Statistical Methods & Applications
https://doi.org/10.1007/s10260-026-00856-x

ORIGINAL PAPER

The logratio Student's *t* distribution: a robust model for compositional data analysis

Gianna Serafina Monti[1] · **Gloria Mateu-Figueras**[2] · **Vera Pawlowsky-Glahn**[2] · **Juan José Egozcue**[3]

Received: 26 September 2025 / Accepted: 13 May 2026

Abstract
We introduce and study the logratio Student's t distribution as a robust and flexible alternative to the classical logratio normal model for compositional data. This distribution preserves the geometric coherence of the Aitchison simplex framework while accommodating heavier tails, enabling improved detection of outliers — an essential feature in many real-world applications. Our main contribution is to formally embed the t distribution within the logratio framework and to demonstrate its equivalence across logratio representations for the purposes of estimation and outlier detection, while clarifying the distinct mathematical properties of each representation beyond distribution fitting. Monte Carlo simulations under scale contamination show that our model, combined with a Leave-One-Out procedure, outperforms traditional robust methods (MCD, COMCoDa) by ensuring higher AUC and near-perfect specificity, even in high dimensions. A real-world application illustrates the model's superior ability to uncover structure and identify outliers in multivariate compositional data. These results position the logratio t distribution as a theoretically sound and practically powerful tool for robust inference in the simplex.

Keywords Multivariate t distribution · Aitchison measure · Logratio normal family · Coordinate representations · Outliers · Atypicality index

1 Introduction

The main parametric families of distributions for random compositions are the Dirichlet distribution, certain generalizations of the Dirichlet family (Monti et al. 2011; Mateu-Figueras et al. 2021), and the logratio normal family (Mateu-Figueras

Gianna Serafina Monti and Gloria Mateu-Figueras contributed equally to this work.

Extended author information available on the last page of the article

Published online: 25 June 2026

et al. 2013). Among these, the additive logistic normal model plays a central role in the analysis of compositional data. This family of distributions, originally defined via the additive logratio transformation (Aitchison 1986; Pawlowsky-Glahn et al. 2015), relies on a multivariate normal distribution applied to logratio coordinates in real space.

Although the additive logratio representation was initially used to construct the model, subsequent developments have shown that other logratio representations, including the centered logratio (clr) (Aitchison 1986) and any isometric logratio (ilr) representation (Egozcue et al. 2003), are consistent with the Aitchison geometry and yield equivalent models for the purposes of parameter estimation and likelihood-based inference. For this reason, the model is more broadly known as the logratio normal, or simply normal on the simplex (Mateu-Figueras et al. 2013), highlighting its broad compatibility with the Aitchison geometry across different coordinate representations.

Despite its advantages, the logratio normal model is sensitive to outliers and may be insufficient in scenarios where data exhibit heavy tails. To address this limitation, robust alternatives based on the Student's t distribution have been proposed, particularly using the additive logratio transformation (Aitchison 1986; Aitchison and Dunsmore 1975). However, while the analytical form of the logratio t density has been previously noted, it has often been treated as a simple algebraic extension rather than a geometrically grounded model. In this paper, we provide a formal development of the logratio t distribution within the Aitchison framework, showing that it constitutes a geometrically coherent robust analogue to the logratio normal model.

The remainder of the article is structured as follows. Section 2 briefly reviews the multivariate t distribution and its properties. Section 3.1 introduces key elements of the Aitchison geometry and the role of coordinate representations in compositional data analysis. In Sect. 3, we formally define the logratio t distribution and examine its main properties. Section 4 presents a simulation study to evaluate the model's performance under scale contamination, providing a comparative analysis with established robust and parametric methods. A real-data application in Sect. 5 illustrates the model's practical utility, particularly in identifying outliers. We conclude in Sect. 6 with a discussion of implications and potential extensions.

2 A review of the multivariate *t* distribution

The multivariate t distribution (MVT) is a generalization of the univariate t distribution to multiple dimensions, providing a robust alternative to the multivariate normal distribution. It is beneficial in situations where data exhibit heavier tails or when the assumption of normality is violated. As an illustrative yet non-exhaustive example, we highlight the application of the MVT in the following contexts, including Bayesian analysis (Gelman et al. 2013), the modeling of random vectors, robust procedures for statistical inference regarding location parameters and scale matrices, as well as their functions, such as regression coefficients (Lange et al. 1989; Liu and Rubin 1995; Zellner 1976). The MVT is also employed in robust hierarchical linear mixed-effects models to account for random effects and within-subject errors (Pinheiro et al.

2001), and in sample selection (Marchenko and Genton 2012) to model the error distribution.

The d-dimensional random vector $\mathrm{X} = (X_1, \ldots, X_d)$ is said to follow a multivariate t distribution (MVT) (Cornish 1954; Dunnett and Sobel 1954), with location parameter $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Upsilon}$, and $\nu \in (0, \infty]$ degrees-of-freedom denoted by $\mathrm{X} \sim t_d(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ if it has probability density function

$$f(\mathrm{x}) = \frac{\Gamma((\nu+d)/2)}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \left(1 + \nu^{-1}(\mathrm{x}-\boldsymbol{\mu})'\boldsymbol{\Upsilon}^{-1}(\mathrm{x}-\boldsymbol{\mu})\right)^{-(\nu+d)/2}, \quad (1)$$

where $\mathrm{x} \in \mathbb{R}^d$. The variance-covariance matrix of X is given by $\nu(\nu-2)^{-1}\boldsymbol{\Upsilon}$, $\nu > 2$. Besides, we assume that $\boldsymbol{\Upsilon}$ is a positive definite matrix to ensure the existence of the density function as expressed in (1). However, this condition can be relaxed to positive semi-definiteness in specific contexts, such as when the distribution is defined on a degenerate subspace-for instance, when applying centered log-ratio transformations, as detailed in the following section.

The distribution (1) is called central if $\boldsymbol{\mu} = \mathbf{0}$, otherwise it is called a non-central MVT distribution. When $d = 1$ and $\boldsymbol{\Upsilon} = 1$, the density (1) corresponds to the density of a univariate Student's t distribution with ν degrees of freedom. As $\nu \to \infty$, the limiting form of (1) becomes the probability density function of a multivariate normal distribution (MVN) with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Upsilon}$, denoted by $N_d(\boldsymbol{\mu}, \boldsymbol{\Upsilon})$ (Kotz and Nadarajah 2004; Lange et al. 1989).

Heavier tails characterize the MVT compared to the MVN; in particular, smaller values of ν result in heavier tails of the probability density, making it particularly appealing for practical applications where extreme observations are likely to occur (Lange et al. 1989).

The density in Equation (1) depends on x only through $\delta^2 = (\mathrm{x}-\boldsymbol{\mu})'\boldsymbol{\Upsilon}^{-1}(\mathrm{x}-\boldsymbol{\mu})$, which is the Mahalanobis squared distance from x to the center $\boldsymbol{\mu}$ with respect to $\boldsymbol{\Upsilon}$. Therefore, the MVT belongs to the class of elliptically contoured distributions (Fernandez and Steel 1999; Fang 1990), and its isodensity contours are ellipsoids with equation $\delta^2 = c$ for some $c > 0$, like the multivariate normal distribution (Box and Tiao 1992).

Property 2.1 (Representation) If Y is a d-variate normal random vector with mean $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{\Upsilon}$, and if νS^2 is the chi-squared random variable with degrees of freedom ν, $\nu S^2 \sim \chi^2_\nu$, independent of Y, then

$$\mathrm{X} = S^{-1}\mathrm{Y} + \boldsymbol{\mu} \sim t_d(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu). \quad (2)$$

This implies that $\mathrm{X}|S^2 = s^2 \sim N(\boldsymbol{\mu}, s^{-2}\boldsymbol{\Upsilon})$.

Property 2.2 (Characterization) Let $\nu S^2 \sim \chi^2_\nu$ and let $X_1, X_2, \ldots, X_d$ be continuous random variates that, conditional on $S^2 = s^2$, are independent and symmetrically distributed with $\mathrm{E}[X_j|S^2 = s^2] = \mu_j$ and $\mathrm{Var}[X_j|S^2 = s^2] = \sigma_j^2/s^2 < \infty$. Then

$$\frac{1}{d}\sum_{j=1}^{d}\frac{(X_j-\mu_j)^2}{\sigma_j^2}\sim F_{d,\nu}, \tag{3}$$

where $F_{d,\nu}$ denotes the F distribution with d and ν degrees of freedom, if and only if $(X_1,\ldots,X_d)\sim t_d(\boldsymbol{\mu},\mathrm{D},\nu)$, where $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_d)$ and D is a diagonal matrix of order d with its jth diagonal element equal to σ_j^2 (see Lin (1972) for a proof.)

The MVT distribution has an interesting property that we will use in the following section, namely, the linear transformation property.

Property 2.3 (Distribution of a Linear Function) If $\mathrm{X}\sim t_d(\boldsymbol{\mu},\boldsymbol{\Upsilon},\nu)$, then, for any $(p\times d)$ full rank matrix C, with $p\leq d$, and for any vector a, the random vector $\mathrm{CX}+\mathrm{a}\sim t_p(\mathrm{C}\boldsymbol{\mu}+\mathrm{a},\mathrm{C}\boldsymbol{\Upsilon}\mathrm{C}',\nu)$ (Nadarajah and Kotz 2005).

Property 2.4 (Marginal Distribution) Let X follow a d-variate t distribution with ν degrees of freedom, mean vector $\boldsymbol{\mu}$, and scale matrix $\boldsymbol{\Upsilon}$, denoted by $\mathrm{X}\sim t_d(\boldsymbol{\mu},\boldsymbol{\Upsilon},\nu)$. Consider the partitions

$$\mathrm{X}=\begin{pmatrix}\mathrm{X}_1\\ \mathrm{X}_2\end{pmatrix},\quad \boldsymbol{\mu}=\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix},\quad \boldsymbol{\Upsilon}=\begin{pmatrix}\boldsymbol{\Upsilon}_{11} & \boldsymbol{\Upsilon}_{12}\\ \boldsymbol{\Upsilon}_{21} & \boldsymbol{\Upsilon}_{22}\end{pmatrix},$$

where X_1 has dimension $(d_1\times 1)$ with a scale matrix $\boldsymbol{\Upsilon}_{11}$. Then $\mathrm{X}_1\sim t_{d_1}(\boldsymbol{\mu}_1,\boldsymbol{\Upsilon}_{11},\nu)$

. Further, X_2 with dimension $(d-d_1)\times 1$ also has a $(d-d_1)$-variate t distribution with ν degrees of freedom, mean vector $\boldsymbol{\mu}_2$ and scale matrix $\boldsymbol{\Upsilon}_{22}$ (Kotz and Nadarajah 2004).

Parameter estimation is performed via Maximum Likelihood (ML). Specifically, we employ the Expectation-Conditional Maximization (ECM) algorithm (Liu and Rubin 1995) to estimate the location vector $\boldsymbol{\mu}$, the scale matrix $\boldsymbol{\Upsilon}$, and the degrees of freedom ν. By updating ν iteratively, the algorithm enables the model to adaptively capture the distribution's tail behavior within the representation space.

Property 2.5 (ML Estimation via ECM) Given the log-likelihood function for the multivariate t distribution, the ML estimators $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Upsilon}}$ satisfy the following system of iterative estimating equations, derived from the score functions (see Appendix A, Eq. A2)

$$\widehat{\boldsymbol{\mu}}=\frac{\sum_{i=1}^{n}\hat{w}_i\mathrm{x}_i}{\sum_{i=1}^{n}\hat{w}_i}, \tag{4}$$

and

$$\widehat{\boldsymbol{\Upsilon}}=\frac{1}{n}\sum_{i=1}^{n}\hat{w}_i(\mathrm{x}_i-\widehat{\boldsymbol{\mu}})(\mathrm{x}_i-\widehat{\boldsymbol{\mu}})', \tag{5}$$

with estimated weights depending on the squared Mahalanobis distance $\hat{\delta}_i^2$

$$\hat{w}_i = \frac{\nu + d}{\nu + \hat{\delta}_i^2}, \quad \text{with} \quad \hat{\delta}_i^2 = (\mathrm{x}_i - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Upsilon}}^{-1} (\mathrm{x}_i - \widehat{\boldsymbol{\mu}}). \tag{6}$$

Equations (4) and (5) define $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Upsilon}}$ as M-estimators (Maronna 1976; Huber 1981). In this iterative framework, the weights $\hat{w}_i$ are updated at each step of the ECM algorithm based on the current estimates of the parameters. Specifically, observations with a large squared distance $\hat{\delta}_i^2$ (potential outliers) are adaptively down-weighted. The degree of down-weighting is controlled by ν: as $\nu \to \infty$, the weights approach unity, recovering the Gaussian case, whereas smaller values of ν assign significantly less weight to atypical observations (Kotz and Nadarajah 2004; Lange et al. 1989).

As described above, the estimation is performed by employing the Expectation-Conditional Maximization (ECM) algorithm (Liu and Rubin 1995). In our analysis, we rely on the implementation provided by the `fitHeavyTail` package, which further enhances this framework by incorporating the Parameter Expansion (PX-EM) technique (Liu et al. 1998). This approach specifically addresses the slow convergence of the degrees-of-freedom parameter ν by expanding the model with auxiliary scale parameters, ensuring computational efficiency and stability even in high-dimensional or small-sample settings. Convergence is monitored via the relative change in parameter estimates, with a tolerance threshold set to 10^{-3} to ensure the stability of the final likelihood maximum.

3 The logratio Student's *t* distribution

3.1 Preliminaries

A D-part composition is a class of D-vectors with positive components, which are considered equivalent when their components are proportional (Aitchison 1992; Barceló-Vidal and Martín-Fernández 2016). Compositions are often represented in the D-part simplex $\mathcal{S}^D = \{\mathrm{x} = (x_1, \ldots, x_D), x_i > 0, \sum_{i=1}^D x_i = k\}$, where $k > 0$ is a constant, usually 1 or 100. The selection of a representative in the simplex is attained through the closure operator $\mathcal{C}$, which normalizes the containing vector so that the sum of its components is k. The appropriate sample space for a D-part random composition is the simplex $\mathcal{S}^D$. The simplex has a particular algebraic-geometric structure, called Aitchison geometry (Pawlowsky-Glahn and Egozcue 2001). It is based on two specific operations, perturbation $\oplus$ and powering $\odot$, that induce a $(D-1)$ -dimensional vector space structure. For compositions $\mathrm{x}, \mathrm{y} \in \mathcal{S}^D$ they are defined as $\mathrm{x} \oplus \mathrm{y} = \mathcal{C}(x_1 y_1, \ldots, x_D y_D)$ and $c \odot \mathrm{x} = \mathcal{C}(x_1^c, \ldots, x_D^c)$ and $c \in \mathbb{R}$.

The Aitchison inner product, $< \mathrm{x}, \mathrm{y} >_a = (1/D) \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} \ln(x_i/x_j) \ln(y_i/y_j)$, induces a $(D-1)$-dimensional Euclidean vector space structure on $\mathcal{S}^D$, enabling the construction and use of orthonormal bases. A composition $\mathrm{x} \in \mathcal{S}^D$ can be represented uniquely by its coordinates with respect to a basis.

The notation $\mathrm{ilr}(\mathrm{x})$ is introduced by Egozcue et al. (2003) for a generic orthonormal coordinates and more recently the notation $\mathrm{olr}(\mathrm{x})$ has also been suggested (Martín-Fernández 2019). In this paper, to avoid confusion, we denote generic coordinates in an orthonormal basis as $h(\mathrm{x})$.

Let $d = D - 1$ be the dimension of the coordinate space. The notation $h^{-1}(\cdot)$ indicates the composition associated with a vector of coordinates. When changing between orthonormal bases, the coordinates are related by an orthogonal matrix. Other frequently used representations of compositions involve some non-orthonormal coordinates, such as the additive logratio coordinates, $\mathrm{alr}(\mathrm{x}) = (\ln(x_1/x_D), \ldots, \ln(x_{D-1}/x_D)) \in \mathbb{R}^d$, or coordinates in a particular generating system called centered logratio coordinates, $\mathrm{clr}(\mathrm{x}) = (\ln(x_1/g(\mathrm{x})), \ldots, \ln(x_D/g(\mathrm{x}))) \in \mathbb{R}^D$, where $g(\mathrm{x})$ denotes the geometric mean of the components of x. It is also common to refer to these vectors as additive or centered logratio transformed vectors. Using the matrix relationship between these coordinates provided by Egozcue et al. (2003) and by Aitchison (1986) we know that $\mathrm{clr}(\mathrm{x}) = \mathrm{U}h(\mathrm{x})$ and $\mathrm{alr}(\mathrm{x}) = \mathrm{B}h(\mathrm{x})$ where the $d \times D$ matrix U contains the clr coordinates of the chosen orthonormal basis in columns and the $d \times d$ matrix B is a non-orthogonal matrix, specifically $\mathrm{B} = \mathrm{FU}$ where $\mathrm{F} = [\mathrm{I}_d : -\mathrm{j}_d]$ a $d \times D$ matrix with j_d denoting the column vector of ones of length d. It can be shown that $h(\mathrm{x}) = \mathbf{U}'\mathrm{clr}(\mathrm{x}) = \mathrm{B}^{-1}\mathrm{alr}(\mathrm{x})$ and $|\mathrm{B}| = \sqrt{D}$ (Egozcue et al. 2003).

We may be interested in a subgroup of parts; this corresponds to the definition of a subcomposition. Geometrically, it can be interpreted as an orthogonal projection. Formally, the formation of a subcomposition can be written as a transformation from the simplex $\mathcal{S}^D$ to a simplex $\mathcal{S}^C$ of lower dimension. In fact, given $\mathrm{x} \in \mathcal{S}^D$, a C-part composition with $C < D$ can be obtained as $\mathcal{C}(\mathrm{Sx})$ where S is a $C \times D$ selection matrix with C elements equal to 1 (one in each row and at most one in each column) and the remaining elements equal to 0 (Aitchison 1986).

A natural measure on the simplex, compatible with its Euclidean vector space structure, is called Aitchison measure and denoted as λ_a (Pawlowsky-Glahn 2003). This measure is absolutely continuous with respect to the Lebesgue measure on real space, denoted here as λ. The relationship between them is

$$|d\lambda_a/d\lambda| = (\sqrt{D}\; x_1 \cdots x_D)^{-1}. \tag{7}$$

The linear operations of the vector space are preserved under any logratio transformation, although these operations refer to different dimensional spaces ($\mathbb{R}^D$ for clr, $\mathbb{R}^{D-1}$ for ilr and alr). This means that $\mathrm{clr}(\mathrm{x} \oplus (c \odot \mathrm{y})) = \mathrm{clr}(\mathrm{x}) + c\mathrm{clr}(\mathrm{y})$, $h(\mathrm{x} \oplus (c \odot \mathrm{y})) = h(\mathrm{x}) + ch(\mathrm{y})$ and $\mathrm{alr}(\mathrm{x} \oplus (c \odot \mathrm{y})) = \mathrm{alr}(\mathrm{x}) + c\mathrm{alr}(\mathrm{y})$. The isometry property $\langle \mathrm{x}, \mathrm{y} \rangle_a = \langle \mathrm{clr}(\mathrm{x}), \mathrm{clr}(\mathrm{y}) \rangle = \langle h(\mathrm{x}), h(\mathrm{y}) \rangle$; however, this does not hold for the alr coordinates. Mateu-Figueras et al. (2013) define the expected value of a random composition as $\mathrm{E}_a(\mathrm{X}) = h^{-1}(\mathrm{E}(h(\mathrm{X})))$ and the metric variance as $\mathrm{Mvar}(\mathrm{X}) = \mathrm{E}(d_a^2(\mathrm{X}, \mathrm{E}_a(\mathrm{X}))) = \mathrm{E}(d^2(h(\mathrm{X}), \mathrm{E}(h(\mathrm{X}))))$, where d_a is the Aitchison distance induced from the Aitchison inner product. Due to the matrix relationships between the different logratio representations, the same composition $\mathrm{E}_a(\mathrm{X})$ is obtained whether using clr or alr coordinates; that is $\mathrm{E}_a(\mathrm{X}) = \mathrm{alr}^{-1}(\mathrm{E}(\mathrm{alr}(\mathrm{X}))) = \mathrm{clr}^{-1}(\mathrm{E}(\mathrm{clr}(\mathrm{X})))$. However, this equivalence does

not extend to the metric variance, as it does not hold for the alr representation. While the metric variance could technically be computed using other representations, for the alr representation would require the inclusion of a metric matrix to account for the non-orthogonal nature of the underlying basis.

According to Mateu-Figueras et al. (2013), a random composition $\mathrm{x} \in \mathcal{S}^D$ has a normal distribution on $\mathcal{S}^D$ with location $\boldsymbol{\mu} \in \mathbb{R}^{D-1}$ and variance-covariance matrix $\boldsymbol{\Upsilon}$, if its density function with respect to the Aitchison measure is

$$f^a(\mathrm{x}) = \frac{dP_a}{d\lambda_a}(\mathrm{x}) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \exp\left(-\frac{1}{2}(h(\mathrm{x}) - \boldsymbol{\mu})'\boldsymbol{\Upsilon}^{-1}(h(\mathrm{x}) - \boldsymbol{\mu})\right), \quad (8)$$

where $h(\cdot)$ stands for a generic orthonormal coordinates, $d = D - 1$, and P_a is the normal probability measure. This model is commonly referred to as the logratio normal or normal on the simplex (Mateu-Figueras et al. 2013).

The key point is to define a density on the simplex with respect to the Aitchison measure λ_a, which is equivalent to a density with respect to the Lebesgue measure in $h(\mathrm{x})$ coordinates. An equivalent distribution is obtained using alr coordinates, under a different parametrization: if $\mathrm{alr}(\mathrm{x}) = \mathrm{B}\, h(\mathrm{x})$, then the location and scale parameters transform as $\mathrm{B}\boldsymbol{\mu}$ and $\mathrm{B}\boldsymbol{\Upsilon}\mathrm{B}'$ respectively. The clr coordinates are constrained to sum to zero ($\sum_i \mathrm{clr}_i(\mathrm{x}) = 0$), so the density is formally defined on the $(D-1)$ -dimensional subspace $\{\mathrm{clr}(\mathrm{x}) \in \mathbb{R}^D : \sum_i \mathrm{clr}_i(\mathrm{x}) = 0\}$ rather than on $\mathbb{R}^D$. Therefore, the clr representation yields an equivalent but degenerate model, with location and scale parameters $\mathrm{U}\boldsymbol{\mu}$ and $\mathrm{U}\boldsymbol{\Upsilon}\mathrm{U}'$ respectively. In particular, the resulting scale matrix is singular, which requires the use of the Moore–Penrose pseudoinverse and pseudo-determinant in the expression of the density function in clr coordinates.

3.2 Early development

The use of the multivariate *t* distribution for random compositions can be traced back to Aitchison (1986), who introduced the logistic Student's *t* density on $\mathcal{S}^D$. This density was obtained through an additive logistic transformation of a multivariate *t* distribution and expressed with respect to the Lebesgue measure in real space λ. The logistic Student's *t* distribution was employed as a predictive density function derived from a logistic-normal distribution. Subsequently, predictive regions and atypicality indices were obtained. All these results were derived from the work of Aitchison and Dunsmore (1975), where predictive distributions in real space are studied for a wide range of scenarios. Among these, the multivariate *t* predictive distribution is derived from a multivariate normal density and a vague prior density function for the parameters (Aitchison and Dunsmore 1975, Table 2.3). The concept of predictive distributions in real space has a long-standing tradition; more detailed derivations can be found in Jeffreys (1983, Chapter 3) or Raiffa and Schlaifer (1961, Chapter 12). Applications of this framework can be seen in Aitchison et al. (1977) or Moran and Murphy (1979).

Later, Katz and King (1999) applied the additive logistic Student's *t* distribution in modeling the composition of multiparty electoral data. Their approach, applied to British district-level elections, demonstrated the utility of the logistic Student's *t*

distribution in analyzing patterns like incumbency advantage and regional vote distributions, further extending its applicability beyond the original framework (Aitchison 1986). More recently, Nguyen (2019) made a similar application to French departmental election data in 2015, which exhibited heavy tail behaviors and also spatial autocorrelation. In particular, a compositional regression model is employed wherein the error vector is modeled using the logratio *t* distribution.

3.3 The model

In this contribution, we focus on the logratio Student's *t* distribution, a robust generalization of the logratio normal model for compositional data on the simplex $\mathcal{S}^D$. The model is defined with respect to the Aitchison measure λ_a on the simplex, which, as the logratio normal model, can be expressed as a density with respect to the Lebesgue measure in the coordinate space. The model is formally defined using generic orthonormal coordinates $h(\mathrm{x})$, which provide an isometric and subcompositionally coherent framework. We further formulate the model in clr and alr coordinates, demonstrating that these representations yield equivalent models, while discussing the particularities of each logratio representation.

This extension introduces heavier tails and increased robustness against outliers, key advantages for real-world compositional data analysis. Furthermore, our formalization clarifies subtle but essential differences in parameter interpretation and estimation arising from the Student's *t* structure.

Hence, the proposed logratio *t* model provides a theoretically sound and practically powerful alternative to existing approaches, preserving the benefits of the Aitchison geometry while enhancing flexibility and inference capabilities.

A random composition $\mathrm{X} \in \mathcal{S}^D$ has a logratio *t* distribution with location $\boldsymbol{\mu} \in \mathbb{R}^d$, with $d = D - 1$, scale matrix $\boldsymbol{\Upsilon}$ and ν degrees of freedom, if it has density function

$$f^a(\mathrm{x}) = \frac{dP_a}{d\lambda_a}(\mathrm{x}) = \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \left(1 + \frac{1}{\nu}(h(\mathrm{x}) - \boldsymbol{\mu})'\boldsymbol{\Upsilon}^{-1}(h(\mathrm{x}) - \boldsymbol{\mu})\right)^{-\left(\frac{\nu+d}{2}\right)}, \quad (9)$$

where $h(\cdot)$ stands for a generic orthonormal coordinates and P_a is the logratio *t* probability measure. This distribution will be denoted by $\mathrm{X} \sim t_{\mathcal{S}}^D(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$. As for the normal case, the key point is to define a density on the simplex with respect to λ_a, which, when expressed in coordinates, yields a standard density with respect to the Lebesgue measure in the space of coordinates.

Note that $\boldsymbol{\mu}$ and $\nu(\nu-2)^{-1}\boldsymbol{\Upsilon}$ (when $\nu > 2$) denote the location and variance-covariance matrix of $h(\mathrm{x})$. A change of logratio representation induces the same probability law on the simplex, with parameters transforming accordingly. Specifically, for alr coordinates, $\mathrm{alr}(\mathrm{x}) = \mathrm{B}\,h(\mathrm{x})$, so the location and scale matrix are transformed as $\boldsymbol{\mu}^* = \mathrm{B}\boldsymbol{\mu}$ and $\boldsymbol{\Upsilon}^* = \mathrm{B}\boldsymbol{\Upsilon}\mathrm{B}'$, while the degrees of freedom remain invariant ($\nu^* = \nu$).

The clr representation yields an equivalent but degenerate model, with location and scale parameters $\mathrm{U}\boldsymbol{\mu}$ and $\mathrm{U}\boldsymbol{\Upsilon}\mathrm{U}'$ respectively. In particular, matrix $\mathrm{U}\boldsymbol{\Upsilon}\mathrm{U}'$ is a $D \times D$ singular matrix of rank d, which requires the use of the Moore–Penrose pseudoinverse and pseudo-determinant in the expression of the density func-

tion in clr coordinates. Here, it should be noted that clr coordinates lie in $\mathbb{R}^D$, but the model is degenerate and supported on the $(D-1)$-dimensional clr subspace $\{\mathrm{clr}(\mathrm{x}) \in \mathbb{R}^D : \sum_i \mathrm{clr}_i(\mathrm{x}) = 0\}$ where the $(D-1)$-dimensional Lebesgue measure should be considered.

We note that all three representations — clr, ilr and alr — are equivalent for the purposes of distribution fitting and outlier detection: the Mahalanobis distances between observations are numerically identical across transformations (maximum difference $< 5 \times 10^{-14}$), as verified on the Kola dataset. The clr has the practical advantage of a unique, simple definition, $\mathrm{clr}_i(\mathrm{x}) = \log(x_i/g(\mathrm{x}))$, with direct component-wise interpretability, whereas ilr coordinates depend on the choice of basis. However, the clr representation has an important limitation: it lacks subcompositional coherence, as its coordinates change non-monotonically when a subcomposition is taken, because the geometric mean in the denominator changes with the number of parts. As a consequence, the clr covariance structure depends on the subcomposition considered — for instance, for a 2-part composition the clr correlation is always -1 by construction, regardless of the actual dependence structure in the data (Egozcue et al. 2003). The ilr, by contrast, can be constructed to be subcompositionally coherent, and the marginal distributions of subcompositions correspond directly to lower-dimensional logratio t distributions (Property 3.6). The choice of representation is therefore guided by the goals of the analysis: the clr is preferable for its simplicity and interpretability, while the ilr is preferable when subcompositional coherence and isometry are required. We note that subcompositional coherence of the ilr depends on the ordering of the parts: only the pivots involving parts of the subcomposition, defined in the same order, are coherent. This can always be achieved by reordering the parts so that the subcompositional parts appear last in the sequential binary partition.

Using the inverse of the Jacobian (7), we can change the measure by applying the Radon-Nikodym chain rule and express this density function with respect to the Lebesgue measure λ in $\mathbb{R}^{D-1}$. The resulting expression is

$$f(\mathrm{x}) = \frac{dP}{d\lambda}(\mathrm{x}) = \frac{\Gamma(\frac{\nu+d}{2})(\sqrt{D}\; x_1 x_2 \cdots x_D)^{-1}}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \left(1 + \frac{1}{\nu}(h(\mathrm{x}) - \boldsymbol{\mu})' \boldsymbol{\Upsilon}^{-1} (h(\mathrm{x}) - \boldsymbol{\mu})\right)^{-\left(\frac{\nu+d}{2}\right)}, \quad (10)$$

Although this change of measure defines the same probability law, it can produce significant changes in characteristic values, such as the expected value or the mode. The effect of changing the measure is also evident in the isodensity curves, where multimodality arises with the Lebesgue measure (see Mateu-Figueras et al. (2021) for a discussion). This leads us always to use the Aitchison measure, or, equivalently, the Lebesgue measure in the space of coordinates

The main properties of this model are as follows. A complete proof of each property can be found in Appendix B. The derivations build on established results in compositional data analysis but are adapted to our new framework. Unless otherwise stated, all properties are formulated in terms of a generic orthonormal coordinates $h(\mathrm{x})$, which provide an isometric and subcompositionally coherent framework. Equivalent results for parameter estimation and Mahalanobis distances are obtained using clr or alr coordinates under the appropriate parametrization (see the considerations following equation (9)). Properties involving inner products, Aitchison dis-

tances, or subcompositions (Properties 3.5 and 3.6) are stated for ilr coordinates and may require additional considerations for clr or alr representations.

Property 3.1 (Location) If $\nu = 1$, the expected value of $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ is undefined. When $\nu > 1$, the mode and the expected value of $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ with respect to the measure λ_a coincide and are

$$\mathrm{mode}_a(\mathrm{X}) = \mathrm{E}_a(\mathrm{X}) = h^{-1}(\boldsymbol{\mu}), \tag{11}$$

independently of the orthonormal coordinates $h(\mathrm{x})$.
Due to the equivariance of the location parameter under linear transformations (see Property 2.3), the same composition $h^{-1}(\boldsymbol{\mu})$ is obtained regardless of the choice of logratio coordinates. Specifically, if $\boldsymbol{\mu}$ is the location in ilr and $\mathrm{B}\boldsymbol{\mu}$ is the location in alr, both map back to the same composition in $\mathcal{S}^D$ via their respective inverse transformations; the same holds for the clr, where $\mathrm{U}\boldsymbol{\mu}$ is a D-vector and $h^{-1}(\boldsymbol{\mu}) = \mathrm{clr}^{-1}(\mathrm{U}\boldsymbol{\mu}) = \mathcal{C}(\exp(\boldsymbol{\mu}))$. However, the mode and the expected value depend on the reference measure; this means that, depending on the measure we consider, λ_a or λ, their values will change. For a composition $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ there is no closed form for $\mathrm{E}(\mathrm{X})$ with respect to the Lebesgue measure λ. Although the integral expression exists, it is not reducible to any simple form, and it is necessary to use numerical integration to obtain it. Note that this result concerns the choice of reference measure (λ_a vs. λ), and holds independently of the logratio representation used (clr, ilr or alr). Importantly, it is worth noting that, in compositional data analysis, the value $\mathrm{E}(\mathrm{X})$ with respect to λ is not used in practice, and the alternative center is recommended. This center coincides with $\mathrm{E}_a(\mathrm{X})$ (Mateu-Figueras et al. 2013). For the mode, a multi-modality can be obtained considering $\mathrm{mode}(\mathrm{X})$ with respect to λ, whereas the unique value $h^{-1}(\boldsymbol{\mu})$ (see equation 11) is always obtained by considering λ_a.

The effect of the change of measure can be observed in the isodensity curves in Figs. 1 and 2 for different parametrizations. The multi-mode is due to the Jacobian, which expresses the ratio of the Lebesgue and the Aitchison measures over the simplex when the Lebesgue measure is considered. On the contrary, considering the Aitchison measure, we observe a unique mode in both cases.

Property 3.2 (Metric variance) Given $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$, when $\nu > 2$, then $\mathrm{Mvar}(\mathrm{X}) = \frac{\nu}{\nu-2}\mathrm{trace}\,(\boldsymbol{\Upsilon})$.

The same metric variance is obtained when considering clr coordinates, since $\mathrm{trace}\left(\mathrm{U}\boldsymbol{\Upsilon}\mathrm{U}'\right) = \mathrm{trace}\,(\boldsymbol{\Upsilon})$. A different metric variance is obtained when non-orthonormal coordinates are considered. Specifically, in the case of alr coordinates $\mathrm{trace}\left(\mathrm{B}\boldsymbol{\Upsilon}\mathrm{B}'\right) \neq \mathrm{trace}\,(\boldsymbol{\Upsilon})$ because B is a non-orthogonal matrix and consequently $\mathrm{B}' \neq \mathrm{B}^{-1}$ (Mateu-Figueras et al. 2021).

Property 3.3 (Perturbation and powering) Let $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$, $\mathrm{b} \in \mathcal{S}^D$ and $c \in \mathbb{R}$, the D-part composition $\mathrm{X}^* = \mathrm{b} \oplus (c \odot \mathrm{X}) \sim t_{\mathcal{S}}^{D}(h(\mathrm{b}) + c\boldsymbol{\mu}, c^2\boldsymbol{\Upsilon}, \nu)$.

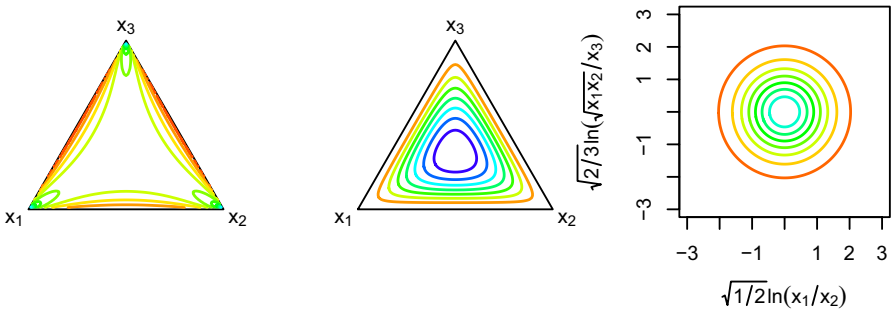

Fig. 1 Isodensity levels of a logratio t distribution with $\boldsymbol{\mu} = (0, 0)$, $\boldsymbol{\Upsilon} = \mathrm{I}$, and $\nu = 10$ with respect to the Lebesgue measure (left), the Aitchison measure (center), and in orthonormal coordinates (right)

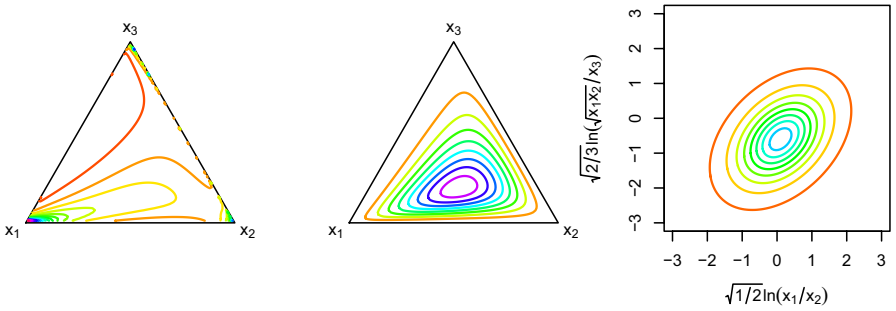

Fig. 2 Isodensity levels of a logratio t distribution with $\boldsymbol{\mu} = (0.1, -0.6)$, $\boldsymbol{\Upsilon} = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}$, and $\nu = 3$ with respect to the Lebesgue measure (left), the Aitchison measure (center), and in orthonormal coordinates (right)

Property 3.4 (Perturbation invariance) Let $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$, and $\mathrm{b} \in \mathcal{S}^{D}$. Then the density of the perturbed composition $\mathrm{b} \oplus \mathrm{X}$ satisfies $f^{a}_{\mathrm{b}\oplus\mathrm{X}}(\mathrm{b} \oplus \mathrm{x}) = f^{a}_{\mathrm{X}}(\mathrm{x})$, that is, the logratio t distribution is equivariant under perturbation with respect to the Aitchison measure.

Property 3.5 (Permutation) Let $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$, and $\mathrm{X}_P = \mathrm{PX}$ the random composition X with the parts reordered by a permutation matrix P. Then, $\mathrm{X}_P \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}_P, \boldsymbol{\Upsilon}_P, \nu)$ with $\boldsymbol{\mu}_P = \mathrm{U}'\mathrm{PU}\boldsymbol{\mu}$; $\boldsymbol{\Upsilon}_P = (\mathrm{U}'\mathrm{PU})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{P}'\mathrm{U})$ where U is a $D \times (D-1)$ matrix with the clr coordinates of an orthonormal basis of $\mathcal{S}^{D}$ as columns.

Property 3.6 (Subcomposition) Let $\mathrm{X} \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$, and $\mathrm{X}_S = \mathcal{C}(\mathrm{SX})$ the C-part subcomposition obtained from the $C \times D$ selection matrix S. Then $\mathrm{X}_S \sim t_{\mathcal{S}}^{D}(\boldsymbol{\mu}_S, \boldsymbol{\Upsilon}_S, \nu)$ with $\boldsymbol{\mu}_S = \mathrm{U}^{*'}\mathrm{SU}\boldsymbol{\mu}$; $\boldsymbol{\Upsilon}_S = (\mathrm{U}^{*'}\mathrm{SU})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{S}'\mathrm{U}^{*})$ where U is a $D \times (D-1)$ contrast matrix with the clr coordinates of an orthonormal basis of $\mathcal{S}^{D}$ as columns and U^{*} is a $C \times (C-1)$ matrix with the clr coordinates of an orthonormal basis of $\mathcal{S}^{C}$ as columns.

The property 3.3 states that the logratio *t* distribution is a closed family of distributions under perturbation and powering. The property 3.4 states that it is equivariant under perturbation. This property does not hold for the density with respect to the Lebesgue measure.

Properties 3.5 and 3.6 show that the logratio *t* family is closed under permutation and subcompositions. These properties are stated for ilr coordinates and involve the matrix U of clr coordinates of an orthonormal basis; additional considerations are required for other representations.

Property 3.7 (Relation with the logratio normal) The logratio *t* distribution includes the logratio normal as a limiting case ($\nu \to \infty$).

The derivation of Property 3.7 follows the classical analogy between multivariate normal and *t* distributions, but its interpretation and implications in the compositional domain are novel.

Given a compositional data set, the MLE of the parameters can be obtained by applying the standard procedure for the multivariate *t* distribution to any chosen logratio coordinates. The estimates for $\boldsymbol{\mu}$ and $\boldsymbol{\Upsilon}$ in different representations are related through the corresponding change-of-basis transformation, while the degrees of freedom ν remain invariant across all logratio transformations.

We emphasize that while the location of the maximum of the log-likelihood is preserved across transformations, its absolute value depends on the choice of coordinates. Specifically, the log-likelihood is invariant under orthogonal transformations (such as changes of ilr basis) or when using the clr with the pseudo-determinant, but differs by a constant when moving to non-orthonormal coordinates such as the alr.

Let $\mathrm{x}_1, \ldots, \mathrm{x}_n$ be an independent sample. Considering ilr coordinates $h(\mathrm{x})$ and alr coordinates, and using the relationship $\mathrm{alr}(\mathrm{x}) = \mathrm{B}h(\mathrm{x})$, the corresponding log-likelihood functions (omitting constant terms) satisfy:

$$\ell(\boldsymbol{\mu}^*, \boldsymbol{\Upsilon}^*, \nu \mid \mathrm{alr}(\mathrm{x})) = -\frac{n}{2}\ln(D) + \ell(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu \mid h(\mathrm{x})) \tag{12}$$

where $\boldsymbol{\mu}^* = \mathrm{B}\boldsymbol{\mu}$ and $\boldsymbol{\Upsilon}^* = \mathrm{B}\boldsymbol{\Upsilon}\mathrm{B}'$, and the term $-\frac{n}{2}\ln(D)$ arises from $|\mathrm{B}| = \sqrt{D}$. In the clr case, the log-likelihood coincides with that of the ilr representation because the pseudo-determinant replaces the standard determinant. Overall, any differences across logratio representations either vanish or reduce to constants depending only on *D*, and thus have no impact on parameter estimation or the comparison of nested models.

3.4 Diagnostic Measures and Outlier Detection Thresholds

In the context of robust compositional data analysis, several approaches have been proposed to mitigate the influence of atypical observations. Methods such as robust principal component analysis (RPCA) (Filzmoser et al. 2009) and robust covariance estimators, including the Minimum Covariance Determinant (MCD) (Rousseeuw and Driessen 1999), are commonly employed to detect outliers by focusing on the

central structure of the data distribution. However, these methods often rely on hard rejection rules or dimensionality reduction that may overlook the structural nature of heavy-tailed distributions. Our proposed Student's *t* model addresses these limitations by providing a flexible probabilistic framework. Unlike RPCA, which identifies outliers based on projection residuals, the Student's *t* approach explicitly models the data's kurtosis through the degrees of freedom parameter. This allows for a more principled distinction between structural heavy-tailedness and true anomalous behavior. By accommodating observations that would be erroneously flagged by rigid normal-based estimators (swamping effect), the Student's *t* model provides a more reliable identification of "strong outliers"-those observations that remain extreme even when compared against a high-kurtosis baseline.

To translate this probabilistic framework into a practical outlier-detection tool, we require a metric that quantifies the distance between an observation and the estimated center of the heavy-tailed distribution. In multivariate analysis, a natural measure for this purpose is the Mahalanobis distance (D_M), which accounts for the data's covariance structure. For a composition x expressed in ilr coordinates $h(\mathrm{x})$, the squared Mahalanobis distance is defined as

$$D_M^2 = (h(\mathrm{x}) - \boldsymbol{\mu})' \boldsymbol{\Upsilon}^{-1} (h(\mathrm{x}) - \boldsymbol{\mu}),$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Upsilon}$ represent the location and scale parameters. Unlike the Euclidean distance, D_M^2 accounts for correlations between parts, remains scale-invariant, and reflects the underlying covariance structure of the compositional data. An equivalent expression is obtained using clr coordinates, replacing $\boldsymbol{\Upsilon}^{-1}$ with the Moore–Penrose pseudoinverse $\boldsymbol{\Upsilon}^{+}$; the resulting distances are numerically identical (maximum difference $< 5 \times 10^{-14}$ on the Kola dataset). For alr coordinates, the distance is computed using $(\boldsymbol{\Upsilon}^{*})^{-1}$ where $\boldsymbol{\Upsilon}^{*} = \mathrm{B}\boldsymbol{\Upsilon}\mathrm{B}'$, and yields identical results by the equivariance of the Mahalanobis distance under linear transformations.

A fundamental property of D_M^2 is its invariance under any logratio transformation: the clr, ilr and alr representations yield numerically identical distances (Filzmoser and Hron 2008; Tolosana-Delgado et al. 2019). This invariance ensures that all likelihood-based inferences — parameter estimates, outlier classifications, and AIC comparisons — are consistent regardless of the chosen coordinate system. The practitioner is therefore free to use whichever representation is most convenient or interpretable for the problem at hand.

In the multivariate normal framework, D_M^2 corresponds to the quadratic form in the exponent of the density function (8) and follows a chi-squared distribution with $d = D - 1$ degrees of freedom. Consequently, an observation is flagged as a potential outlier if $D_M^2 > \chi^2_{d,1-\alpha}$, where α is the significance level (typically 0.05).

We propose here a more robust diagnostic approach based on the Student's *t* model. For the logratio *t* model, the diagnostic statistic is defined as δ_i^2/d, where δ_i^2 is the squared distance calculated using the *t*-distribution parameters. This statistic, according to characterization 2.2, follows an *F*-distribution with *d* and ν degrees of freedom, independently of the logratio representation used:

$$\frac{\delta_i^2}{d} \sim F_{d,\nu}$$

An observation is therefore identified as atypical if $\delta_i^2 > d \cdot F_{d,\nu,1-\alpha}$. This threshold is invariant across logratio representations, since δ_i^2 is numerically identical whether computed via ilr, alr, or clr coordinates.

4 Simulation study

In this section, we present a simulation study designed to evaluate the effectiveness of our proposed outlier detection framework. The primary objective is to investigate the performance of the squared Mahalanobis distance based on the Student's *t* distribution as a diagnostic measure for atypicality in compositional data. By leveraging the heavy-tailed properties of the *t*-model and the theoretical *F*-distribution of its associated quadratic forms, we aim to demonstrate that this approach provides a more reliable balance between sensitivity and specificity compared to traditional normal-based and combinatorial methods, particularly in the presence of scale contamination. As established in Sect. 3.4, all results are invariant under any logratio transformation (clr, ilr, alr); ilr coordinates are used throughout this section as they provide full-rank input required by all competing methods (MCD, COMCoDa, and CN).

Following the experimental framework of Divino et al. (2026), we conducted the study under various conditions by varying three key experimental parameters: the number of observations ($n \in \{100, 1000\}$), the number of components ($D \in \{3, 5\}$), and the proportion of typical observations ($\gamma \in \{0.9, 0.6\}$). Each experimental condition was evaluated over 100 independent repetitions.

4.1 Data Generation and Contamination

Typical compositions were simulated using a multivariate normal density on the simplex with respect to the Aitchison measure λ_a following Mateu-Figueras et al. (2013), with location $\boldsymbol{\mu}$ randomly generated for each repetition using the `Randvec` function from the `Surrogate` package and projected onto the simplex $\mathcal{S}^D$ and variance-covariance matrix $\boldsymbol{\Upsilon}$. Unlike the uniform noise contamination used in Divino et al. (2026), we focus on a pure scale contamination scenario to assess robustness against increased dispersion in the coordinates. Outliers were generated in the ilr-space according to the proportion $(1 - \gamma)$ as

$$\text{ilr}(\text{y}_{\text{out}}) \sim N(\boldsymbol{\mu}, 25 \cdot \boldsymbol{\Upsilon})$$

This setup represents a challenging scenario where anomalous observations share the same location as typical data but exhibit significantly higher variability, a common occurrence in compositional datasets with heavy-tailed behavior.

4.2 Evaluation

In our simulation study, we evaluate the performance of several outlier detection frameworks, with a specific focus on the integration of Leave-One-Out (LOO) strategies for parametric models. This approach is designed to mitigate the *masking effect*, in which outliers exert undue influence on the estimation of location and scale, thereby shrinking their own distance from the center.

Within the parametric framework, we primarily focus on the Student's *t* model and the Normal model, implemented via a Leave-One-Out procedure and hereafter denoted as $\mathrm{T_{LOO}}$ and $\mathrm{N_{LOO}}$, respectively. For the $\mathrm{T_{LOO}}$ approach, for each observation *i*, the location, scale matrix, and ν degrees of freedom are iteratively estimated from the remaining $(n-1)$ observations via the Expectation-Conditional Maximization (ECM) algorithm (Liu and Rubin 1995). The anomaly score is then computed as the Mahalanobis distance of the excluded point from this 'clean' reference. Following the same logic, the $\mathrm{N_{LOO}}$ benchmark is implemented to ensure that, for both models, potential outliers do not distort the parameters used to identify them.

These methods were compared against the following established approaches: Fast Minimum Covariance Determinant (MCD) (Rousseeuw and Driessen 1999), a robust estimator of location and scale applied to the ilr-coordinates; COMCoDa (Palma and Gallo 2016), a robust estimator specifically tailored for compositional data; Contaminated Normal (CN) (Divino et al. 2026), a mixture-based approach that identifies outliers by estimating a two-component model (typical vs. contaminated); and Atypicality Index (Aitchison 1986, Section 7.10).

Crucially, the Atypicality Index is also inherently a LOO diagnostic tool, as it assesses the probability that a specific observation *i* belongs to the reference normal population estimated from the other $(n-1)$ observations. For each point, the index is calculated using the exact Beta distribution:

$$A_i = P\left[\mathrm{Beta}\left(\frac{d}{2}, \frac{n-d}{2}\right) \leq \frac{q_i}{q_i + n - 1}\right],$$

where q_i represents the scaled Mahalanobis distance computed by excluding the *i*-th observation, as defined in the following:

$$q_i = \frac{1}{1+n^{-1}}(\mathbf{y_i} - \hat{\boldsymbol{\mu}}_{-i})^T \hat{\boldsymbol{\Upsilon}}_{-i}^{-1}(\mathbf{y_i} - \hat{\boldsymbol{\mu}}_{-i}),$$

where $\hat{\boldsymbol{\mu}}_{-i}$ and $\hat{\boldsymbol{\Upsilon}}_{-i}$ are the sample mean and scale matrix estimated from the remaining $(n-1)$ observations. This construction ensures that q_i follows the required distribution for the Atypicality Index and prevents the *i*-th observation from masking its own outlyingness.

To ensure consistency, a 95% confidence threshold was adopted. Specifically, for methods assuming multivariate normality in the ilr space (MCD, COMCoDa, and $\mathrm{N_{LOO}}$), observations were classified as outliers if their squared Mahalanobis distances exceeded the 0.95 quantile of the χ^2_d distribution. For the $\mathrm{T_{LOO}}$ the threshold was adjusted using the *F*-distribution ($d \cdot F_{d,\nu,0.95}$), while for the Contaminated Nor-

mal model, classification was based on the posterior probability of belonging to the contaminated component, using a cut-off of 0.5.

4.3 Performance Metrics

To provide a comprehensive evaluation, we calculated several performance metrics: the Area Under the ROC Curve (AUC), Sensitivity (the proportion of true outliers correctly identified), Specificity (the proportion of typical observations correctly identified), Positive Predictive Value (PPV) (the precision of the detector in terms of true outliers among those flagged), and Negative Predictive Value (NPV). Outliers for the $\mathrm{T}_{\mathrm{LOO}}$ method were identified using a threshold based on the F-distribution, $d_i^2 > d \cdot F_{d,\nu,\,1-\alpha}$, while the $\mathrm{N}_{\mathrm{LOO}}$ approach employed the $\chi^2_{d,\,1-\alpha}$ distribution. Finally, the Atypicality Index was evaluated against the $1-\alpha$ quantile of the appropriate Beta distribution.

4.4 Simulation results

Our comparative analysis, summarized in Table 1, demonstrates that the parametric framework-comprising $\mathrm{T}_{\mathrm{LOO}}$, $\mathrm{N}_{\mathrm{LOO}}$, and the Atypicality Index-provides a robust safeguard against false positives, consistently maintaining near-perfect specificity ($Spec \approx 1.000$).

A crucial finding concerns the masking effect under high contamination ($\gamma = 0.6$). In these scenarios, the high proportion of outliers inflates the global scale estimates, pushing detection thresholds upward and reducing the sensitivity of all parametric models. However, $\mathrm{T}_{\mathrm{LOO}}$ shows a consistent edge in ranking accuracy (AUC) and sensitivity compared to the $\mathrm{N}_{\mathrm{LOO}}$ benchmarks as the sample size increases ($n = 1000$). This confirms that modeling heavy-tailed distributions effectively enhances the contrast between the bulk of the data and scale-based anomalies.

The Atypicality Index and $\mathrm{N}_{\mathrm{LOO}}$ show nearly identical performance, confirming that for large samples, the probability-thresholding of the former converges to the standard Mahalanobis distance results. Notably, the Contaminated Normal (CN) model emerges as the most balanced alternative, successfully isolating scale contamination with higher specificity than the MCD.

Overall, $\mathrm{T}_{\mathrm{LOO}}$ excels in high-dimensional settings ($p = 5$). While MCD and COMCoDa show a decline in ranking accuracy, the Student's t approach achieves the highest AUC (0.975 for $n = 1000, p = 5, \gamma = 0.6$). This superior performance stems from the ECM algorithm's ability to dynamically adjust the degrees of freedom (ν), effectively absorbing scale contamination into the heavy tails without distorting the location and scale estimates.

Furthermore, a standout feature of the Student's t and Atypicality Index models is their perfect specificity ($Spec = 1.000$) across nearly all simulated scenarios. In applied contexts such as geochemistry or environmental monitoring, maintaining a zero or near-zero false-alarm rate is often as critical as detecting anomalies. While robust distance-based methods (MCD, COMCoDa) achieve higher sensitivity, their tendency to misclassify typical observations as outliers limits their reliability in

Table 1 Simulation results for outlier detection under scale contamination ($25 \cdot \Upsilon$). Comparison between parametric methods (T_{LOO}, N_{LOO}, and Atypicality Index) and robust/mixture models (MCD, COMCoDa, CN)

n	p	γ	T_{LOO}					N_{LOO}					Atypicality Index				
			AUC	Sens	Spec	PPV	NPV	AUC	Sens	Spec	PPV	NPV	AUC	Sens	Spec	PPV	NPV
100	3	0.9	.969	.667	1.000	1.000	.968	.966	.667	1.000	1.000	.968	.966	.667	1.000	1.000	.968
		0.6	.892	.325	1.000	1.000	.690	.897	.300	1.000	1.000	.682	.897	.300	1.000	1.000	.682
	5	0.9	.965	.778	1.000	1.000	.978	.966	.778	1.000	1.000	.978	.966	.778	1.000	1.000	.978
		0.6	.958	.150	1.000	1.000	.638	.935	.375	1.000	1.000	.706	.935	.300	1.000	1.000	.682
1000	3	0.9	.963	.687	1.000	1.000	.967	.963	.687	1.000	1.000	.967	.963	.687	1.000	1.000	.967
		0.6	.956	.263	1.000	1.000	.670	.955	.280	1.000	1.000	.676	.955	.278	1.000	1.000	.675
	5	0.9	.978	.808	1.000	1.000	.979	.973	.778	1.000	1.000	.976	.973	.778	1.000	1.000	.976
		0.6	.975	.298	1.000	1.000	.681	.958	.350	1.000	1.000	.698	.958	.335	1.000	1.000	.693
			MCD					COMCoDa					CN				
100	3	0.9	.966	.778	.923	.500	.977	.967	.889	.923	.533	.988	.968	.667	.989	.857	.968
		0.6	.899	.800	.983	.970	.881	.894	.725	1.000	1.000	.845	.897	.700	1.000	1.000	.833
	5	0.9	.861	1.000	.725	.265	1.000	.886	1.000	.714	.257	1.000	.967	.778	1.000	1.000	.978
		0.6	.850	.900	.750	.706	.918	.950	.500	1.000	1.000	.750	.911	.975	.733	.709	.978
1000	3	0.9	.963	.878	.970	.763	.986	.963	.899	.958	.701	.989	.963	.808	.997	.964	.979
		0.6	.957	.798	.997	.994	.881	.956	.765	.998	.997	.864	.956	.820	.987	.976	.892
	5	0.9	.872	.949	.715	.268	.992	.874	.949	.713	.266	.992	.980	.879	1.000	1.000	.987
		0.6	.869	.968	.750	.721	.972	.978	.725	1.000	1.000	.845	.982	.915	.997	.995	.946

decision-making processes where false positives carry significant economic or legal consequences. In contrast, the Student's t model offers a more "conservative" and reliable diagnostic tool, providing an optimal balance between anomaly detection and the preservation of typical data integrity.

5 An application to a real data set

The practical applicability of the logratio t distribution can be more effectively demonstrated through an analysis of a real data set. For this purpose, we will consider a well-known dataset in geochemistry and environmental science, used for compositional data analysis: the Kola C-horizon soil data set (Reimann et al. 2008). It was collected as part of the Kola Ecogeochemistry Project, an international research project carried out in the 1990s. The primary aim of the project was to investigate the environmental impacts of heavy industrialization, particularly mining and smelting, in the Kola Peninsula region of Russia, as well as in parts of Finland and Norway. The dataset includes the concentrations of more than 50 chemical elements in about 600 soil samples. The data set is available in the R library `StatDA` (R Core Team 2024).

5.1 Model Selection and Distributional Fit

In this section, we evaluate the effectiveness of the logratio Student's t distribution as a robust model for compositional data. We compare the fitting performance of the Student's t against the classical logratio normal model by analyzing density estimates and using the Akaike Information Criterion (AIC). As established in Sect. 3.4, all logratio transformations yield numerically identical results; we therefore present results for the clr, ilr and alr on equal footing. For the clr, the singular covariance matrix is handled via the Moore–Penrose pseudoinverse and pseudo-determinant (implemented in `fit_mvt_clr()`). For the ilr, we use the sequential Helmert contrast basis produced by the `ilr()` function of the `compositions` package, whose j-th coordinate is defined as

$$h_j(\mathrm{x}) = \sqrt{\frac{j}{j+1}} \log \frac{x_{j+1}}{\left(\prod_{k=1}^{j} x_k\right)^{1/j}}, \quad j = 1, \ldots, D-1,$$

that is, it contrasts the $(j+1)$-th part against the geometric mean of the preceding j parts. For the alr, the last part is used as reference.

We focus our analysis on 12 components that represent the following groups:

- Pollution (P): Co, Cu, Ni
- Seaspray (S): Mg, Na, S
- Contamination (C): As, Bi, Cd, Sb
- Mineralization (M): Ag, Pb

examining how well the logratio *t* distribution fits data exhibiting various patterns and levels of noise (Filzmoser and Hron 2009).

Considering the 3-part subcomposition related to the Pollution elements, Fig. 3 shows the sample points together with the constant-density contours containing 50%, 95%, and 99% of the probability under the logratio normal and logratio *t* distributions. The top row shows results based on the ilr representation — ternary diagram (top-left) and ilr coordinate space (top-right) — while the bottom row shows the corresponding results for the clr representation — ternary diagram (bottom-left) and clr coordinates projected onto the SVD plane (bottom-right). A different fit using the logratio *t* distribution is evident: the inner region is significantly narrower and the

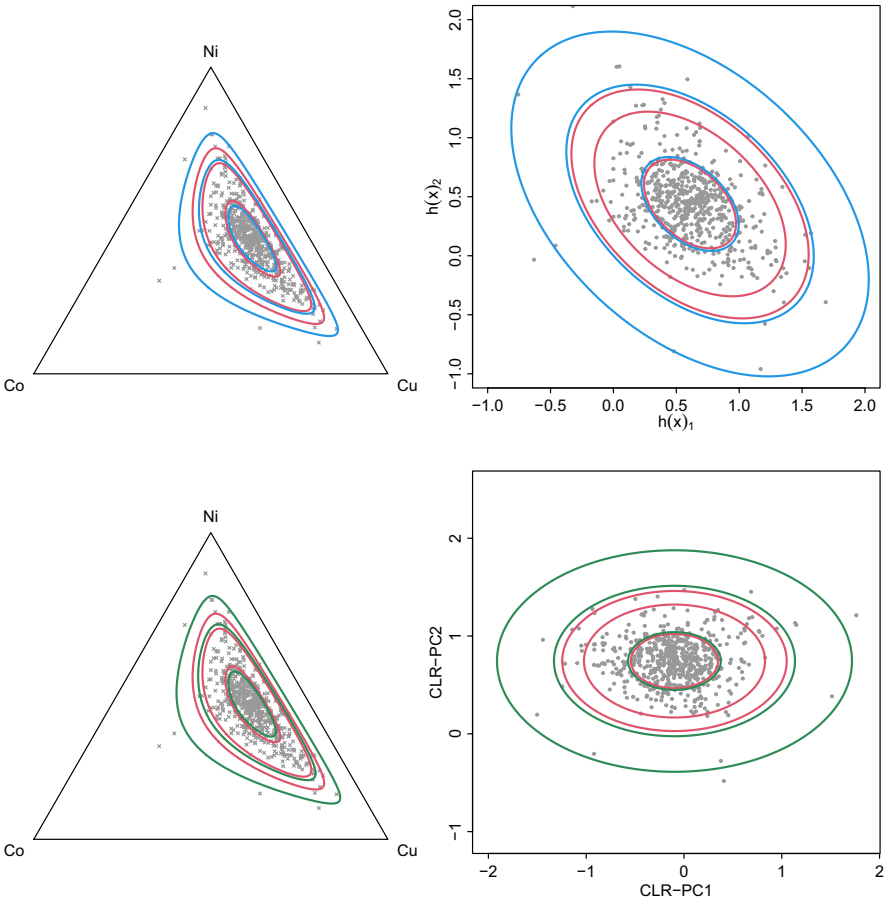

Fig. 3 Analysis of the Pollution subcomposition (Co, Cu, Ni). Top row: ternary diagram (left) and coordinate space (right) for the ilr representation. Bottom row: ternary diagram (left) and coordinate space (right) for the clr representation. All panels show 50%, 95%, and 99% isodensity regions for the logratio Normal (red) and logratio *t* (blue: ilr; seagreen: clr) models. The isodensity curves in the clr coordinate space (bottom-right) are a rigid rotation of those in the ilr space (top-right), reflecting the orthogonal equivalence of the two representations

outer region broader compared to the normal distribution. This pattern is consistent across both representations.

The first coordinate $h(\mathrm{x})_1 = \frac{1}{\sqrt{2}} \log(\mathrm{Cu}/\mathrm{Co})$ represents the log-ratio between Copper(Cu) and Cobalt(Co), while the second coordinate $h(\mathrm{x})_2 = \sqrt{\frac{2}{3}} \log\left(\mathrm{Ni}/\sqrt{\mathrm{Co} \cdot \mathrm{Cu}}\right)$ represents, up to a scalar multiplier, the average of $\log(\mathrm{Ni}/\mathrm{Cu})$ and $\log(\mathrm{Ni}/\mathrm{Co})$. While this interpretation is tractable for this simple 3-part example, it becomes increasingly opaque for larger compositions — a limitation that further motivates the use of the clr, whose coordinates $\mathrm{clr}_i(\mathrm{x}) = \log(x_i/g(\mathrm{x}))$ have a consistent part-specific interpretation regardless of the number of parts, even though they depend on the full composition through the geometric mean.

Isodensity contours for the *d*-dimensional normal distribution and the *d*-dimensional Student's *t* distribution in real space are ellipsoids. This is also true for the logratio normal and logratio *t* distributions on the simplex, with respect to the Aitchison geometry, and holds independently of the logratio representation used. In Fig. 3, the red ellipses correspond to the logratio normal model and have radius $r = \sqrt{\chi^2_{2,\,1-\alpha}}$, where $\chi^2_{2,\,1-\alpha}$ is the $(1-\alpha)$ quantile of a chi-squared distribution with 2 degrees of freedom, enclosing $(1-\alpha) \times 100\%$ of the probability mass. The blue ellipses (ilr panels) and seagreen ellipses (clr panels) correspond to the logratio *t* model and have radius $r = \sqrt{2 f_{2,\nu,\,1-\alpha}}$, where $f_{2,\nu,\,1-\alpha}$ is the $(1-\alpha)$ quantile of an *F* distribution with 2 and ν degrees of freedom, following from characterization 2.2. The isodensity contours in the clr coordinate space (bottom-right panel) are a rigid rotation of those in the ilr space (bottom-left panel), confirming the orthogonal equivalence of the two representations. Note that although clr coordinates are embedded in $\mathbb{R}^3$, the model is degenerate and supported on the 2-dimensional clr subspace, so the same isodensity contours are obtained as in the ilr representation.

Parameters of the two density models, the logratio normal and the logratio *t* distribution, expressed in clr, ilr and alr coordinates, are reported in Tables 2 and 3 alongside likelihood-based goodness-of-fit measures (AIC). The AIC differences between Normal and *t* are consistent across all three representations, confirming the equivalence of the logratio transformations for model comparison purposes.

The superior fit of the logratio Student's *t* distribution is confirmed by the lower AIC values compared to the normal model, consistently across all three representations (Table 3). For the clr, the coordinates represent the log-ratio of each element relative to the geometric mean of the composition: $\mathrm{clr}_i(\mathrm{x}) = \log(x_i/g(\mathrm{x}))$, with a direct geochemical interpretation. For the ilr, the sequential Helmert contrast basis is used (see Sect. 5). For the alr, Nickel (Ni) is used as the reference component, so the coordinates represent $\log(\mathrm{Co}/\mathrm{Ni})$ and $\log(\mathrm{Cu}/\mathrm{Ni})$. As shown in Table 3, the AIC differences between Normal and *t* are identical across all three representations, confirming that model selection is invariant with respect to the choice of logratio transformation. The slight differences in $\hat{\nu}$ across representations are due to the iterative nature of the ECM algorithm, not to any theoretical disagreement. The difference in absolute log-likelihood values between alr and the other representations is accounted for by equation (12).

Table 2 Parameter estimates for the Pollution subcomposition under the logratio Normal and logratio t models

	Normal		Logratio t		
	$\hat{\boldsymbol{\mu}}$	$\widehat{\text{cov}}$	$\hat{\boldsymbol{\mu}}$	$\widehat{\text{cov}}$	$\hat{\nu}$
clr	$\begin{pmatrix} -0.610 \\ 0.252 \\ 0.358 \end{pmatrix}$	$\begin{pmatrix} 0.040 & -0.031 & -0.009 \\ -0.031 & 0.090 & -0.059 \\ -0.009 & -0.059 & 0.068 \end{pmatrix}$	$\begin{pmatrix} -0.610 \\ 0.249 \\ 0.361 \end{pmatrix}$	$\begin{pmatrix} 0.043 & -0.034 & -0.009 \\ -0.034 & 0.092 & -0.058 \\ -0.009 & -0.058 & 0.067 \end{pmatrix}$	5.686
ilr	$\begin{pmatrix} 0.609 \\ 0.438 \end{pmatrix}$	$\begin{pmatrix} 0.096 & -0.043 \\ -0.043 & 0.102 \end{pmatrix}$	$\begin{pmatrix} 0.608 \\ 0.441 \end{pmatrix}$	$\begin{pmatrix} 0.099 & -0.041 \\ -0.041 & 0.098 \end{pmatrix}$	6.292
alr	$\begin{pmatrix} -0.968 \\ -0.107 \end{pmatrix}$	$\begin{pmatrix} 0.126 & 0.105 \\ 0.105 & 0.276 \end{pmatrix}$	$\begin{pmatrix} -0.970 \\ -0.110 \end{pmatrix}$	$\begin{pmatrix} 0.124 & 0.097 \\ 0.097 & 0.266 \end{pmatrix}$	6.508

Table 3 Likelihood-based goodness of fit measures for the Pollution subcomposition under the logratio Normal and logratio t models, for three logratio representations. Comparisons are valid within each coordinate system (Normal vs. t)

	Logratio Normal		Logratio t	
	Log-lik	AIC	Log-lik	AIC
clr	−253.791	511.582	−217.532	441.065
ilr	−253.790	511.580	−217.416	440.833
alr	−586.120	1176.240	−549.786	1105.573

Table 4 Likelihood-based goodness of fit measures for the full Kola composition (12 elements) under the logratio Normal and logratio t models, for three logratio representations. Comparisons are valid within each coordinate system (Normal vs. t)

	Logratio Normal		Logratio t		
	Log-lik	AIC	Log-lik	AIC	$\hat{\nu}$
clr	−4762.549	9547.099	−4517.743	9059.487	8.380
ilr	−4762.545	9547.090	−4517.674	9059.347	8.296
alr	−5502.630	11027.260	−5258.579	10541.160	8.908

When fitting the models to the full composition of 12 elements, the logratio t model consistently provides a better fit than the logratio normal one across all three representations, as shown in Table 4. The AIC differences between the Normal and t models are numerically identical across clr, ilr and alr (approximately 488 units). This confirms that all inferential conclusions — model selection, Mahalanobis distances, and outlier classifications — are invariant with respect to the choice of logratio transformation, as verified empirically on this dataset (maximum difference in Mahalanobis distances $< 5 \times 10^{-14}$). The choice of representation is therefore a matter of interpretability and convenience: the clr offers a unique, directly interpretable definition; the ilr provides full-rank coordinates suitable for algorithms requiring non-singular input; and the alr offers a simple pairwise logratio structure with a chosen reference part.

Figure 4 shows the PCA form biplot (Aitchison and Greenacre 2002) of the whole composition (computed on clr coordinates) with superimposed 50%, 95%, and 99% confidence regions based on the bivariate normal (red lines) and on the bivariate Student's t distributions (blue lines).

The visualization highlights the different tail behavior of the two models: while the 50% confidence regions are practically overlapping, the Student's t ellipses at 95% and 99% are noticeably wider (or more inclusive) than the normal ones. This dual behavior confirms the model's ability to accommodate the heavy-tailed nature of the geochemical data by providing a more flexible fit for distal observations without being distorted by them, a feature that becomes even more evident in the subsequent outlier detection analysis.

To complement the likelihood-based comparison, Fig. 5 presents an envelope analysis for both the Pollution subcomposition ($D = 3$) and the full Kola composition ($D = 12$). For each model, the observed squared Mahalanobis distances are plotted against their theoretical quantiles, together with a 95% simulation envelope obtained from $B = 999$ replications. Under the logratio normal model, the observed

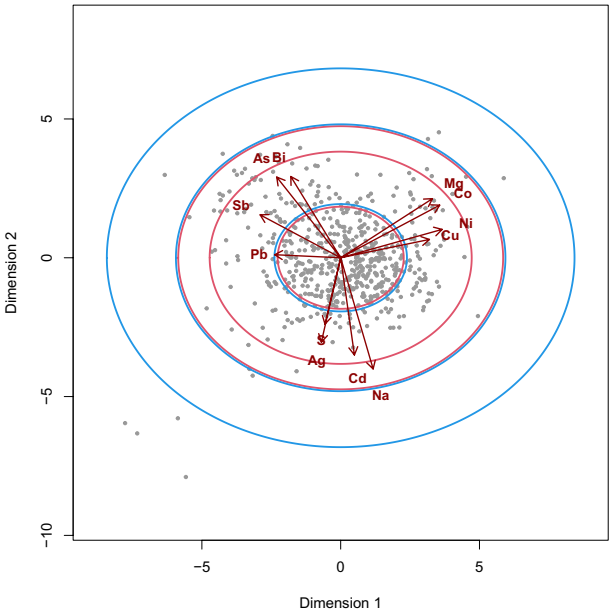

Fig. 4 PCA form biplot of the whole composition with superimposed 50%, 95% and 99% confidence regions based on the normal (red lines) and on the logratio t distributions (blue lines)

distances systematically exceed the envelope in the upper tail, indicating poor fit for extreme observations. Under the logratio t model, the distances fall within the envelope, confirming an adequate fit. This visual assessment corroborates the AIC results and demonstrates that the improvement of the t model is not merely due to the additional degrees of freedom parameter, but reflects a genuine improvement in tail behaviour.

5.2 Robust outlier detection and geochemical validation

Building on the distributional results established in the previous section, we now evaluate the performance of the model in the context of anomaly detection. Regarding the Kola example, considering the whole composition of 12 elements, we have computed the robust distances and the corresponding outlier classification to test the practical reliability of the logratio Student's t distribution. All results reported in this section are invariant with respect to the choice of logratio representation: as verified empirically, the consensus outlier set is identical whether clr, ilr or alr coordinates are used (symmetric difference $= \emptyset$).

Consistent with our simulation study, we differentiated the computational approach based on the properties of each estimator. The $\mathrm{N_{LOO}}$, $\mathrm{T_{LOO}}$, and Atypicality Index were implemented via the LOO scheme to ensure that individual anomalies do not inflate the dispersion estimates. Conversely, for high-breakdown robust methods-MCD, COMCoDa, and the CN model-a full-sample estimation was preferred to leverage their inherent resistance to anomalous clusters. Classification thresholds

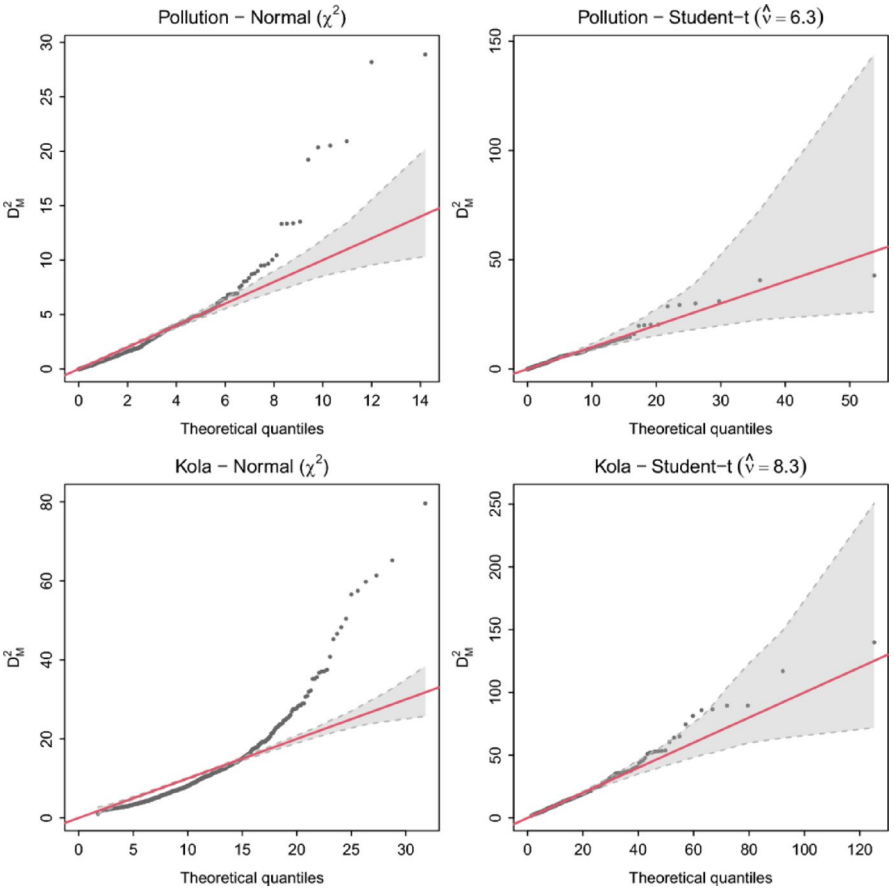

Fig. 5 Envelope analysis for the logratio Normal (left panels) and logratio t (right panels) models. Top row: Pollution subcomposition ($D = 3$); bottom row: full Kola composition ($D = 12$). Observed squared Mahalanobis distances (points) are plotted against theoretical quantiles, with 95% simulation envelopes (shaded bands, $B = 999$ replications)

were established at a 95% significance level. Specifically, we utilized the χ^2 distribution for MCD, COMCoDa, and CN. For $\mathrm{T}_{\mathrm{LOO}}$, the threshold was based on the F-distribution, $d \cdot F_{d,\hat{\nu},0.95}$, to account for the estimated degrees of freedom. This multi-method framework allowed us to identify a consensus subset of outliers, effectively minimizing both masking and swamping effects.

The comparative analysis on the Kola dataset reveals a significant convergence centered around the $\mathrm{T}_{\mathrm{LOO}}$ approach. While methods like MCD and the Contaminated Normal (CN) identified a broader set of atypical observations (141 and 88, respectively), $\mathrm{T}_{\mathrm{LOO}}$ proved to be the most selective, flagging 42 samples (see Table 5). Remarkably, these 42 observations represent a “unanimous consensus" subset:

Table 5 Comparison of outlier detection methods on the Kola dataset. Results are presented for ilr coordinates; identical results are obtained with clr and alr (symmetric difference $= \emptyset$, consensus set of 42 observations unchanged across all representations). The "Unanimous Consensus" column indicates the number of observations simultaneously identified by all five methods. LOO: parameters re-estimated after removing each observation sequentially; Full: robust estimators applied to the entire dataset

Method	Approach	Total Outliers	Unanimous Consensus
MCD	Robust (Full)	141	42
COMCoDa	Robust (Full)	94	42
CN (Contaminated)	Mixture (Full)	88	42
Atypicality Index	Normal (LOO)	63	42
$\mathrm{N_{LOO}}$	Normal (LOO)	70	42
$\mathrm{T_{LOO}}$	Student's *t* (LOO)	**42**	**42**

Table 6 Geochemical characterization of the 42 unanimous outliers compared to the typical population. Results are expressed as group geometric means (G). The table shows the top 5 elements ranked in descending order by their enrichment ratio ($G_{\mathrm{strong}}/G_{\mathrm{typical}}$) to highlight the strongest elemental markers. The outlier set is identical across ilr, clr and alr representations

Element	Symbol	Strong Outlier (G)	Typical (G)	Ratio
Antimony	Sb	0.0360	0.0103	3.51
Arsenic	As	1.59	0.568	2.81
Bismuth	Bi	0.0549	0.0261	2.10
Lead	Pb	3.39	1.89	1.79
Sulfur	S	52.6	29.8	1.76

every sample flagged by $\mathrm{T_{LOO}}$ was simultaneously classified as atypical by all other implemented algorithms. This suggests that $\mathrm{T_{LOO}}$ is particularly effective at filtering out the swamping effect-the misclassification of regular observations as outliers-which appears more pronounced in the MCD results. Consequently, these 42 samples can be considered "strong outliers", representing the most robust geochemical anomalies within the study area, regardless of the specific robust estimator or logratio representation employed.

To gain deeper insight into the geochemical nature of the 42 consensus outliers, we compared their composition with that of the typical population. Given the multiplicative nature of compositional data, we calculated the group geometric means (G), which represent the center of the distribution in the Aitchison simplex. To quantify the geochemical contrast, we computed the enrichment ratio, defined as the ratio of the geometric means of the outlier group and the typical population ($G_{\mathrm{outlier}}/G_{\mathrm{typical}}$). As shown in Table 6, the identified outliers are characterized by a distinct multielemental enrichment. Specifically, Antimony (Sb) and Arsenic (As) exhibit the highest contrast, with concentrations 3.5 and 2.8 times the typical background, respectively. The association with elevated Bismuth (Bi), Lead (Pb), and Sulfur (S) suggests that these anomalies are likely related to localized polymetallic mineralizations or specific environmental contamination events, consistent with established geochemical models of polymetallic mineralization (Grunsky 2010; Grunsky et al. 2013). The

consistency of this geochemical signature across all 42 samples confirms that $\mathrm{T}_{\mathrm{LOO}}$, while being the most conservative method, effectively isolates only those points with such a physical and chemical basis, successfully filtering out stochastic noise. The robustness of the $\mathrm{T}_{\mathrm{LOO}}$ approach is further evidenced by the stability of the degrees of freedom parameter (ν). The estimates are consistent across logratio representations: for the ilr, the full-dataset estimate ($\hat{\nu} \approx 8.296$) is nearly identical to the LOO average ($\bar{\nu}_{\mathrm{LOO}} \approx 8.297$); for the clr, the corresponding values are $\hat{\nu} \approx 8.380$ and $\bar{\nu}_{\mathrm{LOO}} \approx 8.380$, respectively. This remarkable stability indicates that the model's heavy-tailed architecture is inherently resistant to the influence of the 42 identified outliers. Unlike the normal model, where individual extreme observations can significantly distort the covariance structure, the Student's *t* model maintains a consistent distributional shape, ensuring that the detection process is governed by the underlying population characteristics rather than by anomalous clusters.

6 Conclusions

This paper presents the logratio Student's *t* distribution as a robust and geometrically coherent alternative to the classical logratio normal model for compositional data. The model is formally defined in ilr coordinates, which provide an isometric and subcompositionally coherent framework. We demonstrate that, for the purposes of parameter estimation, Mahalanobis distances and outlier detection, equivalent results are obtained using clr coordinates — handled via the Moore–Penrose pseudoinverse and pseudo-determinant — or alr coordinates, under the appropriate parametrization. Our results on the Kola dataset confirm that the Student's *t* model consistently outperforms the normal distribution in accommodating heavy-tailed geochemical anomalies, regardless of the chosen representation.

The main properties of the logratio *t* distribution follow from those of the multivariate *t* distribution (Aitchison and Dunsmore 1975), adapted to the compositional context. The model displays desirable properties when expressed with respect to the Aitchison measure: unimodality centered at the location parameter, elliptical isodensity contours, and avoidance of the multimodality artefacts arising from using the Lebesgue measure. Furthermore, it retains the heavy-tailed behavior and predictive nature of the multivariate *t*.

The practical superiority of this framework has been demonstrated through both controlled simulations and a real-world geochemical case study. Our simulation results confirm that the Student's *t* model consistently outperforms the normal model in terms of sensitivity and false discovery rate, particularly under heavy-tailed contamination. The inferential and predictive advantages — including Mahalanobis distances and outlier detection results — are equivalent across clr, ilr and alr representations for estimation purposes, as the multivariate Student's *t* is affine equivariant.

From a modelling perspective, the logratio *t* framework is compatible with standard CoDa techniques such as PCA, and offers clearer insights into dependence structures. The clr-based biplot provides a natural visualization of all parts, while the

ilr coordinates ensure subcompositional coherence and isometry for formal inference involving inner products and subcompositions. Furthermore, the remarkable stability of the estimated degrees of freedom ($\hat{\nu}$) during the Leave-One-Out procedure underscores the model's resistance to the masking effect, a common pitfall in outlier detection.

On the Kola dataset, the model's selectivity was validated by a methodological consensus; it successfully isolated a core group of 42 "strong outliers" that were unanimously confirmed by multiple robust approaches (MCD, COMCoDa, and Contaminated Normal). Geochemical validation further reinforces these findings, as the identified outliers exhibit extreme enrichment ratios for key tracers, such as Antimony (3.5x) and Arsenic (2.8x), highlighting the model's ability to distinguish significant anomalies from stochastic noise.

In summary, the logratio *t* distribution enhances the theoretical foundations of robust modeling for compositional data and provides practical advantages for interpretation and implementation. Its consistency across logratio representations and robust properties make it a compelling alternative for inference in the simplex. Future work may explore its use in regression, classification, or Bayesian inference in compositional contexts.

Derivation of the score and expected information for MVT

Given n observations $\mathrm{x}_1, \ldots, \mathrm{x}_n$, the loglikelihood of the parameters $(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ for model (1), ignoring constants, is

$$\ell(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu) = \sum_{i=1}^{n} \Big[-\frac{1}{2} \ln |\boldsymbol{\Upsilon}| - \frac{d}{2} \ln(\nu) + \ln \left(\frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \right) - \frac{\nu+d}{2} \ln \left(1 + \frac{\delta_i^2}{\nu} \right) \Big], \tag{A1}$$

where $\delta_i^2 = (\mathrm{x}_i - \boldsymbol{\mu})' \boldsymbol{\Upsilon}^{-1} (\mathrm{x}_i - \boldsymbol{\mu})$. The score functions are:

$$\begin{aligned}
\frac{\partial \ell}{\partial \boldsymbol{\mu}} &= \sum_{i=1}^{n} \frac{\nu+d}{\nu+\delta_i^2} \boldsymbol{\Upsilon}^{-1} (\mathrm{x}_i - \boldsymbol{\mu}), \\
\frac{\partial \ell}{\partial (\boldsymbol{\Upsilon}^{-1})} &= n\boldsymbol{\Upsilon} - \frac{n}{2} \mathrm{diag}\{\boldsymbol{\Upsilon}\} - \sum_{i=1}^{n} \frac{\nu+d}{\nu+\delta_i^2} \Big[(\mathrm{x}_i - \boldsymbol{\mu})(\mathrm{x}_i - \boldsymbol{\mu})' - \frac{1}{2} \mathrm{diag}\{(\mathrm{x}_i - \boldsymbol{\mu})(\mathrm{x}_i - \boldsymbol{\mu})'\} \Big], \\
\frac{\partial \ell}{\partial \nu} &= -\frac{nd}{2\nu} + \frac{n}{2} \left[\psi\left(\frac{\nu+d}{2} \right) - \psi\left(\frac{\nu}{2} \right) \right] + \frac{1}{2} \sum_{i=1}^{n} \left[\frac{\delta_i}{\nu(\nu+\delta_i^2)} - \ln(1 + \frac{\delta_i^2}{\nu}) \right],
\end{aligned} \tag{A2}$$

where $\psi(x) = \mathrm{d} \ln(\Gamma(x))/\mathrm{d}x$ is the digamma function. As shown in Lange et al. (1989), the expected information matrix of the parameters $(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ is block diagonal with $\boldsymbol{\mu}$ in one block and $(\boldsymbol{\Upsilon}, \nu)$ in another. The blocks are

$$
\begin{aligned}
&\mathrm{E}\Big(-\frac{\partial^2\ell}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}'}\Big) = \frac{n(\nu+d)}{\nu+d+2}\boldsymbol{\Upsilon}^{-1},\\
&\mathrm{E}\Big(-\frac{\partial^2\ell}{\partial\varphi^{(ij)}\partial\varphi^{(kl)}}\Big) = \frac{n(\nu+d)}{2(\nu+d+2)}(2-\xi_{ij})\Big[\boldsymbol{\Upsilon}_{ik}\boldsymbol{\Upsilon}_{lj}+(1-\xi_{kl})\boldsymbol{\Upsilon}_{il}\boldsymbol{\Upsilon}_{kj}\Big]\\
&\qquad\qquad -\frac{n}{2(\nu+d+2)}(2-\xi_{ij})(2-\xi_{kl})\boldsymbol{\Upsilon}_{ij}\boldsymbol{\Upsilon}_{kl},\\
&\mathrm{E}\Big(-\frac{\partial^2\ell}{\partial\varphi^{(ij)}\partial\nu}\Big) = \frac{n}{(\nu+d)(\nu+d+2)}(2-\xi_{ij})\boldsymbol{\Upsilon}_{ij},\\
&\mathrm{E}\Big(-\frac{\partial^2\ell}{\partial^2\nu}\Big) = -\frac{n}{2}\left[\frac{1}{2}\psi'\Big(\frac{\nu+d}{2}\Big)-\frac{1}{2}\psi'\Big(\frac{\nu}{2}\Big)+\frac{d}{\nu(\nu+d)}-\frac{1}{\nu+d}+\frac{\nu+2}{\nu(\nu+d+2)}\right]
\end{aligned}
\tag{A3}
$$

where $[\varphi^{(ij)}]_{d\times d} = \boldsymbol{\Upsilon}^{-1}$, $\psi(x)' = \mathrm{d}^2\ln(\Gamma(x))/\mathrm{d}^2x$ is the trigamma function, and

$$
\xi_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j. \end{cases}
$$

Proofs

This appendix contains proofs of properties stated in Sect. 3.3.

Proof of Property 3.1 The expected value and the mode of a random composition are elements of the sample space $\mathcal{S}^D$. We know that $\mathrm{E}_a(\mathrm{X}) = h^{-1}\mathrm{E}(h(\mathrm{X}))$ where $\mathrm{E}(h(\mathrm{X}))$ is the expected value of a MVT distribution on $\mathbb{R}^{D-1}$. The expected value of a standard multivariate Student's *t* random vector is well-defined only when $\nu > 1$. Considering $h(\mathrm{X}) \sim t_d(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ with $\nu > 1$, we know that $\mathrm{E}(h(\mathrm{X})) = \mathrm{mode}(h(\mathrm{X})) = \boldsymbol{\mu}$. Consequently $\mathrm{E}_a(\mathrm{X}) = \mathrm{mode}_a(\mathrm{X}) = h^{-1}(\boldsymbol{\mu})$. □

Proof of Property 3.2 The metric variance is defined as $\mathrm{Mvar}(\mathrm{X}) = \mathrm{E}(d_a^2(\mathrm{X}, \mathrm{E}_a(\mathrm{X})))$. Pawlowsky-Glahn and Egozcue (2002) show that $\mathrm{Mvar}(\mathrm{X}) = \mathrm{totvar}(\mathrm{X})$, the concept of total variance defined as $\mathrm{totvar}(\mathrm{X}) = \mathrm{trace}(\mathrm{Var}(\mathrm{clr}(\mathrm{X})))$. From the matrix relationship between $\mathrm{clr}(\cdot)$ and $h(\cdot)$ logratio vectors (Egozcue et al. 2003) the equality $\mathrm{trace}(\mathrm{Var}(\mathrm{clr}(\mathrm{X}))) = \mathrm{trace}(\mathrm{Var}(h(\mathrm{X})))$ is obtained. $\mathrm{Var}(h(\mathrm{X}))$ is the covariance matrix of a MVT distribution on $\mathbb{R}^{D-1}$, well-defined when $n > 2$. Considering $h(\mathrm{X}) \sim t_d(\boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$ with $\nu > 2$, we know that $\mathrm{Var}(h(\mathrm{X})) = \frac{\nu}{\nu-2}\boldsymbol{\Upsilon}$. Therefore, $\mathrm{Mvar}(\mathrm{X}) = \frac{\nu}{\nu-2}\mathrm{trace}(\boldsymbol{\Upsilon})$. □

Proof of Property 3.3 Given $\mathrm{X}^* = \mathrm{b} \oplus (c \odot \mathrm{X})$, the orthonormal coordinates of the random composition X^* are easily obtained from the orthonormal coordinates of the composition X via $h(\mathrm{X}^*) = h(\mathrm{b}) + ch(\mathrm{X})$. The density function of $h(\mathrm{X})$ is the standard MVT density in real space (1). Thus, the linear transformation property (Property 2.3) can be used to obtain the density function of $h(\mathrm{X})$. Therefore, $h(\mathrm{X}^*) \sim t^D(h(\mathrm{b}) + c\boldsymbol{\mu}, c^2\boldsymbol{\Upsilon}, \nu)$ and consequently $\mathrm{X}^* \sim t_{\mathcal{S}}^D(h(\mathrm{b}) + c\boldsymbol{\mu}, c^2\boldsymbol{\Upsilon}, \nu)$. □

Proof of Property 3.4 From Property 3.3 we know that $\mathrm{b} \oplus \mathrm{X} \sim t_{\mathcal{S}}^{D}(h(\mathrm{b}) + \boldsymbol{\mu}, \boldsymbol{\Upsilon}, \nu)$. It also holds that $h(\mathrm{b} \oplus \mathrm{X}) = h(\mathrm{b}) + h(\mathrm{X})$; therefore

$$
\begin{aligned}
f^{a}_{\mathrm{b}\oplus\mathrm{X}}(\mathrm{b} \oplus \mathrm{x}) = & \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \\
& \left(1 + \frac{1}{\nu}(h(\mathrm{b} \oplus \mathrm{X}) - (h(\mathrm{b}) + \boldsymbol{\mu}))' \boldsymbol{\Upsilon}^{-1} (h(\mathrm{b} \oplus \mathrm{X}) - (h(\mathrm{b}) + \boldsymbol{\mu}))\right)^{-\left(\frac{\nu+d}{2}\right)} \\
= & \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\nu/2)(\nu\pi)^{d/2}|\boldsymbol{\Upsilon}|^{1/2}} \left(1 + \frac{1}{\nu}(h(\mathrm{x}) - \boldsymbol{\mu})' \boldsymbol{\Upsilon}^{-1} (h(\mathrm{x}) - \boldsymbol{\mu})\right)^{-\left(\frac{\nu+d}{2}\right)} \\
= & f^{a}_{\mathrm{X}}(\mathrm{x})
\end{aligned}
$$

□

Proof of Property 3.5 Given $\mathrm{X}_P = \mathrm{PX}$, it can be readily shown that $\mathrm{clr}(\mathrm{X}_P) = \mathrm{Pclr}(\mathrm{X})$ (Aitchison 1986, p.94). From the matrix relationship between $\mathrm{clr}(\cdot)$ and $h(\cdot)$ logratio vectors (Egozcue et al. 2003) the equality $h(\mathrm{X}_P) = \mathrm{U}'\mathrm{P}\mathrm{U}h(\mathrm{X})$ is obtained. Using that $h(\mathrm{X})$ has a MVT distribution and applying Property 2.3 we conclude that $h(\mathrm{X}_P) \sim t^{D}(\mathrm{U}'\mathrm{P}\mathrm{U}\boldsymbol{\mu}, (\mathrm{U}'\mathrm{P}\mathrm{U})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{P}'\mathrm{U}), \nu)$. Therefore $\mathrm{X}_P \sim t_{\mathcal{S}}^{D}(\mathrm{U}'\mathrm{P}\mathrm{U}\boldsymbol{\mu}, (\mathrm{U}'\mathrm{P}\mathrm{U})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{P}'\mathrm{U}), \nu)$. □

Proof of Property 3.6 Given $\mathrm{X}_S = \mathrm{SX}$, it can be formally derived the matrix relationship between $\mathrm{alr}(\mathrm{X}_S)$ and $\mathrm{alr}(\mathrm{X})$ (Aitchison 1986, p.119). From the matrix relationship between the $\mathrm{alr}(\cdot)$, $\mathrm{clr}(\cdot)$ and $h(\cdot)$ logratio vectors (Egozcue et al. 2003) the equality $h(\mathrm{X}_S) = \mathrm{U}^{*'}\mathrm{S}\mathrm{U}h(\mathrm{X})$ is obtained. Using that $h(\mathrm{X})$ has a MVT distribution and applying Property 2.3 we conclude that $h(\mathrm{X}_S) \sim t^{C}(\mathrm{U}^{*'}\mathrm{S}\mathrm{U}\boldsymbol{\mu}, (\mathrm{U}^{*'}\mathrm{S}\mathrm{U})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{S}'\mathrm{U}^{*}), \nu)$. Therefore $\mathrm{X}_S \sim t_{\mathcal{S}}^{C}(\mathrm{U}^{*'}\mathrm{S}\mathrm{U}\boldsymbol{\mu}, (\mathrm{U}^{*'}\mathrm{S}\mathrm{U})\boldsymbol{\Upsilon}(\mathrm{U}'\mathrm{S}'\mathrm{U}^{*}), \nu)$. □

Acknowledgements The authors would like to thank the Editor and the anonymous reviewers for their careful reading of the manuscript and their constructive comments, which led to a substantially improved version of the paper.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement.

Data availability The data set employed here is part of the R package referenced in the text and does not require external sources. R code for the logratio Student's t implementation and the illustrative example is available for reproducibility at the following GitHub repository: https://github.com/giannamonti/Logratio-t.

Declarations

Conflict of interest The authors have no conflict of interest to declare that is relevant to the content of this article.

References

Aitchison J, Dunsmore IR (1975) Statistical Prediction Analysis. Cambridge University Press,

Aitchison J, Greenacre M (2002) Biplots of compositional data. J R Stat Soc: Ser C: Appl Stat 51(4):375–392. https://doi.org/10.1111/1467-9876.00275

Aitchison J, Habbema JDF, Kay JW (1977) A critical comparison of two methods of statistical discrimination. J R Stat Soc: Ser C: Appl Stat 26(1):15–25. https://doi.org/10.2307/2346863

Aitchison J (1986) The Statistical Analysis of Compositional Data. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London. Reprinted in 2003 with additional material by The Blackburn Press. 416 p

Aitchison J (1992) On criteria for measures of compositional difference. Math Geol 24(4): https://doi.org/10.1007/BF00891269

Box GEP, Tiao GC (1992) Bayesian Inference in Statistical Analysis. John Wiley & Sons, New York, p 512

Barceló-Vidal C, Martín-Fernández JA (2016) The mathematics of compositional analysis. Austrian J Stat 45, 57–71 https://doi.org/10.17713/ajs.v45i4.142

Cornish EA (1954) The multivariate t-distribution associated with a set of normal sample deviates. Aust J Phys 7:531–542. https://doi.org/10.1071/PH540531

Divino F, Kärkkäinen S, Maruotti A (2026) Unsupervised outlier detection for compositional data. Stat Prob Lett 227:110510. https://doi.org/10.1016/j.spl.2025.110510

Di Palma M, Gallo M (2016) A co-median approach to detect compositional outliers. J Appl Stat 43(13):2348–2362. https://doi.org/10.1080/02664763.2016.1163525

Dunnett CW, Sobel M (1954) A bivariate generalization of Student's t-distribution, with tables for certain special cases. Biometrika 41(1/2):153–169. https://doi.org/10.1093/biomet/41.1-2.153

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric logratio transformations for compositional data analysis. Math Geol 35:279–300. https://doi.org/10.1023/A:1023818214614

Filzmoser P, Hron K (2008) Outlier detection for compositional data using robust methods. Math Geosci 40(3):233–248. https://doi.org/10.1007/s11004-007-9141-5

Filzmoser P, Hron K (2009) Correlation analysis for compositional data. Math Geosci 41(8):905–919. https://doi.org/10.1007/s11004-008-9196-y

Filzmoser P, Hron K, Reimann C (2009) Principal component analysis for compositional data with outliers. Environmetrics 20(6):621–632. https://doi.org/10.1002/env.966

Fang KT, Kotz S, Ng KW (1990) Symmetric Multivariate and Related Distributions. Chapman and Hall, London

Fernandez C, Steel M (1999) Multivariate Student-*t* regression models: Pitfalls and inference. *Biometrika***86**(1), 153–167 http://www.jstor.org/stable/2673544

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian Data Analysis. Chapman and Hall/CRC,(3rd ed)

Grunsky EC, Drew LJ, Woodruff LG, Friske PWB, Sutphin DM (2013) Statistical variability of the geochemistry and mineralogy of soils in the maritime provinces of canada and part of the northeast united states. *Geochem.: Explor. Environ. Anal.***13**(4), 249–266 https://doi.org/10.1144/geochem2012-138

Grunsky EC (2010) The interpretation of geochemical survey data. *Geochem: Explor Environ Anal***10**(1), 27–74 https://doi.org/10.1144/1467-7873/09-210

Huber PJ (1981) Robust Statistics. John Wiley & Sons, New York

Jeffreys H (1983) Theory of Probability. Oxford University Press, Oxford

Katz JN, King G (1999) A statistical model for multiparty electoral data. Am Polit Sci Rev 93(1):15–32. https://doi.org/10.2307/2585758

Kotz S, Nadarajah S (2004) Multivariate T Distributions and Their Applications. Cambridge University Press, Cambridge

Lin PE (1972) Some characterizations of the multivariate t distribution. J Multivar Anal 2(3):339–344. https://doi.org/10.1016/0047-259X(72)90021-8

Lange KL, Little RJA, Taylor JMG (1989) Robust statistical modeling using the t distribution. J Am Stat Assoc 84(408):881–896. https://doi.org/10.1080/01621459.1989.10478852

Liu C, Rubin DB (1995) ML estimation of the *t* distribution using EM and its extensions, ECM and ECME. *Stat Sin***5**(1), 19–39 http://www.jstor.org/stable/24305551

Liu C, Rubin DB, Wu YN (1998) Parameter expansion to accelerate EM: The PX-EM Algorithm. *Biometrika***85**(4), 755–770 https://www.jstor.org/stable/2337481

Maronna RA (1976) Robust M-estimators of multivariate location and scatter. *Ann Stat***4**(1), 51–67 https://www.jstor.org/stable/2957994

Martín-Fernández JA (2019) Comments on: Compositional data: the sample space and its structure, by Egozcue and Pawlowsky-Glahn. TEST 28:653–657. https://doi.org/10.1007/s11749-019-00672-4

Mateu-Figueras G, Monti GS, Egozcue JJ (2021) Distributions on the simplex revisited. In: Filzmoser P, Hron K, Martín-Fernández JA, Palarea-Albaladejo J (eds) Advances in Compositional Data Analysis: Festschrift in Honour of Vera Pawlowsky-Glahn, pp. 61–82. Springer, Switzerland. Chap. 3

Mateu-Figueras G, Pawlowsky-Glahn V, Egozcue JJ (2013) The normal distribution in some constrained sample spaces. *SORT***37**, 29–56 https://raco.cat/index.php/SORT/article/view/261658

Marchenko YV, Genton MG (2012) A Heckman selection-t model. J Am Stat Assoc 107(497):304–317. https://doi.org/10.1080/01621459.2012.656011

Moran MA, Murphy BJ (1979) A closer look at two alternative methods of statistical discrimination. J R Stat Soc C Appl Stat 28(3):223–232. https://doi.org/10.2307/2347192

Monti GS, Mateu-Figueras G, Pawlowsky-Glahn V (2011) Notes on the scaled Dirichlet distribution. In: Pawlowsky-Glahn V, Buccianti A (eds.) Compositional Data Analysis, pp. 128–138. John Wiley & Sons, Ltd, Chichester, UK . Chap. 10

Nguyen THA (2019) Contribution to the statistical analysis of compositional data with an application to political economy. TSE, University Toulouse 1 Capitole,(PhD thesis)

Nadarajah S, Kotz S (2005) Mathematical properties of the multivariate t distribution. Acta Appl Math 89(1):53–84. https://doi.org/10.1007/s10440-005-9003-4

Pawlowsky-Glahn V (2003) Statistical modelling on coordinates. In: Thió-Henestrosa S, Martín-Fernández JA (eds.) Compositional Data Analysis Workshop – CoDaWork'03, Proceedings. University of Girona, Girona (Spain)

Pawlowsky-Glahn V, Egozcue JJ (2001) Geometric approach to statistical analysis on the simplex. Stoch Environ Res Risk Assess 15:384–398. https://doi.org/10.1007/s004770100077

Pawlowsky-Glahn V, Egozcue JJ (2002) BLU estimators and compositional data. Math Geol 34:259–274. https://doi.org/10.1023/A:1014890722372

Pawlowsky-Glahn V, Egozcue JJ, Tolosana-Delgado R (2015) Modeling and Analysis of Compositional Data. Wiley, Chichester, UK

Pinheiro JC, Liu C, Wu YN (2001) Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. J Comput Graph Stat 10(2):249–276. https://doi.org/10.1198/10618600152628059

R Core Team (2024) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria . R Foundation for Statistical Computing. https://www.R-project.org/

Rousseeuw PJ, Driessen KV (1999) A fast algorithm for the minimum covariance determinant estimator. Technometrics 41(3):212–223. https://doi.org/10.1080/00401706.1999.10485670

Reimann C, Filzmoser P, Garrett RG, Dutter R (2008) Statistical Data Analysis Explained. Applied Environmental Statistics with R. John Wiley & Sons Ltd, Chichester, UK

Raiffa H, Schlaifer R (1961) Applied Statistical Decision Theory. Harvard Business School Publications. Division of Research, Graduate School of Business Adminitration, Harvard University

Tolosana-Delgado R, Mueller U, Boogaart KG (2019) Geostatistics for compositional data: An overview. Math Geosci 51(4):485–526. https://doi.org/10.1007/s11004-018-9769-3

Zellner A (1976) Bayesian and Non-Bayesian analysis of the regression model with multivariate Student-t error terms. J Am Stat Assoc 71(354):400–405. https://doi.org/10.1080/01621459.1976.10480357

Authors and Affiliations

Gianna Serafina Monti[1] · Gloria Mateu-Figueras[2] · Vera Pawlowsky-Glahn[2] · Juan José Egozcue[3]

✉ Gianna Serafina Monti
gianna.monti@unimib.it

✉ Gloria Mateu-Figueras
gloria.mateu@udg.edu

Vera Pawlowsky-Glahn
vera.pawlowsky@udg.edu

Juan José Egozcue
juan.jose.egozcue@upc.edu

[1] Department of Economics, Management and Statistics, University of Milano–Bicocca, Piazza dell'Ateneo Nuovo, 1, Milano 20126, Italy

[2] Department of Computer Science, Applied Mathematics and Statistics, University of Girona, Campus Montilivi, Girona 17003, Spain

[3] Department of Civil and Environmental Engineering, Technical University of Catalonia, C. Jordi Girona, 1–3, Barcelona 08034, Spain