



**HAL**  
open science

# Saddlepoint Approximations of Cumulative Distribution Functions of Sums of Random Vectors

Dadja Anade, Jean-Marie S Gorce, Philippe Mary, Samir M. Perlaza

► **To cite this version:**

Dadja Anade, Jean-Marie S Gorce, Philippe Mary, Samir M. Perlaza. Saddlepoint Approximations of Cumulative Distribution Functions of Sums of Random Vectors. ISIT 2021 - IEEE International Symposium on Information Theory, Jul 2021, Melbourne / Virtual, Australia. pp.1-6, 10.1109/ISIT45174.2021.9518101 . hal-03226009

**HAL Id: hal-03226009**

**<https://inria.hal.science/hal-03226009>**

Submitted on 13 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Saddlepoint Approximations of Cumulative Distribution Functions of Sums of Random Vectors

Dadja Anade, Jean-Marie Gorce, Philippe Mary, and Samir M. Perlaza

**Abstract**—In this paper, a real-valued function that approximates the cumulative distribution function (CDF) of a finite sum of real-valued independent and identically distributed random vectors is presented. The approximation error is upper bounded by an expression that is easy to calculate. As a byproduct, an upper bound and a lower bound on the CDF are obtained. Finally, in the case of lattice and absolutely continuous random variables, the proposed approximation is shown to be identical to the saddlepoint approximation of the CDF.

## I. NOTATION

The real numbers are denoted by  $\mathbb{R}$ , and the natural numbers are denoted by  $\mathbb{N}$ . In particular,  $0 \notin \mathbb{N}$ . The Borel sigma field on  $\mathbb{R}^k$ , with  $k \in \mathbb{N}$ , is denoted by  $\mathcal{B}(\mathbb{R}^k)$ . The Euclidian norm in  $\mathbb{R}^k$  is denoted by  $\|\cdot\|$ . Given a set  $\mathcal{A} \subseteq \mathbb{R}^k$ , the closure of the set  $\mathcal{A}$  is denoted by  $\text{clo}\mathcal{A} \triangleq \{\mathbf{x} \in \mathbb{R}^k : \forall r > 0, \exists \mathbf{y} \in \mathcal{A}, \|\mathbf{x} - \mathbf{y}\| < r\}$ . A diagonal matrix whose diagonal is the vector  $\mathbf{x} \in \mathbb{R}^k$  is denoted by  $\text{diag}(\mathbf{x})$ .

## II. INTRODUCTION

The calculation of cumulative distribution functions (CDFs) of sums of random vectors is omnipresent in the realm of information theory. For instance, the joint decoding error probability in multi-user channels often boils down to the calculation of CDFs of random vectors, c.f., [1]. In the case of the memoryless Gaussian multiple access channel, under certain conditions on the channel inputs, the dependence testing bound corresponds to the CDF of a sum of independent and identically distributed (IID) random vectors [2, Theorem 2]. Unfortunately, the calculation of CDFs of random vectors requires elaborated numerical methods which often lead to errors that are difficult to quantify. From this perspective, approximations to these CDFs, e.g., Gaussian approximations and saddlepoint approximations [3]–[7] have gained remarkable popularity [8]–[14]. In the case of Gaussian approximations, multi-dimensional Berry-Esseen-type theorems provide upper bounds on the approximation errors, c.f., [15]. These bounds are particularly precise around the mean. Alternatively, saddlepoint approximations are known to be more precise than

Dadja Anade and Jean-Marie Gorce are with the Laboratoire CITI, a joint laboratory between the Institut National de Recherche en Informatique et en Automatique (INRIA), the Université de Lyon and the Institut National de Sciences Appliquées (INSA) de Lyon, Villeurbanne, France. {dadja.anade-akpo, jean-marie.gorce}@inria.fr

Samir M. Perlaza is with INRIA, Centre de Recherche de Sophia Antipolis - Méditerranée, Sophia Antipolis, France; and with the Electrical Engineering Department, Princeton University, Princeton, USA; samir.perlaza@inria.fr

Philippe Mary is with the Laboratoire IETR, a joint laboratory between the Centre National de Recherche Scientifique (CNRS), the Université de Rennes and the Institut National de Sciences Appliquées (INSA) de Rennes, Rennes, France. philippe.mary@insa-rennes.fr

This work was partially funded by the French National Agency for Research (ANR) under grant ANR-16-CE25-0001.

Gaussian approximations far apart from the mean. Unfortunately, this claim is often justified only by numerical analysis as formal upper bounds on the error induced by saddlepoint approximations are rather inexistent, c.f., [5] and [6]. This paper contributes in this direction by introducing a real-valued function that approximates the CDF of a finite sum of real-valued IID random vectors. Both Gaussian and saddlepoint approximations are shown to be special cases of the proposed approximation, which is referred to as the exponentially tilted Gaussian approximation. The approximation error is upper bounded and both upper and lower bounds on the CDF are obtained.

## III. PRELIMINARIES

Let  $n$  be a finite integer, with  $n > 1$ , and let  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  be independent random vectors such that each of them induces the probability measure  $P_{\mathbf{Y}}$  on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ , with  $k \in \mathbb{N}$ . Denote by  $K_{\mathbf{Y}} : \mathbb{R}^k \rightarrow \mathbb{R}$  the cumulant generating function (CGF) of each of these random variables. That is, for all  $\mathbf{t} \in \mathbb{R}^k$ ,

$$K_{\mathbf{Y}}(\mathbf{t}) = \ln \left( \mathbb{E}_{P_{\mathbf{Y}}} \left[ \exp \left( \mathbf{t}^T \mathbf{Y} \right) \right] \right). \quad (1)$$

The gradient of the CGF  $K_{\mathbf{Y}}$  is a function denoted by  $K_{\mathbf{Y}}^{(1)} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ . More specifically, for all  $\mathbf{t} \in \mathbb{R}^k$ ,

$$K_{\mathbf{Y}}^{(1)}(\mathbf{t}) = \mathbb{E}_{P_{\mathbf{Y}}} \left[ \mathbf{Y} \exp \left( \mathbf{t}^T \mathbf{Y} - K_{\mathbf{Y}}(\mathbf{t}) \right) \right]. \quad (2)$$

The Hessian of the CGF  $K_{\mathbf{Y}}$  is a function denoted by  $K_{\mathbf{Y}}^{(2)} : \mathbb{R}^k \rightarrow \mathbb{R}^{k \times k}$ . That is, for all  $\mathbf{t} \in \mathbb{R}^k$ ,

$$K_{\mathbf{Y}}^{(2)}(\mathbf{t}) = \mathbb{E}_{P_{\mathbf{Y}}} \left[ \left( \mathbf{Y} - K_{\mathbf{Y}}^{(1)}(\mathbf{t}) \right) \left( \mathbf{Y} - K_{\mathbf{Y}}^{(1)}(\mathbf{t}) \right)^T \exp \left( \mathbf{t}^T \mathbf{Y} - K_{\mathbf{Y}}(\mathbf{t}) \right) \right], \quad (3)$$

Note that  $K_{\mathbf{Y}}^{(1)}(\mathbf{0})$  and  $K_{\mathbf{Y}}^{(2)}(\mathbf{0})$  are respectively the mean vector and the covariance matrix of each of the random vectors  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$ . In the following,  $K_{\mathbf{Y}}^{(2)}(\mathbf{0})$  is assumed to be positive definite.

Let also

$$\mathbf{X}_n \triangleq \sum_{t=1}^n \mathbf{Y}_t \quad (4)$$

be a random vector that induces the probability measure  $P_{\mathbf{X}_n}$  on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ , with cumulative distribution function (CDF) denoted by  $F_{\mathbf{X}_n}$ . When for all  $i \in \{1, 2, \dots, n\}$  the random vector  $\mathbf{Y}_i$  in (4) is absolutely continuous and its corresponding CGF  $K_{\mathbf{Y}}$  is such that the set

$$\mathcal{C}_{\mathbf{Y}} \triangleq \{\boldsymbol{\theta} \in \mathbb{R}^k : K_{\mathbf{Y}}(\boldsymbol{\theta}) < \infty\} \cap ]-\infty, 0[^k \quad (5)$$

is not empty, the CDF  $F_{\mathbf{X}_n}$  can be written as a complex integral [3]. In particular, for all  $\mathbf{x} \in \mathbb{R}^k$ ,

$$F_{\mathbf{X}_n}(\mathbf{x}) = \int_{c-i\infty}^{c+i\infty} \frac{\exp(nK_{\mathbf{Y}}(\boldsymbol{\tau}) - \boldsymbol{\tau}^\top \mathbf{x})}{(2\pi i)^k \prod_{t=1}^k \tau_t} d\boldsymbol{\tau}, \quad (6)$$

where  $i$  is the imaginary unit;  $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_k)$ ; and the constant  $c$  is arbitrarily chosen to satisfy  $c \in \mathcal{C}_{\mathbf{Y}}$ . The complex integral in (6) is a multivariate Laplace inverse transform [9], [16], and can be approximated with high precision, as shown hereunder. Denote by  $\mathcal{D}$  the following set

$$\mathcal{D} \triangleq \left\{ \mathbf{u} \in \mathbb{R}^k : \exists t \in ]-\infty, 0[^k, nK_{\mathbf{Y}}^{(1)}(t) = \mathbf{u} \right\}, \quad (7)$$

and denote by  $\boldsymbol{\tau}_0 \in \mathbb{R}^k$  the unique solution in  $\boldsymbol{\tau}$  to

$$K_{\mathbf{Y}}^{(1)}(\boldsymbol{\tau}) = \frac{1}{n} \mathbf{x}. \quad (8)$$

For all  $\mathbf{x} \in \mathcal{D}$ , a Taylor series expansion of  $nK_{\mathbf{Y}}(\boldsymbol{\tau}) - \boldsymbol{\tau}^\top \mathbf{x}$  in the neighborhood of  $\boldsymbol{\tau}_0$ , leads to the following asymptotic expansion of the integral in (6):

$$F_{\mathbf{X}_n}(\mathbf{x}) = \hat{F}_{\mathbf{X}_n}(\mathbf{x}) + O\left(\frac{\exp(nK_{\mathbf{Y}}(\boldsymbol{\tau}_0) - \boldsymbol{\tau}_0^\top \mathbf{x})}{\sqrt{n}}\right), \quad (9)$$

where the function  $\hat{F}_{\mathbf{X}_n} : \mathcal{D} \rightarrow \mathbb{R}$  is

$$\hat{F}_{\mathbf{X}_n}(\mathbf{x}) = \exp\left(nK_{\mathbf{Y}}(\boldsymbol{\tau}_0) - \boldsymbol{\tau}_0^\top \mathbf{x} + \frac{n\boldsymbol{\tau}_0^\top K_{\mathbf{Y}}^{(2)}(\boldsymbol{\tau}_0)\boldsymbol{\tau}_0}{2}\right) F_{G_n^{(\boldsymbol{\tau}_0)}}(nK_{\mathbf{Y}}^{(2)}(\boldsymbol{\tau}_0)\boldsymbol{\tau}_0), \quad (10)$$

and the function  $F_{G_n^{(\boldsymbol{\tau}_0)}} : \mathbb{R}^k \rightarrow [0, 1]$  is the CDF of a Gaussian random vector with mean vector  $(0, 0, \dots, 0)$  and covariance matrix  $nK_{\mathbf{Y}}^{(2)}(\boldsymbol{\tau}_0)$ .

The vector  $\boldsymbol{\tau}_0$  and the function  $\hat{F}_{\mathbf{X}_n}$  in (10) are respectively referred to as the *saddlepoint* and the *saddlepoint approximation* of the CDF  $F_{\mathbf{X}_n}$ . In [3], it is shown that the approximation  $\hat{F}_{\mathbf{X}_n}$  in (10) also holds for the case in which for all  $i \in \{1, 2, \dots, n\}$  the vector  $\mathbf{Y}_i$  in (4) is a lattice random vector. Moreover, when for all  $i \in \{1, 2, \dots, n\}$  the random vector  $\mathbf{Y}_i$  in (4) is a Gaussian random vector, then the *saddlepoint approximation* is exact. That is,  $\hat{F}_{\mathbf{X}_n}$  and  $F_{\mathbf{X}_n}$  are identical. Using the elements above, the main contributions of this work can be described as follows:

- (a) A real-valued function that approximates the CDF  $F_{\mathbf{X}_n}$  is presented. This approximation is shown to be identical to the saddlepoint approximation  $\hat{F}_{\mathbf{X}_n}$  in (9) when  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  are either absolutely continuous or lattices random vectors; and
- (b) an upper bound on the error induced by the proposed approximation is also presented. The asymptotic behaviour with  $n$  of the proposed upper bound is consistent with the one suggested by (9).

#### IV. GAUSSIAN APPROXIMATION

Let  $\boldsymbol{\mu}_{\mathbf{X}_n} \in \mathbb{R}^k$  and  $\underline{\mathbf{v}}_{\mathbf{X}_n} \in \mathbb{R}^{k \times k}$  be the mean vector and covariance matrix of the random vector  $\mathbf{X}_n$  in (4). The Gaussian approximation of the measure  $P_{\mathbf{X}_n}$  induced by  $\mathbf{X}_n$  is the probability measure induced by a Gaussian

vector with mean vector  $\boldsymbol{\mu}_{\mathbf{X}_n}$  and covariance matrix  $\underline{\mathbf{v}}_{\mathbf{X}_n}$ . The following theorem, known as the multivariate Berry–Esseen theorem [15], introduces an upper bound on the approximation error.

*Theorem 1 ([15, Theorem 1.1]):* Assume that the measure  $P_{\mathbf{Y}}$  induced by each of the random vectors  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  in (4) satisfies,

$$\mathbb{E}_{P_{\mathbf{Y}}}[\mathbf{Y}] = (0, 0, \dots, 0), \text{ and} \quad (11)$$

$$\mathbb{E}_{P_{\mathbf{Y}}}[\mathbf{Y}\mathbf{Y}^\top] = \frac{1}{n} \text{diag}(1, 1, \dots, 1). \quad (12)$$

Let  $P_{\mathbf{Z}_n}$  be the probability measure induced on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$  by a Gaussian random vector  $\mathbf{Z}_n$  with mean vector  $(0, 0, \dots, 0)$  and covariance matrix  $\text{diag}(1, 1, \dots, 1)$ . Then,

$$\sup_{\mathcal{A} \in \mathcal{C}_k} |P_{\mathbf{X}_n}(\mathcal{A}) - P_{\mathbf{Z}_n}(\mathcal{A})| \leq \min\left(1, c(k)n\mathbb{E}_{P_{\mathbf{Y}}}[\|\mathbf{Y}\|^3]\right), \quad (13)$$

where  $\mathcal{C}_k$  is the collection of all convex sets in  $\mathcal{B}(\mathbb{R}^k)$ ; and the function  $c : \mathbb{N} \rightarrow \mathbb{R}$  satisfies for all  $k \in \mathbb{N}$ ,

$$c(k) = 42k^{\frac{1}{4}} + 16. \quad (14)$$

Theorem 1 leads to the following inequalities for all  $\mathbf{x} \in \mathbb{R}^k$ ,

$$\underline{\Sigma}(n, \mathbf{x}) \leq F_{\mathbf{X}_n}(\mathbf{x}) \leq \bar{\Sigma}(n, \mathbf{x}), \quad (15)$$

where,

$$\bar{\Sigma}(n, \mathbf{x}) \triangleq F_{\mathbf{Z}_n}(\mathbf{x}) + \min\left(1, c(k)n\mathbb{E}_{P_{\mathbf{Y}}}[\|\mathbf{Y}\|^3]\right), \text{ and} \quad (16a)$$

$$\underline{\Sigma}(n, \mathbf{x}) \triangleq F_{\mathbf{Z}_n}(\mathbf{x}) - \min\left(1, c(k)n\mathbb{E}_{P_{\mathbf{Y}}}[\|\mathbf{Y}\|^3]\right). \quad (16b)$$

#### V. EXPONENTIALLY TILTED GAUSSIAN APPROXIMATION

This section introduces two central results. First, given a convex set  $\mathcal{A}$  in  $\mathcal{B}(\mathbb{R}^k)$ , the probability  $P_{\mathbf{X}_n}(\mathcal{A})$ , with  $P_{\mathbf{X}_n}$  the probability measure induced by the random vector  $\mathbf{X}_n$  in (4), is approximated by a function that is a measure but not necessary a probability measure. This function, which is parametrized by a vector in  $\mathbb{R}^k$  that can be arbitrarily chosen, is often referred to as the exponentially tilted Gaussian approximation of  $P_{\mathbf{X}_n}$ . Second, using the first result, the CDF of  $\mathbf{X}_n$  is approximated by a function that is not necessarily a CDF. Nonetheless, this function is parametrized by a vector in  $\mathbb{R}^k$  that can be arbitrarily chosen to locally minimize the approximation error.

##### A. Approximation of the Measure

Denote by  $\eta_{\mathbf{Y}} : \mathbb{R}^k \times \mathcal{B}(\mathbb{R}^k) \times \mathbb{N} \rightarrow \mathbb{R}$  the function such that for all  $\boldsymbol{\theta} \in \mathbb{R}^k$

$$\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n) \triangleq \exp\left(n\left(K_{\mathbf{Y}}(\boldsymbol{\theta}) - \boldsymbol{\theta}^\top K_{\mathbf{Y}}^{(1)}(\boldsymbol{\theta}) + \frac{\boldsymbol{\theta}^\top K_{\mathbf{Y}}^{(2)}(\boldsymbol{\theta})\boldsymbol{\theta}}{2}\right)\right) P_{\mathbf{H}_n^{(\boldsymbol{\theta})}}(\mathcal{A}), \quad (17)$$

where the probability measure  $P_{\mathbf{H}_n^{(\boldsymbol{\theta})}}$  is the probability measure induced by a Gaussian random vector  $\mathbf{H}_n^{(\boldsymbol{\theta})}$  with mean vector  $n\left(K_{\mathbf{Y}}^{(1)}(\boldsymbol{\theta}) - K_{\mathbf{Y}}^{(2)}(\boldsymbol{\theta})\boldsymbol{\theta}\right)$  and covariance matrix  $nK_{\mathbf{Y}}^{(2)}(\boldsymbol{\theta})$  on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ .

The objective is to show that  $P_{\mathbf{X}_n}(\mathcal{A})$  can be approximated by  $\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)$  for some  $\boldsymbol{\theta} \in \mathbb{R}^k$ . The intuition behind this is presented hereunder. First, in [17], it is shown that the probability  $P_{\mathbf{X}_n}(\mathcal{A})$  can be written as follows

$$P_{\mathbf{X}_n}(\mathcal{A}) = \exp(nK_{\mathbf{Y}}(\boldsymbol{\theta})) \mathbb{E}_{P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}} \left[ \exp(-\boldsymbol{\theta}^\top \mathbf{S}_n^{(\boldsymbol{\theta})}) \mathbb{1}_{\{\mathbf{S}_n^{(\boldsymbol{\theta})} \in \mathcal{A}\}} \right], \quad (18)$$

where the probability measure  $P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}$  on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$  is induced by the random vector

$$\mathbf{S}_n^{(\boldsymbol{\theta})} \triangleq \sum_{j=1}^n \mathbf{Y}_j^{(\boldsymbol{\theta})}; \quad (19)$$

and for all  $j \in \{1, 2, \dots, n\}$ , the random vector  $\mathbf{Y}_j^{(\boldsymbol{\theta})}$  induces the probability measure  $P_{\mathbf{Y}^{(\boldsymbol{\theta})}}$  on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ . The Radon-Nikodym derivative of  $P_{\mathbf{Y}^{(\boldsymbol{\theta})}}$  with respect to  $P_{\mathbf{Y}}$  satisfies for all  $\mathbf{y} \in \mathbb{R}^k$ ,

$$\frac{dP_{\mathbf{Y}^{(\boldsymbol{\theta})}}}{dP_{\mathbf{Y}}}(\mathbf{y}) = \exp(\boldsymbol{\theta}^\top \mathbf{y} - K_{\mathbf{Y}}(\boldsymbol{\theta})). \quad (20)$$

That is, the probability measure  $P_{\mathbf{Y}^{(\boldsymbol{\theta})}}$  is an exponentially tilted measure with respect to the probability measure  $P_{\mathbf{Y}}$ . The equality in (18) is obtained by a change of measure with which the expectation  $\mathbb{E}_{P_{\mathbf{X}_n}}[\mathbb{1}_{\{\mathbf{X}_n \in \mathcal{A}\}}] = P_{\mathbf{X}_n}(\mathcal{A})$  is calculated with respect to the probability measure  $P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}$  induced by the random variable  $\mathbf{S}_n^{(\boldsymbol{\theta})}$  in (19). Second, in [17], it is shown that  $\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)$  can be written as follows,

$$\begin{aligned} & \eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n) \\ &= \exp(nK_{\mathbf{Y}}(\boldsymbol{\theta})) \mathbb{E}_{P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}} \left[ \exp(-\boldsymbol{\theta}^\top \mathbf{Z}_n^{(\boldsymbol{\theta})}) \mathbb{1}_{\{\mathbf{Z}_n^{(\boldsymbol{\theta})} \in \mathcal{A}\}} \right], \quad (21) \end{aligned}$$

where the probability measure  $P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}$  is induced on the measurable space  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$  by the Gaussian random vector  $\mathbf{Z}_n^{(\boldsymbol{\theta})}$  with the same mean vector and covariance matrix as the random vector  $\mathbf{S}_n^{(\boldsymbol{\theta})}$  in (19). Hence, the approximation of the probability  $P_{\mathbf{X}_n}(\mathcal{A})$  by  $\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)$ , follows from the arbitrary assumption that the probability measure  $P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}$  can be approximated by the probability measure  $P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}$ . Given that  $P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}$  is an exponentially tilted measure with respect to  $P_{\mathbf{X}_n}$  and  $P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}$  is the Gaussian approximation of  $P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}$ , the function  $\eta_{\mathbf{Y}}$  in (21) is referred to as the exponentially tilted Gaussian approximation of  $P_{\mathbf{X}_n}$ . The calculation of an upper bound on the error induced by such approximation is the aim of the following lemma.

*Lemma 2:* Given  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta_{\mathbf{Y}}$ , with

$$\Theta_{\mathbf{Y}} \triangleq \{\mathbf{t} \in \mathbb{R}^k : K_{\mathbf{Y}}(\mathbf{t}) < \infty\}, \quad (22)$$

and a convex set  $\mathcal{A} \in \mathcal{B}(\mathbb{R}^k)$ , it holds that

$$\begin{aligned} & |P_{\mathbf{X}_n}(\mathcal{A}) - \eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)| \\ & \leq \exp(nK_{\mathbf{Y}}(\boldsymbol{\theta}) - \boldsymbol{\theta}^\top \mathbf{a}(\mathcal{A}, \boldsymbol{\theta})) \Delta(P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}, P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}), \quad (23) \end{aligned}$$

where

$$\Delta(P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}, P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}) \triangleq \sup_{\mathcal{B} \in \mathcal{C}_k} |P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}(\mathcal{B}) - P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}}(\mathcal{B})|; \quad (24)$$

the collection  $\mathcal{C}_k$  contains all convex sets in  $\mathcal{B}(\mathbb{R}^k)$ ; and the vector  $\mathbf{a}(\mathcal{A}, \boldsymbol{\theta}) = (a_1(\mathcal{A}, \boldsymbol{\theta}), a_2(\mathcal{A}, \boldsymbol{\theta}), \dots, a_k(\mathcal{A}, \boldsymbol{\theta}))$  is such that for all  $i \in \{1, 2, \dots, k\}$ ,

$$a_i(\mathcal{A}, \boldsymbol{\theta}) \triangleq \begin{cases} 0 & \text{if } \theta_i = 0 \\ \inf_{(b_1, b_2, \dots, b_k) \in \mathcal{A}} b_i & \text{if } \theta_i > 0 \\ \sup_{(b_1, b_2, \dots, b_k) \in \mathcal{A}} b_i & \text{if } \theta_i < 0. \end{cases} \quad (25)$$

*Proof:* The proof of Lemma 2 is presented in [17]. ■

Note that the term  $\Delta(P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}, P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}})$  in (24) can be upper bounded by leveraging the observation that the random vector  $\mathbf{S}_n^{(\boldsymbol{\theta})}$  is the sum of  $n$  independent random vectors and  $\mathbf{Z}_n^{(\boldsymbol{\theta})}$  is a Gaussian random vector with the same mean vector and covariance matrix as  $\mathbf{S}_n^{(\boldsymbol{\theta})}$ . This allows using Theorem 1 for upper bounding  $\Delta(P_{\mathbf{S}_n^{(\boldsymbol{\theta})}}, P_{\mathbf{Z}_n^{(\boldsymbol{\theta})}})$  in (24). For doing so, consider the function  $\xi_{\mathbf{Y}} : \mathbb{R}^k \rightarrow \mathbb{R}$ , such that for all  $\mathbf{t} \in \mathbb{R}^k$ ,

$$\begin{aligned} \xi_{\mathbf{Y}}(\mathbf{t}) \triangleq & \mathbb{E}_{P_{\mathbf{Y}}} \left[ \left( (\mathbf{Y} - K_{\mathbf{Y}}^{(1)}(\mathbf{t}))^\top (K_{\mathbf{Y}}^{(2)}(\mathbf{t}))^{-1} (\mathbf{Y} - K_{\mathbf{Y}}^{(1)}(\mathbf{t})) \right)^{3/2} \right. \\ & \left. \exp(\mathbf{t}^\top \mathbf{Y} - K_{\mathbf{Y}}(\mathbf{t})) \right]. \quad (26) \end{aligned}$$

Using this notation, the following lemma introduces an upper bound on the error induced by the approximation of the probability  $P_{\mathbf{X}_n}(\mathcal{A})$  by  $\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)$ , where  $\mathcal{A} \subseteq \mathbb{R}^k$  is a convex Borel measurable set and  $\boldsymbol{\theta} \in \mathbb{R}^k$  is a fixed parameter.

*Lemma 3:* For all  $\mathcal{A} \in \mathcal{C}_k$ , with  $\mathcal{C}_k$  the collection of all convex sets in  $\mathcal{B}(\mathbb{R}^k)$ , and for all  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k) \in \Theta_{\mathbf{Y}}$ , with  $\Theta_{\mathbf{Y}}$  in (22), it holds that

$$\begin{aligned} & |P_{\mathbf{X}_n}(\mathcal{A}) - \eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}, n)| \\ & \leq \exp(nK_{\mathbf{Y}}(\boldsymbol{\theta}) - \boldsymbol{\theta}^\top \mathbf{a}(\mathcal{A}, \boldsymbol{\theta})) \min\left(1, \frac{c(k)\xi_{\mathbf{Y}}(\boldsymbol{\theta})}{\sqrt{n}}\right). \quad (27) \end{aligned}$$

*Proof:* The proof of Lemma 3 is presented in [17]. ■

## B. Approximation of the CDF

The CDF  $F_{\mathbf{X}_n}$  can be written in the form of the probability of a convex set in  $\mathcal{B}(\mathbb{R}^k)$ . That is, for all  $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ , let the set  $\mathcal{A}_{\mathbf{x}}$  be such that

$$\mathcal{A}_{\mathbf{x}} = \{(t_1, t_2, \dots, t_k) \in \mathbb{R}^k : \forall i \in \{1, 2, \dots, k\}, t_i \leq x_i\}. \quad (28)$$

Then, for all  $\mathbf{x} \in \mathbb{R}^k$ , it holds that

$$F_{\mathbf{X}_n}(\mathbf{x}) = P_{\mathbf{X}_n}(\mathcal{A}_{\mathbf{x}}). \quad (29)$$

This observation allows to use Lemma 3 to approximate the CDF  $F_{\mathbf{X}_n}$  of the random vector  $\mathbf{X}_n$  in (4). Explicitly, for all  $\mathbf{x} \in \mathbb{R}^k$  and for all  $\boldsymbol{\theta} \in \Theta_{\mathbf{Y}}$ , with  $\Theta_{\mathbf{Y}}$  in (22), it holds that

$$\begin{aligned} & |F_{\mathbf{X}_n}(\mathbf{x}) - \eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}_{\mathbf{x}}, n)| \\ & \leq \exp(nK_{\mathbf{Y}}(\boldsymbol{\theta}) - \boldsymbol{\theta}^\top \mathbf{a}(\mathcal{A}_{\mathbf{x}}, \boldsymbol{\theta})) \min\left(1, \frac{c(k)\xi_{\mathbf{Y}}(\boldsymbol{\theta})}{\sqrt{n}}\right). \quad (30) \end{aligned}$$

The upper bound on the approximation error induced by approximating the CDF  $F_{\mathbf{X}_n}$  in (30) by  $\eta_{\mathbf{Y}}(\boldsymbol{\theta}, \mathcal{A}_{\mathbf{x}}, n)$  can be reduced by choosing the parameter  $\boldsymbol{\theta} \in \Theta_{\mathbf{Y}}$  that minimizes

the right-hand side (RHS) of (30). From this standpoint, the parameter  $\theta$  must be searched within a subset of  $\Theta_Y$  in which

$$\theta^\top \mathbf{a}(\mathcal{A}_x, \theta) = \theta^\top \mathbf{x} < +\infty. \quad (31)$$

More specifically, given  $\mathcal{A}_x$  in (28), it follows from (25) that the minimization must be restricted to the set

$$\Theta_Y^- \triangleq \{(t_1, t_2, \dots, t_k) \in \Theta_Y, \forall i \in \{1, 2, \dots, k\}, t_i \leq 0\}. \quad (32)$$

An arbitrary choice of  $\theta$  is the one that minimizes the exponential term  $\exp(nK_Y(\theta) - \theta^\top \mathbf{a}(\mathcal{A}_x, \theta))$ , which depends on  $\mathbf{x}$ . Denote such a choice by  $\theta(\mathbf{x})$ , which is defined in terms of the following quantity:

$$\tau(\mathbf{x}) \triangleq \arg \min_{\mathbf{t} \in \text{clo} \Theta_Y^-} (nK_Y(\mathbf{t}) - \mathbf{t}^\top \mathbf{x}). \quad (33)$$

The uniqueness of  $\tau(\mathbf{x})$  in (33), for a given  $\mathbf{x}$ , follows from the fact that the set  $\Theta_Y^-$  in (32) is convex and the function  $nK_Y(\theta) - \theta^\top \mathbf{a}(\mathcal{A}_x, \theta)$  is strictly convex with respect to  $\theta$ . More specifically, the difference between a strictly convex function, i.e.,  $nK_Y(\theta)$  and a linear function, i.e.,  $\theta^\top \mathbf{a}(\mathcal{A}_x, \theta)$  is strictly convex. The former is strictly convex due to the assumption that the covariance matrix  $K_Y^{(2)}(\mathbf{0})$  is a positive definite matrix, c.f., [5, Section 1.2] and [18, Theorem 7.1]. Given  $\tau(\mathbf{x})$  in (33), the vector  $\theta$  is

$$\theta(\mathbf{x}) \triangleq \begin{cases} \tau(\mathbf{x}) & \text{if } \tau(\mathbf{x}) \in \Theta_Y^- \\ \tau(\mathbf{x}) + \epsilon & \text{otherwise,} \end{cases} \quad (34)$$

where  $\epsilon \in \mathbb{R}^k$  is chosen such that two conditions are simultaneously met: First,  $\|\epsilon\| < r$ , with  $r > 0$  arbitrary small; and second,  $\theta(\mathbf{x}) \in \Theta_Y^-$ . The following lemma, presents some of the properties of  $\theta(\mathbf{x})$ , for all  $\mathbf{x} \in \mathbb{R}^k$ .

*Lemma 4:* For all  $\mathbf{x} \in \mathbb{R}^k$ ,  $\theta(\mathbf{x})$  in (34) satisfies

$$(\mathbf{x} - \mu_{\mathbf{X}_n})^\top \theta(\mathbf{x}) \geq 0, \text{ and } \theta(\mu_{\mathbf{X}_n}) = \mathbf{0}, \quad (35)$$

where,

$$\mu_{\mathbf{X}_n} = (\mu_{X_{n,1}}, \mu_{X_{n,2}}, \dots, \mu_{X_{n,k}})^\top \quad (36)$$

is the mean of the random vector  $\mathbf{X}_n$  in (4).

*Proof:* The proof of Lemma 4 is presented in [17]. ■

Consider the following set

$$\mathcal{E}_{\mathbf{X}_n} \triangleq \{(x_1, x_2, \dots, x_k) \in \mathbb{R}^k : \forall i \in \{1, 2, \dots, k\}, x_i > \mu_{X_{n,i}}\}, \quad (37)$$

where for all  $i \in \{1, 2, \dots, k\}$ ,  $\mu_{X_{n,i}}$  is defined in (36). From (35), it holds that for all  $\mathbf{x} \in \mathcal{E}_{\mathbf{X}_n}$ ,

$$\theta(\mathbf{x}) = \mathbf{0}. \quad (38)$$

This implies that for all  $\mathbf{x} \in \mathcal{E}_{\mathbf{X}_n}$ ,  $\theta(\mathbf{x})$  leads to the Gaussian approximation of the CDF  $F_{\mathbf{X}_n}$ . Hence, for all  $\mathbf{x} \in \mathcal{E}_{\mathbf{X}_n}$  the choice of  $\theta$  in (34) can still be improved. For doing so, for all  $\mathbf{x} \in \mathcal{E}_{\mathbf{X}_n}$ , the objective is to write  $1 - F_{\mathbf{X}_n}(\mathbf{x})$  as a sum of probability measures of convex sets with respect to  $P_{\mathbf{X}_n}$ . The following lemma provides such a result.

*Lemma 5:* For all  $\mathbf{x} = (x_1, x_2, \dots, x_k)^\top \in \mathbb{R}^k$ , with  $k \in \mathbb{N}$ , it holds that

$$1 - F_{\mathbf{X}_n}(\mathbf{x}) = \sum_{i=1}^k P_{\mathbf{X}_n}(\mathcal{B}(\mathbf{x}, i)), \quad (39)$$

where the set  $\mathcal{B}(\mathbf{x}, i)$  is ,

$$\mathcal{B}(\mathbf{x}, i) = \left\{ \mathbf{t} = (t_1, t_2, \dots, t_k) \in \mathbb{R}^k : \forall j \in \{1, 2, \dots, k\}, t_j \leq x_j \text{ if } j < i, \text{ and } t_j > x_j \text{ if } j = i \right\}. \quad (40)$$

*Proof:* The proof of Lemma 5 is presented in [17]. ■

Note that the choice of the sets  $\mathcal{B}(\mathbf{x}, 1), \mathcal{B}(\mathbf{x}, 2), \dots, \mathcal{B}(\mathbf{x}, k)$  in (39) is not unique, c.f., the inclusion-exclusion principle. There exists  $k!$  possible ways of choosing these sets. As shown later, each choice induces a different approximation on  $F_{\mathbf{X}_n}$  and different approximation errors.

For all  $i \in \{1, 2, \dots, k\}$ , let the set  $\Theta_Y^i$  be such that

$$\Theta_Y^i \triangleq \{(\theta_1, \theta_2, \dots, \theta_k) \in \Theta_Y : \forall j \in \{1, 2, \dots, k\}, \theta_j \leq 0 \text{ if } j < i, \theta_j \geq 0 \text{ if } j = i, \text{ and } \theta_j = 0 \text{ otherwise}\}. \quad (41)$$

The probability  $P_{\mathbf{X}_n}(\mathcal{B}(\mathbf{x}, i))$  in (39) can be approximated by using Lemma 3. More specifically, for all  $i \in \{1, 2, \dots, k\}$  and for all  $\theta \in \Theta_Y^i$ ,

$$\begin{aligned} & |P_{\mathbf{X}_n}(\mathcal{B}(\mathbf{x}, i)) - \eta_Y(\theta, \mathcal{B}(\mathbf{x}, i), n)| \\ & \leq \exp(nK_Y(\theta) - \theta^\top \mathbf{x}) \min\left(1, \frac{c^{(k)} \xi_Y(\theta)}{\sqrt{n}}\right). \end{aligned} \quad (42)$$

Let the functions  $\zeta_Y : \mathbb{N} \times \mathbb{R}^k \rightarrow \mathbb{R}$  and  $\delta_Y : \mathbb{N} \times \mathbb{R}^k \rightarrow \mathbb{R}$  be such that for all  $(n, \mathbf{x}) \in \mathbb{N} \times \mathbb{R}^k$ ,

$$\zeta_Y(n, \mathbf{x}) \triangleq \begin{cases} \eta_Y(\theta(\mathbf{x}), \mathcal{A}_x, n) & \text{if } \mathbf{x} \notin \mathcal{E}_{\mathbf{X}_n} \\ 1 - \sum_{i=1}^k \eta_Y(\theta_i(\mathbf{x}), \mathcal{B}(\mathbf{x}, i), n) & \text{otherwise,} \end{cases} \text{ and} \quad (43)$$

$$\delta_Y(n, \mathbf{x}) \triangleq \begin{cases} \exp(nK_Y(\theta(\mathbf{x})) - \theta(\mathbf{x})^\top \mathbf{x}) \min\left(1, \frac{c^{(k)} \xi_Y(\theta(\mathbf{x}))}{\sqrt{n}}\right) & \text{if } \mathbf{x} \notin \mathcal{E}_{\mathbf{X}_n} \\ \sum_{i=1}^k \exp(nK_Y(\theta_i(\mathbf{x})) - \theta_i^\top(\mathbf{x}) \mathbf{x}) \min\left(1, \frac{c^{(k)} \xi_Y(\theta_i(\mathbf{x}))}{\sqrt{n}}\right) & \text{otherwise,} \end{cases}$$

where the vector  $\theta(\mathbf{x})$  is defined (34); and for all  $i \in \{1, 2, \dots, k\}$ , the vector  $\theta_i(\mathbf{x})$  satisfies

$$\theta_i(\mathbf{x}) = \begin{cases} \tau_i(\mathbf{x}) & \text{if } \tau_i(\mathbf{x}) \in \Theta_Y^i \\ \tau_i(\mathbf{x}) + \epsilon & \text{otherwise,} \end{cases} \quad (44)$$

where

$$\tau_i(\mathbf{x}) = \arg \min_{\mathbf{t} \in \text{clo} \Theta_Y^i} (nK_Y(\mathbf{t}) - \mathbf{t}^\top \mathbf{x}); \quad (45)$$

$\epsilon \in \mathbb{R}^k$  is chosen such that two conditions are simultaneously met: First,  $\|\epsilon\| < r$ , with  $r > 0$  arbitrary small; and second,  $\theta_i(\mathbf{x}) \in \Theta_Y^i$ .

Using this notation, the following theorem summarizes the discussion above.

*Theorem 6:* For all  $\mathbf{x} \in \mathbb{R}^k$ , it holds that

$$|F_{\mathbf{X}_n}(\mathbf{x}) - \zeta_Y(n, \mathbf{x})| \leq \delta_Y(n, \mathbf{x}). \quad (46)$$

*Proof:* A detailed proof of Theorem 6 is presented in [17]. ■

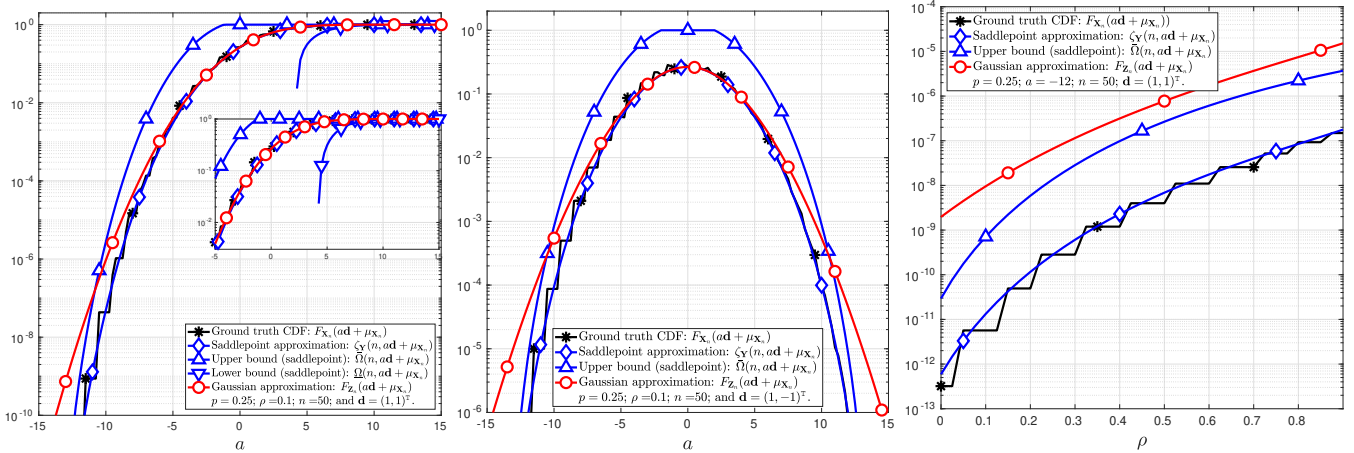


Fig. 1: Sum of the  $n = 50$  independent random vectors  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  such that for all  $i \in \{1, 2, \dots, n\}$ ,  $\mathbf{Y}_i$  satisfies (50).

An immediate result from Theorem 6 is the following upper and lower bounds on  $F_{\mathbf{X}_n}(\mathbf{x})$ , for all  $\mathbf{x} \in \mathbb{R}^k$ ,

$$\underline{\Omega}(n, \mathbf{x}) \leq F_{\mathbf{X}_n}(\mathbf{x}) \leq \bar{\Omega}(n, \mathbf{x}), \quad (47)$$

where,

$$\bar{\Omega}(n, \mathbf{x}) \triangleq \zeta_{\mathbf{Y}}(n, \mathbf{x}) + \delta_{\mathbf{Y}}(n, \mathbf{x}), \quad \text{and} \quad (48)$$

$$\underline{\Omega}(n, \mathbf{x}) \triangleq \zeta_{\mathbf{Y}}(n, \mathbf{x}) - \delta_{\mathbf{Y}}(n, \mathbf{x}). \quad (49)$$

### C. Connexion with the Saddlepoint Approximation

In [17], it is shown that for all  $\mathbf{x} \in \mathcal{D}$ , the saddlepoint approximation  $\hat{F}_{\mathbf{X}_n}(\mathbf{x})$  in (10) is identical to  $\zeta_{\mathbf{Y}}(n, \mathbf{x})$  in (43). That is, the saddlepoint approximation  $\hat{F}_{\mathbf{X}_n}$  can be obtained from Theorem 6 in the special case in which  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  in (4) are either lattice or absolutely continuous random vectors. Indeed, for all  $\mathbf{x} \in \mathcal{D}$ , the parameter  $\tau(\mathbf{x})$  in (34) is the solution in  $\tau$  to (8). Thus, for all  $\mathbf{x} \in \mathcal{D}$ ,  $\theta(\mathbf{x}) = \tau(\mathbf{x}) = \tau_0$ , with  $\tau_0$  the saddlepoint in (9).

## VI. EXAMPLES

Consider the case in which the independent random vectors  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$  in (4), with  $n = 50$ , are such that for all  $i \in \{1, 2, \dots, n\}$ ,

$$\mathbf{Y}_i \triangleq \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad (50)$$

where  $\rho \in [0, 1)$  is the Pearson correlation coefficient between the components of  $\mathbf{Y}_i$ ; and both  $B_1$  and  $B_2$  are independent Bernoulli random variables with parameter  $p = 0.25$ . The mean of  $\mathbf{X}_n$  in (4) is  $\mu_{\mathbf{X}_n} = np \begin{pmatrix} 1, \rho + \sqrt{1-\rho^2} \end{pmatrix}^\top$ .

Figure 1 depicts the CDF  $F_{\mathbf{X}_n}$  of  $\mathbf{X}_n$  in (4); the Gaussian approximation  $F_{\mathbf{Z}_n}$  in (16); the saddlepoint approximation  $\zeta_{\mathbf{Y}}$  in (43); and the saddlepoint upper and lower bounds  $\bar{\Omega}$  in (48) and  $\underline{\Omega}$  in (49); through the line  $a\mathbf{d} + \mu_{\mathbf{X}_n}$ . The plots on the left and the center are respectively for fixed vectors  $\mathbf{d} = (1, 1)^\top$  and  $\mathbf{d} = (1, -1)^\top$ , as a function of  $a$ . The plot on the right is function of  $\rho$  for a fixed point in the line  $a\mathbf{d} + \mu_{\mathbf{X}_n}$ , with  $a = -12$  and  $\mathbf{d} = (1, 1)^\top$ , i.e., tail of the distribution in the direction of the vector  $\mathbf{d} = (1, 1)^\top$ . Note that Gaussian and

saddlepoint approximations are particularly precise near to the mean  $\mu_{\mathbf{X}_n}$ . That is, when  $a = 0$ . Nonetheless, away from the mean, i.e.,  $a < -10$ , the Gaussian approximation induces a large approximation error, in sharp contrast to the saddlepoint approximation.

For  $n = 50$ , the lower bound  $\underline{\Omega}$  is negative, except when  $a > 5$ . Alternatively, the Gaussian upper and lower bounds  $\bar{\Sigma}$  in (16a) and  $\underline{\Sigma}$  in (16b) are trivial. That is, the lower bound is negative and the upper bound is bigger than one, which highlights the lack of formal mathematical arguments to evaluate the Gaussian approximation. For instance, note that when  $a < -10$ , the Gaussian approximation is bigger than the upper bound due to the saddle point approximation. In particular, note that Figure 1 (Right) highlights the fact that the same observation holds for all values of  $\rho$ .

## VII. FINAL REMARKS AND DISCUSSION

Theorem 6 holds for all random vectors for which the CGF exists. Under this condition, the Multivariate Berry-Esseen theorem in [15, Theorem 1.1], and the saddlepoint approximation in (9) are special cases of Lemma 3 for the choice  $\theta = \mathbf{0}$  and  $\theta = \tau_0$ , with  $\tau_0$  in (9), respectively.

The advantages of approximating the probability of a convex set  $\mathcal{A}$  in  $\mathcal{B}(\mathbb{R}^k)$  by using Lemma 3 instead of Theorem 1 are twofold. First, the proposed upper bound on the approximation error depends on the convex set  $\mathcal{A}$ . Second, both the approximation and the upper bound on the approximation error are parametrized by  $\theta \in \Theta_{\mathbf{Y}}$ , with  $\Theta_{\mathbf{Y}}$  in (22). Thus, the vector  $\theta$  in (27) can be tuned to minimize the upper bound on the error induced in (27). Nonetheless, such optimization is not trivial. In this work, a nonnecessarily optimal choice has been made for obtaining Theorem 6. Hence, it may be possible that tighter upper bounds can be obtained from Lemma 3 with another choices of  $\theta$ .

In the case in which  $k = 1$ , Lemma 3 leads to the same approximation as in [19, Theorem 2]. Nonetheless, in this case, the upper bound provided by [19, Theorem 2] is better than the one provided by Lemma 3.

## REFERENCES

- [1] V. Y. F. Tan and O. Kosut, "On the dispersions of three network information theory problems," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 881–903, Feb. 2014.
- [2] E. MolavianJazi and J. N. Laneman, "A random coding approach to Gaussian multiple access channels with finite blocklength," in *Proc. of the 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Illinois, USA, Oct. 2012, pp. 286–293.
- [3] J. E. Kolassa, "Multivariate saddlepoint tail probability approximations," *The Annals of Statistics*, vol. 31, no. 1, pp. 274–286, Feb. 2003.
- [4] J. Kolassa and J. Li, "Multivariate saddlepoint approximations in tail probability and conditional inference," *Bernoulli*, vol. 16, no. 4, pp. 1191–1207, Nov. 2010.
- [5] J. L. Jensen, *Saddlepoint Approximations*. New York, NY, USA: Clarendon press - Oxford, 1995.
- [6] R. Butler, *Saddlepoint approximations with applications*. Cambridge, NY, USA: Cambridge University Press, 2007.
- [7] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. New York, NY, USA: John Wiley and Sons, 1971, vol. 2.
- [8] A. Martinez, J. Scarlett, M. Dalai, and A. G. i Fàbregas, "A complex-integration approach to the saddlepoint approximation for random-coding bounds," in *Proc. of the 11th International Symposium on Wireless Communications Systems (ISWCS)*, Barcelona, Spain, Aug. 2014, pp. 618–621.
- [9] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6854–6883, Dec. 2016.
- [10] J. Font-Segura, G. Vazquez-Vilar, A. Martinez, A. G. i Fàbregas, and A. Lancho, "Saddlepoint approximations of lower and upper bounds to the error probability in channel coding," in *Proc. of the 52nd Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 2018, pp. 1–6.
- [11] G. Vazquez-Vilar, A. G. i Fàbregas, T. Koch, and A. Lancho, "Saddlepoint approximation of the error probability of binary hypothesis testing," in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Vail, CO, USA, 2018, pp. 2306–2310.
- [12] J. Font-Segura, A. Martinez, and A. G. i Fàbregas, "Saddlepoint approximation of the cost-constrained random coding error probability," in *Proc. of the IEEE Information Theory Workshop (ITW)*, Guangzhou, China, Jan. 2018, pp. 1–5.
- [13] J. Font-Segura, A. Martinez, and A. G. i Fàbregas, "Asymptotics of the random coding error probability for constant-composition codes," in *Proc. of the IEEE International Symposium on Information Theory (ISIT)*, Paris, France, Jul 2019, pp. 2947–2951.
- [14] A. Lancho, J. Ostman, G. Durisi, T. Koch, and G. Vazquez-Vilar, "Saddlepoint approximations for short-packet wireless communications," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4831–4846, Jul. 2020.
- [15] M. Raič, "A multivariate Berry-Esseen theorem with explicit constants," *Bernoulli*, vol. 25, no. 4A, pp. 2824–2853, Nov. 2019.
- [16] R. J. Beerends, *Fourier and Laplace transforms*. Cambridge, NY, USA: Cambridge University Press, 2003.
- [17] D. Anade, J.-M. Gorce, P. Mary, and S. M. Perlaza, "Saddlepoint approximations of cumulative distribution functions of sums of random vectors," INRIA Grenoble - Rhône-Alpes, Tech. Rep. 9388, feb. 2021.
- [18] O. E. Nielsen, *Information and exponential families: in statistical theory*. Chichester U.K., NY, USA: John Wiley & Sons, 2014.
- [19] D. Anade, J.-M. Gorce, P. Mary, and S. M. Perlaza, "An upper bound on the error induced by saddlepoint approximations - Applications to information theory," *Entropy*, vol. 22, no. 6, p. 690, Jun. 2020.