Marques, F., Coelho, C. A., and de Carvalho, M. (2015), "On the Distribution of Linear Combinations of Independent Gumbel Random Variables," Statistics and Computing, 25, 683–701.
Stat Comput
DOI 10.1007/s11222-014-9453-5

On the distribution of linear combinations of independent Gumbel random variables

Filipe J. Marques · Carlos A. Coelho · Miguel de Carvalho

Received: 8 August 2012 / Accepted: 7 February 2014

Abstract The distribution of linear combinations of independent Gumbel random variables is of great interest for modeling risk and extremes in the most different areas of application. In this paper we develop near-exact approximations for the distribution of linear combination of independent Gumbel random variables based on a shifted generalized near-integer gamma distribution and on the distribution of the difference of two independent generalized integer gamma distributions. These near-exact distributions are computationally appealing and numerical studies confirm their accuracy, as assessed by a proximity measure used in related studies. We illustrate the proposed approximations on applied problems in networks engineering, computational biology, and flood risk management.

Keywords Generalized integer gamma distribution · Generalized near-integer gamma distribution · Gumbel distribution · Near-exact distribution · Phase type distributions · Risk

Electronic supplementary material The online version of this article (doi:10.1007/s11222-014-9453-5) contains supplementary material, which is available to authorized users.

F. J. Marques (✉) · C. A. Coelho
DM–CMA–FCT, Universidade Nova de Lisboa, Lisbon, Portugal
e-mail: fjm@fct.unl.pt

M. de Carvalho
Pontificia Universidad Católica de Chile, Santiago, Chile

M. de Carvalho
CMA-FCT, Universidade Nova de Lisboa, Lisbon, Portugal

1 Introduction

The Gumbel distribution is a particular case of the Generalized Extreme Value distribution and it has been widely used for modeling risk and extremes (Gumbel 1941; Tiago de Oliveira 1963; Hosking et al. 1985; Balakrishnan et al. 1992; Wang 1995; Arnold et al. 1998; Castillo et al. 2005; Antal et al. 2009). Linear combinations of Gumbel related random variates arise naturally in applications whenever there is the need to model the combination of extremes of several variables, and this has been a topic of considerable attention in diverse applications (Bailey and Gribskov 1997; Cetinkaya et al. 2001; Loaiciga and Leipnik 1999; Burda et al. 2012). Despite the wide range of applications in which the distribution of linear combinations of independent Gumbel random variables may be useful, few results are available on this distribution. Nadarajah (2008) presents the exact distribution of the linear combination of p independent Gumbel random variables, using Fox H and Meijer G functions, but the computational investment required by these functions limits the practical usefulness of this result. This was already remarked by Burda et al. (2012, p. 189), who claimed that the exact distribution proposed by Nadarajah is "extremely complicated to be used."

In this paper we propose three accurate, manageable, and computationally appealing near-exact distributions for the linear combination of independent Gumbel random variables; the first one for positive linear combinations, and the second and third ones can be applied regardless of the sign of the coefficients of the linear combination. Our near-exact distributions have links with phase-type approximations (Aldous and Shepp 1987; O'Cinneide 1990) and, as we discuss below, their accuracy can be controlled effectively through a precision parameter. Our first near-exact

Published online: 14 March 2014

distribution is based on the generalized integer gamma (GIG) and generalized near-integer gamma (GNIG) distributions, which have a wealth of applications in multivariate analysis (Marques and Coelho 2008; Coelho and Marques 2010, 2012; Marques et al. 2011; Coelho et al. 2013). The GIG distribution corresponds to the distribution of the sum of independent Gamma random variables with integer shape parameters (Amari and Misra 1997; Coelho 1998), while the GNIG distribution corresponds to the distribution of the sum of a GIG random variable with an independent Gamma random variable with a non-integer shape parameter (Coelho 2004); further details on the GIG and GNIG distributions are given in Appendix 1. We show that the exact distribution of a positive linear combination of independent Gumbel random variables can be decomposed as the sum of two independent random variables: the first corresponding to a linear combination of independent logarithmized Gamma random variables, and the second to a shifted GIG (SGIG) random variable. Our second near-exact distribution is based on the so-called SDGIG distribution, which corresponds to the distribution of the shifted difference of two independent GIG distributions (Coelho and Mexia 2010, Chap. 2). Hence, in the context of our second near-exact distribution, we show that the linear combination of independent Gumbel random variables can be decomposed as the sum of two independent random variables: the first corresponding to a linear combination of independent logarithmized Gamma random variables, and the second corresponding to a SDGIG random variable. The third near-exact distribution is also based on the previous decomposition, and on the fact that the DGIG distribution can be represented as a particular mixture of integer Gamma distributions. These decompositions are extremely useful as they allow us to construct near-exact distributions by using a shifted version of the GNIG distribution—in the case of positive linear combinations—and by using the SDGIG distribution—in the case where the coefficients of linear combination are arbitrary real numbers.

We illustrate our near-exact approximations by revisiting a problem in network engineering, first addressed by Cetinkaya et al. (2001), a problem in computational biology, earlier considered by Bailey and Gribskov (1997), and by addressing the problem of interval estimation of the location parameter of a Gumbel distribution in a real data application on flood risk management, earlier discussed in Hosking et al. (1985).

The structure of our paper is as follows. In Sect. 2 we introduce the exact and near-exact distributions of interest. In Sect. 3 we conduct numerical experiments to assess the level of accuracy of our near-exact approximations. In Sect. 4 we illustrate our methods in applied modeling issues, and we conclude in Sect. 5.

2 The exact and near-exact distributions

2.1 Exact distribution

Let $X_1, \ldots, X_p$ be p independent Gumbel random variables, with location parameter $\mu_j \in \mathbb{R}$ and scale parameter $\sigma_j \in \mathbb{R}_+^*$, i.e.

$$X_j \overset{\text{ind.}}{\sim} \text{Gumbel}(\mu_j, \sigma_j),$$
$$F_{X_j}(x) = \exp[-\exp\{-(x-\mu_j)/\sigma_j\}], \quad x \in \mathbb{R}, \tag{1}$$

for $j = 1, \ldots, p$. Here and below we use the notations $\mathbb{R}_+^*$ and $\mathbb{A}$ to respectively denote the sets $\{x \in \mathbb{R} : x > 0\}$ and $\{n \in \mathbb{N} : n \geq 2\}$. The characteristic functions of X_j and $W = \sum_{j=1}^p \alpha_j X_j$, for $\alpha_j \in \mathbb{R}$, are respectively defined as

$$\Phi_{X_j}(t) = \Gamma(1 - \mathrm{i}t\sigma_j) \exp\{\mathrm{i}t\mu_j\},$$

$$\Phi_W(t) = \prod_{j=1}^p \Gamma(1 - \mathrm{i}t\sigma_j\alpha_j) \exp\{\mathrm{i}t\mu_j\alpha_j\}, \quad t \in \mathbb{R}.$$

The next theorem provides a characterization of the exact distribution of the linear combination of independent Gumbel random variables.

Theorem 1 *Let* $X_j \overset{\text{ind.}}{\sim} \text{Gumbel}(\mu_j, \sigma_j)$, *with* $\mu_j \in \mathbb{R}$ *and* $\sigma_j \in \mathbb{R}_+^*$. *The exact characteristic function of* $W = \sum_{j=1}^p \alpha_j X_j$, *with* $\alpha_j \in \mathbb{R}$, $j = 1, \ldots, p$, *can be written as* $\Phi_W(t) = \Phi_{W_1}(t)\Phi_{W_2}(t)$, *where for any* $\gamma \in \mathbb{A}$,

$$\Phi_{W_1}(t) = \prod_{j=1}^p \frac{\Gamma(\gamma - \mathrm{i}t\sigma_j\alpha_j)}{\Gamma(\gamma)}, \quad t \in \mathbb{R}, \tag{2}$$

and

$$\Phi_{W_2}(t) = \left\{ \prod_{j=1}^p \prod_{k=0}^{\gamma-2} \left(\frac{1+k}{\sigma_j\alpha_j}\right) \left(\frac{1+k}{\sigma_j\alpha_j} - \mathrm{i}t\right)^{-1} \right\} \times \exp\left\{\mathrm{i}t \sum_{j=1}^p \mu_j\alpha_j\right\}, \quad t \in \mathbb{R}. \tag{3}$$

Proof The proof follows by noticing that we can write the characteristic function of W as

$$\begin{aligned}
\Phi_W(t) &= \prod_{j=1}^p \Gamma(1 - \mathrm{i}t\sigma_j\alpha_j) \exp\{\mathrm{i}t\mu_j\alpha_j\} \\
&= \left\{ \prod_{j=1}^p \frac{\Gamma(\gamma - \mathrm{i}t\sigma_j\alpha_j)}{\Gamma(\gamma)} \frac{\Gamma(\gamma)}{\Gamma(\gamma - \mathrm{i}t\sigma_j\alpha_j)} \right. \\
&\quad \left. \times \frac{\Gamma(1 - \mathrm{i}t\sigma_j\alpha_j)}{\Gamma(1)} \right\} \exp\left\{\mathrm{i}t \sum_{j=1}^p \mu_j\alpha_j\right\} \\
&= \prod_{j=1}^p \frac{\Gamma(\gamma - \mathrm{i}t\sigma_j\alpha_j)}{\Gamma(\gamma)} \left\{ \prod_{j=1}^p \prod_{k=0}^{\gamma-2} (1+k) \right.
\end{aligned}$$

$$\times \left(1 + k - \mathrm{i}t\sigma_j\alpha_j\right)^{-1}\Bigg\} \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\}$$

$$= \prod_{j=1}^{p}\frac{\Gamma(\gamma - \mathrm{i}t\sigma_j\alpha_j)}{\Gamma(\gamma)}\left\{\prod_{j=1}^{p}\prod_{k=0}^{\gamma-2}\left(\frac{1+k}{\sigma_j\alpha_j}\right)\right.$$

$$\left.\times\left(\frac{1+k}{\sigma_j\alpha_j} - \mathrm{i}t\right)^{-1}\right\} \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\}.$$

□

Some comments are in order.

(i) We can write $W = \sum_{j=1}^{p} X'_j$, where $X'_j = \alpha_j X_j \sim \text{Gumbel}(\alpha_j\mu_j, \alpha_j\sigma_j)$, and hence an alternative parameterization can be considered by taking $(\mu'_j, \sigma'_j) = (\alpha_j\mu_j, \alpha_j\sigma_j)$ and setting the corresponding coefficients of the linear combination as $\alpha'_j = 1$; in addition, to simplify the expressions we can consider $\mu_j = 0$, in which case we would be working with a similar distribution apart from a shift.

(ii) Our results can be readily applied to the product of powers of independent Weibull and Fréchet random variables, through simple transformations; actually if $X_j \sim \text{Gumbel}(\mu_j, \sigma_j)$, then $Y_j = \exp\{-X_j\} \sim \text{Weibull}(\exp\{-\mu_j\}, \sigma_j^{-1})$, with distribution function

$$F_{Y_j}(y) = 1 - \exp\left\{-\left(\frac{y}{\exp(-\mu_j)}\right)^{1/\sigma_j}\right\},$$

and thus $\prod_{j=1}^{p} Y_j^{\alpha_j} = \exp\{-\sum_{j=1}^{p}\alpha_j X_j\}$. If $Y_j^* = \exp\{X_j\}$ then $Y_j^* \sim \text{Fréchet}(\exp\{\mu_j\}, \sigma_j^{-1})$, with distribution function

$$F_{Y_j^*}(y) = \exp\left\{-\left(\frac{y}{\exp(\mu_j)}\right)^{-1/\sigma_j}\right\},$$

and thus $\prod_{j=1}^{p}(Y_j^*)^{\alpha_j} = \exp\{\sum_{j=1}^{p}\alpha_j X_j\}$. Using similar transformations it is also possible to apply our results to more complex distributions, such as the Generalized Gamma distribution (Marques 2012).

2.1.1 *Positive linear combinations* ($\alpha_j > 0$)

If all α_j are positive, the exact distribution of W is the same as that of the sum of two independent random variables, W_1 and W_2, where

$$W_1 = -\sum_{j=1}^{p}\sigma_j\alpha_j \log Z_j, \quad Z_j \overset{\text{ind.}}{\sim} \text{Gamma}(\gamma, 1), \tag{4}$$

with $\gamma \in \mathbb{A}$, is a linear combination of p independent logarithmized Gamma random variables and W_2 is distributed according to a shifted sum of $p \times (\gamma - 1)$ independent Exponential distributions with parameters $(1+k)/(\sigma_j\alpha_j)$, for $j = 1, \dots, p$ and $k = 0, \dots, \gamma - 2$, with shift parameter $\sum_{j=1}^{p}\mu_j\alpha_j$.

If some of the Exponential distributions in (3) have the same parameter we can sum them, obtaining in this way Gamma distributions, so that equation (3) can be written as

$$\Phi_{W_2}(t) = \left\{\prod_{j=1}^{\ell}(\lambda_j)^{r_j}(\lambda_j - \mathrm{i}t)^{-r_j}\right\} \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\}, \tag{5}$$

where ℓ is the number of Exponential distributions with different rate parameters, λ_j are the parameters of these distributions, and r_j is the number of such distributions with the same rate parameter λ_j, for $j = 1, \dots, \ell$. We have thus established the following corollary to Theorem 1.

Corollary 1 *Let* $X_j \overset{\text{ind.}}{\sim} \text{Gumbel}(\mu_j, \sigma_j)$, *with* $\mu_j \in \mathbb{R}$ *and* $\sigma_j \in \mathbb{R}_+^*$. *If* $W = \sum_{j=1}^{p}\alpha_j X_j$, *with* $\alpha_j \in \mathbb{R}_+^*$, $j = 1, \dots, p$, *then it holds that* $W = W_1 + W_2$, *with* W_1 *as in* (4) *and*

$$W_2 \sim \text{SGIG}\left(\boldsymbol{r}, \boldsymbol{\lambda}, \ell, \sum_{j=1}^{p}\mu_j\alpha_j\right),$$

where $\boldsymbol{r} = (r_1, \dots, r_\ell)$ *and* $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_\ell)$.

Here and below we use the letter ‘S’ to denote a shifted distribution, and we follow the convention that the last parameter in a shifted distribution is the shift parameter; see Appendix 1 for further details.

It is instructive to consider the case of the sum of p independent Gumbel random variables when $\sigma_j = \sigma$, $j = 1, \dots, p$, for which simple expressions of the characteristic functions are readily available, as a consequence of Corollary 1,

$$\Phi_{W_1}(t) = \left(\frac{\Gamma(\gamma - \mathrm{i}t\sigma)}{\Gamma(\gamma)}\right)^p,$$

$$\Phi_{W_2}(t) = \left\{\prod_{j=0}^{\gamma-2}(\lambda_j)^{r_j}(\lambda_j - \mathrm{i}t)^{-r_j}\right\} \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\right\}, \tag{6}$$

with $r_j = p$, $\lambda_j = (1+j)/\sigma$, for $j = 0, \dots, \gamma - 2$; this implies that in such case

$$W_2 \sim \text{SGIG}\left(p\mathbf{1}_{\gamma-1}^{\mathrm{T}}, \sigma^{-1}(1, \dots, \gamma - 1), \gamma - 1, \sum_{j=1}^{p}\mu_j\right),$$

where $\mathbf{1}_{\gamma-1}$ denotes a $\gamma - 1$ vector of ones. The parameter γ is related with the depth of the SGIG distribution and it may be used as a precision parameter, since, as we will see in Sect. 3, larger values of γ lead to more accurate near-exact approximations.

2.1.2 General linear combinations ($\alpha_j \in \mathbb{R}$)

If we have q positive α_j and $p-q$ negative α_j, the characteristic function in (3) can be written as

$$\Phi_{W_2}(t) = \left\{ \prod_{\{j:\alpha_j>0\}} \prod_{k=0}^{\gamma-2} \left(\frac{1+k}{\sigma_j\alpha_j}\right)\left(\frac{1+k}{\sigma_j\alpha_j} - \mathrm{i}t\right)^{-1} \right\}$$
$$\times \left\{ \prod_{\{j:\alpha_j<0\}} \prod_{k=0}^{\gamma-2} \left(\frac{1+k}{\sigma_j\alpha_j}\right)\left(\frac{1+k}{\sigma_j\alpha_j} + \mathrm{i}t\right)^{-1} \right\}$$
$$\times \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\},$$

so that similarly to (5), we obtain

$$\Phi_{W_2}(t) = \left\{ \prod_{j=1}^{\ell^+}(\lambda_j^+)^{r_j^+}(\lambda_j^+ - \mathrm{i}t)^{-r_j^+} \prod_{j=1}^{\ell^-}(\lambda_j^-)^{r_j^-} \right.$$
$$\left. \times (\lambda_j^- + \mathrm{i}t)^{-r_j^-} \right\} \exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\}. \qquad (7)$$

where $\boldsymbol{r}^+ = (r_1^+, \ldots, r_{\ell^+}^+)$ and $\boldsymbol{\lambda}^+ = (\lambda_1^+, \ldots, \lambda_{\ell^+}^+)$, are respectively the shape and rate parameters corresponding to the positive α_j, and $\boldsymbol{r}^- = (r_1^-, \ldots, r_{\ell^-}^-)$ and $\boldsymbol{\lambda}^- = (\lambda_1^-, \ldots, \lambda_{\ell^-}^-)$ are respectively the shape and rate parameters corresponding to the negative α_j. In this case the exact distribution of W is the distribution of the sum of two independent random variables, W_1 and W_2, where W_1 is as in (4) and W_2 follows a SDGIG distribution. This gives rise to the following corollary.

Corollary 2 *Let $X_j \overset{\text{ind.}}{\sim} \text{Gumbel}(\mu_j, \sigma_j)$, with $\mu_j \in \mathbb{R}$ and $\sigma_j \in \mathbb{R}_+^*$. If $W = \sum_{j=1}^p \alpha_j X_j$, with $\alpha_j \in \mathbb{R}$, $j = 1, \ldots, p$, then it holds that $W = W_1 + W_2$, with W_1 as in (4) and*

$$W_2 \sim \text{SDGIG}\left(\boldsymbol{r}^+, \boldsymbol{r}^-, \boldsymbol{\lambda}^+, \boldsymbol{\lambda}^-, \ell^+, \ell^-, \sum_{j=1}^{p}\mu_j\alpha_j\right), \qquad (8)$$

where $\boldsymbol{r}^+ = (r_1^+, \ldots, r_{\ell^+}^+)$ and $\boldsymbol{\lambda}^+ = (\lambda_1^+, \ldots, \lambda_{\ell^+}^+)$ are respectively the shape and rate parameters corresponding to the positive α_j and $\boldsymbol{r}^- = (r_1^-, \ldots, r_{\ell^-}^-)$ and $\boldsymbol{\lambda}^- = (\lambda_1^-, \ldots, \lambda_{\ell^-}^-)$ are respectively the shape and rate parameters corresponding to the negative α_j.

2.2 Near-exact distributions

2.2.1 First near-exact distribution ($\alpha_j > 0$)

Our first near-exact distribution is based on replacing Φ_{W_1} by an asymptotic approximation $\Phi_{W_1^\star}$, such that for γ sufficiently large

$$\Phi_{W^\star}(t) = \Phi_{W_1^\star}(t)\Phi_{W_2}(t),$$

approximates the exact characteristic function Φ_W; the distribution of the random variable $W^\star$ is said to be a near-exact distribution of W (Coelho 2004). Based on the characterization of the exact distribution of W in Corollary 1, we take

$$\Phi_{W_1^\star}(t) = \left(\frac{l}{l - \mathrm{i}t}\right)^{\rho} \exp\{\mathrm{i}t\theta\}, \qquad (9)$$

which is the characteristic function of a random variable $W_1^\star \sim \text{SGamma}(\rho, l, \theta)$, and replaces asymptotically Φ_{W_1} in (2), for increasing values of γ; see Appendix 1 for details on the shifted Gamma distribution. Our choice is based on the fact that a single logarithmized Gamma random variable may be represented as an infinite sum of independent shifted Exponential random variables (see Appendix 2 for details), and as such the sum of independent logarithmized Gamma random variables, eventually multiplied by a parameter, may be represented as an infinite sum of shifted Gamma distributions. Instead of this infinite sum of shifted Gamma distributions, to avoid computational difficulties, we use a single shifted Gamma distribution, which matches the first three exact moments. Hence, the parameters ρ, l, and θ, are determined by solving the system of equations

$$\left.\frac{\partial^j \Phi_{W_1^\star}(t)}{\partial t^j}\right|_{t=0} = \left.\frac{\partial^j \Phi_{W_1}(t)}{\partial t^j}\right|_{t=0}, \quad j = 1, 2, 3, \qquad (10)$$

so to ensure that the first three of moments of the exact and approximating distributions are equal. The solution to (10) is

$$\begin{aligned}
\rho &= 4(\psi_1\Sigma_2)^3(\psi_2\Sigma_3)^{-2},\\
l &= 2(\psi_1\Sigma_2)|\psi_2\Sigma_3|^{-1},\\
\theta &= -\psi_0\Sigma_1 - 2(\psi_1\Sigma_2)^2|\psi_2\Sigma_3|^{-1}.
\end{aligned} \qquad (11)$$

where we use the following notation throughout the paper

$$\psi_i \equiv \psi_i(\gamma) = \frac{\partial^{i+1}}{\partial\gamma^{i+1}}\log\{\Gamma(\gamma)\},$$
$$\Sigma_i \equiv \Sigma_i(\boldsymbol{\alpha}) = \sum_{j=1}^{p}(\alpha_j\sigma_j)^i, \quad \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p). \qquad (12)$$

The resulting near-exact distribution is established in the next theorem.

Theorem 2 *Let $X_j \overset{\text{ind.}}{\sim} \text{Gumbel}(\mu_j, \sigma_j)$, with $\mu_j \in \mathbb{R}$ and $\sigma_j \in \mathbb{R}_+^*$. If we use as an asymptotic approximation of $\Phi_{W_1}(t)$ in (2) the characteristic function $\Phi_{W_1^\star}(t)$ in (9), we obtain as near-exact distribution for $W = \sum_{j=1}^p \alpha_j X_j$, with $\alpha_j \in \mathbb{R}_+^*$, $j = 1, \ldots, p$, the shifted* GNIG *distribution*

$$\text{SGNIG}\left(\boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1, \theta + \sum_{j=1}^{p}\mu_j\alpha_j\right),$$

with $\boldsymbol{r}^\star = (r_1, \ldots, r_\ell, \rho)$ *and* $\boldsymbol{\lambda}^\star = (\lambda_1, \ldots, \lambda_\ell, l)$, *and where the* r_j, λ_j, *and* ℓ *are given in* (5) *and* ρ, l, *and* θ *are given by* (11).

Proof It is enough to note that for each $t \in \mathbb{R}$, it holds that

$$\begin{aligned}\Phi_{W_1^\star}(t)\Phi_{W_2}(t) &= \left(\frac{l}{l-\mathrm{i}t}\right)^{\rho} \exp\{\mathrm{i}t\theta\}\left\{\prod_{j=1}^{\ell}(\lambda_j)^{r_j}\right.\\ &\quad\left.\times(\lambda_j-\mathrm{i}t)^{-r_j}\right\}\exp\left\{\mathrm{i}t\sum_{j=1}^{p}\mu_j\alpha_j\right\}\\ &= \left\{\prod_{j=1}^{\ell}(\lambda_j)^{r_j}(\lambda_j-\mathrm{i}t)^{-r_j}\right\}\left(\frac{l}{l-\mathrm{i}t}\right)^{\rho}\\ &\quad\times\exp\left\{\mathrm{i}t\left(\theta+\sum_{j=1}^{p}\mu_j\alpha_j\right)\right\}.\end{aligned}$$

□

It is again instructive to consider the particular case addressed in (6), that is when we consider the case of the sum of independent Gumbel random variables with the same scale parameter. In this case we obtain the near-exact distribution

$$\mathrm{SGNIG}\left(\boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \gamma, \theta+\sum_{j=1}^{p}\mu_j\right), \tag{13}$$

where $\boldsymbol{r}^\star = (p\mathbf{1}_{\gamma-1}^{\mathrm{T}}, \rho)$ and $\boldsymbol{\lambda}^\star = \sigma^{-1}(1, \ldots, \gamma-1, l\sigma)$, with ρ, l, and θ given by (11).

Using Corollary 2, and following two interesting recommendations made by an anonymous reviewer, we next develop two near-exact distributions for the case where the sign of the coefficients needs not to be positive.

2.2.2 Second near-exact distribution ($\alpha_j \in \mathbb{R}$)

We now develop a near-exact distribution, that although less accurate, it is computationally fast and can be applied to the case of an arbitrary real α_j. To do so, we approximate the distribution of $W = W_1 + W_2$ in Corollary 2, with the distribution of $W_\star = E(W_1) + W_2$, where $E(W_1) = -\psi_0\Sigma_1$ with ψ_0 and Σ_1 as defined in (12). The distribution of $W_\star$ corresponds to our second near-exact approximation, and as described in the next theorem $W_\star$ follows a SDGIG distribution with shift parameter $E(W_1) + \sum_{j=1}^{p}\mu_j\alpha_j$.

Theorem 3 *Let* $X_j \overset{\text{ind.}}{\sim} \mathrm{Gumbel}(\mu_j, \sigma_j)$, *with* $\mu_j \in \mathbb{R}$ *and* $\sigma_j \in \mathbb{R}_+^*$. *If we replace* W_1 *by* $E(W_1)$ *we obtain as near-exact distribution for* $W = \sum_{j=1}^{p}\alpha_j X_j$, *with* $\alpha_j \in \mathbb{R}$, $j = 1, \ldots, p$, *the shifted* DGIG *distribution*

$$\mathrm{SDGIG}\left(\boldsymbol{r}^+, \boldsymbol{r}^-, \boldsymbol{\lambda}^+, \boldsymbol{\lambda}^-, \ell^+, \ell^-, E(W_1)+\sum_{j=1}^{p}\mu_j\alpha_j\right),$$

where $\boldsymbol{r}^+$, $\boldsymbol{r}^-$, $\boldsymbol{\lambda}^+$, *and* $\boldsymbol{\lambda}^-$ *are as in Corollary* 2.

2.2.3 Third near-exact distribution ($\Sigma_3(\boldsymbol{\alpha}) \neq 0$)

Our third near-exact distribution can be applied when $\Sigma_3 \neq 0$, where Σ_3 is defined in (12); below we assume that $W_1^\star \sim \mathrm{SGamma}(\rho, l, \theta)$ and that $\Sigma_3 \neq 0$, so that either $\mathrm{sign}(\Sigma_3) = 1$ or $\mathrm{sign}(\Sigma_3) = -1$, with $\mathrm{sign}(\cdot)$ denoting the sign function. This near-exact distribution is also based on Corollary 2, but here we approximate the distribution of W_1 in (4) with the distribution of $\mathrm{sign}(\Sigma_3) \times W_1^\star$, whose characteristic function is $\Phi_{\mathrm{sign}(\Sigma_3)\times W_1^\star}(t) = \Phi_{W_1^\star}(\mathrm{sign}(\Sigma_3)\times t)$, and where $\Phi_{W_1^\star}(t)$ is as in (9). Here, the parameters ρ, l, and θ are determined by solving the system of equations

$$\left.\frac{\partial^j \Phi_{W_1^\star}(\mathrm{sign}(\Sigma_3)\times t)}{\partial t^j}\right|_{t=0} = \left.\frac{\partial^j \Phi_{W_1}(t)}{\partial t^j}\right|_{t=0}, \tag{14}$$

for $j = 1, 2, 3$, which has a solution if and only if $\Sigma_3 \neq 0$, in which case

$$\begin{aligned}\rho &= 4(\psi_1\Sigma_2)^3(\psi_2\Sigma_3)^{-2},\\ l &= 2(\psi_1\Sigma_2)|\psi_2\Sigma_3|^{-1},\\ \theta &= -\mathrm{sign}(\Sigma_3)\psi_0\Sigma_1 - 2(\psi_1\Sigma_2)^2|\psi_2\Sigma_3|^{-1}.\end{aligned} \tag{15}$$

The following theorem holds.

Theorem 4 *Let* $X_j \overset{\text{ind.}}{\sim} \mathrm{Gumbel}(\mu_j, \sigma_j)$, *with* $\mu_j \in \mathbb{R}$ *and* $\sigma_j \in \mathbb{R}_+^*$. *If we use as an asymptotic approximation of* $\Phi_{W_1}(t)$ *in* (2) *the characteristic function* $\Phi_{W_1^\star}(\mathrm{sign}(\Sigma_3)\times t)$ *in* (9), *we obtain as near-exact distribution for* $W = \sum_{j=1}^{p}\alpha_j X_j$, *with* $\Sigma_3 \neq 0$, *the distribution of* $\mathrm{sign}(\Sigma_3)\times W_1^\star + W_2$ *where* $W_1^\star \sim \mathrm{SGamma}(\rho, l, \theta)$ *and* W_2 *are as in (8), and where* ρ, l, *and* θ *are given by* (15).

Technical details on the distribution of $\mathrm{sign}(\Sigma_3)\times W_1^\star + W_2$ can be found in the Appendix 1.

3 Numerical studies

3.1 Measuring accuracy

To study the quality of our near-exact approximations we use a measure of proximity between characteristic functions, that is also a measure of the proximity between distribution functions, and which is defined as

$$\Delta = \frac{1}{2\pi}\int_{\mathbb{R}}\left|\frac{\Phi_W(t)-\Phi_{W^\star}(t)}{t}\right|\,\mathrm{d}t\,. \tag{16}$$

This is measure is known to be related with the Berry–Esseen upper bound (Berry 1941; Esseen 1945; Loève 1977; Hwang 1998), and can be shown to verify the inequality

Table 1 Values of Δ for Scenarios I–III

γ	Scenario I $(\boldsymbol{\mu}_{\mathrm{I}}, \boldsymbol{\sigma}_{\mathrm{I}}, \boldsymbol{\alpha}_{\mathrm{I}})$ $p = 2$	Scenario II $(\boldsymbol{\mu}_{\mathrm{II}}, \boldsymbol{\sigma}_{\mathrm{II}}, \boldsymbol{\alpha}_{\mathrm{II}})$ $p = 4$	Scenario III $(\boldsymbol{\mu}_{\mathrm{III}}, \boldsymbol{\sigma}_{\mathrm{III}}, \boldsymbol{\alpha}_{\mathrm{III}})$ $p = 5$
4	1.4×10^{-4}	1.8×10^{-4}	3.4×10^{-4}
10	8.0×10^{-6}	1.0×10^{-5}	2.0×10^{-5}
15	2.3×10^{-6}	2.9×10^{-6}	5.8×10^{-6}
20	9.4×10^{-7}	1.2×10^{-6}	2.4×10^{-6}
50	5.8×10^{-8}	7.4×10^{-8}	1.5×10^{-7}
100	7.1×10^{-9}	9.1×10^{-9}	1.8×10^{-8}
500	5.6×10^{-11}	7.2×10^{-11}	1.4×10^{-10}

$$\|F_W - F_{W^\star}\|_\infty \leq \Delta \leq \frac{1}{2\pi} \int\limits_{\mathbb{R}} \left| \frac{\Phi_{W_1}(t) - \Phi_{W_1^\star}(t)}{t} \right| \mathrm{d}t,$$

where $\|F_W - F_{W^\star}\|_\infty = \sup_{w \in \mathbb{R}} |F_W(w) - F_{W^\star}(w)|$. Here $F_{W^\star}$ denotes a near-exact distribution function, which, for example in the case of our first near-exact distribution is

$$F_{W^\star}(w) = F_{V^\star}\left(w - \theta - \sum_{j=1}^{p} \mu_j \alpha_j; \boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1\right), \qquad (17)$$

where $\boldsymbol{r}^\star$ and $\boldsymbol{\lambda}^\star$ are as in Theorem 2, and $F_{V^\star}$ denotes the distribution function of the random variable $V^\star$ with a GNIG distribution, as defined in (23) in Appendix 1. For our second near-exact distribution all follows analogously, but $\Phi_{W^\star}$ needs to be replaced by $\Phi_{W_\star}$, with $W_\star$ distributed as in Theorem 3, and $F_{W^\star}$ in (17), must be accordingly replaced with the distribution function

$$F_{V_\star}\left(w - E(W_1) - \sum_{j=1}^{p} \mu_j \alpha_j; \boldsymbol{r}^+, \boldsymbol{r}^-, \boldsymbol{\lambda}^+, \boldsymbol{\lambda}^-, \ell^+, \ell^-\right).$$

Here $\boldsymbol{r}^+$, $\boldsymbol{r}^-$, $\boldsymbol{\lambda}^+$, and $\boldsymbol{\lambda}^-$ are defined as in Theorem 3, and $F_{V_\star}$ is the distribution function of a random variable $V_\star$ with a DGIG distribution. For our third near-exact distribution, which can be applied when $\Sigma_3 \neq 0$, $\Phi_{W^\star}$ should be replaced by $\Phi_{\mathrm{sign}(\Sigma_3) \times W_1^\star}$ and $F_{W^\star}$ replaced by the distribution function of $\mathrm{sign}(\Sigma_3) \times W_1^\star + W_2$ in Theorem 4 (see expressions (25) and (26) in Appendix 1 for details on this distribution function).

We note that when $\gamma \to \infty$, we have $\Delta \to 0$ and $W^\star \rightsquigarrow W$, where '$\rightsquigarrow$' is used to denote weak convergence. Parenthetically, we further note that to be ensured that we accurately approximate the tail of the exact distribution, we need to keep increasing the precision parameter γ as we move towards higher quantiles; further details on the measure Δ can be found in Grilo and Coelho (2007), Marques and Coelho (2008), and Coelho and Marques (2010, 2012).

3.2 Numerical results

3.2.1 *First near-exact distribution* ($\alpha_j > 0$)

In Tables 1 and 2 we report numerical results conducted according to the following Scenarios:

- Scenario I: $\boldsymbol{\mu}_{\mathrm{I}} = (2, 3)$, $\boldsymbol{\sigma}_{\mathrm{I}} = (5, 6)$, and $\boldsymbol{\alpha}_{\mathrm{I}} = \mathbf{1}_2^{\mathrm{T}}$;
- Scenario II: $\boldsymbol{\mu}_{\mathrm{II}} = (-4, -1, 2, 3)$, $\boldsymbol{\sigma}_{\mathrm{II}} = (0.1, 0.2, 0.3, 0.4)$, and $\boldsymbol{\alpha}_{\mathrm{II}} = (1, 2, 3, 4)$;
- Scenario III: $\boldsymbol{\mu}_{\mathrm{III}} = (-10, 10, 20, 30, 40)$, $\boldsymbol{\sigma}_{\mathrm{III}} = (1, 2, 3, 4, 5)$, and $\boldsymbol{\alpha}_{\mathrm{III}} = (1/2, 1, 3/4, 5, 1)$.

In Table 1 it can be observed that the values of Δ are quite low—indicating a good approximation—and that the parameter γ is inversely related to Δ. In addition, it can also be noticed that Δ is unresponsive to changes in μ_j, and the same happens if we multiply all the σ_j by the same constant. The quality of the near-exact approximations is patent from the extremely reduced values of Δ.

The parameter γ may be chosen according to the desired precision. Higher values of γ entail however a higher computational investment, and hence the selection of this parameter involves a precision–burden tradeoff. In Table 2 we present the computation time, in seconds, for the calculation of the

Table 2 Computation time (in seconds) for the near-exact cumulative distribution functions for Scenarios I–III

γ	Scenario I $(\boldsymbol{\mu}_{\mathrm{I}}, \boldsymbol{\sigma}_{\mathrm{I}}, \boldsymbol{\alpha}_{\mathrm{I}})$ $p = 2$			Scenario II $(\boldsymbol{\mu}_{\mathrm{II}}, \boldsymbol{\sigma}_{\mathrm{II}}, \boldsymbol{\alpha}_{\mathrm{II}})$ $p = 4$			Scenario III $(\boldsymbol{\mu}_{\mathrm{III}}, \boldsymbol{\sigma}_{\mathrm{III}}, \boldsymbol{\alpha}_{\mathrm{III}})$ $p = 5$		
	p values			p values			p values		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
4	0.02	0.02	0.02	0.05	0.03	0.03	0.05	0.03	0.03
10	0.08	0.08	0.08	0.39	0.41	0.33	0.38	0.39	0.41
15	0.17	0.17	0.20	1.17	1.28	1.14	1.48	1.62	1.08
20	0.30	0.31	0.36	2.87	3.09	2.82	2.93	2.90	2.95
50	2.62	2.50	3.00	64.2	70.9	65.8	72.6	68.3	70.7

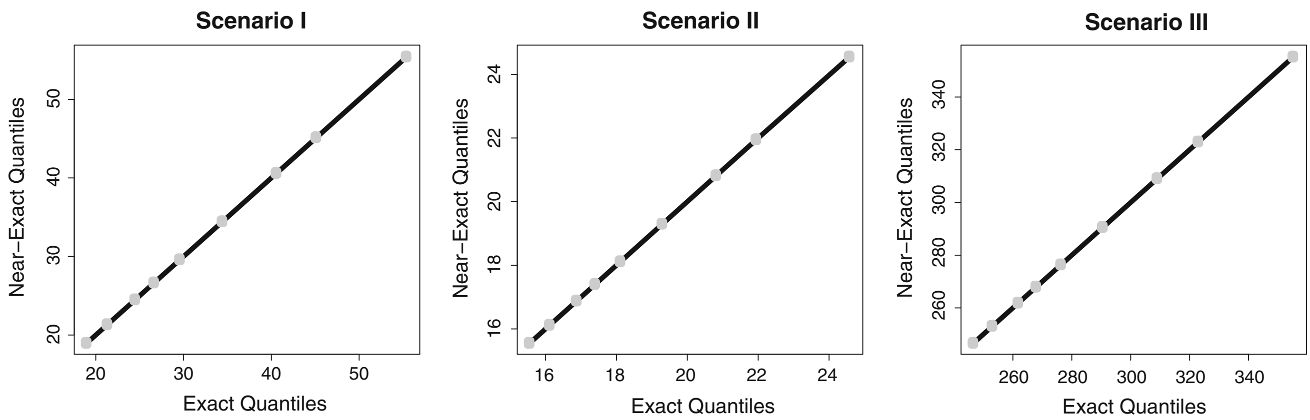

Fig. 1 QQ-plots for Scenarios I–III. The near-exact quantiles, 0.80, 0.85, 0.90, 0.925, 0.95, 0.975, 0.99, 0.995 and 0.999, were computed using the near-exact distribution function in (17) for $\gamma = 10$ and the corresponding exact quantiles were computed using the Gil-Pelaez (1951) inversion formulas and the bisection method

Table 3 Values of Δ for Scenario A

γ	$p = 2$	$p = 10$	$p = 20$	$p = 30$	$p = 50$
4	1.4×10^{-4}	1.5×10^{-5}	6.7×10^{-6}	4.2×10^{-6}	2.4×10^{-6}
10	8.0×10^{-6}	8.1×10^{-7}	3.5×10^{-7}	2.2×10^{-7}	1.3×10^{-7}
15	2.3×10^{-6}	2.3×10^{-7}	9.9×10^{-8}	6.3×10^{-8}	3.6×10^{-8}
20	9.4×10^{-7}	9.4×10^{-8}	4.1×10^{-8}	2.6×10^{-8}	1.5×10^{-8}
50	5.8×10^{-8}	5.8×10^{-9}	2.5×10^{-9}	1.6×10^{-9}	8.9×10^{-10}
100	7.1×10^{-9}	7.1×10^{-10}	3.1×10^{-10}	1.9×10^{-10}	1.1×10^{-10}
500	5.6×10^{-11}	5.6×10^{-12}	2.4×10^{-12}	1.5×10^{-12}	8.8×10^{-13}

p values 0.10, 0.05 and 0.01, using the near-exact quantiles. These calculations were done using an Intel i7 2GHz processor; for values of γ larger than 50 the computation times start to increase steadily. In most computations below we use the value $\gamma = 10$ as a reference value, as it provides a sensible computation time/Δ ratio for our first near-exact distribution.

It is also possible to observe from Table 2 that, as expected, when we increase p the computation time also grows, being this growth less steep for small to moderate values of γ.

To compare the exact and near-exact quantiles we present in Fig. 1 QQ-plots for Scenarios I–III. The extreme closeness between the exact and near-exact quantiles is sustained by the fact that all the points are extremely close to the line of equation $y = x$; the exact quantiles were computed using the Gil-Pelaez (1951) inversion formulas and the bisection method which is very time consuming, numerically unstable, and hence inappropriate for a regular use.

It is also interesting that the near-exact approximations tend to slightly improve with an increasing number of variables, as can be seen from Table 3, where we consider a Scenario A with parameters $\boldsymbol{\mu}_{\text{A}} = ((-1)^j 2j,\ j = 1, \ldots, p)$, $\boldsymbol{\sigma}_{\text{A}} = (5/j,\ j = 1, \ldots, p)$ and $\boldsymbol{\alpha}_{\text{A}} = (2j + 1,\ j = 1, \ldots, p)$ for $p = 2,\ 10,\ 30,\ 50$.

3.2.2 *Second near-exact distribution* ($\alpha_j \in \mathbb{R}$)

In Tables 4 and 5 we report numerical results conducted according to the following scenarios:

- Scenario IV: $\boldsymbol{\mu}_{\text{IV}} = \boldsymbol{\mu}_{\text{I}} = (2, 3)$, $\boldsymbol{\sigma}_{\text{IV}} = \boldsymbol{\sigma}_{\text{I}} = (5, 6)$, and $\boldsymbol{\alpha}_{\text{IV}} = (1, -1)$;
- Scenario V: $\boldsymbol{\mu}_{\text{V}} = \boldsymbol{\mu}_{\text{II}} = (-4, -1, 2, 3)$, $\boldsymbol{\sigma}_{\text{V}} = \boldsymbol{\sigma}_{\text{II}} = (0.1, 0.2, 0.3, 0.4)$, and $\boldsymbol{\alpha}_{\text{V}} = (1, -2, 3, -4)$;
- Scenario VI: $\boldsymbol{\mu}_{\text{VI}} = \boldsymbol{\mu}_{\text{III}} = (-10, 10, 20, 30, 40)$, $\boldsymbol{\sigma}_{\text{VI}} = \boldsymbol{\sigma}_{\text{III}} = (1, 2, 3, 4, 5)$, and $\boldsymbol{\alpha}_{\text{VI}} = (1/2, -1, -3/4, -5, 1)$.

From Tables 4 and 5 we can observe that the near-exact approximation obtained using the result in Theorem 3 is not as accurate as the one obtained with Theorem 2, although it presents faster computation times for the same values of γ. To achieve in Scenarios IV–VI similar performances as the ones obtained for Scenarios I–III we need to consider at least $\gamma = 500$ as can be seen in Table 4. Again, for Scenarios IV–VI, it can be ascertained from Table 5, that for higher values of p we obtain a higher computational cost.

From the QQ-plots in Fig. 2 it can be noticed that, for Scenarios IV–VI, the near-exact quantiles approximate rea-

Table 4 Values of Δ for Scenarios IV–VI

γ	Scenario IV ($\boldsymbol{\mu}_{\text{IV}}, \boldsymbol{\sigma}_{\text{IV}}, \boldsymbol{\alpha}_{\text{IV}}$) $p = 2$	Scenario V ($\boldsymbol{\mu}_{\text{V}}, \boldsymbol{\sigma}_{\text{V}}, \boldsymbol{\alpha}_{\text{V}}$) $p = 4$	Scenario VI ($\boldsymbol{\mu}_{\text{VI}}, \boldsymbol{\sigma}_{\text{VI}}, \boldsymbol{\alpha}_{\text{VI}}$) $p = 5$
Second near-exact distribution			
4	4.7×10^{-2}	4.6×10^{-2}	5.5×10^{-2}
10	1.5×10^{-2}	1.5×10^{-2}	1.7×10^{-2}
15	9.8×10^{-3}	9.8×10^{-3}	1.1×10^{-2}
20	7.2×10^{-3}	7.2×10^{-3}	8.3×10^{-3}
50	2.8×10^{-3}	2.7×10^{-3}	3.2×10^{-3}
100	1.3×10^{-3}	1.4×10^{-3}	1.6×10^{-3}
500	2.7×10^{-4}	2.7×10^{-4}	3.1×10^{-4}
Third near-exact distribution			
4	5.3×10^{-4}	4.0×10^{-4}	3.9×10^{-4}
10	3.0×10^{-5}	2.2×10^{-5}	2.3×10^{-5}
15	8.5×10^{-6}	6.4×10^{-6}	6.6×10^{-6}
20	3.5×10^{-6}	2.6×10^{-6}	2.7×10^{-6}
50	2.1×10^{-7}	1.6×10^{-7}	1.7×10^{-7}
100	2.6×10^{-8}	2.0×10^{-8}	2.1×10^{-8}
500	2.1×10^{-10}	1.6×10^{-10}	1.6×10^{-10}

sonably well the exact ones. In these QQ-plots we consider $\gamma = 100$, given that it provides a reasonable computation time/Δ ratio for our second near-exact distribution.

From Table 6 it can be ascertained that the accuracy of our second near-exact approximation also tends to improve as the number of variables increases, although in this case the decrements in Δ occur at a much slower rate; in Table 6 we considered a Scenario B with $\boldsymbol{\mu}_{\text{B}} = (j/2, j = 1, \ldots, p)$, $\boldsymbol{\sigma}_{\text{B}} = (3j - 1, j = 1, \ldots, p)$, and $\boldsymbol{\alpha}_{\text{B}} = ((-1)^{j+1}\frac{j}{3}, j = 1, \ldots, p)$ for $p = 2, 10, 20, 30, 50$.

3.2.3 Third near-exact distribution ($\Sigma_3(\boldsymbol{\alpha}) \neq 0$)

We assess the performance of the third near-exact distribution on Scenarios IV–VI. Tables 4 and 6 reveal that our third near-exact distribution possesses similar asymptotic properties as our second approach, although—as reflected by its lower values of Δ—it is much more precise. The computation time of our third near-exact distribution increases however as a function of γ, in some cases beyond the realms of practicality. From Table 5 it is possible to observe that our third near-exact distribution presents higher computing times than our second one, and thus we propose $\gamma = 4$ as a reference, as it provides a sensible computation time/Δ ratio, to be used in practical applications. Note that for $\gamma = 4$, the value of Δ is slightly lower than the one for our second approach with $\gamma = 100$, but even with this difference the near-exact quantiles of both approaches would be virtually indistinguishable if plotted simultaneously in Fig. 2.

Our third near-exact distribution can also be applied to the case of positive linear combination coefficients. In practice we have found that although both our first and third near-exact distributions have tantamount precision, the third approach requires a higher computational investment.

4 Examples and illustrations

All examples in this section entail positive linear combinations, and hence for conciseness only our first near-exact approximation is used.

4.1 Network engineering

The real time management of massive data streams in large-scale networks leads to a number of challenging problems in computational statistics (Domingos and Hulten 2003). One of such problems entails achieving at least a minimum level of quality-of-service, and a well-known method for achieving this goal is the so-called *egress admission control* algorithm (Cetinkaya et al. 2001). A full description of this algorithm is beyond the scope of our paper. What is relevant for our purposes is that their algorithm is based on the sum of two independent Gumbel distributed random variables, and quoting the authors (Cetinkaya et al. 2001, p. 76):

> "Approximating the sum of two Gumbel distributed random variables by a Gumbel random variable, the admission control test follows."

Thus, Cetinkaya et al. (2001) inadequately use a single Gumbel distribution to approximate the sum of two independent Gumbel distributions, as already remarked in Nadarajah and Kotz (2008). In Fig. 3 we illustrate the reliability of our SGNIG-based near-exact approximation, introduced in Sect. 2.2, and the inadequacy of the approach in Cetinkaya et al. (2001), as assessed by the pointwise difference to the exact density obtained using the inversion formulas in Gil-Pelaez (1951).

Fig. 3 clearly provides evidence to support the claim that the egress admission control algorithm could benefit from using our near-exact approximation. To give a more complete view of the comparison between our approach and the one in Cetinkaya et al. (2001), we revisit Scenario I from Sect. 3, and to assess the performance of both approaches we again use the measure Δ, as defined in (16). The results are reported in Fig. 4, and again provide evidence suggesting that our near-exact approximation would yield more precise and reliable egress admission control algorithms.

Parenthetically, we note that Nadarajah and Kotz (2008) present an expression for the exact distribution of the sum of two independent Gumbel random variables, but only for the cases where the ratio between the scale parameters is a

Table 5 Computation time (in seconds) for the near-exact cumulative distribution functions for Scenarios IV–VI

γ	Scenario IV $(\boldsymbol{\mu}_{\text{IV}}, \boldsymbol{\sigma}_{\text{IV}}, \boldsymbol{\alpha}_{\text{IV}})$ $p = 2$			Scenario V $(\boldsymbol{\mu}_{\text{V}}, \boldsymbol{\sigma}_{\text{V}}, \boldsymbol{\alpha}_{\text{V}})$ $p = 4$			Scenario VI $(\boldsymbol{\mu}_{\text{VI}}, \boldsymbol{\sigma}_{\text{VI}}, \boldsymbol{\alpha}_{\text{VI}})$ $p = 5$		
	p values			p values			p values		
	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
Second near-exact distribution									
4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
10	0.00	0.02	0.00	0.02	0.02	0.03	0.03	0.03	0.02
15	0.02	0.02	0.02	0.05	0.05	0.05	0.08	0.06	0.08
20	0.03	0.02	0.02	0.09	0.08	0.09	0.13	0.13	0.14
50	0.16	0.13	0.12	0.75	0.73	0.73	1.22	1.23	1.22
100	0.89	0.87	0.86	5.76	5.77	5.76	12.4	12.5	12.4
500	44.6	44.9	45.2	526.3	628.5	630.2	2540.0	2550.0	2545.0
Third near-exact distribution									
4	0.23	0.20	0.17	0.22	0.22	0.19	0.44	0.41	0.34
10	2.56	1.75	1.28	2.62	2.59	2.12	12.50	11.65	8.42
15	7.72	5.69	4.17	13.44	6.37	5.68	46.84	47.11	39.61
20	15.88	13.74	11.29	29.03	12.59	11.45	173.46	190.15	166.89
50	130.96	102.31	72.76	260.52	241.05	110.67	*	*	*

* Above 1 h. The same applies for $\gamma = 100$ and $\gamma = 500$

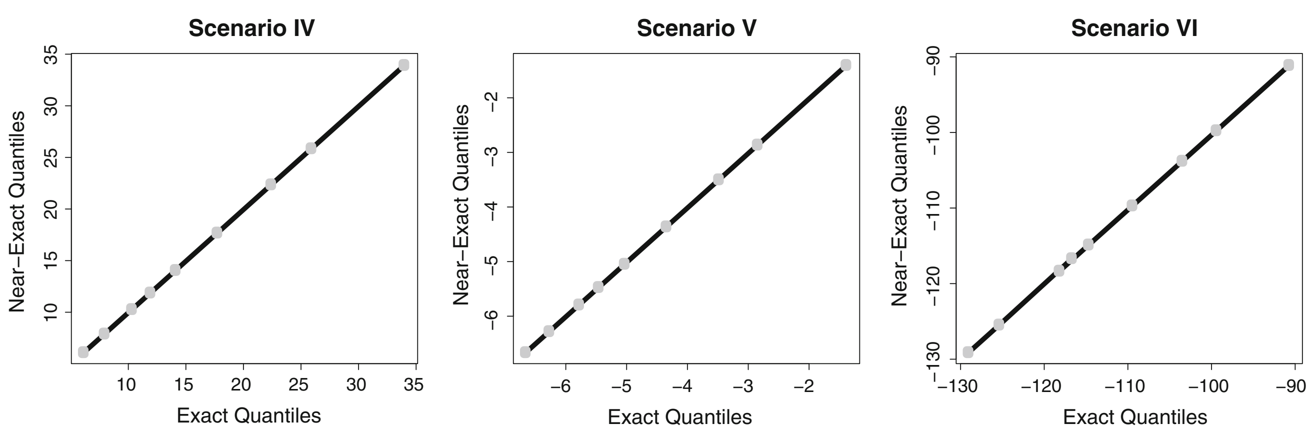

Fig. 2 QQ-plots for Scenarios IV–VI. The near-exact quantiles, 0.80, 0.85, 0.90, 0.925, 0.95, 0.975, 0.99, 0.995 and 0.999, were computed using the near-exact distribution function in (17) for $\gamma = 100$ and the corresponding exact quantiles were computed using the Gil-Pelaez (1951) inversion formulas and the bisection method

rational number. However, the expression they use for their function $J(\cdot, \cdot, \cdot, \cdot)$ is not valid when its first and third arguments are symmetrical, which means that their expressions for the cumulative distribution and probability density functions simply do not work.

4.2 Computational biology

Our second example is on motif discovery in biological sequences. Some interesting computational and statistical issues arising in modeling these problems are documented in Keich and Nagarajan (2006), and the huge literature on the topic is reviewed by Sandve and Drabløs (2006). Our analysis focuses on a method proposed by Bailey and Gribskov (1997) for calculating p values for the test of simultaneous matching of p DNA sequences in a database. More precisely, the authors consider the test statistic

$$W_p(n) = \sum_{i=1}^{p} X_i(n), \tag{18}$$

where for each i, $X_i(n)$ is a sequence of random variables converging in distribution to a standard Gumbel distribution X_i, as $n \to \infty$; here n should be understood as the number of DNA sequences in the database. Under the assumptions in

Table 6 Values of Δ for scenario B

γ	$p = 2$	$p = 10$	$p = 20$	$p = 30$	$p = 50$
Second near-exact distribution					
4	6.0×10^{-2}	3.7×10^{-2}	3.3×10^{-2}	3.2×10^{-2}	3.1×10^{-2}
10	1.9×10^{-2}	1.3×10^{-2}	1.2×10^{-2}	1.1×10^{-2}	1.1×10^{-2}
15	1.2×10^{-2}	8.2×10^{-3}	7.6×10^{-3}	7.3×10^{-3}	7.1×10^{-3}
20	8.7×10^{-3}	6.1×10^{-3}	5.6×10^{-3}	5.4×10^{-3}	5.3×10^{-3}
50	3.4×10^{-3}	2.4×10^{-3}	2.2×10^{-3}	2.1×10^{-3}	2.1×10^{-3}
100	1.7×10^{-3}	1.2×10^{-3}	1.1×10^{-3}	1.0×10^{-3}	1.0×10^{-3}
500	3.3×10^{-4}	2.3×10^{-4}	2.1×10^{-4}	2.1×10^{-4}	2.0×10^{-4}
Third near-exact distribution					
4	4.2×10^{-4}	1.7×10^{-4}	7.7×10^{-5}	4.8×10^{-5}	2.7×10^{-5}
10	2.6×10^{-5}	9.3×10^{-6}	4.0×10^{-6}	2.5×10^{-6}	1.4×10^{-6}
15	7.4×10^{-6}	2.6×10^{-6}	1.1×10^{-6}	7.1×10^{-7}	4.0×10^{-7}
20	3.1×10^{-6}	1.1×10^{-6}	4.7×10^{-7}	2.9×10^{-7}	1.6×10^{-7}
50	1.9×10^{-7}	6.7×10^{-8}	2.9×10^{-8}	1.8×10^{-8}	1.0×10^{-8}
100	2.3×10^{-8}	8.2×10^{-9}	3.6×10^{-9}	2.2×10^{-9}	1.2×10^{-9}
500	1.8×10^{-10}	6.5×10^{-11}	2.8×10^{-11}	1.7×10^{-11}	9.8×10^{-12}

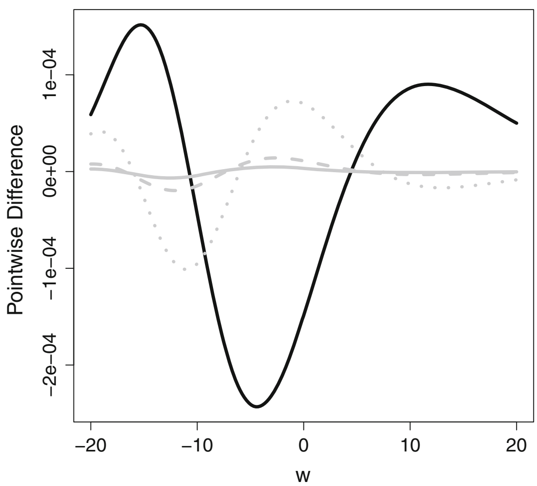

Fig. 3 Pointwise difference between the exact density of the sum of two independent Gumbel random variables $\big((\mu_1, \sigma_1) = (0, 1)$ and $(\mu_2, \sigma_2) = (0, 10)\big)$ and the densities obtained through our near-exact approximation (*gray lines*), as well as the difference between the exact density and the approximation in Cetinkaya et al. (2001) (*black line*). The exact density was obtained using the inversion formulas in Gil-Pelaez (1951), and for our approach we take a precision parameter of $\gamma = 4, 7, 10$, respectively corresponding to the *dotted*, *dashed*, and *solid gray lines*

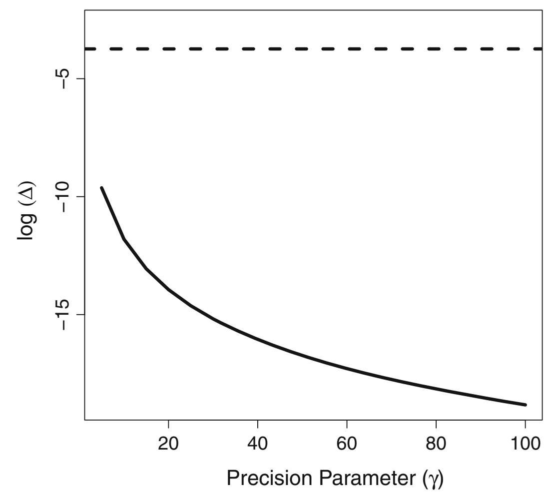

Fig. 4 Comparing our near-exact approximation with the approximation in Cetinkaya et al. (2001), on the basis of the measure Δ, as defined in (16), over several values of the precision parameter γ; the *solid* and *dashed lines* respectively correspond to our near-exact approximation and the approach in Cetinkaya et al. (2001) for Scenario I

Bailey and Gribskov (1997), $X_1, \ldots, X_p$ is thus a sequence of independent standard Gumbel random variables, so that the limiting distribution of the test statistic (18) is

$$W_p(n) \rightsquigarrow W = \sum_{i=1}^{p} X_i.$$

The authors then propose

$$F_W(w) = P(W \le w) \approx \frac{(p-1)! - \exp\{-w\} w^{p-1}}{(p-1)!} = \widetilde{P}(W \le w), \quad (19)$$

as an approximation to the distribution function of W. In opposition, our near-exact approximation for W can be obtained from (13), and it is based on the shifted GNIG distribution

$$\mathrm{SGNIG}\Big(\mathbf{r}^\star = (p\mathbf{1}_{\gamma-1}^{\mathrm{T}}, \rho), \boldsymbol{\lambda}^\star = (1, \ldots, \gamma-1, l), \gamma, \theta\Big),$$

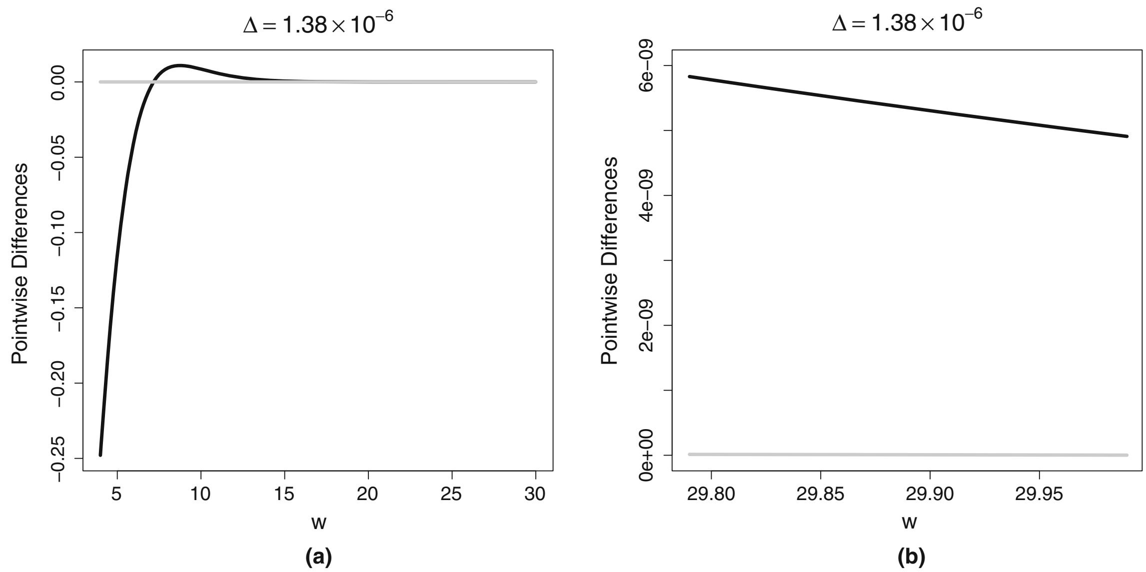

Fig. 5 Pointwise difference between the exact distribution function of the sum of six independent Gumbel random variables $\big((\mu_1, \sigma_1) = \cdots = (\mu_6, \sigma_6) = (0, 1)\big)$ and the distribution function obtained through our near-exact approximation (*gray line*), as well as the difference between the exact density and the approximation in Bailey and Gribskov (1997) (*black line*). The exact distribution function was obtained using the inversion formulas in Gil-Pelaez (1951), and for our approach we take a precision parameter of $\gamma = 10$. Here **a**, **b** correspond to different windows of interest

with ρ, l, θ given in (11). In Fig. 5 we consider the case $p = 6$ with $\gamma = 10$. As it can be observed the pointwise differences between the exact distribution function and the distribution function (19) corresponding to the Bailey and Gribskov approximation, are much larger in absolute value than the ones provided by our near-exact approximation.

To make direct comparisons with the results obtained by Bailey and Gribskov, we use the 'percent error', which they define as

$$\mathrm{err}_{\%}(w) = 100 \times \frac{\widetilde{P}(W \geq w) - P(W \geq w)}{P(W \geq w)}, \tag{20}$$

and where we replace $\widetilde{P}(W \geq w)$ by $P(W^\star \geq w)$ corresponding to our near-exact approximation, which is given by

$$\begin{aligned} P(W^\star \geq w) &= 1 - F_{W^\star}(w) \\ &= 1 - F_{V^\star}\big(w - \theta; \mathbf{r}^\star, \boldsymbol{\lambda}^\star, \gamma\big), \end{aligned} \tag{21}$$

where $\mathbf{r}^\star = (p\mathbf{1}^{\mathrm{T}}_{\gamma-1}, \rho)$ and $\boldsymbol{\lambda}^\star = (1, \ldots, \gamma - 1, l)$, and where $F_{V^\star}$ is the distribution function of a random variable $V^\star = W^\star - \theta$ with a GNIG distribution, as defined in (23) in Appendix 1. To evaluate $P(W \geq w)$ we use the Gil-Pelaez inversion formulas.

The resulting 'percent error' is plotted in Fig. 6. Comparing this figure with Fig. 5 in Bailey and Gribskov (1997) it is possible to observe the differences of scales in the vertical axis which show that the percent errors for the near-exact approximations are extremely low when compared to the ones obtained for the approximation in Bailey and Gribskov (1997). These results reinforce the proximity, already assessed in Sect. 3, between the near-exact distributions developed and the exact distribution.

Although the computation times for the Bailey and Gribskov approximation are comparable with those for the near-exact approximations developed for $\gamma = 4$ or $\gamma = 6$, the precision obtained has no comparison.

4.3 Flood risk management

In this section we first show how our results can be used to obtain simple confidence intervals for the location parameter of a Gumbel distribution, and then we apply our results to a real data set of annual maximum floods of the river Nidd in Yorkshire, UK, used by Hosking et al. (1985). Let $(X_1, \ldots, X_n)$ be a random sample from a population with Gumbel(μ, σ) distribution, so that

$$E(X_i) = \sigma\gamma^* + \mu, \quad \mathrm{var}(X_i) = \frac{\pi^2}{6}\sigma^2,$$

where γ^* is the Euler–Mascheroni constant. There are two cases to be considered: I) σ is known, and II) σ is unknown.

4.3.1 Case I

If σ is known, the classical moment estimator of μ,

$$\widehat{\mu} = \overline{X} - \sigma\gamma^*,$$

is unbiased and consistent in quadratic mean, and as such a good candidate to build confidence intervals for μ. Based on

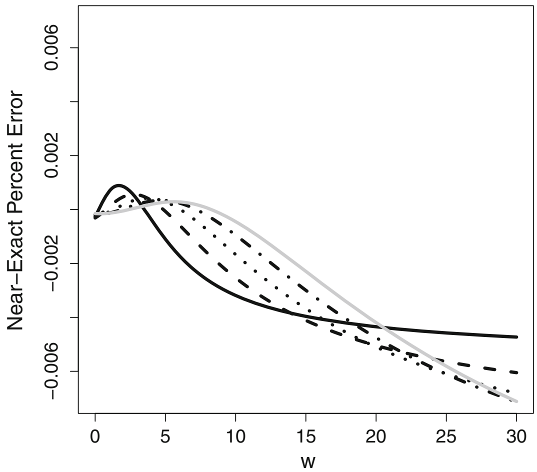

Fig. 6 Percent error, as defined in (20), for near-exact distributions. The near-exact distribution functions were obtained using (21), where we take a precision parameter of $\gamma = 10$. The *solid*, *dashed*, *dotted*, and *dashed-dotted black lines* respectively correspond to $p = 2, 3, 4, 5$; the *solid gray line* corresponds to $p = 6$

our near-exact approximations in Sect. 2, we can compute for a given level of confidence α, the near-exact quantiles $q_{\alpha/2}$ and $q_{1-\alpha/2}$ of $\widehat{\mu} - \mu$, such that

$$1 - \alpha = P\left(q_{\alpha/2} < \widehat{\mu} - \mu < q_{1-\alpha/2}\right) = P\left(\widehat{\mu} - q_{1-\alpha/2} < \mu < \widehat{\mu} - q_{\alpha/2}\right),$$

and thus

$$\left[\widehat{\mu} - q_{1-\alpha/2}, \widehat{\mu} - q_{\alpha/2}\right]$$

is a near-exact level α confidence interval for μ. Note that the α quantile of $\widehat{\mu} - \mu$ is the α quantile of $n^{-1}\sum_{i=1}^{n} X_i^*$, where $X_i^* \sim \text{Gumbel}(-\sigma\gamma^*, \sigma)$, and a close approximation to this may be obtained through the near-exact quantiles determined using the near-exact distribution function in (17).

4.3.2 Case II

If σ is unknown, for $S^2 = (n-1)^{-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$, we have

$$E(S^2) = \text{var}(X_i) = \frac{\sigma^2\pi^2}{6}, \quad S^2 \xrightarrow{p} \text{var}(X_i) = \frac{\sigma^2\pi^2}{6},$$

so that

$$\frac{\sqrt{6}}{\pi}\sqrt{S^2} \xrightarrow{p} \sigma\,.$$

Thus, we propose using the estimator

$$\widehat{U} = \overline{X} - \frac{\sqrt{6}}{\pi}\sqrt{S^2}\gamma^*,$$

to build confidence intervals for μ, given that $\widehat{U}$ is consistent for μ, since

$$\widehat{U} = \underbrace{\overline{X}}_{\xrightarrow{p}\sigma\gamma^*+\mu} - \underbrace{\frac{\sqrt{6}}{\pi}\sqrt{S^2}}_{\xrightarrow{p}\sigma}\gamma^* \xrightarrow{p} \mu\,.$$

As such, an approximate confidence interval for μ is given by

$$\left[\widehat{U} - q_{1-\alpha/2}, \widehat{U} - q_{\alpha/2}\right], \tag{22}$$

where $q_{\alpha/2}$ and $q_{1-\alpha/2}$ are respectively the $\alpha/2$ and the $1-\alpha/2$ quantiles of $\widehat{U} - \mu$, where we have that $\widehat{U} - \mu = n^{-1}\sum_{i=1}^{n} X_i^{**}$, with

$$X_i^{**} \sim \text{Gumbel}\left(-\frac{\sqrt{6}}{\pi}\sqrt{S^2}\gamma^*, \frac{\sqrt{6}}{\pi}\sqrt{S^2}\right).$$

Based on a first impression, one could be tempted to infer from (22) that changes in the value of S^2 would not affect the width of the confidence interval, but we note that $\sqrt{S^2}$ appears multiplying in both parameters of the Gumbel distribution of X_i^{**}, and hence the larger the S^2 the wider the confidence interval.

To show that these confidence intervals yield the due coverage probabilities, we performed some simulation studies for coverage probabilities of 0.90, 0.95 and 0.99; the results are reported in Tables 7–8. For each case we have simulated 50 batches of 100 samples of size 5 for the case of σ known and of size 10 for the case of σ unknown and we counted the number of times, out of 100, that the true value of μ fell into the respective confidence interval. The near-exact quantiles, needed to determine the confidence intervals, were calculated using the near-exact distribution function in (17) taking $\gamma = 4$ for both cases of known and unknown σ.

For the case of known σ, using $\mu = 5$ and $\sigma = 5.6$, confidence intervals for the proportion of times that the true value of μ fell into the corresponding confidence interval, based on the asymptotic distribution of the maximum likelihood estimator of the proportion p^* in a Binomial (100, p^*) distribution, for a sample of size 50, gave, respectively for the nominal coverage probabilities of 0.90, 0.95 and 0.99,

$$[0.8944, 0.9108], \quad [0.9414, 0.9538], \quad [0.9863, 0.9921],$$

being clear that in each case the nominal coverage probability falls in the respective confidence interval.

For the case of unknown σ, we also used $\mu = 5$ and $\sigma = 5.6$ to simulate the samples, and then we estimated σ as described above. A similar procedure as described above, gave the following confidence intervals for the proportion of times that the true value of μ fell into the corresponding confidence interval

$$[0.8975, 0.9137], \quad [0.9482, 0.9598], \quad [0.9875, 0.9929],$$

Table 7 Number of times, out of 100, that the true value of μ fell into the corresponding confidence interval, in the case of known σ

Coverage probability	Number of times
0.90	93, 90, 84, 90, 85, 93, 89, 88, 98, 91, 87, 94, 91, 88, 97, 88, 90, 87, 92, 95, 90, 93, 86, 93, 89, 88, 92, 90, 89, 90, 85, 92, 86, 92, 91, 92, 91, 93, 84, 90, 87, 89, 93, 91, 87, 97, 93, 91, 89, 90
0.95	96, 94, 96, 95, 95, 95, 95, 91, 93, 97, 98, 98, 95, 95, 95, 89, 93, 97, 97, 98, 92, 94, 93, 97, 95, 97, 94, 92, 96, 97, 94, 99, 97, 92, 95, 91, 95, 96, 92, 92, 98, 97, 93, 94, 89, 96, 92, 95, 95, 97
0.99	100, 100, 97, 99, 99, 99, 99, 99, 100, 98, 100, 100, 98, 99, 100, 98, 100, 96, 97, 99, 100, 96, 100, 99, 99, 100, 99, 99, 100, 98, 98, 99, 100, 100, 99, 99, 100, 99, 97, 98, 99, 98, 99, 100, 100, 99, 98, 99, 100, 98

Table 8 Number of times, out of 100, that the true value of μ fell into the corresponding confidence interval, in the case of unknown σ

Probability	Number of times
0.90	91, 93, 91, 91, 92, 84, 92, 84, 91, 89, 96, 96, 93, 92, 89, 89, 90, 89, 93, 89, 93, 97, 86, 90, 94, 89, 89, 86, 91, 87, 90, 88, 89, 91, 95, 93, 91, 90, 89, 95, 87, 93, 87, 95, 88, 87, 94, 91, 92, 87
0.95	92, 94, 96, 99, 96, 99, 93, 97, 97, 95, 98, 94, 97, 97, 95, 97, 93, 95, 97, 93, 95, 96, 96, 92, 98, 92, 93, 97, 97, 96, 92, 95, 95, 96, 94, 98, 97, 94, 97, 99, 90, 97, 97, 97, 96, 97, 92, 92, 94, 95
0.99	100, 99, 98, 99, 98, 97, 96, 100, 100, 99, 98, 99, 97, 99, 99, 100, 97, 99, 98, 100, 99, 99, 100, 100, 98, 100, 99, 99, 100, 100, 99, 100, 100, 99, 100, 100, 99, 99, 100, 100, 99, 99, 100, 98, 99, 99, 99, 100, 97, 99

being once again clear that in each case the nominal coverage probability falls indeed in the respective confidence interval. The above results show the adequacy of the confidence intervals proposed even for very small sample sizes.

To illustrate the utility of the interval estimation procedure developed above, we consider 35 annual maximum annual maximum floods of the river Nidd in Yorkshire, UK, taken from the Natural Environment Research Council NERC (1975, p. 235). As mentioned by Hosking et al. (1985, p. 258) these data "may reasonably be assumed to come from a Gumbel distribution." For these data we have as estimates for the parameters μ and σ respectively $\widehat{\mu} = 109.33$ and $\widehat{\sigma} = 47.34$. The near-exact quantiles, $q_{\alpha/2}$ and $q_{1-\alpha/2}$, were determined using the near-exact distribution function in (17) and taking $\gamma = 4$. Hence for $1-\alpha = 0.90, 0.95, 0.99$ we have

$$\begin{aligned} q_{0.05} &= 11.03, & q_{0.95} &= 44.76, \\ q_{0.025} &= 8.15, & q_{0.975} &= 48.38, \\ q_{0.005} &= 2.69, & q_{0.995} &= 55.68, \end{aligned}$$

and thus the confidence intervals for μ, and for $1-\alpha = 0.90, 0.95, 0.99$, are respectively given by

$$[91.91, 125.64]\,, \quad [88.29, 128.52]\,, \quad [80.99, 133.98]\,.$$

The coverage probabilities obtained for samples of size 5 and 10, show that the confidence intervals obtained in this way may be applied even for small sample sizes; these are usually situations in which maximum-likelihood estimation procedures, are not always satisfactory, even for moderate sample sizes as pointed out by Hosking et al. (1985).

5 Discussion

In this paper we develop precise, tractable, and computationally appealing near-exact approximations for the distribution of the linear combination of independent Gumbel random variables. The precision parameter γ plays a key role in modulating the desired reliability of our approximations, with larger values of γ leading to a higher accuracy. The value of γ can hence be chosen according to the targeted level of precision, but this entails a precision–burden tradeoff as a higher value of γ requires a larger computational investment. Although our illustrations focused mostly on the case of sums of independent Gumbel variates, our approaches are tailored for linear combinations in general, and their accuracy seems to be mildly uniform over a different set of weights and several combinations of shape and scale parameters. From the point of view of modeling extremes, more complex structures of dependence—other than exact independence—are certainly of interest, as well as tails which are heavier than the Gumbel. As discussed by Albrecher et al. (2011) simple and manageable models—such as the Cramér–Lundberg model—are based on restrictive independence assumptions, but still can be used as a natural starting point for modeling.

Although not explored here, our near-exact approximations have the potential to be used as a baseline model—say as a centering distribution in a Bayesian nonparametric setting (Müller and Quintana 2004)—and from that point of view it can be understood as a computationally appealing starting point for modeling linear combinations of heavy-tailed data with more complex structures of dependence. In this context, it seems for example natural 'centering' a Dirichlet process $\mathrm{DP}(M, F_{W^\star})$ at our near-exact approximation $F_{W^\star}$, where $M > 0$ controls the variability of the random distributions

F generated according to the DP prior, such that we have $F \sim \text{Beta}(MF_{W^\star}, M(1-F_{W^\star}))$. Since $E(F) = F_{W^\star}$ random realizations of the DP process would on average coincide with our near-exact distribution, and the role played by the parameter M can be better understood by noticing that $\text{var}(F) = F_{W^\star}(1-F_{W^\star})/(M+1)$. Hence, by taking a small value of M the contribution to the inference of the parametric model $F_{W^\star}$ would be large, whereas larger values of M will give more priority to the data, which may hopefully be informative on the tails and on the structure of dependence to be revealed.

6 Supplementary material

The supplemental files include additional numerical reports, and `Mathematica` programs which can be used to implement the methods described in the article.

Acknowledgments We thank the Editor, the Associate Editor, and two Reviewers for their careful reading and constructive comments. We also thank Anthony Davison, Barry Arnold, and Vanda de Carvalho for helpful discussions and recommendations on an earlier version of this paper. Part of this work was developed during an academic visit of F. Marques and C. Coelho to the Ecole Polytechnique Fédérale de Lausanne where M. de Carvalho was a Post-Doctoral Fellow. This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project PEst-OE/MAT/UI0297/2014 (Centro de Matemática e Aplicações), and under the Fondecyt Project 11121186.

Appendix

Appendix 1: Results and definitions on distributions of interest

Part I: The GIG and GNIG distributions

Let $X_j \overset{\text{ind.}}{\sim} \text{Gamma}(r_j, \lambda_j)$ with shape parameters $r_j \in \mathbb{N}$ and rate parameters $\lambda_j \in \mathbb{R}_+^*$, all different, for $j = 1, \ldots, \ell$. The GIG distribution of depth $\ell \in \mathbb{N}$, introduced by Coelho (1998), is defined as the distribution of $Y = \sum_{j=1}^{\ell} X_j$, and we denote this by $Y \sim \text{GIG}(\boldsymbol{r}, \boldsymbol{\lambda}, \ell)$, for $\boldsymbol{r} = (r_1, \ldots, r_\ell)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_\ell)$. The density and distribution functions of Y are

$$f_Y(y; \boldsymbol{r}, \boldsymbol{\lambda}, \ell) = K \sum_{j=1}^{\ell} \pi_j(y) \exp\{-\lambda_j y\},$$

and

$$F_Y(y; \boldsymbol{r}, \boldsymbol{\lambda}, \ell) = 1 - K \sum_{j=1}^{\ell} \Pi_j(y) \exp\{-\lambda_j y\},$$

where $y > 0$, $K = \prod_{i=1}^{p} \lambda_i^{r_i}$,

$$\begin{cases} \pi_j(y) = \sum_{k=1}^{r_j} c_{j,k} y^{k-1}, \\ \Pi_j(y) = \sum_{k=1}^{r_j} c_{j,k} (k-1)! \sum_{i=0}^{k-1} \frac{y^k}{i! \lambda_j^{k-i}}, \end{cases}$$

and the $c_{j,k}$ are given in (11)–(13) in Coelho (1998). The GNIG distribution of depth $(\ell + 1) \in \mathbb{N}$, introduced by Coelho (2004), is defined as the distribution of $Y^\star = X^\star + \sum_{j=1}^{\ell} X_j$, where $X^\star$ is independent of $\sum_{j=1}^{\ell} X_j$, and $X^\star \sim \text{Gamma}(\rho, l)$, with $\rho \in \mathbb{R}_+^* \backslash \mathbb{N}$. We denote this by $Y^\star \sim \text{GNIG}(\boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1)$, where $\boldsymbol{r}^\star = (\boldsymbol{r}, \rho)$ and $\boldsymbol{\lambda}^\star = (\boldsymbol{\lambda}, l)$, and the corresponding density and distribution functions are

$$\begin{aligned} &f_{Y^\star}(y; \boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1) \\ &= K l^\rho \sum_{j=1}^{\ell} \exp\{-\lambda_j y\} \\ &\times \sum_{k=1}^{r_j} \left\{ c_{j,k} \frac{\Gamma(k)}{\Gamma(k+\rho)} y^{k+\rho-1} {}_1F_1(\rho, k+\rho, -(l-\lambda_j)y) \right\}, \end{aligned}$$

and

$$\begin{aligned} &F_{Y^\star}(y; \boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1) \\ &= \frac{l^\rho y^\rho}{\Gamma(\rho+1)} {}_1F_1(\rho, \rho+1, -ly) - K l^\rho \sum_{j=1}^{\ell} \exp\{-\lambda_j y\} \\ &\times \sum_{k=1}^{r_j} c_{j,k}^* \sum_{i=0}^{k-1} \frac{y^{r+i} \lambda_j^i}{\Gamma(\rho+1+i)} {}_1F_1(\rho, \rho+1+i, -(l-\lambda_j)y), \end{aligned} \tag{23}$$

for $y > 0$ and where $c_{j,k}^* = (c_{j,k} \lambda_j^k)/\Gamma(k)$; in the above expressions ${}_1F_1(\cdot)$ denotes the Kummer confluent hypergeometric function.

The random variable $X^* = X + \theta$ is a shifted Gamma distribution with rate $\lambda \in \mathbb{R}_+^*$, shape $r \in \mathbb{R}_+^*$, and shift $\theta \in \mathbb{R}$, if $X \sim \text{Gamma}(r, \lambda)$, and we denote this by $X^* \sim \text{SGamma}(r, \lambda, \theta)$; the shifted GIG and GNIG distributions are analogously defined and denoted by $\text{SGIG}(\boldsymbol{r}, \boldsymbol{\lambda}, \ell, \theta)$ and $\text{SGNIG}(\boldsymbol{r}^\star, \boldsymbol{\lambda}^\star, \ell + 1, \theta)$.

Part II: The DGIG distribution and the sum (and the difference) of a DGIG random variable with an independent Gamma random variable

Let $X_1 \sim \text{GIG}(\boldsymbol{r}_1, \boldsymbol{\lambda}_1, p_1)$, with $\boldsymbol{r}_1 = (r_{11}, \ldots, r_{1p_1})$ and $\boldsymbol{\lambda}_1 = (\lambda_{11}, \ldots, \lambda_{1p_1})$, and $X_2 \sim \text{GIG}(\boldsymbol{r}_2, \boldsymbol{\lambda}_2, p_2)$, with $\boldsymbol{r}_2 = (r_{21}, \ldots, r_{2p_2})$ and $\boldsymbol{\lambda}_2 = (\lambda_{21}, \ldots, \lambda_{2p_2})$ be two independent random variables with GIG distributions. Let us then consider the random variable $Y = X_1 - X_2$. Y has a DGIG distribution whose density and distribution functions are given by (2.12) and (2.15) in Coelho and Mexia (2010), and we denote this by $Y \sim \text{DGIG}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, p_1, p_2)$. The shifted SDGIG distribution, with shift $\theta \in \mathbb{R}$, is denoted by $Y \sim \text{SDGIG}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, p_1, p_2, \theta)$. Next we obtain results on the distribution of the sum (and the difference) of a DGIG with an independent Gamma random variable; these results are relevant for our third near-exact distribution. One useful way to look at the distribution of Y is to see it as a particular mixture of integer Gamma or Erlang distributions. Indeed, after some rearrangements the density and distribution functions of Y may be respectively written as

$$f_Y(y) = \begin{cases} \sum_{j=1}^{p_1} \sum_{k=1}^{r_{1j}} \sum_{i=0}^{k-1} p_{jki}\, f_{Y_{jki}}(y), & y \geq 0, \\ \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki}\, f_{Y^*_{jki}}(-y), & y < 0, \end{cases}$$

and

$$F_Y(y) = \begin{cases} \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki} \\ \quad + \sum_{j=1}^{p_1} \sum_{k=1}^{r_{1j}} \sum_{i=0}^{k-1} p_{jki}\, F_{Y_{jki}}(y), & y \geq 0, \\ \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki} \\ \quad - \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki}\, F_{Y^*_{jki}}(-y), & y < 0, \end{cases}$$

where, for $j=1, \ldots, p_1; k=1, \ldots, r_{1j}; i=0, \ldots, k-1$,

$$p_{jki} = \frac{K_1 K_2}{\lambda_{1j}^{k-i}}\, c_{jk} \sum_{\ell=1}^{p_2} \sum_{h=1}^{r_{2\ell}} d_{\ell h}\, \frac{(k-1)!}{i!}\, \frac{(h+i-1)!}{(\lambda_{1j}+\lambda_{2\ell})^{h+i}}$$

and, for $j = 1, \ldots, p_2; k = 1, \ldots, r_{2j}; i = 0, \ldots, k-1$,

$$p^*_{jki} = \frac{K_1 K_2}{\lambda_{2j}^{k-i}}\, d_{jk} \sum_{\ell=1}^{p_1} \sum_{h=1}^{r_{1\ell}} c_{\ell h}\, \frac{(k-1)!}{i!}\, \frac{(h+i-1)!}{(\lambda_{1j}+\lambda_{2\ell})^{h+i}},$$

with

$$K_1 = \prod_{j=1}^{p_1} \lambda_{1j}^{r_{1j}}, \quad K_2 = \prod_{j=1}^{p_2} \lambda_{2j}^{r_{2j}},$$

and c_{jk} $(j = 1, \ldots, p_1; k = 1, \ldots, r_{1j})$ given by (2.9)–(2.11) in Coelho and Mexia (2010), with p replaced by p_1 and r_j replaced by r_{1j} and d_{jk} $(j = 1, \ldots, p_2; k = 1, \ldots, r_{2j})$ defined in a similar manner, replacing p_1 by p_2 and r_{1j} by r_{2j}, and where, for $y \geq 0$,

$$f_{Y_{jki}}(y) = \frac{\lambda_{1j}^{k-i}}{\Gamma(k-i)}\, y^{k-i-1} e^{-\lambda_{1j} y},$$

and

$$F_{Y_{jki}}(y) = 1 - \sum_{t=0}^{k-i-1} \frac{\lambda_{1j}^t}{t!}\, y^t\, e^{-\lambda_{1j} y}, \tag{24}$$

are respectively the density and distribution functions of $Y_{jki} \sim \text{Gamma}(k-i, \lambda_{1j})$, while $f_{Y^*_{jki}}(\cdot)$ and $F_{Y^*_{jki}}(\cdot)$ are the density and distribution functions of $Y^*_{jki} \sim \text{Gamma}\,(k-i, \lambda_{2j})$.

The weights p_{jki} and p^*_{jki} verify the relation

$$\sum_{j=1}^{p_1} \sum_{k=1}^{r_{1j}} \sum_{i=0}^{k-1} p_{jki} + \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki} = 1\,.$$

Let now $W \sim \text{Gamma}(\rho, \lambda)$, where ρ is a positive non-integer real, be independent of Y. We will consider the random variables $Z_1 = Y + W$ and $Z_2 = Y - W$ and derive their distribution functions. The distribution function of Z_1, will be given by

$$F_{Z_1}(z) = \int_0^{+\infty} F_Y(z-w)\, f_W(w)\, \mathrm{d}w,$$

which, for $z \geq 0$, using the notation introduced above for the GNIG distribution function, with $\boldsymbol{r}^\star = (k-i, \rho)$ and $\boldsymbol{\lambda}_1^\star = (\lambda_{1j}, \lambda)$, may be written as

$$\begin{aligned} F_{Z_1}(z) &= \int_0^{z} F_Y(\underbrace{z-w}_{\geq 0})\, f_W(w)\, \mathrm{d}w \\ &\quad + \int_z^{+\infty} F_Y(\underbrace{z-w}_{\leq 0})\, f_W(w)\, \mathrm{d}w \\ &= \sum_{j=1}^{p_2} \sum_{k=1}^{r_{2j}} \sum_{i=0}^{k-1} p^*_{jki} \int_0^{z} f_W(w)\, \mathrm{d}w \\ &\quad + \sum_{j=1}^{p_1} \sum_{k=1}^{r_{1j}} \sum_{i=0}^{k-1} p_{jki} \underbrace{\int_0^{z} F_{Y_{jki}}(z-w)\, f_W(w)\, \mathrm{d}w}_{\substack{\text{distribution function of} \\ G_1 \sim \text{GNIG}(\boldsymbol{r}^\star, \boldsymbol{\lambda}_1^\star, 2)}} \end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\left(1-\int_0^z f_W(w)\,\mathrm{d}w\right)\\
&-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\underbrace{\int_z^{+\infty} F_{Y^*_{jki}}(w-z)\, f_W(w)\,\mathrm{d}w}_{1-F_{W-Y^*_{jki}}(z)}\\
=&\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\\
&+\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki} F_{G_1}(z,\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_1,2)\\
&-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\left(1-F_{W-Y^*_{jki}}(z)\right)\\
=&\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki} F_{G_1}(z,\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_1,2)\\
&+\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y^*_{jki}}(z),
\end{aligned}
$$

while for $z<0$ we have

$$
\begin{aligned}
F_{Z_1}(z) &= \int_0^{+\infty} F_Y(\underbrace{z-w}_{\le 0})\, f_W(w)\,\mathrm{d}w\\
&=\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\underbrace{\int_0^{+\infty} f_W(w)\,\mathrm{d}w}_{=1}\\
&\quad-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\underbrace{\int_0^{+\infty} F_{Y^*_{jki}}(w-z)\, f_W(w)\,\mathrm{d}w}_{=1-F_{W-Y^*_{jki}}(z)}\\
&=\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y^*_{jki}}(z)\,.
\end{aligned}
$$

We thus have

$$
F_{Z_1}(z)=\begin{cases}
\displaystyle\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki} F_{G_1}(z;\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_1,2) & \\
\displaystyle\quad+\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y^*_{jki}}(z), & z\ge 0,\\
\displaystyle\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y^*_{jki}}(z), & z<0.
\end{cases}
\tag{25}
$$

Concerning $Z_2=Y-W$ we have, for $z<0$, using the notation introduced above for the GNIG distribution function, with $\boldsymbol{r}^\star=(k-i,\rho)$ and $\boldsymbol{\lambda}^\star_2=(\lambda_{2j},\lambda)$,

$$
\begin{aligned}
F_{Z_2}(z) &= P(Y-W\le z)=P(Y\le W+z)\\
&=\int_0^{-z} F_Y(\underbrace{w+z}_{\le 0})\, f_W(w)\,\mathrm{d}w\\
&\quad+\int_{-z}^{+\infty} F_Y(\underbrace{w+z}_{\ge 0})\, f_W(w)\,\mathrm{d}w\\
&=\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\int_0^{-z} f_W(w)\,\mathrm{d}w\\
&\quad-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\underbrace{\int_0^{-z} F_{Y^*_{jki}}(-w-z)\, f_W(w)\,\mathrm{d}w}_{\substack{\text{distribution function of}\\ G_2\sim\text{GNIG}(\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_2,2)}}\\
&\quad+\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\int_{-z}^{+\infty} f_W(w)\,\mathrm{d}w\\
&\quad+\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki}\underbrace{\int_{-z}^{+\infty} F_{Y_{jki}^*}(w+z)\, f_W(w)\,\mathrm{d}w}_{1-F_{W-Y_{jki}}(-z)}\\
&=\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}+\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki}\\
&\quad-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{G_2}(-z;\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_2,2)\\
&\quad-\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y_{jki}}(-z)\\
&=1-\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y_{jki}}(-z)\\
&\quad-\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{G_2}(-z;\boldsymbol{r}^\star,\boldsymbol{\lambda}^\star_2,2)
\end{aligned}
$$

while for $z\ge 0$ we have

$$
\begin{aligned}
F_{Z_2}(z) &= P(Y-W\le z)=P(Y\le W+z)\\
&=\int_0^{+\infty} F_Y(\underbrace{w+z}_{\ge 0})\, f_W(w)\,\mathrm{d}w\\
&=\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\int_0^{+\infty} f_W(w)\,\mathrm{d}w\\
&\quad+\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki}\underbrace{\int_0^{+\infty} F_{Y_{jki}}(w+z)\, f_W(w)\,\mathrm{d}w}_{1-F_{W-Y_{jki}}(-z)}\\
&=1-\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki} F_{W-Y_{jki}}(-z),
\end{aligned}
$$

so that

$$F_{Z_2}(z) = \begin{cases} 1-\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p_{jki}\, F_{W-Y_{jki}}(-z), & z\geq 0, \\ 1-\sum_{j=1}^{p_1}\sum_{k=1}^{r_{1j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{W-Y_{jki}}(-z) \\ -\sum_{j=1}^{p_2}\sum_{k=1}^{r_{2j}}\sum_{i=0}^{k-1} p^*_{jki}\, F_{G_2}(-z; \boldsymbol{r}^\star, \boldsymbol{\lambda}_2^\star, 2), & z<0\,. \end{cases} \tag{26}$$

It remains now to obtain the distribution function of random variables of the type of $Z^* = W - Y^*$, where $W \sim \text{Gamma}(\rho, \lambda)$ and $Y^* \sim \text{Gamma}(r, \lambda_1)$, where ρ, λ_1 and λ_2 are positive reals and r is a positive integer. The distribution function of Z^* is given by

$$\begin{aligned} F_{Z^*}(z) &= P(W-Y^*\leq z) = P(-Y^*\leq z-W) \\ &= 1-P(Y^*\leq W-z) \\ &= 1-\int_0^{+\infty} F_{Y^*}(w-z)\, f_W(w)\,\mathrm{d}w \end{aligned}$$

which for $z \geq 0$, using the expression in (24) for the distribution function of an integer Gamma or Erlang distribution, yields

$$\begin{aligned} F_{Z^*}(z) &= 1-\int_0^z \underbrace{F_{Y^*}(\underbrace{w-z}_{\leq 0})}_{=0}\, f_W(w)\,\mathrm{d}w \\ &\quad -\int_z^{+\infty} F_{Y^*}(w-z)\, f_W(w)\,\mathrm{d}w \\ &= 1-\int_z^{+\infty}\{1-P(Y^*>w-z)\}\, f_W(w)\,\mathrm{d}w \\ &= 1-\int_z^{+\infty} f_W(w)\,\mathrm{d}w \\ &\quad +\int_z^{+\infty} P(Y^*>w-z)\, f_W(w)\,\mathrm{d}w \\ &= F_W(z)+\frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\right. \\ &\quad \left.\int_z^{+\infty}(w-z)^t w^{\rho-1}\, e^{-w(\lambda+\lambda_1)}\,\mathrm{d}w\right\} \\ &= F_W(z)+\frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}(-z)^k\right. \\ &\quad \left.\int_z^{+\infty} w^{t+\rho-k-1} e^{-w(\lambda+\lambda_1)}\mathrm{d}w\right\} \\ &= 1-\frac{\Gamma(\rho,\lambda z)}{\Gamma(\rho)}+\frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}\right. \\ &\quad \left.(-z)^k(\lambda+\lambda_1)^{-t-\rho+k}\Gamma(t+\rho-k,\,(\lambda+\lambda_1)z)\right\} \end{aligned}$$

while for $z < 0$ it yields

$$\begin{aligned} F_{Z^*}(z) &= 1-\int_0^{+\infty} F_{Y^*}(w-z)\, f_W(w)\,\mathrm{d}w \\ &= 1-\int_0^{+\infty}\{1-P(Y^*>w-z)\}\, f_W(w)\,\mathrm{d}w \\ &= 1-\int_0^{+\infty} f_W(w)\,\mathrm{d}w \\ &\quad +\int_0^{+\infty} P(Y^*>w-z)\, f_W(w)\,\mathrm{d}w \\ &= \frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}(-z)^k\right. \\ &\quad \left.\int_0^{+\infty} w^{t+\rho-k-1} e^{-w(\lambda+\lambda_1)}\mathrm{d}w\right\} \\ &= \frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}(-z)^k\right. \\ &\quad \left.(\lambda+\lambda_1)^{-t-\rho+k}\Gamma(t+\rho-k)\right\}, \end{aligned}$$

and as such

$$F_{Z^*}(z) = \begin{cases} 1-\frac{\Gamma(\rho,\lambda z)}{\Gamma(\rho)}+\frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}\right. \\ \left.(-z)^k(\lambda+\lambda_1)^{-t-\rho+k}\Gamma(t+\rho-k,\,(\lambda+\lambda_1)z)\right\}, & z\geq 0, \\ \frac{\lambda^\rho}{\Gamma(\rho)}\, e^{\lambda_1 z}\left\{\sum_{t=0}^{r-1}\frac{\lambda_1^r}{t!}\sum_{k=0}^{t}\binom{t}{k}(-z)^k\right. \\ \left.(\lambda+\lambda_1)^{-t-\rho+k}\Gamma(t+\rho-k)\right\}, & z<0. \end{cases}$$

Appendix 2: Representation of a logarithmized Gamma distribution as an infinite sum of shifted exponential distributions

If $X \sim \text{Gamma}(r, \lambda)$ its hth moment is given by

$$E\left(X^h\right) = \frac{\Gamma(r+h)}{\Gamma(r)}\lambda^{-h}. \qquad (27)$$

Then, the random variable $Y = -\log X$ has what we call a logarithmized Gamma distribution and its characteristic function may be obtained from (27) in the following way

$$\Phi_Y(t) = E(Y^{-\mathrm{i}t}) = \frac{\Gamma(r-\mathrm{i}t)}{\Gamma(r)}\lambda^{\mathrm{i}t}, \quad t \in \mathbb{R},$$

Using the equality

$$\Gamma(z) = \frac{1}{z}\prod_{n=1}^{\infty}\left[\left(1+\frac{1}{n}\right)^z\left(1+\frac{z}{n}\right)^{-1}\right], \quad z \in \mathbb{C},$$

we have

$$\begin{aligned}\Phi_Y(t) &= \frac{1}{\Gamma(r)}\frac{1}{r-\mathrm{i}t}\prod_{n=1}^{\infty}\left[\left(1+\frac{1}{n}\right)^{r-\mathrm{i}t}\right.\\ &\quad\times\left.\left(1+\frac{r-\mathrm{i}t}{n}\right)^{-1}\right]\exp\{\log\lambda^{it}\}\\ &= \left\{\frac{r}{r-\mathrm{i}t}\exp\{\log\lambda^{it}\}\right\}\left[\prod_{n=1}^{\infty}\frac{n+r}{n+r-\mathrm{i}t}\right.\\ &\quad\times\left.\exp\left\{it\left(-\log\left(1+\frac{1}{n}\right)\right)\right\}\right].\end{aligned}$$

Hence Φ_Y is also the characteristic function of an infinite sum of independent shifted Exponential distributions. This shows that a logarithmized Gamma random variable may be represented as an infinite sum of independent shifted Exponential random variables.

References

Albrecher, H., Constantinescu, C., Loisel, S.: Explicit ruin formulas for models with dependence among risks. Ins. Math. Econ. **48**, 265–270 (2011)

Aldous, D., Shepp, L.: The least variable phase-type distribution is Erlang. Stoch. Mod. **3**, 467–473 (1987)

Amari, S.V., Misra, R.B.: Closed-form expressions for distribution of sum of exponential random variables. IEEE Trans. Rel. **46**, 519–522 (1997)

Antal, T., Sylos Labini, F., Vasilyev, N.L., Baryshev, Y.V.: Galaxy-distribution and extreme-value statistics. Europhy. Lett. **58**, 59001 (2009)

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: Records. Wiley, New York (1998)

Bailey, T.L., Gribskov, M.: Score distributions for simultaneous matching to multiple motifs. J. Comput. Bio. **4**, 45–59 (1997)

Balakrishnan, N., Ahsanullah, M., Chan, P.S.: Relations for single and product moments of record values from Gumbel distribution. Stat. Prob. Lett. **15**, 223–227 (1992)

Berry, A.: The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Am. Math. Soc. **49**, 122–136 (1941)

Burda, M., Harding, M., Hausman, J.: A Poisson mixture model of discrete choice. J. Econom. **166**, 184–203 (2012)

Castillo, E., Hadi, A.S., Balakrishnan, N., Sarabia, J.M.: Extreme Value and Related Models with Applications in Engineering and Science. Wiley, Hoboken (2005)

Cetinkaya, C., Kanodia, V., Knightly, E.W.: Scalable services via egress admission control. IEEE Trans. Multimed. **3**, 69–81 (2001)

Coelho, C.A.: The generalized integer Gamma distribution—a basis for distributions in multivariate statistics. J. Mult. Anal. **64**, 86–102 (1998)

Coelho, C.A.: The generalized near-integer Gamma distribution: A basis for 'near-exact' approximations to the distribution of statistics which are the product of an odd number of independent Beta random variables. J. Mult. Anal. **89**, 191–218 (2004)

Coelho, C.A., Arnold, B.C., Marques, F.J.: The distribution of the product of powers of independent uniform random variables. J. Mult. Anal. **113**, 19–36 (2013)

Coelho, C.A., Marques, F.J.: Near-exact distributions for the independence and sphericity likelihood ratio test statistics. J. Mult. Anal. **101**, 583–593 (2010)

Coelho, C.A., Marques, F.J.: Near-exact distributions for the likelihood ratio test statistic to test equality of several variance-covariance matrices in elliptically contoured distributions. Comput. Stat. **27**, 627–659 (2012)

Coelho, C.A., Mexia, J.T.: Product and Ratio of Generalized Gamma-Ratio Random Variables: Exact and Near-exact Distributions-Applications. Lambert Academic Publishing AG & Co, Saarbrücken (2010)

Domingos, P., Hulten, G.: A general framework for mining massive data streams. J. Comput. Graph. Stat. **12**, 945–949 (2003)

Esseen, C.-G.: Fourier analysis of distribution functions, a mathematical study of the Laplace–Gaussian law. Acta Math. **77**, 1–125 (1945)

Gil-Pelaez, J.: Note on the inversion theorem. Biometrika **38**, 481–482 (1951)

Gumbel, E.J.: The return period of flood flows. Ann. Math. Stat. **12**, 163–190 (1941)

Grilo, L.M., Coelho, C.A.: Development and study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables. J. Stat. Plann. Infer. **5**, 1560–1575 (2007)

Hwang, H.-K.: On convergence rates in the central limit theorems for combinatorial structures. Euro. J. Combin. **19**, 329–343 (1998)

Hosking, J.R.M., Wallis, J.R., Wood, E.F.: Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics **27**, 251–261 (1985)

Keich, U., Nagarajan, N.: A fast and numerically robust method for exact multinomial goodness-of-fit test. J. Comput. Graph. Stat. **15**, 779–802 (2006)

Loaiciga, H.A., Leipnik, R.B.: Analysis of extreme hydrology events with Gumbel distributions: marginal and additive cases. Stoch. Environ. Res. Risk Ass. **13**, 251–259 (1999)

Loève, M.: Probability Theory, 4th edn. Springer, New York (1977)

Marques, F.J.: On the product of independent Generalized Gamma random variables. Discussion Paper 19–2012, CMA-FCT-Universidade Nova de Lisboa (2012)

Marques, F.J., Coelho, C.A.: Near-exact distributions for the sphericity likelihood ratio test statistic. J. Stat. Plann. Infer. **138**, 726–741 (2008)

Marques, F.J., Coelho, C.A., Arnold, B.C.: A general near-exact distribution theory for the most common likelihood ratio test statistics used in multivariate analysis. Test **20**, 180–203 (2011)

Müller, P., Quintana, F.: Nonparametric Bayesian data analysis. Stat. Sci. **19**, 95–110 (2004)

Nadarajah, S.: Exact distribution of the linear combination of *p* Gumbel random variables. Int. J. Comput. Math. **85**, 1355–1362 (2008)

Nadarajah, S., Kotz, S.: Comments on scalable services via egress admission control. IEEE Trans. Multimed. **10**, 160–161 (2008)

Natural Environment Research Council: Flood Studies Report, Vol. 4, London (1975)

O'Cinneide, C.A.: Characterization of phase-type distributions. Stoch. Models **6**, 1–57 (1990)

Sandve, G.K., Drabløs, F.: A survey of motif discovery methods in an integrated framework. Biol. Dir. **1**, 11 (2006)

Tiago de Oliveira, J.: Decision results for the parameters of the extreme value (Gumbel) distribution based on the mean and the standard deviation. Trabajos de Estadística y de Investigación Operativa **14**, 61–81 (1963)

Wang, J.Z.: Selection of the *K* largest order statistics for the domain of attraction of the Gumbel distribution. J. Am. Stat. Assoc. **90**, 1055–1061 (1995)