $x_1 > 0$ and $x_2 > 0$, where $z = \min(x_1, x_2)$.

Note that, if $a_i = a$, $\beta_i = \beta$ and $\lambda_i = \lambda$, $i = 1, 2, 3$, then $(X_1, X_2) \sim BEMWE(\theta, a, \beta, \lambda)$ given by El-Gohary et al. [16].

**Extended bivariate inverse Weibull model.** The cdf of the IW distribution (e.g., see Keller et al. [29]) is defined by

$$F_{IW}(u; \theta, \lambda) = e^{-\lambda u^{-\theta}}, \text{ for } u > 0.$$

If $U_i \sim IW(\theta_i, \lambda_i)$ $i = 1, 2, 3$, then the GBD model with IWs baseline distribution vector is an extended BIW model with $\theta = (\theta_1, \theta_2, \theta_3)$ and $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ parameter vectors, denoted as $(X_1, X_2) \sim EBIW(\theta, \lambda)$, and its joint cdf can be written as

$$F_{EBIW}(x_1, x_2) = e^{-\lambda_1 x_1^{-\theta_1} - \lambda_2 x_2^{-\theta_2} - \lambda_3^{-\theta_3}}, \text{ for } x_1 > 0, x_2 > 0,$$

where $z = \min(x_1, x_2)$. 
In particular, \((X_1, X_2) \sim BIW(\theta, \lambda)\) studied by Muhammed [17] and Kundu and Gupta [18], when \(\theta_i = \theta\) for \(i = 1, 2, 3\).

Extended bivariate Dagum model. It is said that a random variable \(U\) follows a Dagum distribution [30], \(U \sim D(\theta, \lambda, \gamma)\), if its cdf is given by

\[
F_D(u; \theta, \lambda, \gamma) = (1 + \lambda u^{-\gamma} - \gamma)^{-\theta}, \quad \text{for} \ u > 0.
\]

If \(U_i \sim D(\theta_i, \lambda_i, \gamma_i)\) \(i = 1, 2, 3\), then the GBD model with Dagum baseline distribution vector is an extended BD model with \(\theta = (\theta_1, \theta_2, \theta_3), \lambda = (\lambda_1, \lambda_2, \lambda_3)\) and \(\gamma = (\gamma_1, \gamma_2, \gamma_3)\) parameter vectors, denoted as \((X_1, X_2) \sim EBD(\theta, \lambda, \gamma)\), having joint cdf

\[
F_{EBD}(x_1, x_2) = F_D(x_1; \theta_1, \lambda_1, \gamma_1)F_D(x_2; \theta_2, \lambda_2, \gamma_2)F_D(z; \theta_3, \lambda_3, \gamma_3), \quad \text{for} \ x_1 > 0, x_2 > 0,
\]

where \(z = \min(x_1, x_2)\).

Note that, when \(\lambda_i = \lambda\) and \(\gamma_i = \gamma\) for \(i = 1, 2, 3\), it is simplified to the model \((X_1, X_2) \sim BD(\theta, \lambda, \gamma)\) defined by Muhammed [19].

Extended bivariate generalized Rayleigh model. The cdf of the GR distribution, also called Burr type X model [31], is

\[
F_{GR}(u; \theta, \lambda) = \left(1 - e^{-(\lambda u)^2}\right)^{\theta}, \quad \text{for} \ u > 0.
\]

If \(U_i \sim GR(\theta_i, \lambda_i)\) \(i = 1, 2, 3\), then the GBD model with a GR baseline distribution vector is an extended BGR model with \(\theta = (\theta_1, \theta_2, \theta_3)\) and \(\lambda = (\lambda_1, \lambda_2, \lambda_3)\) parameter vectors, \((X_1, X_2) \sim EBG(\theta, \lambda)\), with joint cdf

\[
F_{EBGR}(x_1, x_2) = F_{GR}(x_1; \theta_1, \lambda_1)F_{GR}(x_2; \theta_2, \lambda_2)F_{GR}(z; \theta_3, \lambda_3), \quad \text{for} \ x_1 > 0, x_2 > 0,
\]

where \(z = \min(x_1, x_2)\).

Hence, if \(\lambda_i = \lambda, i = 1, 2, 3\), it is obtained that \((X_1, X_2) \sim BGR(\theta, \lambda)\) given by Sarhan [20].

Extended bivariate Gumbel-G model. Alzaatrech et al. [32] proposed a transformed-transformer method for generating families of continuous distributions. From such method, it is said that a random variable \(U\) follows a Gumbel-G model, \(U \sim GU-G(\theta, \alpha, \lambda)\) if its cdf can be expressed as

\[
F_{GU-G}(u; \theta, \alpha, \lambda) = \exp \left(-\theta \left(\frac{1 - G(u; \lambda)}{G(u; \lambda)}\right)^{\alpha}\right), \quad \text{for} \ u > 0
\]

where \(G\) is the transformer distribution with parameter vector \(\lambda\). If \(U_i \sim GU-G(\theta_i, \alpha_i, \lambda_i)\) \(i = 1, 2, 3\), then the GBD model with Gu-Gs baseline distribution vector is an extended BGu-G model, \((X_1, X_2) \sim EBG(\theta, \alpha, \lambda_G)\), with parameters \(\theta = (\theta_1, \theta_2, \theta_3), \alpha = (\alpha_1, \alpha_2, \alpha_3), \) and \(\lambda_G = (\lambda_1, \lambda_2, \lambda_3)\), where \(\lambda_G\) encompasses all parameter vectors of \(G\) in each baseline component. Thus, its joint cdf is given by

\[
F_{EBGu-G}(x_1, x_2) = F_{Gu-G}(x_1; \theta_1, \alpha_1, \lambda_1)F_{Gu-G}(x_2; \theta_2, \alpha_2, \lambda_2)F_{Gu-G}(z; \theta_3, \alpha_3, \lambda_3), \quad \text{for} \ x_1 > 0, x_2 > 0, \text{ where } z = \min(x_1, x_2).
\]

In particular, when \(\alpha_i = \alpha\) and \(\lambda_i = \lambda\) for \(i = 1, 2, 3\), \((X_1, X_2) \sim BGu-G(\theta, \alpha, \lambda)\) given by Eliwa and El-Morsheyed [21].

Extended bivariate generalized inverted Kumaraswamy model. A random variable \(U\) is said to be a GIK distribution defined by Iqbal et al. [33], if its cdf is given by

\[
F_{GIK}(u; \theta, \alpha, \gamma) = (1 - (1 + u^\gamma)^{-\alpha})^\theta, \quad \text{for} \ u > 0.
\]
If \( U_i \sim GIK(\theta_i, \alpha_i, \gamma_i) \) \( i = 1, 2, 3 \), then the GBD model with GIKs baseline distribution vector is an extended BG1K model, \((X_1, X_2) \sim EBG1K(\theta, \alpha, \gamma)\), with parameters \( \theta = (\theta_1, \theta_2, \theta_3) \), \( \alpha = (\alpha_1, \alpha_2, \alpha_3) \), and \( \gamma = (\gamma_1, \gamma_2, \gamma_3) \), and its joint cdf can be written as

\[
F_{EBG1K}(x_1, x_2) = F_{G1K}(x_1; \theta_1, \alpha_1, \gamma_1)F_{G1K}(x_2; \theta_2, \alpha_2, \gamma_2)F_{G1K}(z; \theta_3, \alpha_3, \gamma_3), \quad \text{for } x_1 > 0, x_2 > 0,
\]

where \( z = \min(x_1, x_2) \).

It is straightforward to see that \((X_1, X_2) \sim BGIK(\theta, \alpha, \gamma)\) analyzed by Muhammed [10] when \( \alpha = \alpha_i \) and \( \gamma = \gamma_i \) for \( i = 1, 2, 3 \).

**Extended bivariate Burr type X-G model.** From the transformed-transformer method of Alzaatreh et al. [32], it is said that a random variable \( U \) follows a Burr X-G model, \( U \sim BX-G(\theta, \lambda) \) if its cdf can be expressed as

\[
F_{BX-G}(u; G, \theta, \lambda) = \left(1 - \exp \left(- \frac{G(u; \lambda)}{1 - G(u; \lambda)} \right)^2 \right)^\theta, \quad \text{for } u > 0
\]

where \( \lambda \) is the parameter vector of the transformer distribution \( G \).

If \( U_i \sim BX-G(\theta_i, \lambda_i) \) \( i = 1, 2, 3 \), then the GBD model with BX-Gs baseline distribution vector is an extended BBX-G model, \((X_1, X_2) \sim EBBX-G(\theta, \lambda_G)\), with parameters \( \theta = (\theta_1, \theta_2, \theta_3) \), and \( \lambda_G = (\lambda_1, \lambda_2, \lambda_3) \), where \( \lambda_G \) encompasses all parameter vectors of \( G \) in each baseline component. Then, its joint cdf can be expressed as

\[
F_{EBBX-G}(x_1, x_2) = F_{BX-G}(x_1; \theta_1, \lambda_1)F_{BX-G}(x_2; \theta_2, \lambda_2)F_{BX-G}(z; \theta_3, \lambda_3), \quad \text{for } x_1 > 0, x_2 > 0,
\]

where \( z = \min(x_1, x_2) \).

In particular, if \( \lambda = \lambda_i \) for \( i = 1, 2, 3 \), then \((X_1, X_2) \sim BBX-G(\theta, \lambda)\) introduced by El-Morshedey et al. [12].

**GBD models from different baseline components.** In addition, a GBD model can be derived from baseline components \( U_i \)s belonging to different distribution families, which allows one to generate new bivariate distributions.

For illustrative purposes, Figure 1a–d display 3D surfaces of different joint pdfs given by Theorem 3, along with their contour plots. Here, \( U_1 \) and \( U_2 \) are taken identically distributed \( GE(\theta, \lambda) \) with different shape and scale parameter values, and \( U_3 \) having a Weibull distribution with scale parameter \( \lambda_3 \) and shape parameter \( \alpha = 6, W(\lambda_3, 6) \).

Figure 1 shows that some of these GBD models are multi-modal bivariate models. It indicates a variety of shapes for the GBD family depending on the different baseline distribution components and for different parameter values.

![Figure 1.](image-url)
(c) \( U_1, U_2 \sim GE(6,1) \) and \( U_3 \sim W(1,6) \).

(d) \( U_1, U_2 \sim GE(3,0.5) \) and \( U_3 \sim W(0.5,6) \).

**Figure 1.** Surface and contour plots of the joint pdf of GBD models \((X_1, X_2)\) with different components \((U_1, U_2, U_3)\).

4. Distributional Properties

Here, we derive the marginal and conditional distributions of the GBD family, and the order statistics. Furthermore, some properties for particular baseline distribution vectors are provided.

4.1. Marginal and Conditional Distributions

From Theorem 1, it is easy to obtain the marginal cdfs of the components \(X_i\)'s, which can be written as

\[
F_{X_i}(x_i) = F_{U_i}(x_i) F_{U_3}(x_i), \quad \text{with } i = 1, 2, \tag{5}
\]

and, when the pdf \(f_{U_i}\) of \(U_i\) exists, \(i = 1, 2, 3\), the corresponding pdfs are given by

\[
f_{X_i}(x_i) = f_{U_i}(x_i) F_{U_3}(x_i) + F_{U_i}(x_i) f_{U_3}(x_i), \quad \text{with } i = 1, 2. \tag{6}
\]

For instance, we shall now suppose that \(U_i\)s have PRH distributions, in order to provide some preservation results of the PRH property on the marginals, and its closure under exponentiation of the underlying distributions.

**Proposition 1.** If \((X_1, X_2)\) has a GBD model formed by \(U_i \sim PRH(\theta_i)\) with baseline distribution \(F_{B_i}\) \((i = 1, 2, 3)\), then \(X_i \sim PRH(\theta_i + \theta_3)\) with base distribution \(F_{B_i} = F_{B_i}^{\theta_i/\theta_3} F_{B_3}^{\theta_3/\theta_i} F_{B_3}^{\theta_3/\theta_3} F_{B_3}^{\theta_3/\theta_3} F_{B_3}^{\theta_3/\theta_3}\). Moreover, when the base distribution is common, \(F_{B} = F_{B_1}\), then \(X_i\)s also have the same base distribution \(F_{B}\).

**Proof.** It immediately follows from (5) and the EBPRH model, since \(F_{U_i} = F_{B_i}^{\theta_i/\theta_3}\).

**Corollary 1.** If \(U_i \sim PRH(\theta_i)\) with base distribution \(F_{B_i}\), having \(F_{B} \sim PRH(\lambda_i)\) with base distribution \(F_{B_i} = F_{B_i}^{\theta_i/\theta_i + \lambda_3} F_{B_3}^{\theta_3/\theta_i + \lambda_3} F_{B_3}^{\theta_3/\theta_i + \lambda_3}\). Moreover, if \(F_{B} = F_{B_i}\) \((i = 1, 2, 3)\), then \(X_i\)s also have the same base distribution \(F_{B}\).

In addition, Figure 2 displays the plots of the marginal pdfs of the GBD models depicted in Figure 1a–d.

Note that Figure 2a–d show some bimodal shapes for the marginal pdfs given by (6) of the GBD models represented in Figure 1a–d, which also exhibit some multi-modal shapes of the joint pdfs.

In this setting, Proposition 1 might be used to generate bimodal distributions from the marginals of the GBD family by mixing different baseline distribution components as in Figure 1.
Furthermore, we provide some results about the conditional distributions of a GBD model whose proof can be found in Appendix A.

**Theorem 4.** If \((X_1, X_2)\) has a GBD model with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\), then

1. The conditional distribution of \(X_i\) given \(X_j \leq x_j (i \neq j)\), say \(F_{i|X_j \leq x_j}(x_i)\), is an absolutely continuous cdf given by
   \[
   F_{i|X_j \leq x_j}(x_i) = \begin{cases} 
   F_{U_i}(x_i) \frac{F_{U_3}(x_i)}{F_{U_3}(x_j)}, & \text{if } x_i < x_j \\
   F_{U_i}(x_i), & \text{if } x_i \geq x_j 
   \end{cases}
   \]

2. The conditional pdf of \(X_i\) given \(X_j = x_j (i \neq j)\), say \(f_{i|X_j = x_j}(x_i)\), is a convex combination of an absolutely continuous cdf and a degenerate cdf given by
   \[
   f_{i|X_j = x_j}(x_i) = \alpha_j I_{x_j}(x_i) + (1 - \alpha_j) f_{i|X_j,x_{ij},ac}(x_i),
   \]

where \(I_{x_j}\) is the indicator function of the given point \(x_j\), and \(f_{i|X_j,x_{ij},ac}\) is the absolutely continuous part

\[
 f_{i|X_j = x_j,x_{ij},ac}(x_i) = \frac{1}{1 - \alpha_j} \begin{cases} 
   f_{X_i}(x_i) \frac{f_{U_3}(x_i)}{f_{X_j}(x_j)}, & \text{if } x_i < x_j \\
   f_{U_i}(x_i), & \text{if } x_i > x_j \\
   0, & \text{if } x_i = x_j
   \end{cases}
\]
and the mixing weight \( a_j \) is constant with respect to \( x_i \)

\[
a_j = F_{U_1}(x_j) F_{U_2}(x_j) \frac{f_{U_3}(x_j)}{f(x_j)}.
\]

4.2. Minimum and Maximum Order Statistics

Now, we provide the cdfs of the maximum and minimum order statistics of a GBD model, which may be interpreted as the lifetimes of parallel and series systems based on the components of \((X_1, X_2)\).

**Theorem 5.** If \( T_1 = \min(X_1, X_2) \) and \( T_2 = \max(X_1, X_2) \) of a GBD model \((X_1, X_2)\) with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\), then their cdfs are given by

\[
F_{T_1}(x) = F_{U_3}(x) F_{U_{1:2}}(x) \quad \text{and} \quad F_{T_2}(x) = F_{U_{3:3}}(x)
\]

where \( U_{1:2} = \min(U_1, U_2) \) and \( U_{3:3} = \max(U_1, U_2, U_3) \).

**Proof.** It is trivial from (1) and (5) by taking into account that \( F_{T_2}(x) = F(x, x) \) and \( F_{T_1}(x) = F_{X_1}(x) + F_{X_2}(x) - F_{T_2}(x) \). \( \Box \)

The pdfs \( f_{T_1} \) and \( f_{T_2} \) of the minimum and maximum statistics can be readily obtained by differentiation of (7).

Furthermore, the PRH property is preserved by the maximum order statistic of a GBD model, which is immediately derived from Theorem 5.

**Corollary 2.** If \( U_i \sim PRH(\theta_i) \) with baseline distribution \( F_{B_i} \) \((i = 1, 2, 3)\), then \( T_2 \sim PRH(\theta) \) with base \( F_{B_{2:2}} = F_{B_1}^{\theta_1/\theta} F_{B_2}^{\theta_2/\theta} F_{B_3}^{\theta_3/\theta} \) and \( \theta = \theta_1 + \theta_2 + \theta_3 \). Moreover, when \( F_B = F_{B_i} \) \((i = 1, 2, 3)\), then \( T_2 \) also has the same base distribution \( F_B \).

5. Dependence and Stochastic Properties

In this section, we study various dependence and stochastic properties on the GBD family, its marginals and order statistics, and its copula representation. Notions of dependence and ageing for bivariate distributions can be found in Lai and Xie [34] and Balakrishnan and Lai [4]; see also Shaked and Shantikumar [35] for univariate and multivariate stochastic orders.

5.1. GBD Model

**Proposition 2.** If \((X_1, X_2) \sim GBD \) model, then \((X_1, X_2)\) is positive quadrant dependent (PQD).

**Proof.** From (1) and (5), it is readily obtainable that \( F(x_1, x_2) \geq F_{X_1}(x_1) F_{X_2}(x_2) \), which is equivalent to say that all random vector \((X_1, X_2)\), having a GBD model, is PQD. \( \Box \)

An immediate consequence of the PQD property is that \( \text{Cov}(X_1, X_2) > 0 \). Other important bivariate dependence properties are the following, whose proofs are provided in Appendix A.

**Proposition 3.** Let \((X_1, X_2)\) be a random vector having a GBD model:

1. \((X_1, X_2)\) is left tail decreasing (LTD).
2. \((X_1, X_2)\) is left corner set decreasing (LCSD).
3. Its joint cdf \( F \) is totally positive and of order 2 (TP2).

**Proof.** Note that \( F \) is TP2 equivalent to \((X_1, X_2)\) is LCSD, which implies LTD (e.g., see Balakrishnan and Lai [4]). Thereby, we only have to prove (3). From the definition of TP2 property, it is equivalent to check that the following inequality holds:
\[
\frac{F(x)F(x')}{F(x \lor x')F(x \land x')} \leq 1,
\]
for all \(x \) and \(x'\), where \(x \lor x' = (\max(x_1, x'_1), \max(x_2, x'_2))\), and \(x \land x' = (\min(x_1, x'_1), \min(x_2, x'_2))\). Hence, from (1), the inequality (8) can be expressed as
\[
\frac{F_{U_i}(u)F_{U_j}(v)}{F_{U_i}(w)F_{U_j}(y)} \leq 1,
\]
where \(u = x_1 \land x_2, v = x'_1 \land x'_2, w = (x_1 \lor x'_1) \land (x_2 \lor x'_2)\) and \(y = u \land v\). Moreover, one can observe that \(y \leq u \lor v = \max(u, v) \leq w\).

Therefore, when \(u \leq v\), i.e., \(y = u \leq v \leq w\), the inequality (8) can be simplified as follows:
\[
\frac{F_{U_i}(v)}{F_{U_i}(w)} \leq 1,
\]
which is trivial, since \(v \leq w\) and \(F_{U_i}\) is a cdf. An analogous development follows for \(u > v\), which completes the proof. \(\square\)

Let us see now some results related to the reversed hazard gradient of a random vector from the GBD family, which is defined as an extension of the univariate case, see Domma [36],
\[
r(x) = (r_1(x), r_2(x)) = \left( \frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2} \right) \ln F(x_1, x_2)
\]
where each \(r_i(x)\) represents the reversed hazard function of \((X_i | X_j \leq x_j)\), \(i \neq j = 1, 2\), and assuming that \(F\) is differentiable. In addition, it is said that \((X_1, X_2)\) has a bivariate decreasing (increasing) reversed hazard gradient, BDRHG (BIRHG), if all components \(r_i\)'s are decreasing (increasing) functions in the corresponding variables.

**Proposition 4.** If \((X_1, X_2)\) has a GBD model with baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\), then its reversed hazard gradient \(r(x)\) is given by
\[
r_i(x) = \begin{cases} 
  r_{U_i}(x_i) + r_{U_j}(x_i), & \text{if } x_i < x_j \\
  r_{U_i}(x_i), & \text{if } x_i \geq x_j
\end{cases}
\]
for \(i \neq j = 1, 2\), when the reversed hazard function of \(U_i, r_{U_i} = f_{U_i}/F_{U_i}\) exists, \(i = 1, 2, 3\).

**Proof.** The proof is straightforward from the definition of reversed hazard rate function corresponding to the conditional cdf \(F_{i|X_i \leq x_i}\) given by (1) of Theorem 4. \(\square\)

**Theorem 6.** Let \((X_1, X_2)\) be a random vector having a GBD model. If \(U_i\)'s have decreasing reversed hazard functions (DRH), then \((X_1, X_2) \in \text{BDRHG}\).

**Proof.** It is straightforward from Proposition 4. \(\square\)

Note that Theorem 6 provides the closure of the DRH property under the formation of a GBD model. Thus, the bivariate extension of a DRH distribution \(F_{U_l}\) generated by the GBD family is BDRHG.

Nevertheless, it does not hold for the increasing reversed hazard (IRH) property, since both \(r_i(x)\) given in Proposition 4 have a negative jump discontinuity at \(x_i = x_j\) for \(i \neq j = 1, 2\). Therefore, if \(U_i \in \text{IRH}\), then \((X_1, X_2)\) cannot be BIRHG.

Finally, we present some interesting stochastic ordering results between bivariate random vectors of GBD type.
**Theorem 7.** Let \( X = (X_1, X_2) \) and \( Y = (Y_1, Y_2) \) have GBD models with baseline distribution vectors \((F_{U_1}, F_{U_2}, F_{U_3})\) and \((F_{V_1}, F_{V_2}, F_{V_3})\), respectively. If \( U_i \leq_{st} V_i \) \((i = 1, 2, 3)\), then \( X \leq_{lo} Y \).

**Proof.** The result immediately follows from the stochastic ordering between components and (1), since \( U_i \leq_{st} V_i \) is equivalent to \( F_{U_i}(x) \geq F_{V_i}(x) \), and the lower orthant ordering is defined by the inequality \( F_X(x) \geq F_Y(x) \) for all \( x = (x_1, x_2) \). \( \square \)

**Corollary 3.** Let \( X \sim EBPRH(\theta, \lambda) \) and \( Y \sim EBPRH(\theta^*, \lambda^*) \) with base distributions \( F_{B_i} \) and \( F_{B^*_i} \) \((i = 1, 2, 3)\), respectively. If \( \theta_i \leq \theta^*_i \) and \( F_{B_i} \leq_{st} F_{B^*_i} \) \((i = 1, 2, 3)\), then \( X \leq_{lo} Y \).

**Proof.** It is obvious that \( r_{B_i}^{\theta_i}(x_i) \geq r_{B^*_i}^{\theta^*_i}(x_i) \), \( i = 1, 2, 3 \), i.e., \( U_i \leq_{st} V_i \), and then the proof readily follows from Theorem 7. \( \square \)

**Remark 1.** From Corollary 3, if both EBPRH models are based on a common base distribution vector, \( F_{B_i} = F_{B^*_i} \) \((i = 1, 2, 3)\), then it is only necessary that \( \theta_i \leq \theta^*_i \) to hold the lower orthant ordering.

### 5.2. Marginals and Order Statistics

Now, we study some stochastic properties of the marginals and the minimum and maximum order statistics of the GBD model.

Firstly, from (5) and (6), the reversed hazard function of the marginal \( X_i \) is can be expressed as

\[
    r_{X_i}(x) = \frac{f_{X_i}(x)}{F_{X_i}(x)} = r_{U_i}(x) + U_{X_i}(x), \quad i = 1, 2.
\]

Therefore, the DRH (IRH) property is preserved to the marginals.

**Theorem 8.** If \( (X_1, X_2) \) has a GBD model formed by \( U_i \in DRH \) \((i = 1, 2, 3)\), then \( X_i \in DRH \) \((i = 1, 2)\).

**Remark 2.** Note that the IRH distributions have upper bounded support [37]. Thus, if any \( U_i \) is not upper bounded, its reversed hazard function is always decreasing at the end, and then the marginal cannot be IRH. Therefore, it is necessary that \( U_i \in IRH \) \((i = 1, 2, 3)\) and they have the same upper bounds to be \( X_i \in IRH \) \((i = 1, 2)\).

**Example 1.** Suppose \( U_i \) have extreme value distributions of type 3 with a common support, \( U_i \sim EV3(\beta, \lambda_i, k_i) \), whose cdf is defined by

\[
    F_{U_i}(u) = \exp \left( -\lambda_i(\beta - u)^{k_i} \right), \quad \text{for} \ u \leq \beta
\]

and \( F_{U_i}(u) = 1 \) otherwise. Its reversed hazard function is given by

\[
    r_{U_i}(u) = \lambda_i k_i (\beta - u)^{k_i-1}, \quad \text{for} \ u \in (-\infty, \beta],
\]

which is increasing (decreasing) in its support for \( k_i \leq (\geq) 1 \). Thus, if \( k_i \leq (\geq) 1 \) \((i = 1, 2, 3)\), then \( U_i \in IRH(DRH) \), and, consequently, \( X_i \in IRH(DRH) \) \((i = 1, 2)\).

**Example 2.** If \( (X_1, X_2) \) has an EBGE model, then its marginals are DRH, since \( r_{X_i} \) given by (9) is the sum of two decreasing functions because of each \( U_i \sim GE(\lambda_i, \lambda_i) \) is a PRH(\( \theta_i \)) with exponential baseline distribution

\[
    r_{U_i}(u) = \theta r_{\text{Exp}(\lambda_i)}(u) = \frac{\theta_i \lambda_i}{e^{\lambda_i u} - 1},
\]

which is evidently a decreasing function. Here, \( \text{Exp}(\lambda) \) denotes an exponential random variable with mean \( 1/\lambda \).
Remark 3. From (9), when the \( U_i \)'s have a common distribution \( F_{U_i} \), then the marginals \( X_i \sim PRH(2) \) with base distribution \( F_B \). Therefore, \( r_X(x) = 2r_U(x) \) has the same monotonicity. In particular, if \( F_{U_i} \in DRH (IRH) \), then \( X_i \in DRH (IRH) \).

Remark 4. From (9), if \( U_i \sim PRH(\theta_i) \) with the same base distribution \( F_B \), then \( X_i \sim PRH(\theta_i + \theta_3) \) with base \( F_B \), i.e., \( r_X(x) = (\theta_i + \theta_3) r_B(x) \). Thus, Remark 3 also holds by using \( F_B \) instead of \( F_{U_1} \).

Secondly, the mean inactivity time (MIT), also called mean waiting time [37], of a random variable \( X \) is defined as

\[
m_X(x) = E(x - X | X \leq x) = \int_x^\infty \frac{F_X(y)}{F_X(x)} dy.
\]

Thus, from (5), the MIT of the marginal \( X_i \)'s of a GBD model can be derived by

\[
m_{X_i}(x) = \frac{1}{F_{U_i}(x) F_{U_i}(x)} \int_x^\infty F_{U_i}(y) F_{U_i}(y) dy, \quad i = 1, 2.
\]  

(10)

Here, we shall focus on two particular cases of GBD models, having baseline components with monotonous MIT, which is preserved by the marginals.

Example 3. Suppose \( U_i \sim Exp(\lambda) \), then its MIT can be expressed as

\[
m_{U_i}(u) = \frac{u}{1 - e^{-\lambda u}} - \frac{1}{\lambda},
\]

which is an increasing MIT function (IMIT), i.e., \( U_i \in IIMIT \). From (10), we obtain the MIT function of the marginals \( X_i \)'s for the bivariate exponential version of GBD type,

\[
m_{X_i}(x) = \frac{2\lambda x - 3 + 4e^{-\lambda x} - e^{-2\lambda x}}{2\lambda (1 - e^{-\lambda x})^2}.
\]

Then, upon differentiation, \( m'_{X_i}(x) \) has the same sign as the expression \( 1 - e^{-2\lambda x} - 2\lambda xe^{-\lambda x} \), which is positive, and therefore \( X_i \in IIMIT \) (i = 1, 2).

Example 4. Suppose \( U_i \sim EV3(\beta, \lambda_i, k = 2) \), then its MIT can be expressed as

\[
m_{U_i}(u) = \frac{1}{e^{-\lambda_i (\beta - u)^2}} \int_0^u e^{-\lambda_i (\beta - y)^2} dy = \frac{\Phi(u; \mu, \sigma)}{\Phi(u; \mu, \sigma)}, \quad for \ u \leq \beta
\]

where \( \Phi(u; \mu, \sigma) \) and \( \phi(u; \mu, \sigma) \) are the cdf and pdf of a normal model with \( \mu = \beta \) and \( \sigma = \frac{1}{\sqrt{2\lambda_i}} \), respectively. Moreover, taking into account that a random variable and its standardized version have PRH functions, and the standard normal distribution has the DRH property [38], we obtain that \( U_i \in IIMIT \).

Upon considering the cdf of \( U_i \)'s and (5), the marginal \( X_i \sim EV3(\beta, \lambda_i + \lambda_3, 2) \) (i = 1, 2). Thus, their MIT can be written as

\[
m_{X_i}(x) = \frac{\Phi(x; \mu, \sigma_i)}{\phi(x; \mu, \sigma_i)}, \quad for \ x \leq \beta
\]

where \( \sigma_i = \frac{1}{\sqrt{2(\lambda_i + \lambda_3)}} \) for i = 1, 2, and, consequently, \( X_i \in IIMIT \).

On the other hand, the following stochastic orderings among the three baseline components of two GBD models are preserved by their corresponding marginals. The proof immediately follows from the definitions of the stochastic orderings.

Theorem 9. Let \( (X_1, X_2) \) and \( (Y_1, Y_2) \) have GBD models with base distribution vectors \( (F_{U_1}, F_{U_2}, F_{U_3}) \) and \( (F_{V_1}, F_{V_2}, F_{V_3}) \), respectively.
1. If \( U_i \leq_{st} V_i \) \((i = 1, 2, 3)\), then \( X_i \leq_{st} Y_i \) \((i = 1, 2)\).
2. If \( U_i \leq_{rh} V_i \) \((i = 1, 2, 3)\), then \( X_i \leq_{rh} Y_i \) \((i = 1, 2)\).

Finally, we discuss some stochastic properties of the minimum and maximum order statistics of the GBD family. In this setting, from (7), the reversed hazard function of the maximum statistic \( T_2 \) of \((X_1, X_2)\) of GBD type is determined by the sum of the reversed hazard rates of the baseline distribution vector:

\[
r_{T_2}(x) = r_{U_1}(x) + r_{U_2}(x) + r_{U_3}(x)
\]

when the pdf \( f_{U_i} \) of \( U_i \) exists, \( i = 1, 2, 3 \). Hence, it is immediate the following result.

**Theorem 10.** If \( U_i \in DRH(IRH) \) \((i = 1, 2, 3)\), then \( T_2 \in DRH(IRH) \).

**Example 5.** Suppose \( U_i \sim EV3(\beta, \lambda_i, k_i) \) \((i = 1, 2, 3)\). Then, the reversed hazard function of \( T_2 \) is given by

\[
r_{T_2}(x) = \sum_{i=1}^{3} \lambda_i k_i (\beta - x)^{k_i-1},
\]

and, therefore, if every \( k_i \leq (\geq) 1, i = 1, 2, 3 \), then \( r_{T_2} \) is increasing (decreasing) in \( x \), i.e., \( T_2 \in IRH(DRH) \).

**Example 6.** If \( U_i \sim GE(\theta, \lambda_i) \), then the maximum statistic of the EBGE model is \( DRH, T_2 \in DRH \), since (11) is the sum of three decreasing functions.

**Remark 5.** When \( U_i \)'s have a common distribution \( F_{U_i} \), the GBD model has a maximum statistic whose cdf is \( F_{U_i} \) cube, and (11) can be written as \( r_{T_2}(x) = 3r_{U}(x) \). In particular, if \( F_{U_i} \in DRH(1RH) \), then \( T_2 \in DRH(1RH) \).

**Remark 6.** From Corollary 2, if \( U_i \sim PRH(\theta_i) \) with the same base distribution \( F_B \), \( T_2 \sim PRH(\theta) \) with base \( F_B \) and \( \theta = \theta_1 + \theta_2 + \theta_3 \), i.e., \( r_{T_2}(x) = \theta r_{B}(x) \). Thus, \( T_2 \in DRH(1RH) \) if and only if \( F_B \in DRH(1RH) \).

Furthermore, the MIT of the maximum statistic of a GBD model \((X_1, X_2)\) can be derived by

\[
m_{T_2}(x) = \frac{1}{F_{U_1}(x)F_{U_2}(x)F_{U_3}(x)} \int_{-\infty}^{x} F_{U_1}(y)F_{U_2}(y)F_{U_3}(y)dy,
\]

for each specific baseline distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\), when the integral exists. For instance, we will consider a particular case, similar to one used in Example 4.

**Example 7.** Suppose \((X_1, X_2)\) has a GBD model with \( U_i \sim PRH(\theta_i) \) and base distributions \( F_{B_i} \sim EV3(\beta, \lambda_i, k = 2) \) for \( i = 1, 2, 3 \), then each component \( U_i \sim EV3(\beta, \theta_i \lambda_i, 2) \), and consequently, \( U_i \in IMIT \) for \( i = 1, 2, 3 \). Moreover, from Corollary 2, the maximum statistic \( T_2 \sim EV3(\beta, \theta^*, 2) \) with \( \theta^* = \theta_1 \lambda_1 + \theta_2 \lambda_2 + \theta_3 \lambda_3 \). Thus, \( T_2 \in IMIT \) which is obtained along the same line as Example 4, since

\[
m_{T_2}(x) = \frac{\Phi(x; \beta, (2\theta^*)^{-1/2})}{\phi(x; \beta, (2\theta^*)^{-1/2})}, \text{ for } x \leq \beta.
\]

Regarding the minimum statistic \( T_1 \) of \((X_1, X_2)\) of GBD type, some preservation results are also obtained based on its reversed hazard rate \( r_{T_1} \), the proofs are given in Appendix A, and from (7) \( r_{T_1} \) can be written as

\[
r_{T_1}(x) = r_{U_1}(x) + r_{U_2}(x).
\]

**Theorem 11.** If \( U_i \in DRH(i = 1, 2, 3) \) and \( U_{12} \leq_{rh} U_i \) \((i = 1, 2)\), then \( T_1 \in DRH \).
Corollary 4. If \( U_i \in DRH \) \((i = 1, 2, 3)\) and \( U_1 =_{st} U_2 \), then \( T_1 \in DRH \).

**Example 8.** Suppose \( U_i \sim GE(\theta, \lambda) \) for \( i = 1, 2 \) and \( U_3 \sim GE(\theta_3, \lambda_3) \), then \( U_i \in DRH \), and, consequently, \( T_1 \in DRH \) from Corollary 4.

**Remark 7.** Note that, when \( U_i \)'s have a common distribution \( F_U \), (12) can be expressed as \( r_{U_1}(x) = r_{U_1}(x)(3 - 2/(2 - F_U(x))) \), and from Corollary 4, it is immediate to have that, if \( F_U \in DRH \), then \( T_1 \in DRH \).

**Theorem 12.** Let \((X_1, X_2)\) be a GBD model. Then, \( T_1 \leq_{st} T_2 \).

**Proof.** From (11) and (12), the statement is equivalent to \( r_{U_{12}}(x) \leq r_{U_{13}}(x) \), which readily follows from Theorem 1.B.56 of Shaked and Shanthikumar [35], since the baseline components \( U_i \)'s are independent.

### 5.3. Copula and Related Association Measures

Let us see now the copula representation of the GBD family and some related dependence measures of interest in the analysis of two-dimensional data.

It is well known that the dependence between the random variables \( X_1 \) and \( X_2 \) is completely described by the joint cdf \( F(x_1, x_2) \), and it is often represented by a copula which describes the dependence structure in a separate form from the marginal behaviour. In this setting, from Sklar’s theorem (e.g., see [39]), if its marginal cdfs \( F_{X_i} \)'s are absolutely continuous, then the joint cdf has a unique copula representation for

\[
F(x_1, x_2) = C \left( F_{X_1}(x_1), F_{X_2}(x_2) \right),
\]

and reciprocally, if \( F_{X_i}^{-1} \) is the inverse function of \( F_{X_i} \) \((i = 1, 2)\), then there exists a unique copula \( C \) in \([0, 1]^2\), such that

\[
C(u_1, u_2) = F \left( F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2) \right).
\]

Now, we can derive the copula representation for the joint cdf of the GBD family as a function of its base distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\). In order to do this, by using (5), the joint cdf (1) can be expressed as

\[
F(x_1, x_2) = F_{X_1}(x_1) F_{X_2}(x_2) \frac{F_{U_3}(\min(x_1, x_2))}{F_{U_3}(x_1) F_{U_3}(x_2)}
\]

and taking \( u_i = F_{X_i}(x_i) \), the associated copula for an arbitrary base distribution vector \((F_{U_1}, F_{U_2}, F_{U_3})\) can be written as

\[
C(u_1, u_2) = u_1 u_2 \frac{\min(A_1(u_1), A_2(u_2))}{A_1(u_1) A_2(u_2)},
\]

where

\[
A_i(u_i) = \frac{F_{U_i}(x_i)}{F_{U_i}(x_i) - F_{U_i}(x_i)}, \quad i = 1, 2,
\]

which allows us to give an additional result.

**Theorem 13.** Let \( X = (X_1, X_2) \) and \( Y = (Y_1, Y_2) \) be two GBD models with baseline distribution vectors \((F_{U_1}, F_{U_2}, F_{U_3})\) and \((F_{V_1}, F_{V_2}, F_{V_3})\), respectively. If \( X \) and \( Y \) have the same associated copula and \( U_i \leq_{st} V_i \), then \( X \leq_{st} Y \).

**Proof.** It is immediate by using Theorem 6.B.14 of Shaked and Shanthikumar [35] and (5), since \( U_i \leq_{st} V_i \) implies \( X_i \leq_{st} Y_i \).

**Corollary 5.** Let \( X = (X_1, X_2) \) and \( Y = (Y_1, Y_2) \) be two GBD models with common baseline distributions, \( F_U \) and \( F_V \), respectively. If \( U \leq_{st} V \), then \( X \leq_{st} Y \).
Note that (13) provides a general formula to establish the specific copula upon considering two particular continuous and increasing bijective functions $A_1$ and $A_2$ from $[0, 1]$ onto $[0, 1]$. Fang and Li [40] analyzed some stochastic orderings for an equivalent copula representation to (13) with interesting applications in network security and insurance. In the last section, we shall use the bivariate copula representation (13) to discuss the multivariate extension of the GBD family.

Furthermore, (13) may be considered a generalization of the Marshall–Olkin copula, as displayed in the following results whose proofs are omitted.

**Corollary 6.** If $(X_1, X_2)$ has a GBD model with a common base distribution $F_{\perp\perp}$, then the copula representation of its joint cdf is

$$C(u_1, u_2) = \min(u_1 u_2^{1/2}, u_1^{1/2} u_2).$$

**Corollary 7.** If $(X_1, X_2)$ has a GBD model with PRHs baseline distribution vector of the same base $F_B$, i.e., $(X_1, X_2) \sim \text{BPRH}(\theta_1, \theta_2, \theta_3)$, then the copula representation of its joint cdf is

$$C(u_1, u_2) = \min \left( u_1 u_2^{\theta_2/(\theta_2 + \theta_3)}, u_1^{\theta_1/(\theta_1 + \theta_3)} u_2 \right).$$

Some association measures for a bivariate random vector $(X_1, X_2)$ of GBD type can be derived from the dependence structure described by the general expression (13) for each particular pair of continuous and increasing bijective functions $A_1$ and $A_2$ determined by the specific baseline distribution vector. For instance, for the special GBD models given in Corollaries 6 and 7, the measures of dependence namely Kendall’s tau, Spearman’s rho, Blomqvist’s beta, and tail dependence coefficients, see Nelsen [39] among others, can be calculated as follows.

**Kendall’s tau.** The Kendall’s $\tau$ is defined as the probability of concordance minus the probability of discordance between two pairs of independent and identically distributed random vectors, $(X_1, X_2)$ and $(Y_1, Y_2)$, as follows:

$$\tau = P((X_1 - Y_1)(X_2 - Y_2) > 0) - P((X_1 - Y_1)(X_2 - Y_2) < 0),$$

and it can be calculated through its copula representation $C(u_1, u_2)$ by

$$\tau = 4E(C(U_1, U_2)) - 1 = 1 - 4 \int_{[0,1]^2} \frac{\partial C(u_1, u_2)}{\partial u_1} \frac{\partial C(u_1, u_2)}{\partial u_2} du_1 du_2$$

(14)

with $U_8$ uniform $[0, 1]$ random variables whose joint cdf is $C$.

For example, if $(X_1, X_2)$ has a GBD model with a common baseline $F_{\perp\perp}$, upon substituting from the copula of Corollary 6 in (14), it is easy to check that Kendall’s $\tau = 1/3$.

Analogously, from the copula given in Corollary 7 of the GBD model for PRH($\theta_1$) components with a common base $F_B$, the Kendall’s $\tau$ coefficient (14) can be written as

$$\tau = \frac{\theta_3}{\theta_1 + \theta_2 + \theta_3}.$$

**Spearman’s rho.** The Spearman’s $\rho$ coefficient measures the dependence by three pairs of independent and identically distributed random vectors, $(X_1, X_2)$, $(Y_1, Y_2)$ and $(Z_1, Z_2)$. It is defined as

$$\rho = 3(P((X_1 - Y_1)(X_2 - Z_2) > 0) - P((X_1 - Y_1)(X_2 - Z_2) < 0)), $$

which can be computed by its copula representation $C(u_1, u_2)$ by

$$\rho = 12E(U_1 U_2) - 3.$$
Thus, if there is a common base distribution as in Corollary 6, the Spearman’s \( \rho \) coefficient between \( X_1 \) and \( X_2 \) is \( \rho = 3/7 \).

In the case of \( U_i \sim PRH(\theta_i) \) with a common base distribution \( F_B \), from (15) and Corollary 7, this association measure is
\[
\rho = \frac{3\theta_3}{2\theta_1 + 2\theta_2 + 3\theta_3}
\]
which coincides with one obtained by Kundu et al. [9] for this specific GBD model, \( (X_1, X_2) \sim BPRH(\theta_1, \theta_2, \theta_3) \). As remarked by Kundu et al. [9] for the BPRH model, both coefficients, \( \tau \) and \( \rho \), vary between 0 and 1 as \( \theta_3 \) varies from 0 to \( \infty \).

**Blomqvist’s Beta.** The Blomqvist’s \( \beta \) coefficient, also called the medial correlation coefficient, is defined as the probability of concordance minus the probability of discordance between \( (X_1, X_2) \) and its median point, say \( (m_1, m_2) \), taking the following form:
\[
\beta = P((X_1 - m_1)/(X_2 - m_2) > 0) - P((X_1 - m_1)/(X_2 - m_2) < 0) = 4F(m_1, m_2) - 1,
\]
and from the copula of its joint cdf \( F \), it can be expressed as
\[
\beta = 4C(1/2, 1/2) - 1. \quad (16)
\]

In the case of Corollary 6, it is immediate that the medial correlation coefficient between \( X_1 \) and \( X_2 \) is \( \beta = \sqrt{2} - 1 \) when it follows a GBD model with a common baseline distribution.

In the other case, from Corollary 7, the Blomqvist’s \( \beta \) coefficient (16) is also readily obtainable between the marginals of a BPRH model:
\[
\beta = \begin{cases} 
2\theta_3/(\theta_2 + \theta_3), & \text{if } \theta_1 \leq \theta_2 \\
2\theta_3/(\theta_1 + \theta_3), & \text{if } \theta_1 > \theta_2.
\end{cases}
\]
which takes values between 0 and 1 as \( \theta_3 \) varies from 0 to \( \infty \).

**Tail Dependence.** The tail dependence measures the association of extreme events in both directions, the upper (lower) tail dependence \( \lambda_U (\lambda_L) \) provides an asymptotical association measurement in the upper (lower) quadrant tail of a bivariate random vector, given by (if it exists)
\[
\lambda_U (\lambda_L) = \lim_{u \to 1^- (0^+)} P \left( X_2 > (\leq) F_{X_2}^{-1}(u) | X_1 > (\leq) F_{X_1}^{-1}(u) \right).
\]

Similar to the above association coefficients, the tail dependence indexes can be calculated from the copula representation \( C(u_1, u_2) \) of the joint cdf of \( (X_1, X_2) \), as follows:
\[
\lambda_U = 2 - \lim_{u \to 1^-} \frac{1 - C(u, u)}{1 - u} \quad \text{and} \quad \lambda_L = \lim_{u \to 0^+} \frac{C(u, u)}{u}. \quad (17)
\]

In particular, if \( (X_1, X_2) \) follows a GBD model with a common baseline distribution, upon substituting from the copula of Corollary 6 in (17), it is easy to check that \( \lambda_L = 0 \) and \( \lambda_U = 1/2 \).

In the case of \( U_i \sim PRH(\theta_i) \) with the same base, from (17) and Corollary 7, it is clear that the tail dependence indexes of the BPRH model are \( \lambda_L = 0 \) and
\[
\lambda_U = \begin{cases} 
\frac{\theta_3}{\theta_2 + \theta_3}, & \text{if } \theta_1 \leq \theta_2 \\
\frac{\theta_3}{\theta_1 + \theta_3}, & \text{if } \theta_1 > \theta_2.
\end{cases}
\]
which takes values between 0 and 1 as \( \theta_3 \) varies from 0 to \( \infty \).
6. Maximum Likelihood Estimation

In this section, we address the problem of computing the maximum likelihood estimations (MLEs) of the unknown parameters based on a random sample. The problem can be formulated as follows. Suppose \( \{(x_{1i}, x_{2i}); i = 1, \ldots, n\} \) is a random sample of size \( n \) from a GBD model, where it is assumed that, for \( j = 1, 2, 3, U_j \) has the pdf \( f_{U_j}(u; \theta_j) \) and \( \theta_j \) is of dimension \( p_j \). The objective is to estimate the unknown parameter vector \( \theta = (\theta_1, \theta_2, \theta_3) \). We use the following partition of the sample:

\[
I_1 = \{i : x_{1i} < x_{2i}\}, \quad I_2 = \{i : x_{1i} > x_{2i}\}, \quad I_0 = \{i : x_{1i} = x_{2i} = x_i\}.
\]

Based on the above observations, the log-likelihood function becomes

\[
\ell(\theta) = \sum_{i \in I_0} \ln f_0(x_i; \theta) + \sum_{i \in I_1} \ln f_1(x_{1i}, x_{2i}; \theta) + \sum_{i \in I_2} \ln f_2(x_{1i}, x_{2i}; \theta),
\]

where \( f_0(x_i; \theta), f_1(x_{1i}, x_{2i}; \theta), f_2(x_{1i}, x_{2i}; \theta) \) have been defined in Theorem 3.

Here, it is difficult to compute the MLEs of the unknown parameter vector \( \theta \) by solving a \( p_1 + p_2 + p_3 \) optimization problem. To avoid that, we suggest using the EM algorithm, and the basic idea is based on considering a random sample of size \( n \) from \( (U_1, U_2, U_3) \), instead of the random sample of size \( n \) from \( (X_1, X_2) \). From the observed sample \( \{(x_{1i}, x_{2i})\} \), the sample \( \{(u_{1i}, u_{2i}, u_{3i}); i = 1, \ldots, n\} \) has missing values as shown in Table 1. It is immediate that the MLEs of \( \theta_1, \theta_2 \) and \( \theta_3 \) can be obtained by solving the following three optimization problems of dimensions \( p_1, p_2 \) and \( p_3 \), respectively,

\[
\ell_j(\theta_j) = \sum_{i=1}^{n} \ln f_{U_j}(u_{ji}; \theta_j); \quad j = 1, 2, 3,
\]

which are computationally more tractable.

| \( I_k \) | Ordering of \( U_j \) | \( X_1 \) | \( X_2 \) | Missing |
|---|---|---|---|---|
| \( I_0 \) | \( u_{1i} < u_{2i} < u_{3i} \) | \( u_{3i} \) | \( u_{3i} \) | \( u_{1i}, u_{2i} \) |
| \( I_0 \) | \( u_{2i} < u_{1i} < u_{3i} \) | \( u_{3i} \) | \( u_{3i} \) | \( u_{1i}, u_{2i} \) |
| \( I_1 \) | \( u_{1i} < u_{3i} < u_{2i} \) | \( u_{3i} \) | \( u_{3i} \) | \( u_{1i}, u_{2i} \) |
| \( I_1 \) | \( u_{3i} < u_{1i} < u_{2i} \) | \( u_{1i} \) | \( u_{2i} \) | \( u_{1i} \) |
| \( I_2 \) | \( u_{2i} < u_{3i} < u_{1i} \) | \( u_{1i} \) | \( u_{2i} \) | \( u_{3i} \) |
| \( I_2 \) | \( u_{3i} < u_{2i} < u_{1i} \) | \( u_{1i} \) | \( u_{2i} \) | \( u_{3i} \) |

From Table 1, if \( i \in I_0 \), then \( u_{3i} \) is known, and \( u_{1i} \) and \( u_{2i} \) are unknown. Similarly, if \( i \in I_1 \) (\( i \in I_2 \)), then \( u_{2i} \) (\( u_{1i} \)) and \( \max\{u_{1i}, u_{3i}\} \) (\( \max\{u_{2i}, u_{3i}\} \)) are known. Hence, in the E-step of the EM algorithm, the ‘pseudo’ log-likelihood function is formed by replacing the missing \( u_{ji} \) by its expected value, \( u_{ji}(\theta_j) \), for \( i = 1, \ldots, n \) and \( j = 1, 2, 3 \):

1. If \( i \in I_0 \), then
   \[
   u_{ji}(\theta_j) = E(U_j|U_j < x_i) = \frac{1}{F_{U_j}(x_i)} \int_{-\infty}^{x_i} u f_{U_j}(u) du, \quad j = 1, 2.
   \]
2. If \( i \in I_1 \) and \( j, k \in \{1, 3\}, j \neq k \), then
   \[
   u_{ji}(\theta_j) = E(U_j|\max\{U_1, U_3\} = x_{1i})
   = x_{1i} P(U_j > U_k) + P(U_j < U_k) \frac{1}{F_{U_j}(x_{1i})} \int_{-\infty}^{x_{1i}} u f_{U_j}(u) du
   = x_{1i} \int_{-\infty}^{\infty} f_{U_j}(u) F_{U_k}(u) du + \frac{1}{F_{U_j}(x_{1i})} \int_{-\infty}^{x_{1i}} f_{U_k}(u) f_{U_j}(u) du \int_{-\infty}^{x_{1i}} u f_{U_j}(u) du.
   \]
If \( i \in I_2 \) and \( j, k \in \{2, 3\}, j \neq k \), then

\[
\ell_{3i}^{(k)}(\theta) = \sum_{i \in I_1} \ln f_{u_1}(u_{1im}(\theta^{(k)}); \theta_1) + \sum_{i \in I_1} \ln f_{u_1}(u_{1im}(\theta^{(k)}); \theta_1) + \sum_{i \in I_1} \ln f_{u_1}(u_{1im}; \theta_1)
\]

\[
\ell_{2i}^{(k)}(\theta) = \sum_{i \in I_1} \ln f_{u_2}(u_{2im}(\theta^{(k)}); \theta_2) + \sum_{i \in I_1} \ln f_{u_2}(u_{2im}(\theta^{(k)}); \theta_2) + \sum_{i \in I_1} \ln f_{u_2}(u_{2im}; \theta_2)
\]

\[
\ell_{3i}^{(k)}(\theta_3) = \sum_{i \in I_1} \ln f_{u_3}(u_{3im}(\theta^{(k)}); \theta_3) + \sum_{i \in I_1} \ln f_{u_3}(u_{3im}(\theta^{(k)}); \theta_3) + \sum_{i \in I_1} \ln f_{u_3}(u_{3im}; \theta_3)
\]

M-step

\[ \theta^{(k+1)} = (\theta^{(k+1)}_1, \theta^{(k+1)}_2, \theta^{(k+1)}_3) \] can be obtained by maximizing \( \ell_{1i}^{(k)}(\theta_1), \ell_{2i}^{(k)}(\theta_2) \) and \( \ell_{3i}^{(k)}(\theta_3) \) with respect to \( \theta_1, \theta_2 \) and \( \theta_3 \), respectively.

Mainly for illustrative purposes, two particular GBD models will be applied in the next section to show the usefulness of the above EM algorithm. Firstly, we shall consider a GBD model with baseline components having the same distribution type and different underlying parameters. Secondly, we shall use a GBD model with baseline components from different distribution families. The technical details of both of them can be found in Appendix B.

7. Data Analysis

In this section, we present the analysis of two-dimensional data sets in order to show how the proposed EM algorithm can be applied to fit particular GBD models. For that, we shall suppose the following two models described in Appendix B: Model I is the GBD model with the exponential baseline distributions and different underlying parameters, \( U_j \sim \text{Exp}(\lambda_j) \) (\( j = 1, 2, 3 \)). Model II is the GBD model with baseline components from Weibull and generalized exponential distributions, \( U_1 \sim W(\lambda_1, \alpha_1), U_2 \sim W(\lambda_2, \alpha_2) \) and \( U_3 \sim \text{GE}(\alpha_3, \lambda_3) \).

7.1. Soccer Data

We have analyzed a UEFA Champion’s League data set [41], played during the seasons 2004–2005 and 2005–2006. This set represents the soccer data where at least one goal has been scored by a kick goal (penalty kick, foul kick or any other direct kick) by any team and one goal has been scored by the home team. Here, in the bivariate data, \((X_1, X_2), X_1\) represents the time in minutes of the first kick goal and \(X_2\) represents the time in minutes scored by the home team. Clearly, all possibilities exist in the data set, namely \(X_1 < X_2, X_1 > X_2\) and \(X_1 = X_2\).
Meintanis [41] analyzed this data set using the Marshall–Olkin bivariate exponential model. The marginals of the Marshall–Olkin bivariate exponential distribution are exponential, and then they have constant hazard functions. A preliminary data analysis indicated that the empirical hazard function of both the marginals are increasing and their reversed hazard functions are decreasing. Hence, it may not be proper to use the Marshall–Olkin bivariate exponential model to analyze this data.

**Example 9.** In order to use Model I, we have started the initial guess as $\hat{\lambda}_1^{(0)} = \lambda_1^{(0)} = \lambda_3^{(0)} = 1$. The algorithm stops after eight iterations, the final estimates and the associated 95% confidence intervals with diabetic retinopathy. For each patient, one eye was randomly selected for laser photocoagulation Kolmogorov–Smirnov (KS) distances with the corresponding $p$-values between the empirical and fitted cdfs for the marginals of the Marshall–Olkin bivariate exponential distribution are exponential, and then the corresponding ones to the subsequent examples.

One natural question is whether Model I fits the bivariate data or not. We have computed the Kolmogorov–Smirnov (KS) distances with the corresponding $p$-values between the empirical and fitted cdfs for the marginals and the maximum statistic. The results are reported in Table 2, and, from them, we cannot reject the null hypothesis that this data are coming from the GBD model with exponential baseline distributions.

**Example 10.** Let us consider now Model II. We have started the EM algorithm with the initial guesses $\hat{\alpha}_1 = \hat{\alpha}_2 = \hat{\alpha}_3 = 1$, $\hat{\lambda}_1^{(0)} = 0.03$, $\hat{\lambda}_2^{(0)} = 0.05$ and $\hat{\lambda}_3^{(0)} = 0.04$. The algorithm converges in nineteen iterations, the final estimates and the associated 95% confidence intervals are $\hat{\lambda}_1 = 1.2987 \pm 0.3124$, $\hat{\lambda}_2 = 0.8047 \pm 0.2823$, $\hat{\lambda}_3 = 0.0993 \pm 0.0021$, $\hat{\alpha}_3 = 1.0037 \pm 0.2879$, $\hat{\lambda}_3 = 0.0369 \pm 0.008$, with $-201.1141$ being the pseudo log-likelihood value.

The KS distances with the corresponding $p$-values for the marginals and the maximum statistic are reported in Table 2. Thus, based on the $p$-values, we can say that the GBD model with two baseline Weibull distributions and the third GE one fits the data reasonably well.

Summarizing, it is clear that both of the GBD models provide a good fit to the given data set and the EM algorithm also works quite effectively in both the cases. Now, to compare Models I and II of Examples 9 and 10, which provide a better fit, we compute the Akaike’s information criterion (AIC) and Bayesian information criterion (BIC) values and they are also presented in Table 2. Therefore, based on the AIC and BIC values, it is clear that Model I provides a better fit than Model II to the UEFA Champion’s League data set.

**Table 2.** Goodness-of-fit results for UEFA Champion’s League data.

| GBD Model | KS ($p$-Value) | X1 | X2 | max{X1, X2} | AIC | BIC |
|-----------|----------------|----|----|-------------|-----|-----|
| Model I   |                | 0.1491 (0.3830) | 0.1099 (0.7622) | 0.1530 (0.3517) | 604.8663 | 609.6990 |
| Model II  |                | 0.0976 (0.8719) | 0.0839 (0.9565) | 0.1139 (0.7228) | 708.5430 | 718.2085 |

**7.2. Diabetic Retinopathy Data**

Let us consider now the diabetic retinopathy data set [42], available in the R package “SurvCor” [43]. Such data were investigated by the National Eye Institute to assess the effect of laser photocoagulation in delaying the onset of severe visual loss such as blindness in 197 patients with diabetic retinopathy. For each patient, one eye was randomly selected for laser photocoagulation and the other was given no treatment, being used as the control. The times to blindness in both eyes were recorded in months and the censoring was caused by death, dropout, or the end of the study.

For illustrative purposes, we have considered those patients for which complete data are available. Here, $X_1$ denotes the time to the blindness of the untreated or control eye and $X_2$ denotes the time...
to blindness of the treated eye. Out of 197 patients, we have complete information of \( X_1 \) and \( X_2 \) for 38 patients.

**Example 11.** As in Example 9, we have used Model I to analyze the data set. In this case, we have also used the same initial guess as \( \lambda_1^{(0)} = \lambda_2^{(0)} = \lambda_3^{(0)} = 1 \). We have used the proposed EM algorithm, the iteration stops after 14 iterations, and the estimates of unknown parameters and the corresponding 95% confidence intervals are \( \hat{\lambda}_1 = 0.0653 \) (±0.0175), \( \hat{\lambda}_2 = 0.0737 \) (±0.0210) and \( \hat{\lambda}_3 = 0.1345 \) (±0.3879), with \(-172.2314\) being the associated pseudo log-likelihood value.

The KS distances with the corresponding \( p \)-values between the empirical and fitted cdfs for the marginals and the maximum statistic are presented in Table 3.

**Example 12.** As in Example 10, we have analyzed the data set by using Model II. We have started the EM algorithm with the initial guesses \( \alpha_1^{(0)} = \alpha_2^{(0)} = \alpha_3^{(0)} = 1, \lambda_1^{(0)} = 0.06, \lambda_2^{(0)} = 0.07 \) and \( \lambda_3^{(0)} = 0.13 \). The algorithm stops after 27 iterations, the final estimates and the corresponding 95% confidence intervals are \( \hat{\alpha}_1 = 1.0937 \) (±0.2563), \( \hat{\lambda}_1 = 0.0447 \) (±0.0146), \( \hat{\alpha}_2 = 0.5851 \) (±0.1345), \( \hat{\lambda}_2 = 0.2369 \) (±0.0763), \( \hat{\alpha}_3 = 0.8995 \) (±0.2787), \( \hat{\lambda}_3 = 0.1896 \) (±0.0478), with \(-125.4519\) being the associated pseudo log-likelihood value.

The KS distances with the corresponding \( p \)-values for the marginals and the maximum order statistic are presented in Table 3.

From Table 3, we can also say that the estimated GBD models fit the diabetic retinopathy data reasonably well in both the cases. Moreover, we also present the AIC and BIC values of the two models in Table 3. Therefore, based on the AIC and BIC values, it is clear that Model I provides a better fit than Model II for the diabetic retinopathy data.

| GBD Model | KS \((p\)-Value\) | \( X_1 \) | \( X_2 \) | \( \max \{ X_1, X_2 \} \) | AIC | BIC |
|-----------|------------------|--------|--------|------------------|-----|-----|
| Model I   |                  | 0.1033 | 0.1848 | 0.1229 (0.6310)  | 585.9757 | 590.8884 |
| Model II  |                  | 0.0920 | 0.0952 | 0.1152 (0.6778)  | 592.1515 | 601.9770 |

### 8. Discussion and Conclusions

In this paper, we have presented the generalized bivariate distribution family by a generator system based on the maximization process from any three-dimensional baseline continuous distribution vector with independent components, providing bivariate models with dependence structure.

For the proposed GBD family, several distributional and stochastic properties have been established. The preservation of the PRH property for the marginals and the maximum order statistic has been obtained. The positive dependence has been shown between both marginals of the GBD models, some results about stochastic orders and on the preservation of the monotonicity of the reversed hazard function and of the mean inactivity time. Furthermore, the copula representation of the GBD model has been discussed, providing a general formula, and some related dependence measures have been also calculated for specific copulas of particular bivariate distributions of the GBD family. In addition, new bivariate distributions can be generated by combining independent baseline components from different distribution families, and several bivariate distributions given in the literature are derived as particular cases of the GBD family.

Note that, even in the simple case, the MLEs cannot be obtained in explicit forms, and it is required solving a multidimensional nonlinear optimization problem. We have proposed using an EM algorithm to compute the MLEs of the unknown parameters, and it is observed that the proposed
EM algorithm perform quite satisfactorily in the two data analyses by using two different models of the GBD family. The experimental results summarized in Table A1 disclose such efficiency of the EM algorithm with respect to a conventional numerical iterative procedure of the Newton-type. In more detail, Table A1 presents the experimental results obtained by the Broyden–Fletcher–Goldfarb–Shanno algorithm for maximizing the log-likelihood function, available in the R package “maxLik” [44].

It is worth mentioning that the bivariate copula representation (13) allows us to discuss its multivariate extension. Let \( U_i \) for \( i = 1, \ldots, q + 1 \) be a set of \( q + 1 \) mutually independent random variables with any continuous distribution functions, denoting by \( F_{U_i} \) the cdf of each \( U_i \). Similarly to (1), the joint cdf of the \( q \)-dimensional random vector \((X_1, \ldots, X_q)\) with \( X_i = \max(X_1, X_2) \) is given by

\[
F(x_1, \ldots x_q) = F_{U_q+1}(\min(x_1, \ldots, x_q)) \prod_{i=1}^{q} F_{U_i}(x_i)
\]

which can be considered as a generator of \( q \)-dimensional distribution models, called generalized multivariate distribution (GMD) family with baseline distribution vector \((F_{U_1}, \ldots, F_{U_{q+1}})\). Hence, the \( q \)-dimensional copula representation of this GMD family can be expressed as

\[
C(u_1, \ldots, u_q) = \left( \prod_{i=1}^{q} u_i \right) \frac{\min_{u_1=1, \ldots, q} A_i(u_i)}{\prod_{i=1}^{q} A_i(u_i)},
\]

where

\[
A_i(u_i) = F_{U_{q+1}} \left( (F_{U_i} \times F_{U_{q+1}})^{-1}(u_i) \right), \quad \text{for } i = 1, \ldots, q.
\]

From these \( q \)-dimensional joint cdf and copula, many distributional and stochastic properties established for the GBD family are extensible to the GMD family. Furthermore, by using this generator of multivariate distributions, the special bivariate models given in Section 3 can be easily extended to the multivariate case, which contain multivariate versions of bivariate distributions given in the literature.

Author Contributions: Conceptualization, M.F., J.-M.V., and D.K.; methodology, M.F. and J.-M.V.; software, M.F. and D.K.; validation, M.F. and D.K.; formal analysis, M.F., J.-M.V., and D.K.; investigation, M.F., J.-M.V., and D.K.; writing—original draft preparation, M.F. and J.-M.V.; writing—review and editing, M.F., J.-M.V., and D.K.; supervision, M.F.; project administration, M.F. and J.-M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Spanish Ministry of Economy, Industry and Competitiveness, the European Regional Development Fund Program through grant TIN2017-85949-C2-1-R.

Acknowledgments: The authors would like to thank the editors and the anonymous reviewers for their comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Theorem 2. First, taking into account the event \( A = (U_3 > \max(U_1, U_2)) \), the joint cdf can be expressed as

\[
F(x_1, x_2) = P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq \min(x_1, x_2) | A) P(A)
+ P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq \min(x_1, x_2) | A') P(A')
\]

where \( A' \) is the complementary event of \( A \). For \( z = \min(x_1, x_2) \), note that

\[
P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq z | A) = P(U_1 \leq x_1, U_2 \leq x_2, U_3 \leq z | U_1 < U_3, U_2 < U_3)
= P(U_1 \leq U_3, U_2 \leq U_3, U_3 \leq z) = \int_{-\infty}^{z^2} F_{U_1}(u) F_{U_2}(u) dF_{U_3}(u).
\]
Hence, it is immediate that \( F(x_1, x_2) \) given by (3) is a singular cdf as its mixed second partial derivatives are zero when \( x_1 \neq x_2 \).

Thus, \( \alpha = P(A) \) may be established as follows:

\[
\alpha = P(U_3 > \max(U_1, U_2)) = \int_{-\infty}^{\infty} P(U_1 < u, U_2 < u) dF_{U_3}(u)
\]

\[
= \int_{-\infty}^{\infty} F_{U_1}(u) F_{U_2}(u) dF_{U_3}(u),
\]

and, consequently, the bivariate cdf \( F(x_1, x_2) \) can be rewritten as (2), where the absolutely continuous part \( F_{ac}(x_1, x_2) \) can be obtained by subtraction:

\[
F_{ac}(x_1, x_2) = \frac{1}{1 - \alpha} (F(x_1, x_2) - \alpha F_s(x_1, x_2))
\]

\[
= \frac{1}{1 - \alpha} \left( \int_{-\infty}^{\infty} F_{U_1}(u) F_{U_2}(u) dF_{U_3}(u) \right),
\]

which completes the proof of the theorem. \( \square \)

**Proof of Theorem 3.** Let \( \mu, \mu_s \) and \( \mu_{ac} \) be the measures associated with \( F, F_s \) and \( F_{ac} \), respectively. Obviously, \( \mu_{ac} \) is an absolutely continuous measure with respect to the two-dimensional Lebesgue measure since

\[
\mu_{ac} \left( (-\infty, x_1] \times (-\infty, x_2] \right) = F_{ac}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{ac}(u,v) du dv
\]

where the pdf associated with \( F_{ac} \) in (4), \( f_{ac}(u,v) = \frac{\partial^2}{\partial u \partial v} F_{ac}(u,v) \), can be written as

\[
f_{ac}(x_1, x_2) = \begin{cases} 
\frac{1}{1-\alpha} f_1(x_1, x_2), & \text{if } x_1 < x_2 \\
\frac{1}{1-\alpha} f_2(x_1, x_2), & \text{if } x_1 > x_2 \\
0, & \text{if } x_1 = x_2 = x.
\end{cases}
\]

On the other hand, \( \mu_s \) is given by

\[
\mu_s \left( (-\infty, x_1] \times (-\infty, x_2] \right) = F_s(x_1, x_2) = F_s(z,z) = \frac{1}{\alpha} \int_{-\infty}^{z} F_{U_1}(u) F_{U_2}(u) dF_{U_3}(u)
\]

where \( z = \min(x_1, x_2) \), and so it can be expressed as an absolutely continuous measure \( \mu_s^* \) with respect to the one-dimensional Lebesgue measure on the projection onto the line \( \mathbb{R} \) of the intersection between \( (-\infty, x_1] \times (-\infty, x_2] \) and the line \( x_1 = x_2 \):

\[
\mu_s \left( (-\infty, x_1] \times (-\infty, x_2] \right) = \mu_s^* \left( (-\infty, z] \right) = \int_{-\infty}^{z} f_1^*(u) du,
\]

where \( f_1^*(u) = \frac{1}{\alpha} F_{U_1}(u) F_{U_2}(u) f_{U_3}(u) \), which can be also written as \( f_1^*(u) = \frac{1}{\alpha} f_0(u) \).

Furthermore, it is trivial that the line \( x_1 = x_2 \) is a null set under the two-dimensional Lebesgue measure, and hence with respect to \( \mu_{ac} \). In addition, its complement \( \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \neq x_2 \} \) is a null set with respect to \( \mu_s \), since its projection onto the line \( \mathbb{R} \) is the empty set,

\[
\mu_s \left( \{(x_1, x_2) \in \mathbb{R}^2 | x_1 \neq x_2 \} \right) = \mu_s^* (\emptyset) = 0,
\]

and, consequently, the measures \( \mu_s \) and \( \mu_{ac} \) are mutually singular. Therefore, the measure associated with \( F \).
\[ \mu \left( (\infty, x_1] \times (\infty, x_2]) \right) = F(x_1, x_2) = a \mu_s \left( (\infty, x_1] \times (\infty, x_2] \right) + (1 - a) \mu_{ac} \left( (\infty, x_1] \times (\infty, x_2] \right) \]

allows us to have the pdf of a GBD model with respect to \( \mu \), given by

\[ f(x_1, x_2) = a f_s^* (x_1) I_{(x_1=x_2)} (x_1, x_2) + (1 - a) f_{ac} (x_1, x_2) \]

where \( I_{(x_1=x_2)} \) is the indicator function of \( x_1 = x_2 \). Hence, it is easy to check that

\[ \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(u, v) d\mu = F(x_1, x_2) \]

for all \((x_1, x_2) \in \mathbb{R}^2\). \( \square \)

**Proof of Theorem 4.** From (1) and (5), the proof of (1) of Theorem 4 is straightforward.

In order to prove (2) of Theorem 4, from the joint pdf of a GBD model given in Theorem 3 and its marginal pdf (6), the conditional pdf \( f_{i|x_j=x_j} \) can be expressed as

\[
\begin{aligned}
f_{i|x_j=x_j}(x_i) &= \begin{cases} 
\frac{f_{x_i}(x_i)f_{u_i}(x_i)}{f_{x_j}(x_j)}, & \text{if } x_i < x_j \\
f_{u_i}(x_i), & \text{if } x_i > x_j \\
\frac{f_{u_i}(x_i)f_{u_j}(x_i)f_{u_k}(x_i)}{f_{x_j}(x_j)}, & \text{if } x_i = x_j,
\end{cases}
\end{aligned}
\]

by using the notation \( \alpha_j = f_{i|X_j=x_j}(x_i) \), this conditional pdf can be readily rewritten as in the statement of Theorem 4. \( \square \)

**Proof of Theorem 11.** The reversed hazard function (12) of the minimum statistic can be rewritten as

\[ r_{T_1}(x) = r_{U_1}(x)g_2(x) + r_{U_2}(x)g_1(x) + r_{U_3}(x), \]

where each \( g_i \) is a positive function \((i = 1, 2)\) defined by

\[ g_i(x) = 1 - \frac{F_{U_i}(x)}{F_{U_{i+1}}(x)}. \]

Here, observe that \( U_{1:2} \leq r_{th} U_i \) implies the decreasing monotonicity of \( g_i(x) \), and therefore \( r_{T_1} \) is a sum of three decreasing functions, which completes the proof. \( \square \)

**Proof of Corollary 4.** The proof readily follows along the same line as Theorem 11, taking into account that (12) can be simplified by using

\[ r_{U_{i+2}}(x) = 2r_{U_i}(x)g_i(x) \]

where \( g_i(x) = 1 - \frac{1}{2r_{U_i}(x)} \) decreases in \( x \). \( \square \)

**Appendix B**

For practical implementation of the EM algorithm in the data analysis applications, we give the technical details of the EM algorithm for two particular GBD models, first with baseline component vector with the same distribution (Model I), and then with different baseline distributions (Model II).

**Model I.**

Suppose \( U_1 \sim \text{Exp}(\lambda_1), U_2 \sim \text{Exp}(\lambda_2) \) and \( U_3 \sim \text{Exp}(\lambda_3) \). To compute the MLEs of the unknown parameter vector \( \theta = (\lambda_1, \lambda_2, \lambda_3) \), one needs to solve a three-dimensional optimization problem.

For implementation of the EM algorithm, we need the following expected values:
1. If \( i \in I_0 \), then
\[
\begin{align*}
  u_{1im}(\theta) &= E(U_1 | U_1 < x_i) = H(x_i; \lambda_1) \\
  u_{2im}(\theta) &= E(U_2 | U_2 < x_i) = H(x_i; \lambda_2),
\end{align*}
\]
where
\[
H(x; \lambda) = \frac{1}{\lambda} - \frac{xe^{-\lambda x}}{1 - e^{-\lambda x}}.
\]

2. If \( i \in I_1 \), then
\[
\begin{align*}
  u_{1im}(\theta) &= E(U_1 | \max\{U_1, U_3\} = x_{1i}) = \frac{\lambda_3}{\lambda_1 + \lambda_3} x_{1i} + \frac{\lambda_1}{\lambda_1 + \lambda_3} H(x_{1i}; \lambda_1) \\
  u_{3im}(\theta) &= E(U_3 | \max\{U_1, U_3\} = x_{1i}) = \frac{\lambda_3}{\lambda_1 + \lambda_3} x_{1i} + \frac{\lambda_3}{\lambda_1 + \lambda_3} H(x_{1i}; \lambda_3).
\end{align*}
\]

3. If \( i \in I_2 \), then
\[
\begin{align*}
  u_{2im}(\theta) &= E(U_2 | \max\{U_2, U_3\} = x_{2i}) = \frac{\lambda_3}{\lambda_2 + \lambda_3} x_{2i} + \frac{\lambda_2}{\lambda_2 + \lambda_3} H(x_{2i}; \lambda_2) \\
  u_{3im}(\theta) &= E(U_3 | \max\{U_2, U_3\} = x_{2i}) = \frac{\lambda_3}{\lambda_2 + \lambda_3} x_{2i} + \frac{\lambda_3}{\lambda_2 + \lambda_3} H(x_{2i}; \lambda_3).
\end{align*}
\]

Hence, the ‘pseudo’ log-likelihood function in this case becomes
\[
\ell^{(k)}_s(\lambda_1, \lambda_2, \lambda_3) = \ell^{(k)}_{1s}(\lambda_1) + \ell^{(k)}_{2s}(\lambda_2) + \ell^{(k)}_{3s}(\lambda_3),
\]
where
\[
\begin{align*}
  \ell^{(k)}_{1s}(\lambda_1) &= n \ln \lambda_1 - \lambda_1 \left[ \sum_{i \in I_0 \cup I_1} u^{(k)}_{1im} + \sum_{i \in I_2} x_{1i} \right] \\
  \ell^{(k)}_{2s}(\lambda_2) &= n \ln \lambda_2 - \lambda_2 \left[ \sum_{i \in I_0 \cup I_2} u^{(k)}_{2im} + \sum_{i \in I_1} x_{2i} \right] \\
  \ell^{(k)}_{3s}(\lambda_3) &= n \ln \lambda_3 - \lambda_3 \left[ \sum_{i \in I_0 \cup I_2} u^{(k)}_{3im} + \sum_{i \in I_0} x_{1i} \right],
\end{align*}
\]
and the \( u^{(k)}_{im} \)s are obtained from \( u_{im}(\theta) \), \( j = 1, 2, 3 \), by replacing \( \theta = (\lambda_1, \lambda_2, \lambda_3) \) with \( \theta^{(k)} = (\lambda_1^{(k)}, \lambda_2^{(k)}, \lambda_3^{(k)}) \). Therefore,
\[
\begin{align*}
  \lambda_1^{(k+1)} &= \frac{n}{\sum_{i \in I_0 \cup I_1} u^{(k)}_{1im} + \sum_{i \in I_2} x_{1i}} \\
  \lambda_2^{(k+1)} &= \frac{n}{\sum_{i \in I_0 \cup I_2} u^{(k)}_{2im} + \sum_{i \in I_1} x_{2i}} \\
  \lambda_3^{(k+1)} &= \frac{n}{\sum_{i \in I_0 \cup I_2} u^{(k)}_{3im} + \sum_{i \in I_0} x_{1i}}.
\end{align*}
\]

Note that, in this case, the maximization can be performed analytically at each M-Step.
Model II.

Suppose \( U_1 \sim W(\lambda_1, a_1), U_2 \sim W(\lambda_2, a_2) \) and \( U_3 \sim GE(\lambda_3, \alpha_3) \). The pdf of a Weibull distribution \( W(\lambda, a) \) with scale parameter \( \lambda > 0 \) and the shape parameter \( a > 0 \) can be written as
\[
f_W(u; \lambda, a) = a\lambda e^{a-1} e^{-\lambda u^a}, \quad \text{for } u > 0,
\]
and zero otherwise. Similarly, the \( GE(\alpha, \lambda) \) model defined in Section 3 has the pdf
\[
f_{GE}(u; \alpha, \lambda) = \alpha \lambda e^{-\lambda u}(1 - e^{-\lambda u})^{\alpha - 1}; \quad \text{for } u > 0,
\]
and zero otherwise. Hence, one needs to solve a six-dimensional optimization problem to compute the MLEs of the unknown parameter vector \( \theta = (\theta_1, \theta_2, \theta_3) \) where each \( \theta_i \) represents the parameter vector of \( U_i \).

We need the following expected values for implementation of the EM algorithm:
1. If \( i \in I_0 \), then
   \[
u_{ijm}(\theta) = E(U_j | U_j < x_i) = H_W(x_i; a_j, \lambda_j), \quad j = 1, 2,
   \]
   \[H_W(x; a, \lambda) = \frac{1}{1 - e^{-\lambda x^a}} \int_0^{\lambda x^a} \left( \frac{u}{\lambda} \right)^{1/a} e^{-u} du.
   \]
2. If \( i \in I_1 \), then
   \[
u_{1im}(\theta) = E(U_1 | \max\{U_1, U_3\} = x\{i\}) = p_{13}x_{1i} + (1 - p_{13})H_W(x_{1i}; a_1, \lambda_1)
   \]
   \[u_{3im}(\theta) = E(U_3 | \max\{U_1, U_3\} = x_{2i}) = (1 - p_{13})x_{1i} + p_{13}H_G(x_{1i}; a_3, \lambda_3),
   \]
   where \( p_{13} = P(U_1 > U_3) = K(a_1, \lambda_1) \) and
   \[K(a, \lambda) = \int_0^\infty a \lambda x^{a-1} e^{-\lambda x} (1 - e^{-\lambda x})^a dx, \quad H_G(x; a, \lambda) = x - \frac{1}{\lambda (1 - e^{-\lambda x})^a} \int_0^{1-e^{-\lambda x}} \frac{t^a}{1-t} dt
   \]
3. If \( i \in I_2 \), then
   \[
u_{2im}(\theta) = E(U_2 | \max\{U_2, U_3\} = x_{2i}) = p_{23}x_{2i} + (1 - p_{23})H_W(x_{2i}; a_2, \lambda_2)
   \]
   \[u_{3im}(\theta) = E(U_3 | \max\{U_2, U_3\} = x_{2i}) = (1 - p_{23})x_{2i} + p_{23}H_G(x_{2i}; a_3, \lambda_3),
   \]
   where \( p_{23} = P(U_2 > U_3) = K(a_2, \lambda_2) \).

In this case, the terms of the ‘pseudo’ log-likelihood function \( \ell_3^{(k)}(\theta) \) can be written as
\[
\ell_1^{(k)}(a_1, \lambda_1) = n \ln a_1 + n \ln \lambda_1 + (a_1 - 1) \left[ \sum_{i \in I_0 \cup I_1} \ln u_{1im}^{(k)} + \sum_{i \in I_2} \ln x_{1i} \right] - \lambda_1 \left[ \sum_{i \in I_0 \cup I_1} (u_{1im}^{(k)})^{a_1} + \sum_{i \in I_2} x_{1i}^{a_1} \right] \tag{A1}
\]
\[
\ell_2^{(k)}(a_2, \lambda_2) = n \ln a_2 + n \ln \lambda_2 + (a_2 - 1) \left[ \sum_{i \in I_0 \cup I_2} \ln u_{2im}^{(k)} + \sum_{i \in I_1} \ln x_{2i} \right] - \lambda_2 \left[ \sum_{i \in I_0 \cup I_2} (u_{2im}^{(k)})^{a_2} + \sum_{i \in I_1} x_{2i}^{a_2} \right] \tag{A2}
\]
\[
\ell_3^{(k)}(a_3, \lambda_3) = n \ln a_3 + n \ln \lambda_3 + (a_3 - 1) \left[ \sum_{i \in I_0 \cup I_2} \ln (1 - e^{-\lambda_3 u_{3im}^{(k)}}) + \sum_{i \in I_0} \ln (1 - e^{-\lambda_3 x_{3i}}) \right] - \lambda_3 \left[ \sum_{i \in I_0} x_{3i} + \sum_{i \in I_1 \cup I_2} u_{3im}^{(k)} \right]. \tag{A3}
\]
Therefore, \( u_{1im}^{(k)}, u_{2im}^{(k)}, u_{3im}^{(k)} \) can be obtained from \( u_{1im}^{(1)}(\theta), u_{2im}^{(1)}(\theta) \) and \( u_{3im}^{(1)}(\theta) \) by replacing \( \theta = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \alpha_3, \lambda_3) \) with \( \theta^{(k)} = (\alpha_1^{(k)}(\lambda_1^{(k)}), \alpha_2^{(k)}(\lambda_2^{(k)}), \alpha_3^{(k)}(\lambda_3^{(k)})) \). Thus, \( \theta_1^{(k+1)} = (\alpha_1^{(k+1)}, \lambda_1^{(k+1)}) \), \( \theta_2^{(k+1)} = (\alpha_2^{(k+1)}, \lambda_2^{(k+1)}) \) and \( \theta_3^{(k+1)} = (\alpha_3^{(k+1)}, \lambda_3^{(k+1)}) \) can be obtained by maximizing (A1)–(A3), respectively. Hence, we obtain them as follows:

\[
\begin{align*}
\lambda_1^{(k+1)} &= \frac{n}{\sum_{i \in I_0 \cup I_1} (u_{1im}^{(k)})^{\alpha_1} + \sum_{i \in I_2} x_{1i}^{\alpha_1}}, \\
\lambda_2^{(k+1)} &= \frac{n}{\sum_{i \in I_0 \cup I_2} (u_{2im}^{(k)})^{\alpha_2} + \sum_{i \in I_1} x_{2i}^{\alpha_2}}, \\
\lambda_3^{(k+1)} &= \frac{n}{\sum_{i \in I_1 \cup I_2} (u_{3im}^{(k)})^{\alpha_3} + \sum_{i \in I_0} x_{3i}^{\alpha_3}}.
\end{align*}
\]

where

\[
\begin{align*}
p_1(\alpha_1) &= n \ln \alpha_1 - n \ln \left[ \sum_{i \in I_0 \cup I_1} (u_{1im}^{(k)})^{\alpha_1} + \sum_{i \in I_2} x_{1i}^{\alpha_1} \right] \\
&\quad + (\alpha_1 - 1) \left[ \sum_{i \in I_0 \cup I_1} \ln u_{1im}^{(k)} + \sum_{i \in I_2} \ln x_{1i} \right], \\
p_2(\alpha_2) &= n \ln \alpha_2 - n \ln \left[ \sum_{i \in I_0 \cup I_2} (u_{2im}^{(k)})^{\alpha_2} + \sum_{i \in I_1} x_{2i}^{\alpha_2} \right] \\
&\quad + (\alpha_2 - 1) \left[ \sum_{i \in I_0 \cup I_2} \ln u_{2im}^{(k)} + \sum_{i \in I_1} \ln x_{2i} \right], \\
p_3(\alpha_3) &= n \ln \alpha_3 - n \ln \left[ \sum_{i \in I_1 \cup I_2} (u_{3im}^{(k)})^{\alpha_3} + \sum_{i \in I_0} x_{3i}^{\alpha_3} \right] \\
&\quad - \lambda_3 \left[ \sum_{i \in I_0} x_i + \sum_{i \in I_1 \cup I_2} u_{3im}^{(k)} \right] - \left[ \sum_{i \in I_0} \ln (1 - e^{-\lambda_3 u_{3im}^{(k)}}) + \sum_{i \in I_0} \ln (1 - e^{-\lambda_3 x_i}) \right].
\end{align*}
\]

Note that, in this case, one needs to solve three one-dimensional optimization problems numerically at each M-Step.
Appendix C

Table A1. Summary of fitted GBD models for the two real data. EM rows are the parameters estimated with the EM algorithm for maximizing the pseudo log-likelihood function, along with the log-likelihood, AIC and BIC values, and BFGS rows correspond to the results obtained by applying the Broyden–Fletcher–Goldfarb–Shanno algorithm for maximizing the log-likelihood function.

| GBD Model | $\theta$ | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $\ell(\theta)$ | AIC | BIC |
|-----------|----------|-------------|-------------|-------------|----------------|-----|-----|
| Soccer data | Model I | EM | 0.03126 | 0.04630 | 0.04269 | −299.4331 | 604.8663 | 609.6990 |
| | | BFGS | 0.03116 | 0.04636 | 0.04283 | −299.4328 | 604.8656 | 609.6984 |
| | Model II | EM | 1.2987 | 0.0097 | 0.8047 | 0.0093 | 1.0037 | 0.0369 | −348.2715 | 708.5430 | 718.2085 |
| | | BFGS | 1.3808 | 0.00698 | 0.5652 | 0.2546 | 1.53813 | 0.05219 | −295.3057 | 602.6114 | 612.2770 |
| Diabetic retinopathy data | Model I | EM | 0.0653 | 0.0737 | 0.1345 | −289.9878 | 585.9757 | 590.8884 |
| | | BFGS | 0.06290 | 0.07181 | 0.14282 | −289.9144 | 585.8288 | 590.7415 |
| | Model II | EM | 1.0937 | 0.0447 | 0.5851 | 0.2369 | 0.8995 | 0.1898 | −290.0758 | 592.1515 | 601.9770 |
| | | BFGS | 1.1477 | 0.03920 | 0.7917 | 0.13923 | 0.41913 | 0.08272 | −285.5795 | 583.1590 | 592.9846 |

References

1. Gumbel, E.J. Bivariate exponential distributions. *J. Am. Stat. Assoc.* 1960, 55, 698–707, doi:10.1080/01621459.1960.10483368.
2. Freund, J.E. A bivariate extension of the exponential distribution. *J. Am. Stat. Assoc.* 1961, 56, 971–977, doi:10.1080/01621459.1961.10482138.
3. Marshall, A.W.; Olkin, I. A multivariate exponential distribution. *J. Am. Stat. Assoc.* 1967, 62, 30–44, doi:10.1080/01621459.1967.10482885.
4. Balakrishnan, N.; Lai, C.D. *Continuous Bivariate Distributions*, 2nd ed.; Springer: New York, NY, USA, 2009; doi:10.1007/b101765.
5. Franco, M.; Vivo, J.M. A multivariate extension of Sarhan and Balakrishnan’s bivariate distribution and its ageing and dependence properties. *J. Multivar. Anal.* 2010, 101, 491–499, doi:10.1016/j.jmva.2009.08.008.
6. Kundu, D.; Gupta, R.D. Modified Sarhan–Balakrishnan singular bivariate distribution. *J. Stat. Plan. Inference* 2010, 40, 526–538, doi:10.1016/j.jspi.2009.07.026.
7. Franco, M.; Kundu, D.; Vivo, J.M. Multivariate extension of the modified Sarhan-Balakrishnan bivariate distribution. *J. Stat. Plan. Inference* 2011, 141, 3400–3412, doi:10.1016/j.jspi.2011.03.017.
8. Gupta, R.C.; Kirlmani, S.N.U.A.; Balakrishnan, N. On a class of generalized Marshall–Olkin bivariate distributions and some reliability characteristics. *Probab. Engrg. Inform. Sci.* 2013, 27, 261–275, doi:10.1017/S0269964813000107.
9. Kundu, D.; Franco, M.; Vivo, J.M. Multivariate distributions with proportional reversed hazard marginals. *Comput. Stat. Data Anal.* 2014, 77, 98–112, doi:10.1016/j.csda.2014.02.004.
10. Muhammed, H.Z. On a bivariate generalized inverted Kumaraswamy distribution. *Physica A* 2020, 553, 124281, doi:10.1016/j.physa.2020.124281.
11. Franco, M.; Vivo, J.M.; Kundu, D. A generalized Freund bivariate model for a two-component load sharing system. *Reliab. Eng. Syst. Saf.* 2020, 203, 107096, doi:10.1016/j.ress.2020.107096.
12. El-Morsheyed, M.; Ali-Alhussain, Z.; Atta, D.; Almetwally, E.M.; Eliwa, M.S. Bivariate Burr X generator of distributions: Properties and estimation methods with applications to complete and type-II censored samples. *Mathematics* 2020, 8, 264, doi:10.3390/math8020264.
13. Kundu, D.; Gupta, R.D. Bivariate generalized exponential distribution. *J. Multivar. Anal.* **2009**, *100*, 581–593, doi:10.1016/j.jmva.2008.06.012.

14. Sarhan, A.M.; Hamilton, D.C.; Smith, B.; Kundu, D. The bivariate generalized linear failure rate distribution and its multivariate extension. *Comput. Stat. Data Anal.* **2011**, *55*, 644–654, doi:10.1016/j.csda.2010.06.006.

15. Elsherpieny, E.A.; Ibrahim, S.A.; Bedar, R.E. A New Bivariate Distribution with Log-Exponentiated Kumaraswamy Marginals. *Chil. J. Stat.* **2014**, *5*, 55–69. Available online: http://www.soche.cl/chjs/volumes/05/02/Elsherpieny_etal(2014).pdf (accessed on 31 August 2020).

16. El-Gohary, A.; El-Bassiouny, A.H.; El-Morshedy, M. Bivariate exponentiated modified Weibull extension distribution. *J. Stat. Appl. Probab.* **2016**, *5*, 67–78, doi:10.1214/16-BJPS313.

17. Muhammed, H.Z. Bivariate inverse Weibull distribution. *J. Stat. Comput. Simul.* **2016**, *86*, 2335–2345, doi:10.1080/00949655.2015.1110585.

18. Kundu, D.; Gupta, A.K. On bivariate inverse Weibull distribution. *Braz. J. Probab. Stat.* **2017**, *31*, 275–302, doi:10.1214/16-BJPS313.

19. Muhammed, H.Z. Bivariate Dagum Distribution. *Int. J. Reliab. Appl.* **2017**, *18*, 65–82. Available online: https://www.koreascience.or.kr/article/JAKO201715565837044.pdf (accessed on 31 August 2020).

20. Sarhan, A.M. The bivariate generalized Rayleigh distribution. *J. Math. Sci. Model.* **2019**, *2*, 99–111, doi:10.33187/jmsm.439873.

21. Eliwa, M.S.; El-Morshedy, M. Bivariate Gumbel-G family of distributions: Statistical properties, bayesian and non-bayesian estimation with application. *Ann. Data Sci.* **2019**, *6*, 39–60, doi:10.1007/s40745-018-00190-4.

22. Gupta, R.C.; Gupta, P.L.; Gupta, R.D. Modeling failure time data by Lehman alternatives. *Commun. Stat. Theory Methods* **1998**, *24*, 887–904, doi:10.1080/03610929808832134.

23. Di Crescenzo, A. Some results on the proportional reversed hazards model. *Stat. Probab. Lett.* **2000**, *50*, 313–321, doi:10.1016/S0167-7152(00)00127-9.

24. Kundu, D.; Gupta, R.D. A class of bivariate models with proportional reversed hazard marginals. *Sankhya B* **2010**, *72*, 236–253, doi:10.1007/s13571-011-0012-1.

25. Gupta, R.D.; Kundu, D. Generalized exponential distribution. *Aust. N. Z. J. Stat.* **1999**, *41*, 173–188, doi:10.1111/1467-842X.00072.

26. Sarhan, A.M.; Kundu, D. Generalized linear failure rate distribution. *Commun. Stat. Theory Methods* **2009**, *38*, 642–660, doi:10.1080/03610920802272414.

27. Lemonte, A.J.; Cordeiro, G.M.; Barreto-Souza, W. The exponentiated Kumaraswamy distribution and its log-transform. *Braz. J. Probab. Stat.* **2013**, *27*, 31–53, doi:10.1214/11-BJPS149.

28. Sarhan, A.M.; Apaloo, J. Exponentiated modified Weibull extension distribution. *Reliab. Eng. Syst. Saf.* **2013**, *112*, 137–144, doi:10.1016/j.ress.2012.10.013.

29. Keller, A.Z.; Giblin, M.T.; Farnworth, N.R. Reliability analysis of commercial vehicle engines. *Reliab. Eng.* **1985**, *10*, 15–25, doi:10.1016/0143-8174(85)90039-3.

30. Dagum, C. A new model of personal income distribution: Specification and estimation. *Econ. Appl.* **1977**, *30*, 413–437.

31. Burr, I.W. Cumulative frequency functions. *Ann. Math. Stat.* **1942**, *13*, 215–232, doi:10.1214/aoms/1177731607.

32. Alzaatreh, A.; Lee, C.; Famoye, F. A new method for generating families of continuous distributions. *Metron* **2013**, *71*, 63–79, doi:10.1007/s40300-013-0007-y.

33. Iqbal, Z.; Tahir, M.M.; Ali, S.A.; Ahmad, M. Generalized inverted Kumaraswamy distribution: Properties and application. *Open J. Stat.* **2017**, *7*, 645–662, doi:10.4236/ojs.2017.74045.

34. Lai, C.D.; Xie, M. *Stochastic Ageing and Dependence for Reliability*; Springer: New York, NY, USA, 2006, doi:10.1007/0-387-34322-X.

35. Shaked, M.; Shanthikumar, J.G. *Stochastic Orders*; Springer: New York, NY, USA, 2007, doi:10.1007/978-0-387-34675-5.

36. Domma, F. Bivariate reversed hazard rate, notions, and measures of dependence and their relationships. *Commun. Stat. Theory Methods* **2011**, *40*, 989–999, doi:10.1080/03610920903511777.

37. Finkelstein, M.S. On the reversed hazard rate. *Reliab. Eng. Syst. Saf.* **2002**, *78*, 71–75, doi:10.1016/S0951-8320(02)00113-8.
38. Gupta, R.C.; Balakrishnan, N. Log-concavity and monotonicity of hazard and reversed hazard functions of univariate and multivariate skew-normal distributions. *Metrika* **2012**, *75*, 181–191, doi:10.1007/s00184-010-0321-9.

39. Nelsen, R.B. *An Introduction to Copulas*, 2nd ed.; Springer: New York, NY, USA, 2006, doi:10.1007/0-387-28678-0.

40. Fang, R.; Li, X. A note on bivariate dual generalized Marshall–Olkin distributions with applications. *Probab. Engrg. Inform. Sci.* **2013**, *27*, 367–374, doi:10.1017/S0269964813000089.

41. Meintanis, S.G. Test of fit for Marshall–Olkin distribution with applications. *J. Stat. Plan. Inference* **2007**, *137*, 3954–3963, doi:10.1016/j.jspi.2007.04.013.

42. Huster, W.J.; Brookmeyer, R.; Self, S.G. Modelling paired survival data with covariates. *Biometrics* **1989**, *45*, 145–156, doi:10.2307/2532041.

43. Ploner, M.; Kaider, A.; Heinze, G. SurvCorr: Correlation of Bivariate Survival Times. R Package Version 1.0. 2015. Available online: [https://CRAN.R-project.org/package=SurvCorr](https://CRAN.R-project.org/package=SurvCorr) (accessed on 31 August 2020).

44. Henningsen, A.; Toomet, O. maxLik: A package for maximum likelihood estimation in R. *Comput. Stat.* **2011**, *26*, 443–458, doi:10.1007/s00180-010-0217-1.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license ([http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/)).