

ON THE ASYMPTOTICS OF ROSENBLATT-TYPE
TRANSFORMATIONS IN A GAUSSIAN MIXTURE
IDENTIFICATION PROBLEME.A. SAVINOV 

PRESENTED BY N.S. ARKASHOV

Abstract: We show that a cross independence (CI) transformation of some Gaussian mixture has an asymptotically Gaussian distribution connected with the Gaussian core of the mixture by the same type of transform. We suggest using this fact for testing the fit of high-dimensional samples to a mixture of Gaussian distributions. In addition we study a behavior of extreme values in related triangular arrays.

Keywords: Gaussian mixture, multivariate copula, multivariate t distribution, extreme values, mixture identification

1 Introduction

Consider a random vector $X_n = (X_1, X_2, \dots, X_n)$ with an absolutely continuous distribution function $F(x_1, \dots, x_n)$. Denote $f_{1\dots n}(x_1, \dots, x_n)$ the

probability density function of the vector X_n , and

$$F_{i|1\ldots\hat{i}\ldots n}(x_i|x_1, \ldots, \hat{x}_i, \ldots, x_n) = \frac{\int_{-\infty}^{x_i} f_{1\ldots n}(x_1, \ldots, u_i, \ldots, x_n) du_i}{\int_{-\infty}^{+\infty} f_{1\ldots n}(x_1, \ldots, u_i, \ldots, x_n) du_i}$$

denote the conditional distribution function of the random variable X_i given all others (where the symbol $\hat{}$ indicates the omission of the corresponding element). In his work [1], Murray Rosenblatt suggested a transformation using conditional distribution functions

$$\begin{aligned} Y_1 &= F_1(X_1) \\ Y_2 &= F_{2|1}(X_2|X_1) \\ &\vdots \\ Y_n &= F_{n|1,2,\ldots,n-1}(X_n|X_1, \ldots, X_{n-1}), \end{aligned} \tag{1}$$

which, in this case, results in a uniform distribution of the vector $Y = (Y_1, Y_2, \ldots, Y_n)$ on the unit cube and, consequently, in the independence of the random variables Y_i .

A similar transformation was specifically considered by S.Ya. Shatskikh, particularly in [2]

$$\begin{aligned} X_1^* &= F_1^{-1} [F_{1|2,3,\ldots,n}(X_1|X_2, X_3, \ldots, X_n)] \\ X_2^* &= F_2^{-1} [F_{2|1,3,\ldots,n}(X_2|X_1, X_3, \ldots, X_n)] \\ &\vdots \\ X_n^* &= F_n^{-1} [F_{n|1,2,\ldots,n-1}(X_n|X_1, X_2, \ldots, X_{n-1})]. \end{aligned} \tag{2}$$

Both the original system (vector) of random variables $\{X_i\}$ and the new system $\{X_i^*\}$ consist, generally speaking, of dependent random variables. However, both systems can be called 'cross-independent' in a certain sense, since it is not hard to show (similar to [1]) that for each i , the random variable X_i^* is independent of the system $\{X_1, \ldots, \hat{X}_i, \ldots, X_n\}$, which includes all original random variables except the one with the same index (the symbol $\hat{}$ again indicates the omission of the corresponding element).

Recall one of the definitions of the so-called copula: a multivariate distribution function with univariate marginal distributions that are uniform on the segment $[0, 1]$. According to Sklar's theorem (see, for example, [3]), for any multivariate distribution function, there exists a copula $C(u_1, \ldots, u_n)$ connecting it with its univariate marginal distributions:

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

Thus, a separation occurs between univariate distributions and the dependence structure characterized by the copula.

In [4] it was shown, in particular, that the transformation (2) (without applying inverse functions) applied to any absolutely continuous vector with

a dependence structure described by the copula C produces a new copula that depends only on C and, thus, can be regarded as a copula transformation.

In this paper, it is shown how this transformation can be utilized in the problem of fitting a sample to a mixture of distributions that have a Gaussian structure.

Let's outline the problem formulation. Consider the following general problem of identifying mixture components. Consider a system of random variables $X_1, X_2, \dots, X_n$, modeling the lifetimes of n different components of a complex system operating in a random environment. We assume that, given the state of the environment (t), the components are dependent, have different characteristics, and the joint distribution function of their lifetimes is defined as $G_n^{(t)}(x_1, x_2, \dots, x_n)$. Thus, the lifetimes of the components operating in random environments are described by a mixture

$$F_n(x_1, x_2, \dots, x_n) = \int_{-\infty}^{\infty} G_n^{(t)}(x_1, x_2, \dots, x_n) \mu(dt), \quad (3)$$

where the environment parameter is assumed to be univariate and the measure μ describes its probabilistic behavior.

Among the studies focused on estimating the components of such continuous (scale) mixtures, works on estimating the weight distribution can primarily be highlighted (see, for example, [5], [6], [7]), as well as studies on the parameters of mixture components in the case of conditionally independent random variables ([8]).

In this paper, we will be interested in the question of whether something can be inferred about the distribution $G_n^{(t)}$ from a sample drawn from the distribution F_n , without any knowledge of μ . Theorem 1 in Section 3 of this paper establishes that the CI-transformation allows for the identification of the Gaussian structure (type of dependence) of the functions $G_n^{(t)}$ in the mixture (3), provided that the latter is one of the variants of the multivariate Student's t distribution (Kshirsagar's Multivariate t Distribution, see [9], p. 87) with r degrees of freedom. Furthermore, the modeling experiments described in Section 4 support the assumption that the Gaussian structure of components can be identified for a broader class of Gaussian mixtures of the type (3). At the same time, there exist samples from multivariate distributions whose CI-transformations do not lead to a Gaussian structure through the described procedure, which negates their extraction from any Gaussian mixture.

In Section 5, we investigate the behavior of the maximum of Gaussian random variables in a triangular array, where the joint distributions of each row are not Gaussian but are generated by a mixture of Gaussian distributions whose copulas have undergone the CI-transformation. The results of this section are not directly related to the procedure for determining the Gaussianity of the mixture. Here, we raise the question of whether, given the conclusions of Theorem 1, the asymptotically Gaussian random vectors obtained

in the rows of the triangular array will behave like Gaussian vectors in terms of the asymptotic behavior of the maxima of their components. In the case of strong dependence considered here, it has been shown that the behavior of the maxima in such triangular arrays is analogous to that in Gaussian schemes. Note that a similar problem for the case of weak dependence was addressed in [10].

2 Basic concepts.

We introduce the concept of CI-transformation of a copula.

Consider a random vector X_n with an absolutely continuous distribution function given by

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

on the probability space $\{\Omega, \mathfrak{B}, \mathbf{P}\}$. Here, $F_i(x_i)$ are the marginal distributions and C is the copula. We consider the random variables

$$X_{i,n}^* = F_{i|1\dots\hat{i}\dots n}(X_i|X_1, \dots, \widehat{X_i}, \dots, X_n). \quad (4)$$

As shown in [4], their joint distribution function is a copula, and the following definition is valid.

Definition 1. We denote the mapping $C \mapsto C^{ci}$ as

$$C^{ci}(u) = \mathbf{P}\{X_{1,n}^* \leq u_1, \dots, X_{n,n}^* \leq u_n\}, \quad u = (u_1, \dots, u_n),$$

where $\mathbf{X}_n = (X_1, \dots, X_n)$ is an absolutely continuous vector with copula C . We will refer to this transformation as the CI-transformation (Cross-Independence) of the absolutely continuous copula C . The copula $C^{ci}(u)$ will be called the CI-copula or the CI-image of the copula C .

Next, we introduce a measure on the Hilbert space solely for the purpose of obtaining a consistent family of copulas through the projections of the measure onto a certain orthonormal basis.

Consider a measurable space $\{\mathbb{H}, \mathcal{B}(\mathbb{H})\}$, where $\mathbb{H}$ a real separable Hilbert space with a countable orthonormal basis $\{e_i\}_{i=1}^\infty$, Borel σ -algebra and inner product $\langle \cdot, \cdot \rangle$. We will consider a countably additive measure μ on it with a characteristic functional

$$\Psi_\mu(y) = \int_0^\infty \exp\left\{-\frac{t}{2}\langle By, y \rangle\right\} g_r(t) dt, \quad y \in \mathbb{H}, \quad (5)$$

where B is a linear self-adjoint positive definite nuclear operator with eigenvectors $\{e_i\}_{i=1}^\infty$ and

$$g_r(t) = \frac{r^{r/2}}{2^{r/2}\Gamma(r/2)} t^{-r/2-1} \exp\left\{-\frac{r}{2t}\right\}, \quad t > 0.$$

It should be noted that this measure can be referred to as the Student's measure on $\{\mathbb{H}, \mathcal{B}(\mathbb{H})\}$ with r degrees of freedom. Indeed, let $\{f_k\}$ be an arbitrary orthonormal basis in the Hilbert space $\mathbb{H}$, and consider the random

variables $X_i = \langle \cdot, f_i \rangle$ on the probability space $\{\mathbb{H}, \mathcal{B}(\mathbb{H}), \mu\}$. It is clear that the corresponding projections of the measure μ (distributions of the random vectors $\mathbf{X}_n = (X_1, \dots, X_n)$) have characteristic functions given by

$$\psi_{\mu_n}(y_1, \dots, y_n) = \int_0^\infty \exp \left\{ -\frac{t}{2} \sum_{i,j=1}^n y_i y_j \langle B f_i, f_j \rangle \right\} g_r(t) dt. \quad (6)$$

We will show that the characteristic function (6) corresponds to one of the variants of the Kshirsagar's Multivariate t distribution (see [9], p. 87).

Lemma 1. *Assume that the vector Y has an n -variate Gaussian distribution with zero mean and covariance matrix $C = (\langle B f_i, f_j \rangle)$, and that S_r^2 is a random variable with a $\chi^2(r)$ distribution, independent of Y ($r \in \mathbb{N}$). Then the random vector $T = Y/\sqrt{S^2/r}$ has the characteristic function given by (6).*

Proof. The characteristic function of the random vector T is given by

$$\begin{aligned} \varphi_T(y) &= \mathbb{E} \exp \left\{ i \frac{\langle Y, y \rangle}{\sqrt{S_r^2/r}} \right\} = \mathbb{E} \left[\mathbb{E} \left(\exp \left\{ i \frac{\langle Y, y \rangle}{\sqrt{S_r^2/r}} \right\} \middle| \frac{1}{S_r^2/r} \right) \right] \\ &= \mathbb{E} \left[\mathbb{E} \left(\exp \left\{ i \langle Y, y \rangle \sqrt{U} \right\} \middle| U \right) \right], \end{aligned}$$

where $U \sim \text{inv}\Gamma\left(\frac{r}{2}, \frac{r}{2}\right)$ (the inverse gamma distribution with parameters $(r/2, r/2)$). Then

$$\begin{aligned} \varphi_T(y) &= \int_0^\infty \mathbb{E} \left(\exp \left\{ i \langle Y, y \sqrt{t} \rangle \right\} \right) g_r(t) dt = \int_0^\infty \varphi_Y(y \sqrt{t}) g_r(t) dt \\ &= \int_0^\infty \exp \left\{ -\frac{t}{2} \langle C y, y \rangle \right\} g_r(t) dt. \end{aligned}$$

□

Of course, this representation of the Student's t distribution as a mixture of Gaussians is well known, and its dependence structure (as well as that of other normal mixtures in the case of dimension two) has been studied using copulas, for example, in [11].

3 Convergence of CI-copulas

Let's introduce some auxiliary notations.

Let $B_n = \pi_n B \pi_n$, where π_n is the orthogonal projector $\mathbb{H} \rightarrow \mathbb{H}_n = \text{span}\{f_1, \dots, f_n\}$.

We denote by μ^{tB} (for $t = 1$, simply μ^B) the Gaussian measure on $\mathbb{H}, \mathcal{B}(\mathbb{H})$ with the characteristic functional

$$\Psi_{\mu^{tB}}(y) = \exp \left\{ -\frac{t}{2} \langle B y, y \rangle \right\}, \quad y \in \mathbb{H}.$$

Its projections onto $\mathbb{H}_n$ have characteristic functions

$$\psi_{\mu_n^{tB}}(y_1, \dots, y_n) = \exp \left\{ -\frac{t}{2} \sum_{i,j=1}^n y_i y_j \langle B_n f_i, f_j \rangle \right\}.$$

Let us denote the corresponding correlation matrix (it does not depend on t) $R_n = (r_{i,j}^{(n)})$, where

$$r_{i,j}^{(n)} = \frac{\langle B_n f_i, f_j \rangle}{[\langle B_n f_i, f_i \rangle \langle B_n f_j, f_j \rangle]^{1/2}}, \quad 1 \leq i, j \leq n.$$

Let us introduce notations for the following quadratic forms:

$$\begin{aligned} s_n^2 &:= s_n^2(h) = \frac{1}{n} \langle B_n^{-1} \pi_n h, \pi_n h \rangle, \\ s_\infty^2 &:= s_\infty^2(h) = \lim_{n \rightarrow \infty} s_n^2(h), \end{aligned}$$

and for the random variables:

$$\zeta_{i,n} := \zeta_{i,n}(h) = \frac{\langle B_n^{-1} f_i, h \rangle}{s_\infty \langle B_n^{-1} f_i, f_i \rangle^{1/2}}, \quad i = 1, \dots, n, \quad (7)$$

Lemma 2. *With respect to the measure μ , the random variables $\{\zeta_{i,n}\}_{i=1}^n$ are jointly Gaussian with covariances (which obviously coincide with the correlation coefficients) given by*

$$c_{ij}^{(n)} := \text{cov}(\zeta_{i,n}, \zeta_{j,n}) = \frac{\langle B_n^{-1} f_i, f_j \rangle}{[\langle B_n^{-1} f_i, f_i \rangle \langle B_n^{-1} f_j, f_j \rangle]^{1/2}}. \quad (8)$$

Proof. Indeed, let us introduce the random variables

$$\tilde{\zeta}_{i,n} := \zeta_{i,n} s_\infty = \frac{\langle B_n^{-1} f_i, h \rangle}{\langle B_n^{-1} f_i, f_i \rangle^{1/2}}, \quad i = 1..n$$

and consider the distribution function of the random vector $\zeta_{\cdot,n}$ with respect to the measure μ

$$\mu \{ \zeta_{1,n} \leq u_1, \dots, \zeta_{n,n} \leq u_n \} = \int_0^\infty \mu^{tB} \{ \zeta_{1,n} \leq u_1, \dots, \zeta_{n,n} \leq u_n \} g_r(t) dt.$$

Moreover (see the proof of Lemma 7 in [2]), we have $\mu^{tB} \{ s_\infty^2(h) = t \} = 1$, then

$$\begin{aligned} &\mu \{ \zeta_{1,n} \leq u_1, \dots, \zeta_{n,n} \leq u_n \} \\ &= \int_0^\infty \mu^{tB} \left\{ \tilde{\zeta}_{1,n}/\sqrt{t} \leq u_1, \dots, \tilde{\zeta}_{n,n}/\sqrt{t} \leq u_n \right\} g_r(t) dt. \end{aligned}$$

Note that $\tilde{\zeta}_{i,n}$ are linear continuous functionals on $\mathbb{H}$, and

$$\mu^{tB} \left\{ \tilde{\zeta}_{1,n}/\sqrt{t} \leq u_1, \dots, \tilde{\zeta}_{n,n}/\sqrt{t} \leq u_n \right\} = \mu^B \left\{ \tilde{\zeta}_{1,n} \leq u_1, \dots, \tilde{\zeta}_{n,n} \leq u_n \right\},$$

since the corresponding characteristic functions are equal:

$$\begin{aligned} & \exp \left\{ - \left\langle tB \sum_{i=1}^n y_i \tilde{\zeta}_{i,n} / \sqrt{t}, \sum_{j=1}^n y_j \tilde{\zeta}_{j,n} / \sqrt{t} \right\rangle \right\} \\ &= \exp \left\{ - \left\langle B \sum_{i=1}^n y_i \tilde{\zeta}_{i,n}, \sum_{j=1}^n y_j \tilde{\zeta}_{j,n} \right\rangle \right\}. \end{aligned}$$

Thus,

$$\mu \{ \zeta_{1,n} \leq u_1, \dots, \zeta_{n,n} \leq u_n \} = \mu^B \{ \tilde{\zeta}_{1,n} \leq u_1, \dots, \tilde{\zeta}_{n,n} \leq u_n \}.$$

Since the vector $\tilde{\zeta}_{\cdot,n}$ is a vector of linear continuous functionals on $\mathbb{H}$ and is Gaussian with respect to the measure μ^B , the vector $\zeta_{\cdot,n}$ is also Gaussian with respect to the measure μ . Moreover,

$$\text{cov}(\zeta_{i,n}, \zeta_{j,n}) = \text{cov}_{\mu^B}(\tilde{\zeta}_{i,n}, \tilde{\zeta}_{j,n}) = \left\langle B \tilde{\zeta}_{i,n}, \tilde{\zeta}_{j,n} \right\rangle = \frac{\langle BB_n^{-1} f_i, B_n^{-1} f_j \rangle}{[\langle B_n^{-1} f_i, f_i \rangle \langle B_n^{-1} f_j, f_j \rangle]^{1/2}}. \quad (9)$$

Let's consider the numerator and establish the identity:

$$\langle BB_n^{-1} f_i, B_n^{-1} f_j \rangle = \langle B_n^{-1} f_i, f_j \rangle. \quad (10)$$

To this end, let us denote:

$$h := B_n^{-1} f_i \in \mathbb{H}_n, \quad g := B_n^{-1} f_j \in \mathbb{H}_n. \quad (11)$$

Then

$$\begin{aligned} f_j &= B_n g = \pi_n B g = \pi_n \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \langle g, e_{\tilde{j}} \rangle e_{\tilde{j}} = \pi_n \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \left[\sum_{\tilde{i}=1}^n \langle g, f_{\tilde{i}} \rangle \langle f_{\tilde{i}}, e_{\tilde{j}} \rangle \right] e_{\tilde{j}} \\ &= \sum_{\tilde{k}=1}^n \sum_{\tilde{i}=1}^n \langle g, f_{\tilde{i}} \rangle \left[\sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \langle f_{\tilde{i}}, e_{\tilde{j}} \rangle \langle e_{\tilde{j}}, f_{\tilde{k}} \rangle \right] f_{\tilde{k}}. \end{aligned}$$

Consequently,

$$\sum_{\tilde{i}=1}^n \langle g, f_{\tilde{i}} \rangle a_{\tilde{i}, \tilde{k}} = \delta_{j \tilde{k}} = \begin{cases} 1, & j = \tilde{k} \\ 0, & j \neq \tilde{k} \end{cases}, \quad \text{where} \quad a_{\tilde{i}, \tilde{k}} = \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \langle f_{\tilde{i}}, e_{\tilde{j}} \rangle \langle e_{\tilde{j}}, f_{\tilde{k}} \rangle. \quad (12)$$

Now, let's calculate

$$\begin{aligned} \langle Bh, g \rangle &= \left\langle \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \left[\sum_{\tilde{i}=1}^n \langle h, f_{\tilde{i}} \rangle \langle f_{\tilde{i}}, e_{\tilde{j}} \rangle \right] e_{\tilde{j}}, \sum_{m=1}^n \langle g, f_m \rangle f_m \right\rangle \\ &= \sum_{\tilde{i}=1}^n \sum_{m=1}^n \langle h, f_{\tilde{i}} \rangle \langle g, f_m \rangle a_{\tilde{i}, m} \end{aligned}$$

$$= \sum_{i=1}^n \langle h, f_i \rangle \left[\sum_{m=1}^n \langle g, f_m \rangle a_{i,m} \right] = \sum_{i=1}^n \langle h, f_i \rangle \delta_{i,j} = \langle h, f_j \rangle,$$

which is the identity (10). From (9) and (10), it follows (8). $\square$

Let $R_n^- = (c_{ij}^{(n)})$ be the matrix composed of the correlations given by (8).

Let us also introduce notations for the matrices $M_n = (\langle B_n f_i, f_i \rangle)$ and $M_n^- = (\langle B_n^{-1} f_i, f_i \rangle)$.

Since the Gaussian copula is completely defined by the correlation matrix, let us introduce notations for the corresponding Gaussian copulas: C_{R_n} and $C_{R_n^-}$.

Theorem 1. *Let C_n be the Student's t copulas of the distributions with characteristic functions (6). Let $C_{R_n^-}$ be the Gaussian copula defined above. Then for any $k \in \mathbb{N}$ and $(u_1, \dots, u_k) \in (0, 1)^k$, the convergence holds:*

$$C_n^{\text{ci}}(u_1, \dots, u_k, 1, \dots, 1) \rightarrow C_{R_k^-}(u_1, \dots, u_k), \quad n \rightarrow \infty.$$

Proof. Without loss of generality, we can assume C_n to be the copulas of the Student's vector family $\mathbf{X}_n = \{X_1, X_2, \dots, X_n\}$, $n = 1, 2, \dots$, defined on the measurable space $\{\mathbb{H}, \mathcal{B}(\mathbb{H})\}$, where $\mathbb{H}$ is a real separable Hilbert space with a countable orthonormal basis $\{f_i\}_{i=1}^\infty$, Borel σ -algebra, and inner product $\langle \cdot, \cdot \rangle$, a countably additive Student's measure μ with r degrees of freedom and characteristic functional (5), $X_i = \langle \cdot, f_i \rangle$.

Let's denote $x_i := \Phi^{-1}(u_i)$. Then

$$\begin{aligned} C_n^{\text{ci}}(u_1, \dots, u_k, 1, \dots, 1) &= C_n^{\text{ci}}(\Phi(x_1), \dots, \Phi(x_k), 1, \dots, 1) \\ &= \mu \{ \Phi^{-1}(X_{1,n}^*) \leq x_1, \dots, \Phi^{-1}(X_{k,n}^*) \leq x_k \}. \end{aligned} \quad (13)$$

It follows from item 3 of Lemma 3 in [10] that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu \{ \Phi^{-1}(X_{1,n}^*) \leq x_1, \dots, \Phi^{-1}(X_{k,n}^*) \leq x_k \} \\ = \lim_{n \rightarrow \infty} \mu \{ \zeta_{1,n} \leq x_1, \dots, \zeta_{k,n} \leq x_k \}. \end{aligned} \quad (14)$$

Note that for any $i \leq k < n$, the following holds:

$$B_n^{-1} f_i = B_k^{-1} f_i. \quad (15)$$

Indeed, let us set

$$h := B_n^{-1} f_i \in \mathbb{H}_n, \quad h' := B_k^{-1} f_i \in \mathbb{H}_k,$$

therefore,

$$\begin{aligned}
 f_i &= B_k h' = \pi_k B h' = \pi_k \sum_{\tilde{i}=1}^k \langle h', f_{\tilde{i}} \rangle B f_{\tilde{i}} = \sum_{\tilde{k}=1}^k \sum_{\tilde{i}=1}^k \langle h', f_{\tilde{i}} \rangle \langle B f_{\tilde{i}}, f_{\tilde{k}} \rangle f_{\tilde{k}} \Rightarrow \\
 &\sum_{\tilde{i}=1}^k \langle h', f_{\tilde{i}} \rangle \langle B f_{\tilde{i}}, f_{\tilde{k}} \rangle = \delta_{i\tilde{k}}, \\
 B_n h' &= \pi_n B h' = \sum_{\tilde{k}=1}^n \sum_{\tilde{i}=1}^k \langle h', f_{\tilde{i}} \rangle \langle B f_{\tilde{i}}, f_{\tilde{k}} \rangle f_{\tilde{k}} = \sum_{\tilde{k}=1}^n \delta_{i\tilde{k}} f_{\tilde{k}} = f_i = B_n h \Rightarrow \\
 B_n(h - h') &= 0 \Rightarrow h = h',
 \end{aligned}$$

which proves the equality (15).

It follows from (15) that $c_{ij}^{(n)} = c_{ij}^{(k)}$ for $i, j \leq k < n$, and, taking into account (7) and (8), we obtain

$$\begin{aligned}
 \mu \{ \zeta_{1,n} \leq x_1, \dots, \zeta_{k,n} \leq x_k \} &= \mu \{ \zeta_{1,k} \leq x_1, \dots, \zeta_{k,k} \leq x_k \} = \\
 &= C_{R_k^-}(\Phi(x_1), \dots, \Phi(x_k)) = C_{R_k^-}(u_1, \dots, u_k),
 \end{aligned} \tag{16}$$

and the proof of the theorem follows from equalities (13), (14), and (16). $\square$

Next, we will show that the limiting copula $C_{R_n^-}$ is connected to the Gaussian core of the mixture (6) by the CI-transformation. First, we will prove the following auxiliary result.

Lemma 3.

$$M_n^- = M_n^{-1}$$

Proof. Let us consider the element m_{ij} of the matrix $M_n^- \cdot M_n$

$$m_{ij} = \sum_{\ell=1}^n \langle B_n^{-1} f_i, f_\ell \rangle \langle B_n f_\ell, f_j \rangle.$$

Using the notations (11), one can show, analogous to (12), that

$$\sum_{\tilde{i}=1}^n \langle h, f_{\tilde{i}} \rangle a_{\tilde{i},k} = \delta_{i\tilde{k}}, \quad \text{where} \quad a_{\tilde{i},k} = \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \langle f_{\tilde{i}}, e_{\tilde{j}} \rangle \langle e_{\tilde{j}}, f_{\tilde{k}} \rangle. \tag{17}$$

Тогда

$$\begin{aligned}
 \langle B_n f_\ell, f_j \rangle &= \sum_{\tilde{j}=1}^{\infty} \lambda_{\tilde{j}}^2 \langle f_\ell, e_{\tilde{j}} \rangle \langle e_{\tilde{j}}, f_j \rangle = a_{\ell j}, \\
 m_{ij} &= \sum_{\ell=1}^n \langle h, f_\ell \rangle a_{\ell j} = \delta_{ij},
 \end{aligned}$$

which is what was to be proved. $\square$

Corollary 1 (from Theorem 1).

$$C_n^{ci}(u_1, \dots, u_k, 1, \dots, 1) \rightarrow C_{R_k}^{ci}(u_1, \dots, u_k), \quad n \rightarrow \infty.$$

Proof. An example of the CI-transformation of an n -dimensional Gaussian copula in [4] shows that for any covariance matrix M_k corresponding to a Gaussian distribution with copula C_{R_k} , the CI-transformation $C_{R_k}^{ci}$ of this copula coincides with the copula of the Gaussian distribution with covariance matrix M_k^{-1} . Therefore, by Lemma 3

$$C_{R_k}^{ci}(u_1, \dots, u_k) = C_{R_k^-}(u_1, \dots, u_k).$$

□

4 Modeling and testing the fit of a high-dimensional sample with a normal mixture

To test the fit of a certain n -dimensional sample (for sufficiently large n) with the family of distributions (3), a procedure is proposed, consisting of sequential CI-transformation, normalization (in the sense of bringing the coordinates of the elements of the obtained sample to a standard Gaussian distribution), reduction (projecting the sample onto a lower dimension), and checking the resulting sample for compliance with a multivariate normal distribution.

For participation in the experiment, samples of dimension $n = 20$ and 40 with a volume of $N = 200$ were simulated from the following six distributions: the Kshirsagar's Multivariate t distribution of the form (6) with non-diagonal matrices $C = (\langle Bf_i, f_j \rangle)$ with $r = 1, 2, 4$ and 8 degrees of freedom, a Gaussian mixture of the form (3), where the measure μ is defined by an exponential distribution with density e^{-t} , $t > 0$ (thus yielding one of the variants of the multivariate Laplace distribution), and a multivariate distribution where the univariate Gaussian components are linked by D-vine copulas (see, for example, [12], [13], [14]), in which bivariate conditional copulas were chosen to be bivariate Gaussian and Clayton copulas (with different parameters).

In calculating the CI-transformation for estimates of conditional distribution functions, an approach that utilizes dimensionality reduction was employed (see [15]).

To check for multivariate normality, the criterion described in [16] was used, with a significance level of $\alpha = 0.05$. All twelve samples collected in Tables 1 and 2 were first transformed to one-dimensional marginal normal distributions and exhibit the corresponding dependency structure indicated by the specified distribution (for example, $\tilde{t}(8)$ is a sample from the distribution with the Student's $t(8)$ copula and marginal $N(0, 1)$, while $ci \tilde{t}(8)$ represents the distribution with the CI-image of the Student's t copula and the same marginal normal distributions). Thus, the multivariate normality test effectively checks for the presence of a Gaussian dependency structure.

The cells in the tables contain p-values and the results of normality checks (True/False) for the projections of the corresponding dimensions indicated in the column headings.

As can be seen, the projections of the original samples of any dimension demonstrated a deviation from the Gaussian structure. At the same time, the projections of low-dimensional (2-3) CI-images showed agreement with the Gaussian distribution for all Gaussian mixtures, including the case with exponential mixing (Laplace), while there was a predominant deviation from the Gaussian structure for the projections of the distribution based on the CI-image of the D-vine copula.

Figures 1–11 show two-dimensional projections of kernel density estimations for all pairs of distributions, before and after the ci-transformation, which are listed in Tables 1 and 2.

Table 1.

n=200 d=20	2	3	5	8
$\tilde{t}(8)$	3.0e-01 (True)	1.8e-01 (True)	2.5e-02 (False)	2.0e-11 (False)
$\tilde{t}(4)$	9.0e-02 (True)	1.1e-02 (False)	3.9e-09 (False)	2.2e-46 (False)
$\tilde{t}(2)$	4.1e-03 (False)	1.8e-13 (False)	3.7e-69 (False)	1.8e-299 (False)
$\tilde{t}(1)$	3.3e-09 (False)	4.8e-15 (False)	5.2e-95 (False)	0.0e+00 (False)
<i>Laplace</i>	8.6e-04 (False)	8.1e-07 (False)	2.2e-26 (False)	0.0e+00 (False)
<i>vine</i>	1.3e-04 (False)	6.8e-05 (False)	1.5e-29 (False)	2.4e-178 (False)
<i>ci</i> $\tilde{t}(8)$	9.0e-01 (True)	8.2e-01 (True)	5.2e-01 (True)	1.2e-01 (True)
<i>ci</i> $\tilde{t}(4)$	6.9e-02 (True)	1.9e-01 (True)	1.9e-01 (True)	2.9e-02 (False)
<i>ci</i> $\tilde{t}(2)$	9.3e-01 (True)	1.9e-01 (True)	2.2e-01 (True)	7.4e-02 (True)
<i>ci</i> $\tilde{t}(1)$	1.5e-01 (True)	3.6e-01 (True)	1.3e-02 (False)	9.4e-05 (False)
<i>ci</i> <i>Laplace</i>	8.2e-01 (True)	5.3e-01 (True)	4.1e-02 (False)	1.7e-23 (False)
<i>ci</i> <i>vine</i>	2.5e-02 (False)	4.8e-01 (True)	2.4e-02 (False)	4.7e-13 (False)

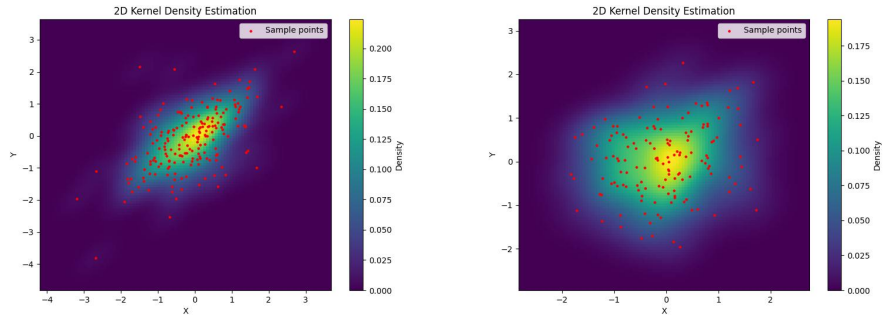


FIG. 1. 2D projected density estimation for $\tilde{t}(8)$ (left) and $ci \tilde{t}(8)$ (right) for base dimension 20

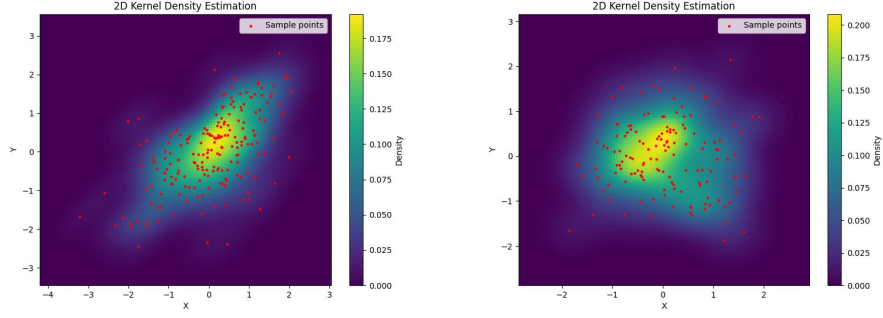


FIG. 2. 2D projected density estimation for $\tilde{t}(4)$ (left) and $ci \tilde{t}(4)$ (right) for base dimension 20

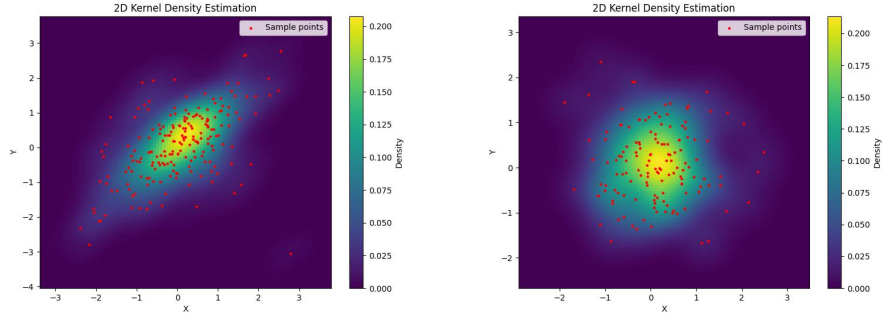


FIG. 3. 2D projected density estimation for $\tilde{t}(2)$ (left) and $ci \tilde{t}(2)$ (right) for base dimension 20

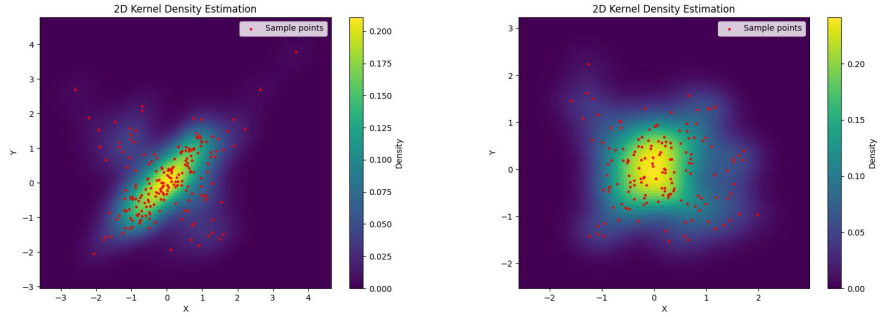


FIG. 4. 2D projected density estimation for $\tilde{t}(1)$ (left) and $ci \tilde{t}(1)$ (right) for base dimension 20

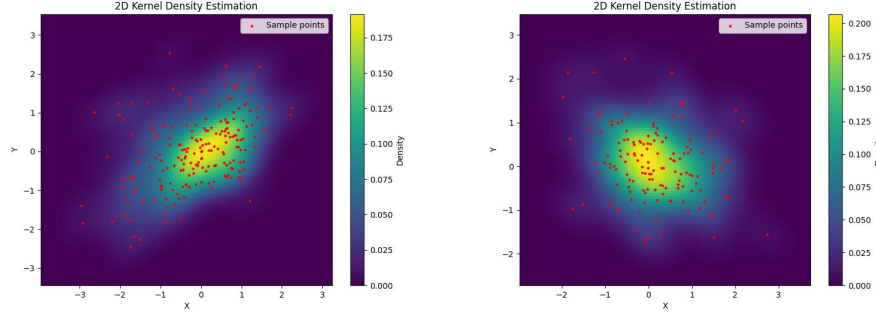


FIG. 5. 2D projected density estimation for $\widetilde{Laplace}$ (left) and $ci \widetilde{Laplace}$ (right) for base dimension 20

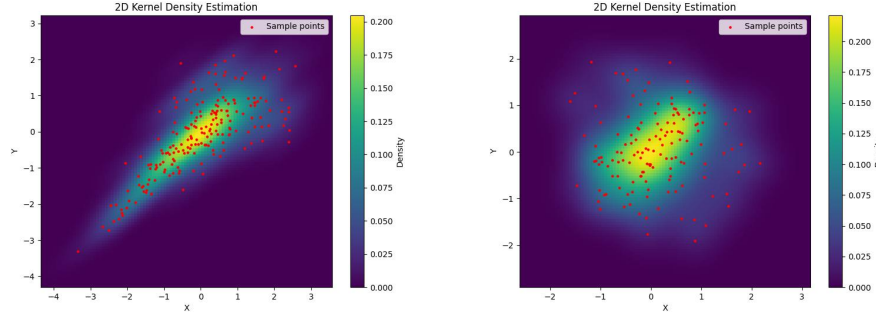


FIG. 6. 2D projected density estimation for $vine$ (left) and $ci \ vine$ (right) for base dimension 20

Table 2.

n=200 d=40	2	3	5	8
$\hat{t}(8)$	1.3e-01 (True)	1.3e-01 (True)	3.8e-01 (True)	1.2e-01 (True)
$\hat{t}(4)$	1.7e-01 (True)	3.2e-02 (False)	7.7e-05 (False)	1.6e-13 (False)
$\hat{t}(2)$	2.6e-01 (True)	1.2e-05 (False)	4.0e-19 (False)	6.9e-157 (False)
$\hat{t}(1)$	1.6e-05 (False)	5.2e-15 (False)	8.8e-93 (False)	0.0e+00 (False)
$\widetilde{Laplace}$	6.1e-03 (False)	6.2e-07 (False)	3.0e-35 (False)	0.0e+00 (False)
$ci \ \hat{t}(8)$	9.3e-01 (True)	9.7e-01 (True)	7.2e-01 (True)	4.9e-01 (True)
$ci \ \hat{t}(4)$	9.8e-01 (True)	9.5e-01 (True)	9.3e-01 (True)	5.8e-01 (True)
$ci \ \hat{t}(2)$	7.3e-01 (True)	9.0e-01 (True)	6.7e-01 (True)	7.6e-01 (True)
$ci \ \hat{t}(1)$	7.0e-01 (True)	1.1e-01 (True)	1.2e-01 (True)	3.3e-01 (True)
$ci \ \widetilde{Laplace}$	8.2e-02 (True)	7.9e-02 (True)	6.6e-10 (False)	6.9e-48 (False)

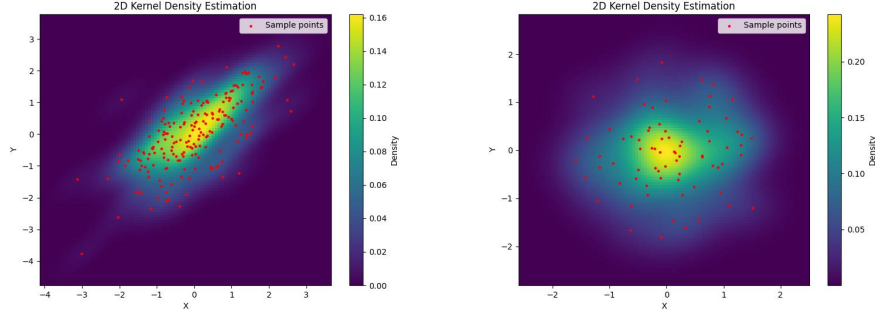


FIG. 7. 2D projected density estimation for $\tilde{t}(8)$ (left) and $ci \tilde{t}(8)$ (right) for base dimension 40

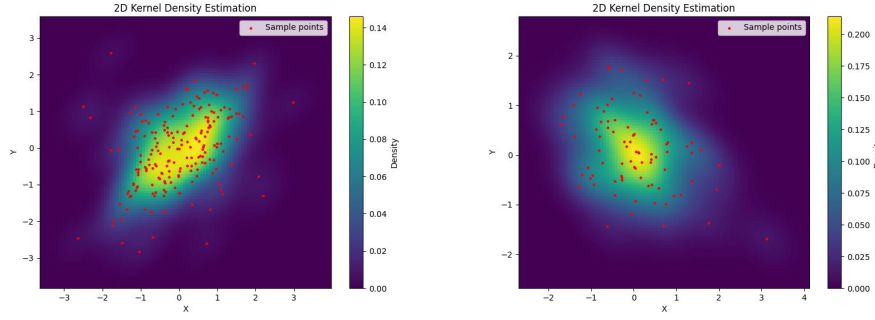


FIG. 8. 2D projected density estimation for $\tilde{t}(4)$ (left) and $ci \tilde{t}(4)$ (right) for base dimension 40

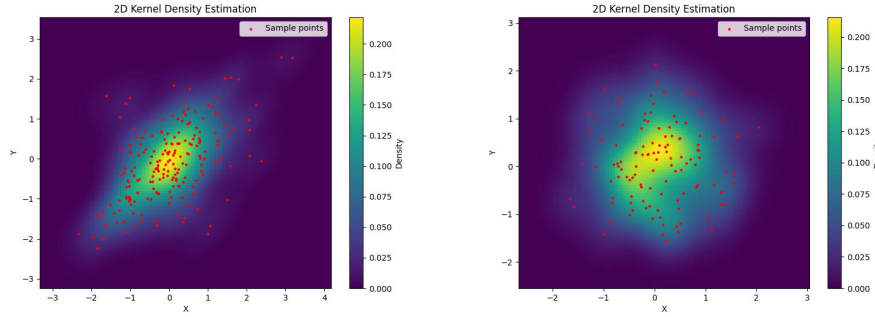


FIG. 9. 2D projected density estimation for $\tilde{t}(2)$ (left) and $ci \tilde{t}(2)$ (right) for base dimension 40

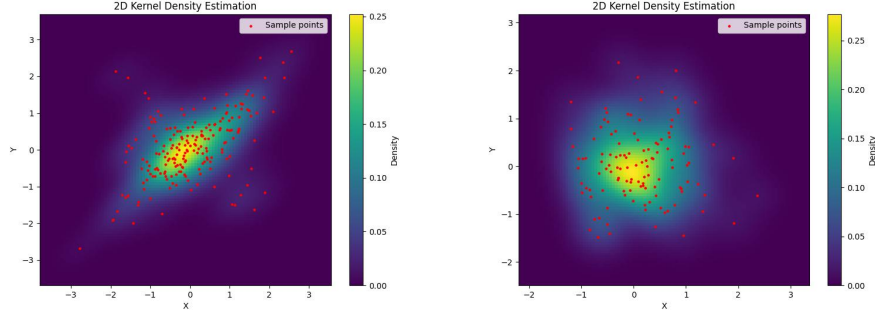


FIG. 10. 2D projected density estimation for $\tilde{t}(1)$ (left) and $ci \tilde{t}(1)$ (right) for base dimension 40

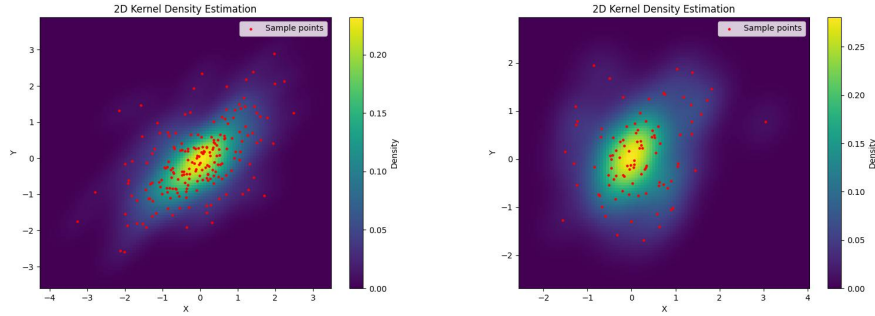


FIG. 11. $\widetilde{Laplace}$ projected density estimation for $\widetilde{Laplace}$ (left) and $ci \widetilde{Laplace}$ (right) for base dimension 40

Let's note that it is not surprising that sometimes we obtain 'True' for $\tilde{t}(8)$, given that $t(8)$ itself resembles a multivariate normal distribution, much less $\tilde{t}(8)$.

The relatively small sample size ($N = 200$) is explained by the fact that, with a fixed base dimension n , increasing the sample size leads to an increase in the power of the test, which begins to show a deviation from normality in the projections of all CI-images. This is natural since these images are only asymptotically Gaussian.

5 Extreme limit theorem

Theorem 1 and Corollary 1 show that the CI-transformation of the t distribution turns out to be an asymptotically Gaussian copula, which, in turn, is the result of the CI-transformation of the Gaussian core of the original mixture.

We are now interested in the following question: is it true that the extreme values of random variables with t distribution CI-copulas and Gaussian marginal distributions behave similarly to the extremes of Gaussian vectors in a triangular array?

It is worth noting that the extreme values in triangular array of dependent random variables associated with various families of copulas have been studied in the work [17]. This paper also presents generalized results on extreme indices for triangular arrays with Archimedean copulas. Additionally, questions regarding series of random lengths were explored in [18].

We use classical numerical sequences from extreme value theory (see [19]) ($n \geq 2$)

$$\alpha_n = (2 \ln n)^{1/2}, \quad \beta_n = (2 \ln n)^{1/2} - \frac{1}{2}(2 \ln n)^{-1/2}(\ln \ln n + \ln(4\pi)),$$

arising naturally when analyzing the asymptotics of probabilities of the form $\mathbf{P}\{M_n \leq u_n(x)\}$, where M_n is the maximum of n independent standard normal random variables, and the sequence $u_n(x)$ is required to have a limiting non-degenerate distribution function $G(x)$ under linear normalization of M_n . It turns out that this is only possible when $u_n(x) = x/\alpha_n + \beta_n$ for the specified α_n and β_n .

Theorem 2. *Let us consider a family of t -distributed random vectors. $\mathbf{X}_n = \{X_1, X_2, \dots, X_n\}$ $n = 1, 2, \dots$, defined on a probability space $\{\mathbb{H}, \mathcal{B}(\mathbb{H}), \mu\}$ with characteristic functions given by (6). Let C_n be the corresponding t -copulas, $\Phi(x)$ and $\varphi(x)$ denote the standard normal cumulative distribution function and density function, respectively. Let us introduce a triangular array of random variables $\{X_i^{(n)}\}_{i=1}^n$, $n = 1, 2, \dots$, defined on a suitable probability space $\{\Omega_0, \mathfrak{B}_0, \mathbf{P}_0\}$ and having a joint distribution function*

$$\mathbf{P}_0 \left\{ X_1^{(n)} \leq x_1, \dots, X_n^{(n)} \leq x_n \right\} = C_n^{ci}(\Phi(x_1), \dots, \Phi(x_n)).$$

If the basis $\{f_k\}$ is such that

$$\delta := \limsup_{n \rightarrow \infty} \max_{i \neq j} |c_{ij}^{(n)}| < 1, \quad (18)$$

and for $\gamma > 0$ there exists $0 < \alpha < \frac{1-\delta}{1+\delta}$ such that

$$\max_{n^\alpha < j-i < n} |c_{ij}^{(n)} \ln(j-i) - \gamma| \rightarrow 0, \quad n \rightarrow \infty, \quad (19)$$

then for any $x \in \mathbb{R}$ the convergence holds:

$$\mathbf{P}_0 \left\{ \alpha_n \left(\max_{1 \leq i \leq n} X_i^{(n)} - \beta_n \right) \leq x \right\} \rightarrow \int_{-\infty}^{\infty} \exp \left\{ -e^{-x-\gamma+\sqrt{2\gamma}z} \right\} \varphi(z) dz. \quad (20)$$

Proof.

In the proof, Theorem 1 from the work [20] and Lemmas 2, 3, and 4 from the work [10] are used.

Without loss of generality, we can assume that the probability space $\{\Omega_0, \mathfrak{B}_0, \mathbf{P}_0\}$ is $\{\mathbb{H}, \mathcal{B}(\mathbb{H}), \mu\}$, and

$$X_i^{(n)} = \Phi^{-1}(X_{i,n}^*). \quad (21)$$

Let us denote

$$M_n := \max_{1 \leq i \leq n} X_i^{(n)}. \quad (22)$$

We will now use the following notations:

$$\begin{aligned} a_i(n) &:= \frac{X_i^{(n)}}{\zeta_{i,n}}, \quad b_i(n) := \zeta_{i,n}, \\ a_*(n) &:= \min_{1 \leq i \leq n} \frac{X_i^{(n)}}{\zeta_{i,n}}, \quad a^*(n) := \max_{1 \leq i \leq n} \frac{X_i^{(n)}}{\zeta_{i,n}}, \quad b^*(n) := \max_{1 \leq i \leq n} \zeta_{i,n}, \\ d(n) &:= a_*(n) \mathbf{1}_{\{b^*(n) \geq 0\}} + a^*(n) \mathbf{1}_{\{b^*(n) < 0\}}, \\ e(n) &:= a^*(n) \mathbf{1}_{\{b^*(n) \geq 0\}} + a_*(n) \mathbf{1}_{\{b^*(n) < 0\}}, \end{aligned}$$

where $\zeta_{i,n}$ defined in (7) from the paper [10].

Since μ -a.s. $a_*(n) > 0$ (see Lemma 3, item 2 in [10]) and μ -a.s.

$$d(n)b^*(n) \leq M_n \leq e(n)b^*(n),$$

(see Lemma 4 in [10]), then μ -a.s.

$$\alpha_n [d(n)b^*(n) - \beta_n] \leq \alpha_n (M_n - \beta_n) \leq \alpha_n [e(n)b^*(n) - \beta_n]. \quad (23)$$

Note (see Lemma 3, item 3 in [10]) that μ -a.s. $a_*(n) \rightarrow 1$ and $a^*(n) \rightarrow 1$, ; therefore, due to

$$a_*(n) \leq d(n) \leq a^*(n), \quad a_*(n) \leq e(n) \leq a^*(n),$$

we have

$$d(n) \rightarrow 1, \quad e(n) \rightarrow 1, \quad \mu - \text{a.s.} \quad (24)$$

Let's consider the left side of inequality (23)

$$\alpha_n [d(n)b^*(n) - \beta_n] = d(n)\alpha_n [b^*(n) - \beta_n] + \alpha_n \beta_n [d(n) - 1]. \quad (25)$$

According to Lemma 2, the random variables $\zeta_{i,n}$ for $i = 1, 2, \dots, n$ are jointly Gaussian with covariances

$$\text{cov}(\zeta_{i,n}, \zeta_{j,n}) = c_{ij}^{(n)},$$

defined by (8). Therefore, due to (24) and Theorem 1 from [20], for all $x \in \mathbb{R}$

$$\mu\{d(n)\alpha_n [b^*(n) - \beta_n] \leq x\} \rightarrow \int_{-\infty}^{\infty} \exp\left\{-e^{-x-\gamma+\sqrt{2\gamma}z}\right\} \varphi(z) dz. \quad (26)$$

Let us use the previously obtained results, noting that for $n \geq 2$

$$\alpha_n \beta_n = 2 \ln n - \frac{1}{2} (\ln \ln n + \ln 4\pi) > 0,$$

μ -a.s. (see Lemma 3, item 1 in [10])

$$\begin{aligned}\alpha_n \beta_n A^{(n)} &\leq \alpha_n \beta_n [a_*(n) - 1] \leq \\ &\leq \alpha_n \beta_n [d(n) - 1] \leq \\ &\leq \alpha_n \beta_n [a^*(n) - 1] \leq \alpha_n \beta_n \max_{1 \leq i \leq n} B_i^{(n)}.\end{aligned}$$

Thus, (see Lemma 2, items 2 and 3 in [10]) μ -a.s. $\alpha_n \beta_n [d(n) - 1] \rightarrow 0$. Hence, taking into account (25) and (26), we have

$$\mu\{\alpha_n [d(n)b^*(n) - \beta_n] \leq x\} \rightarrow \int_{-\infty}^{\infty} \exp\left\{-e^{-x-\gamma+\sqrt{2\gamma}z}\right\} \varphi(z) dz.$$

The convergence on the right side of (23) is proven similarly, from which the statement of the theorem follows. $\square$

We note that the limit in (20) has the same form as that for the case of a stationary normal sequence (see [19], p. 137), which is not surprising, as Theorem 1 implies the asymptotic normality of the vectors $(X_1^{(n)}, \dots, X_n^{(n)})$, whose distribution functions are linked by the copulas C_n^{ci} . It is important to emphasize that, unlike the aforementioned result, our case involves a triangular array, where its rows are neither stationary nor Gaussian.

6 Conclusion

Mixture distributions are widely studied in various contexts, such as reliability theory and other similar cases where we deal with a set of observations consisting of heterogeneous subgroups. As demonstrated in this work, by applying methods for estimating conditional distribution functions and utilizing the CI-transformation, it is possible to construct some criteria for assessing the fit of a sample to a Gaussian mixture (more precisely, to a mixture of distributions with a Gaussian structure), albeit currently in a relatively specific case of the family of t distributions.

To illustrate, experiments were conducted on model data, which showed that there is indeed hope to extend this approach to Gaussian mixtures with other weight distributions.

It should be noted that the result of the work [4] allows us to discuss the use of this approach for testing infinite exchangeability (i.e., membership, due to de Finetti's theorem, in mixtures of independent random variables).

Finally, it was shown that the maximum component of the vector from the transformed multivariate sample of such a Gaussian mixture behaves analogously to the maximum component of the vector from the transformed multivariate Gaussian sample.

Since the results of Theorem 1 provide insights into the limiting dependence structure and its connection to the original structure, subsequent research may focus on estimating the parameters of the distributions $G_k^{(t)}$

and their asymptotic properties, including for a broader class of continuous mixtures.

We express our sincere gratitude to the reviewer and the editor for their valuable suggestions and extremely helpful comments. Their feedback significantly contributed to the improvement of the work.

References

- [1] M. Rosenblatt, *Remarks on Multivariate Transformation*, Ann. Math. Stat., **23**:3 (1952), 470–472.
- [2] S.Ya. Shatskikh, *The Strong Law of Large Numbers for a Triangular Array Scheme of Conditional Distributions of Stable Elliptically Contoured Measures*, Theory Probab. Appl., **50**:2 (2006), 248–264.
- [3] R. Nelsen, *An Introduction to Copulas*, Second ed., Springer-Verlag, New York, 2006.
- [4] E. Savinov, V. Shamraeva, *On a Rosenblatt-type transformation of multivariate copulas*, Econometrics and Statistics, **25** (2023), 39–48.
- [5] T. Eltoft, T. Kim, T.W. Lee, *Multivariate Scale Mixture of Gaussians Modeling*, in: Rosca, J., Erdogmus, D., Principe, J.C., Haykin, S. (Eds.), Independent Component Analysis and Blind Signal Separation, Springer Berlin Heidelberg, Berlin, Heidelberg, (2006), 799–806.
- [6] R. Orellana, R. Carvajal, J.C. Agüero, *Maximum Likelihood Infinite Mixture Distribution Estimation Utilizing Finite Gaussian Mixtures*, IFAC-Papers OnLine, **51**:15 (2018), 706–711.
- [7] L. Melkumova, S.Ya. Shatskikh, *Maximum Likelihood Method in de Finetti's Theorem*, Theory Probab. Appl., **63**:4 (2019), 808–816.
- [8] D.K. Dey, *Estimation of scale parameters in mixture distributions*, The Canadian Journal of Statistics / La Revue Canadienne de Statistique, **18**:2 (1990), 171–178.
- [9] S. Kotz, S. Nadarajah, *Multivariate t Distributions and Their Applications*, Cambridge University Press, Cambridge, 2004.
- [10] E.A. Savinov, *Limit Theorem for the Maximum of Random Variables Connected by IT-Copulas of Student's t-Distribution*, Theory Probab. Appl., **59**:3 (2015), 508–516.
- [11] A. Heinen, A. Valdesogo, *Spearman rank correlation of the bivariate Student t and scale mixtures of normal distributions*, J. of Multivariate Analysis, **179** (2020), 104650.
- [12] H. Joe, *Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters*, Lecture notes-monograph series, **28** (1996), 120–141.
- [13] T. Bedford, R.M. Cooke, *Probability Density Decomposition for Conditionally Dependent Random Variables Modeled by Vines*, Annals of Mathematics and Artificial Intelligence, **32** (2001), 245–268.
- [14] T. Bedford, R.M. Cooke, *Vines – a new graphical model for dependent random variables*, The Annals of Statistics, **30**:4 (2002), 1031–1068.
- [15] P. Hall, Q. Yao, *Approximating conditional distribution functions using dimension reduction*, The Annals of Statistics, **33**:3 (2005), 1404–1421.
- [16] N. Henze, B. Zirkler, *A class of invariant consistent tests for multivariate normality*, Communications in Statistics - Theory and Methods, **19**:10 (1990), 3595–3617.
- [17] A.V. Lebedev, *Extremal indices in a series scheme and their applications*, Informatics and Applications, **9**:3 (2015), 39–54.
- [18] A.A. Goldaeva, A.V. Lebedev, *On extremal indices greater than one for a scheme of series*, Lithuanian Mathematical Journal, **58** (2018), 384–398.
- [19] M. Leadbetter, G. Lindgren, H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983.

- [20] E. Savinov, *On Gaussian triangular arrays in the case of strong dependence*, Extremes, (2024)

EVGENIY SAVINOV
FINANCIAL UNIVERSITY UNDER THE GOVERNMENT OF THE RUSSIAN FEDERATION,
49 LENINGRADSKY PROSPEKT, GSP-3,
125993, MOSCOW, RUSSIAN FEDERATION
Email address: egasavinov@fa.ru