
A NOTE ON THE FOLDING TEST OF UNIMODALITY: LIMITATION AND IMPROVED ALTERNATIVE

A PREPRINT

 **Colombe Becquart ***
Toulouse School of Economics,
Université Toulouse Capitole
France
Université de Toulouse
France

 **Aurore Archimbaud**
TBS Business School
France

 **Anne M. Ruiz**
Toulouse School of Economics,
Université Toulouse Capitole
France

 **Zaineb Smida**
Univ Lyon, INSA Lyon, UJM, UCBL, ECL,
CNRS UMR 5208, ICI
France

May 14, 2026

ABSTRACT

This note addresses a key limitation of the Folding Test of Unimodality (FTU). In specific univariate mixture settings, the folding-based criterion can systematically fail, misclassifying clearly multimodal distributions as unimodal. We fully characterize these failures for Dirac mixtures and extend the analysis to Gaussian mixtures. We then introduce a double-folding procedure that captures complementary information, leading to a new test, the Double Folding Test of Unimodality. It resolves the FTU failures and improves multimodality detection power in simulations.

Keywords Dip test · Mixture distribution · Multimodality · Unimodality test

1 Introduction

Determining whether a distribution is unimodal or multimodal is a fundamental problem in statistical analysis, particularly in the context of clustering, where the presence of multiple modes is often interpreted as evidence of underlying cluster structure. This connection has been widely exploited, both within clustering algorithms [Kalogeratos and Likas, 2012, Shahbaba and Beheshti, 2014] and as a preprocessing step [Adolfsson et al., 2019, Krause and Liebscher, 2005]. In particular, the interplay between dimension reduction and unimodality provides valuable tools for assessing clusterability [Adolfsson et al., 2019] and arises naturally in projection pursuit [Krause and Liebscher, 2005] and tandem clustering [Arabie and Hubert, 1994, Alfons et al., 2024].

According to Khintchine [1938], a real random variable is said to be unimodal, with mode a , if its cumulative distribution function is convex on $(-\infty, a)$ and concave on $(a, +\infty)$. A standard reference for testing unimodality is Hartigan’s dip test [Hartigan and Hartigan, 1985], which is based on the cumulative distribution function. Another commonly used test is Silverman’s test [Silverman, 1981], relying on kernel density estimation.

The Folding Test of Unimodality (FTU) proposed by Siffer et al. [2018] offers an appealing alternative: it is computationally efficient, easy to implement, and particularly well suited to streaming data. Its rationale is that folding a multimodal distribution greatly reduces its variance, whereas the reduction is much smaller for a unimodal distribution. In the univariate case, for a real random variable X with finite second moment, folding consists of applying the

*Corresponding author: colombe.becquart@tse-fr.eu

transformation $X \mapsto |X - s|$, where s is the folding pivot. The associated variance reduction is measured by the so-called folding ratio:

$$\phi^*(X) = \min_{s \in \mathbb{R}} \frac{\text{Var}|X - s|}{\text{Var}X}. \quad (1)$$

The smaller $\phi^*(X)$, the more effective the folding is at reducing the variance. The standardized folding ratio (SFR), called the folding statistic in Siffer et al. [2018], compares this ratio to that of a reference unimodal distribution. Following Hartigan and Hartigan [1985], Siffer et al. [2018] use a random variable U following a uniform distribution $\mathcal{U}([l, u])$ on any bounded interval $[l, u]$ as the least concentrated unimodal case. Then, the SFR is defined by

$$\Phi^*(X) = \frac{\phi^*(X)}{\phi^*(U)}.$$

By Definition 7.2 in Siffer [2019], if $\Phi^*(X) \geq 1$, the distribution is considered unimodal; otherwise, it is deemed multimodal. The underlying heuristic of folding a distribution and measuring the resulting variance reduction is not equivalent to Khintchine’s definition. This motivated Siffer [2019] to study the behavior of the SFR under well-known unimodal distributions. His results provide evidence that variance reduction reflects unimodality for Beta and Gamma distributions. However, the limitation of this heuristic is illustrated by a counterexample repeatedly discussed in the literature [Siffer, 2019, Chasani and Likas, 2022, Kolyvakis and Likas, 2023]: the balanced mixture of three symmetric Dirac masses at $-a$, 0 , and a . We refer to this case as the “pathological” example, together with the corresponding Gaussian mixture with small noise. According to Khintchine’s definition, the latter is clearly multimodal, whereas the folding-based criterion concludes unimodality. This example reveals a fundamental blind spot of folding ratios and highlights their limitation in capturing multimodality.

This paper addresses this limitation through a detailed analysis of the SFR. We make two main contributions. First, we characterize the limitation of the SFR beyond the pathological example, by identifying structural conditions under which mixture models are misclassified. Second, we introduce a double-folding procedure that yields a revised unimodality criterion based on two complementary folding ratios, computed on the original and folded data, and an associated test, the Double Folding Test of Unimodality (DFTU), which resolves these failures.

Section 2 provides a deeper understanding of the folding mechanism through the study of Dirac and Gaussian mixtures. Section 3 introduces the double-folding procedure and the DFTU, while Section 4 reports the simulation results. Section 5 concludes. All proofs, technical details, tables from the simulation studies, and additional illustrations are collected in the Appendix. Replication files are available from: <https://github.com/cbecquart/DFTU-Replication>.

2 Revisiting the standardized folding ratio

Let X be a real random variable with finite fourth moment. An exact pivot (not necessarily unique), denoted s^* , is derived from expression (1) and defined as $s^* = \underset{s \in \mathbb{R}}{\text{argmin}} \text{Var}|X - s|$. This minimization problem is not necessarily

convex; a solution always exists but may not be unique [see p.119 in Siffer et al., 2018, and Appendix A.1]. To simplify the problem, Siffer et al. [2018] introduced the approximate pivot, denoted s^{**} , given by $s^{**} = \underset{s \in \mathbb{R}}{\text{argmin}} \text{Var}[(X - s)^2]$.

In the univariate case, s^{**} exists and is unique [see Siffer et al., 2018, Theorem 3.1]. Its associated folding ratio is $\phi^{**}(X) = \text{Var}|X - s^{**}|/\text{Var}X$. In this setting, both $\phi^*(U)$ and $\phi^{**}(U)$ equal $1/4$ (see Appendix A.3). Thus the two SFR associated to s^* and s^{**} are, respectively, $\Phi^*(X) = 4 \text{Var}|X - s^*|$ and $\Phi^{**}(X) = 4 \text{Var}|X - s^{**}|$.

2.1 Mixture of Dirac distributions

In the discrete case, let us consider a k -Dirac mixture with proportions $\varepsilon_i > 0$ and locations μ_i for $i \in \{1, \dots, k\}$, $k > 1$:

$$X \sim \sum_{i=1}^k \varepsilon_i \delta_{\mu_i} \quad (2)$$

where $\sum_{i=1}^k \varepsilon_i = 1$. Using the affine invariance property of $\Phi^*(X)$ and $\Phi^{**}(X)$ (see Appendix A.4), let us consider without loss of generality (w.l.o.g.) that X is standardized ($\mathbb{E}[X] = 0$ and $\text{Var}[X] = 1$) with ordered locations $\mu_1 < \dots < \mu_k$ in model (2).

For Dirac mixtures, one expects $\Phi^*(X) < 1$ and $\Phi^{**}(X) < 1$, indicating multimodality. However, there is at least the aforementioned pathological example where this is not the case. This balanced Dirac mixture ($\mu_1 = -a$, $\mu_2 = 0$, $\mu_3 = a$) results in $s^{**} = 0$ and $\Phi^{**}(X) > 1$, while $s^* = \pm a/4$ yields $\Phi^*(X) = 1$. In this case, the SFR are considered

to “fail” since they are not lower than one despite clear multimodality. The key question is whether similar failures occur in broader mixture models and whether they can be characterized.

For a 2-Dirac mixture, $\Phi^*(X) = \Phi^{**}(X) = 0 < 1$ for any choice of proportions and locations, meaning the SFR always correctly detect multimodality. In the general case, Propositions 1 and 2 provide explicit expressions for $\Phi^*(X)$ and $\Phi^{**}(X)$ respectively. They also give the conditions under which the ratios exceed 1, thus falsely indicating unimodality.

Suppose that model (2) holds. For any $g \in \{1, \dots, k-1\}$, let $\alpha_g = \sum_{i=1}^g \varepsilon_i \mu_i$ and $\eta_g = \sum_{i=1}^g \varepsilon_i$. Proofs of the following propositions are available in Appendices A.5 and A.6.

Proposition 1 *For the exact pivot approach, we have:*

1. $s^* = \frac{\alpha_{g^*}}{2} \left(\frac{1}{\eta_{g^*}} - \frac{1}{1-\eta_{g^*}} \right)$, where $g^* = \underset{g \in \{1, \dots, k-1\}}{\operatorname{argmin}} \min_{s \in [\mu_g, \mu_{g+1}]} \operatorname{Var}[X - s]$.
2. The associated SFR is $\Phi^*(X) = 4 \left(1 - \frac{\alpha_{g^*}^2}{\eta_{g^*}(1-\eta_{g^*})} \right)$.
3. The equation $\Phi^*(X) = 1$ defines an ellipse in the $(\alpha_{g^*}, \eta_{g^*})$ -plane with center $(0, \frac{1}{2})$, semi-minor axis $\frac{\sqrt{3}}{4}$, and semi-major axis $\frac{1}{2}$.
4. The SFR fails, i.e. $\Phi^*(X) \geq 1$, if and only if $\alpha_{g^*} \in \left[-\frac{\sqrt{3}}{2} \sqrt{\eta_{g^*}(1-\eta_{g^*})}, 0 \right]$.

Proposition 2 *For the approximate pivot approach, we have:*

1. $s^{**} = \frac{\gamma}{2}$, with $\gamma = \mathbb{E}[X^3]$.
2. The associated SFR is $\Phi^{**}(X) = 4 \left(1 - 4\alpha_{g^{**}}^2 - 2\alpha_{g^{**}}(1 - 2\eta_{g^{**}})\gamma + \eta_{g^{**}}(1 - \eta_{g^{**}})\gamma^2 \right)$, where $g^{**} = \sum_{i=1}^k \mathbb{1}_{\{\mu_i \leq s^{**}\}}$.
3. The equation $\Phi^{**}(X) = 1$ corresponds to a line in the $(\alpha_{g^{**}}, \eta_{g^{**}})$ -plane, given by $4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma + \sqrt{\gamma^2 + 3} = 0$.
4. The SFR fails, i.e. $\Phi^{**}(X) \geq 1$, if and only if $4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma + \sqrt{\gamma^2 + 3} \geq 0$.

Remark 1 *Since $\Phi^*(X) \leq \Phi^{**}(X)$, if the exact pivot approach fails ($\Phi^*(X) \geq 1$), then the approximate pivot also fails ($\Phi^{**}(X) \geq 1$).*

Propositions 1 and 2 show that the SFR fail in certain cases. These cases are not necessarily characterized by balance, symmetry, or a three-group structure, as in the pathological example. Example 1 illustrates two failure cases that differ from the pathological example. It also highlights a situation where only the approximate pivot fails (case (a)).

Example 1 (a) *Let $X \sim 0.2 \cdot \delta_{-2} + 0.4 \cdot \delta_0 + 0.4 \cdot \delta_2$. In this unbalanced case, only the approximate pivot fails: $\Phi^{**}(X) = 1.41 > 1$ while $\Phi^*(X) = 0.95 < 1$.*

(b) *Let $X \sim 0.2 \cdot \delta_{-3} + 0.2 \cdot \delta_{-1.5} + 0.2 \cdot \delta_{2.5} + 0.2 \cdot \delta_4 + 0.2 \cdot \delta_{11}$. In this asymmetric case, both pivots fail: $\Phi^{**}(X) = 1.38 > 1$ and $\Phi^*(X) = 1.07 > 1$.*

2.2 Mixture of Gaussian distributions

In the continuous setting, let us consider a k -Gaussian mixture displaying a within-variance term, with proportions $\varepsilon_i > 0$, means μ_i , and variances σ_i^2 , for $i \in \{1, \dots, k\}$, $k > 1$:

$$X \sim \sum_{i=1}^k \varepsilon_i \mathcal{N}(\mu_i, \sigma_i^2), \quad (3)$$

where $\sum_{i=1}^k \varepsilon_i = 1$, and $\mathbb{E}[X] = 0$. In this setting, we cannot obtain an explicit expression for the exact and approximate pivots. The SFR, can however, be derived from the expression:

$$\Phi(X, s) = 4 \left[\operatorname{Var}[X] + s^2 - \left(\sum_{i=1}^k \varepsilon_i h(\mu_i, \sigma_i, s) \right)^2 \right], \quad (4)$$

where $h(\mu, \sigma, s) = (s - \mu) \left(2F\left(\frac{s-\mu}{\sigma}\right) - 1 \right) + 2\sigma f\left(\frac{s-\mu}{\sigma}\right)$, and where f and F denote, respectively, the probability density function and the cumulative distribution function of the standard normal distribution $\mathcal{N}(0, 1)$. Equation (4) is easily derived from the variance of the folded normal distribution [Leone et al., 1961].

We have $\Phi^*(X) = \Phi(X, s^*)$ and $\Phi^{**}(X) = \Phi(X, s^{**})$. When all variances σ_i are set to 0 for $i = \{1, \dots, k\}$, the expressions $\Phi^*(X)$ and $\Phi^{**}(X)$ obtained from equation (4) reduce to the Dirac case (Propositions 1 and 2).

Note that one can standardize X such that $\text{Var}[X] = 1 + W$, where $W = \sum_{i=1}^k \varepsilon_i \sigma_i^2$ represents the within-variance. Accordingly, the signal-to-noise ratio (SNR) is defined as $\text{SNR} = 1/W$. Setting the between-variance to 1 ensures consistency with Subsection 2.1, which corresponds to the special case where $W = 0$. Since theoretical computations beyond this point are challenging, we consider further the 2-Gaussian mixture case and highlight the impact of the within-variance term. The results differ from the Dirac case, where the SFR always detects multimodality for $k = 2$, as shown in Example 2.

Example 2 Let $X \sim 0.3 \cdot \mathcal{N}(-2.8, \sigma^2) + 0.7 \cdot \mathcal{N}(1.2, \sigma^2)$. Fig. 1a shows $\Phi^*(X)$ as a function of σ^2 , while Fig. 1b and 1c illustrate the distribution for some values of σ^2 . When σ^2 is small (high SNR), the distribution resembles a 2-Dirac mixture and $\Phi^*(X)$ is close to zero. As σ^2 increases, the two modes become less distinct (Fig. 1b), the distribution looks more unimodal (Fig. 1c) and $\Phi^*(X)$ rises (Fig. 1a). In addition, when σ^2 exceeds 1.34 (Fig. 1a), $\Phi^*(X)$ becomes greater than 1, indicating the transition from multimodality to unimodality according to Definition 7.2 of unimodality in Siffer [2019]. Note, however, that the distribution with $\sigma^2 = 1.34$ would still be considered multimodal under the definition of Khintchine [1938], as seen from the multiple changes of convexity in Fig. 1c.

The remarks of Example 2 are not specific to this particular example. For Gaussian mixtures (3) with low SNR that nearly satisfy Khintchine’s definition of unimodality, the SFR may conclude unimodality ($\Phi^*(X) \geq 1$). In that situation, the ratio cannot be considered to have failed. In other cases, however, $\Phi^*(X) \geq 1$ reflects a genuine failure caused by the mixture structure, as in the Dirac case (see Appendix B for details).

In the following section, we propose a revised unimodality criterion and an associated test, designed to correctly identify multimodality in the aforementioned failure cases.

3 Enhancing the standardized folding ratio with a double-folding procedure

3.1 The double-folding algorithm

The limitation of the SFR in detecting multimodality is addressed through the development of an iterative algorithm. The first step consists of computing the exact SFR $\Phi_1^*(X)$ with associated pivot s_1^* . If $\Phi_1^*(X) < 1$, the procedure stops and the algorithm returns a multimodal conclusion. Otherwise, a second step is performed. The data are folded using the approximate pivot s_1^{**} , and the exact SFR $\Phi_2^*(X)$, with associated pivot s_2^* , is then computed on the folded data. The procedure concludes unimodality if $\Phi_2^*(X) \geq 1$, and multimodality otherwise.

The use of exact SFR in both steps is justified by Remark 1. The rationale for using the approximate pivot to transform the data is its ability to break the symmetry of the data when necessary. Indeed, folding the pathological example’s distribution using an exact pivot yields another pathological mixture, with the gap between groups reduced by half (see Appendix C.1). The second step will thus also falsely conclude unimodality. In contrast, using the approximate pivot breaks the initial symmetry, enabling the detection of multimodality in the second step. An alternative would be to use an exact pivot shifted by a small delta value.

The second step is not always necessary. According to our empirical experience, the SFR do not falsely conclude multimodality; the only issue arises when they incorrectly conclude unimodality. If the second step were applied even when the first step ratio detects multimodality, it could override the correct detection. For instance, folding a 2-Dirac mixture may lead to a unimodal distribution.

If the data are unimodal, both steps should correctly conclude unimodality ($\Phi_1^*(X) \geq 1$ and $\Phi_2^*(X) \geq 1$). If they are multimodal, either the first step detects it or, if not, we conjecture that the second step reveals it ($\Phi_1^*(X) < 1$ or $\Phi_2^*(X) < 1$). In particular, the double-folding procedure yields the correct conclusion for the “pathological” example. When folding with the approximate pivot ($s_1^{**} = 0$), the resulting distribution becomes a mixture of two Dirac masses at 0 and a , with respective weights $1/3$ and $2/3$. This leads to $s_2^* = a/2$ and $\Phi_2^* = 0$, indicating clear multimodality. The following subsection analyzes the general case of a 3-Dirac mixture to further clarify the motivation.

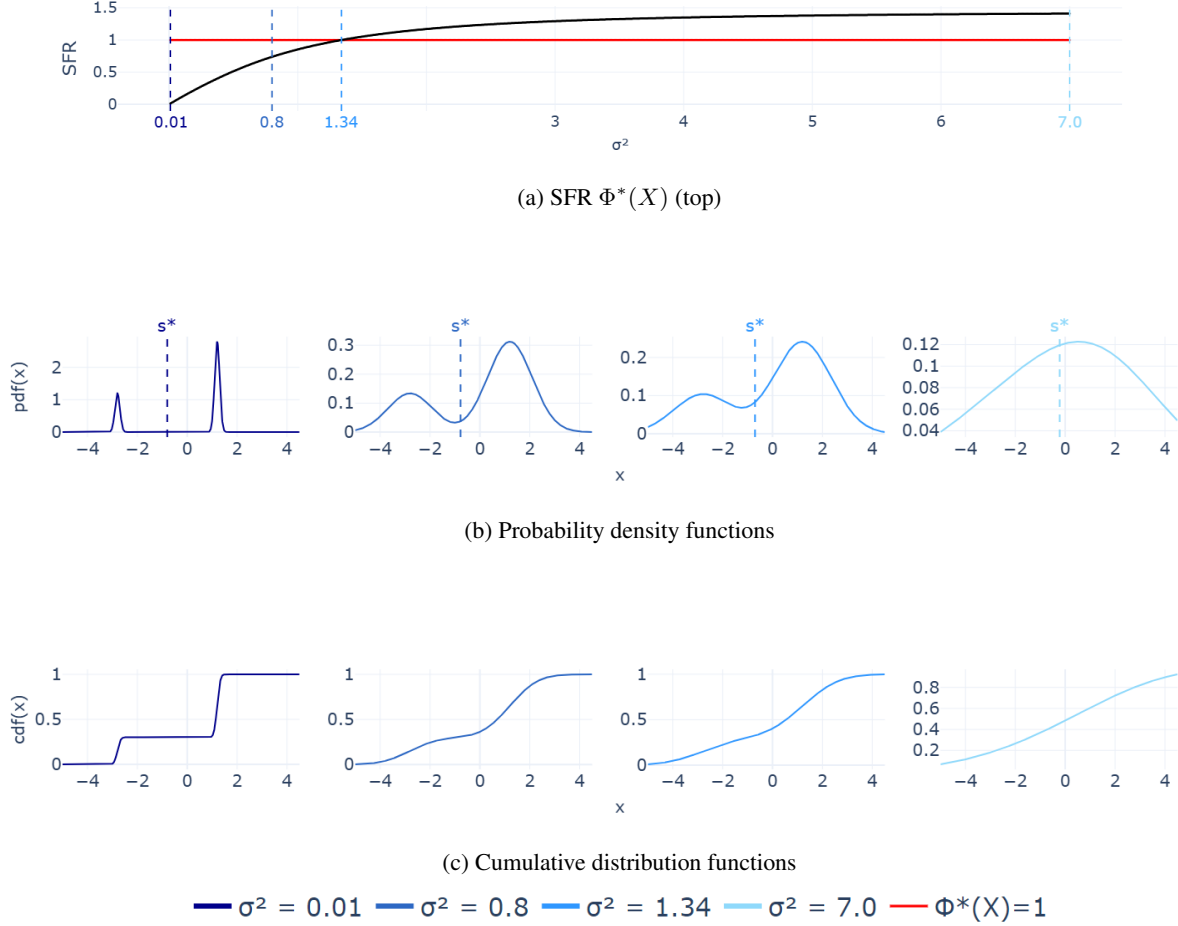


Figure 1: 2-Gaussian mixture with $\varepsilon_1 = 0.3$, $\mu_1 = -2.8$, $\varepsilon_2 = 0.7$, $\mu_2 = 1.2$, and varying σ^2 : (a) values of $\Phi^*(X)$ (black) as a function of σ^2 , (b) probability density functions and associated exact pivots s^* , (c) cumulative distribution functions.

3.2 Motivation through the 3-Dirac mixture case

Let us consider the standardized model (2) with $k = 3$ and ordered locations. Throughout this section, when the minimizer of $\text{Var}|X - s|$ is not unique, we take s^* to be the smallest minimizer. We provide a new characterization of Proposition 1, in which the exact SFR failure conditions depend only on μ_1 , ε_1 , and ε_3 . It allows us to analyze the second step of the double-folding procedure when the first-step ratio fails. Then, based on numerical evidence, we argue that in this setting the double-folding correctly detects multimodality. This characterization is given in Proposition 3 (see Appendix A.7). Note that, by the affine invariance of the SFR, we may assume w.l.o.g. that $s^* \in [\mu_1, \mu_2]$.

Proposition 3 1. The existence of $s^* \in [\mu_1, \mu_2]$ is equivalent to $\mu_1 \leq U(\varepsilon_1, \varepsilon_3)$ with $U(\varepsilon_1, \varepsilon_3) := \max\{U_1, U_2\}$ where

$$U_1 = -\frac{\sqrt{2}}{2} \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} + \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}} \text{ and } U_2 = -\frac{2\varepsilon_2+4\varepsilon_1\varepsilon_3}{\sqrt{4\varepsilon_1\varepsilon_3(\varepsilon_2+4\varepsilon_1\varepsilon_3)}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}}.$$

2. When $s^* \in [\mu_1, \mu_2]$, $\Phi^*(X) \geq 1$ if and only if $\mu_1 \geq F(\varepsilon_1)$, where the failure bound is given by $F(\varepsilon_1) = -\frac{\sqrt{3}}{2} \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}}$.

According to the double-folding algorithm, whenever the first-step ratio detects multimodality, the procedure stops. When it is not the case (Proposition 3), the distribution is folded using the approximate pivot, the locations are reordered

and standardized, and the bounds in Proposition 3 are updated accordingly. We analyze these updated quantities via numerical optimization. First, the results show that s_2^* lies between the first and second folded locations, since the inequality in Proposition 3 (1.) is satisfied for any admissible parameter values. Consequently, the inequality in Proposition 3 (2.) can be used to assess whether the second-step ratio fails. Numerical optimization indicates that this inequality does not hold, supporting the view that the double-folding procedure does not fail in this model (see Appendix C.2 for details).

Motivated by this result, we introduce a new unimodality test based on the double-folding procedure. The test is described in the following subsection.

3.3 The Double Folding Test of Unimodality

In Siffer et al. [2018], the SFR is used to test the null hypothesis $H_0 : \Phi^*(X) \geq 1$ at significance level α , in accordance with their definition of unimodality. The introduction of a double-folding criterion provides a novel approach to testing unimodality, leading to the Double Folding Test of Unimodality (DFTU), which tests:

$$\mathcal{H}_0 : \Phi_1^*(X) \geq 1 \text{ and } \Phi_2^*(X) \geq 1 \quad \text{vs.} \quad \mathcal{H}_1 : \Phi_1^*(X) < 1 \text{ or } \Phi_2^*(X) < 1,$$

where $\Phi_1^*(X) = 4 \text{Var}|X - s_1^*|/\text{Var}X$ with s_1^* an exact pivot for the original data X , and $\Phi_2^*(X) = 4 \text{Var}|\tilde{X} - s_2^*|/\text{Var}\tilde{X}$ with $\tilde{X} = |X - s_1^*|$, s_1^* being the approximate pivot for X , and s_2^* an exact pivot for the folded data $\tilde{X}$.

Regarding the statistical significance of the DFTU, the challenge lies in the fact that $\mathcal{H}_0$ involves two folding statistics $\Phi_1^*(X)$ and $\Phi_2^*(X)$. The two folding statistics, computed sequentially, are associated with significance levels α_1 and α_2 , and corresponding critical values q_1 and q_2 . Significance levels are $\alpha_1 = \mathbb{P}(\Phi_1^*(X) < q_1 | \mathcal{H}_0)$ and $\alpha_2 = \mathbb{P}(\Phi_2^*(X) < q_2 | \Phi_1^*(X) \geq q_1; \mathcal{H}_0)$. In this case, $\mathcal{H}_0$ is rejected if either $\Phi_1^*(X) < q_1$, or if $\Phi_1^*(X) \geq q_1$ and $\Phi_2^*(X) < q_2$. Consequently, the overall significance level of the test α is given by $\alpha = \alpha_1 + (1 - \alpha_1)\alpha_2$. Given predefined values of α_1 and α_2 , the corresponding critical values q_1 and q_2 can be estimated via Monte Carlo simulation. Following Siffer [2019], we generate observations from a uniform distribution. The critical values q_1 and q_2 are then taken as the empirical quantiles of order α_1 and α_2 of $\Phi_1^*(X)$ and $\Phi_2^*(X)$, respectively. For details about the implementation of the DFTU, see Algorithm C.3.1 in Appendix C.3.

4 Simulations

We want to compare the proposed DFTU with the FTU [Siffer et al., 2018] based on both pivots, and with the dip test [Hartigan and Hartigan, 1985]. For each distribution, we generate 100 datasets of 1,000 observations and apply all four tests at a 5% significance level. For the DFTU, α_1 is set to 3%, although similar results are obtained for other choices of α_1 . Table 1 reports the majority decision for each test (see Table D.1 in Appendix D for detailed results). The first three rows illustrate cases in which all tests reach the correct conclusion. The standard Gaussian distribution and the 2-Gaussian mixture with low SNR are both classified as unimodal, whereas the well-separated mixture of a Gaussian and a uniform distribution is identified as multimodal by all tests. The last four rows present multimodal examples in which the FTU does not detect multimodality, yet the DFTU accurately identifies it. Finally, the mixture $0.6 \cdot \mathcal{N}(0, 0.5) + 0.4 \cdot \mathcal{U}([1, 4])$ is noteworthy, as folding-based tests and the dip test lead to different conclusions. According to the definition of unimodality based on the cumulative distribution function [Khintchine, 1938], this distribution can be considered unimodal (see Table D.2 in Appendix D). However, from a clustering perspective, it may be preferable to conclude multimodality, which matches the DFTU's conclusion, as well as the FTU**'s and the FTU***'s.

Distribution	FTU*	FTU**	DFTU	dip test
$\mathcal{N}(0, 1)$	Uni	Uni	Uni	Uni
$0.5 \cdot (\mathcal{N}_0 + \mathcal{N}_1)$	Uni	Uni	Uni	Uni
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([4, 8])$	Multi	Multi	Multi	Multi
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([1, 4])$	Multi	Multi	Multi	Uni
$1/3 \cdot \delta_{-2} + 1/3 \cdot \delta_0 + 1/3 \cdot \delta_2$	Uni	Uni	Multi	Multi
$1/3 \cdot \mathcal{N}_{-2} + 1/3 \cdot \mathcal{N}_0 + 1/3 \cdot \mathcal{N}_2$	Uni	Uni	Multi	Multi
$0.2 \cdot (\delta_{-3} + \delta_{-1.5} + \delta_{2.5} + \delta_4 + \delta_{11})$	Uni	Uni	Multi	Multi
$0.2 \cdot (\mathcal{N}_{-3} + \mathcal{N}_{-1.5} + \mathcal{N}_{2.5} + \mathcal{N}_4 + \mathcal{N}_{11})$	Uni	Uni	Multi	Multi

Table 1: Simulation results. FTU* is the FTU with exact pivot, FTU** is the FTU with approximate pivot. $\mathcal{N}_a = \mathcal{N}(a, 0.5)$. Uni and Multi stand respectively for unimodal and multimodal.

5 Concluding remarks

In this paper, we investigate a key limitation of the FTU, namely its inability to detect multimodality in certain mixture distributions, which stems from the behavior of the SFR. By analyzing Dirac and Gaussian mixtures, we identify distributional structures under which the SFR fails to capture multimodality.

To address this issue, we propose the DFTU, a two-step test that applies the FTU to the original data and, when unimodality is not rejected, to the folded data. Computations and simulations show that it successfully detects multimodality in cases where the original FTU fails. These results suggest that combining information across successive transformations, as done through double folding, can substantially improve the detection of multimodality.

Although an iterative procedure with more than two steps is conceivable, we found no examples where this was necessary. Future work could focus on generalizing the DFTU to the multivariate case, as initially intended for the FTU, or on replacing the variance with a more robust estimator such as Q_n [Rousseeuw and Croux, 1993], since the current test may be sensitive to outliers. Another promising direction is the application of the proposed approach to ICS tandem clustering. While informative directions are traditionally identified through departures from normality, deviations from unimodality may be more directly connected to group separation, which can be exploited for component selection. All replication files are available at: <https://github.com/cbecquart/DFTU-Replication>.

A Proofs

A.1 Proof of the existence of s^*

This subsection proves the existence of s^* . The proof follows p. 119 in Siffer [2019], adapted to the univariate case.

Proposition A.1.1 *Let X be a real random variable (r.v.) with finite second moment. Then the function $s \mapsto \text{Var}|X - s|$ is continuous and bounded on $\mathbb{R}$. In particular, $\forall s \in \mathbb{R}, 0 \leq \text{Var}|X - s| \leq \text{Var}X$.*

Proof We have

$$\text{Var}|X - s| = \text{Var}X + \mathbb{E}^2[X - s] - \mathbb{E}^2[|X - s|]$$

Define $v : \mathbb{R} \rightarrow \mathbb{R}$ by $v(s) = \text{Var}|X - s|$. Since it is a composition of continuous functions, v is continuous. Moreover, $v(s) \geq 0$ for all $s \in \mathbb{R}$, as it is a variance.

By Jensen's inequality, $|\mathbb{E}[X - s]| \leq \mathbb{E}[|X - s|]$, which implies $\mathbb{E}^2[X - s] \leq \mathbb{E}^2[|X - s|]$ and $\text{Var}|X - s| \leq \text{Var}X$.

Proposition A.1.2 *Let X be a real r.v. with finite second moment. Then the function $s \mapsto \text{Var}|X - s|$ attains an absolute minimum on $\mathbb{R}$.*

Proof We have

$$\begin{aligned} v(s) &= \text{Var}X + \mathbb{E}^2[X - s] - \mathbb{E}^2[|X - s|] \\ &= \text{Var}X + s^2 \left(1 - \frac{\mathbb{E}[X]}{s}\right)^2 - s^2 \left(\mathbb{E}\left[\left|1 - \frac{X}{s}\right|\right]\right)^2. \end{aligned}$$

As $s \rightarrow \pm\infty$, we use the second-order expansion

$$\mathbb{E}[|1 + \alpha X|] = 1 + \alpha \mathbb{E}[X] + o(\alpha^2) \quad \text{as } \alpha \rightarrow 0.$$

which yields

$$\begin{aligned} v(s) &= \text{Var}X + s^2 \left(1 - \frac{\mathbb{E}[X]}{s}\right)^2 - s^2 \left(1 - \frac{\mathbb{E}[X]}{s} + o\left(\frac{1}{s^2}\right)\right)^2 \\ &= \text{Var}X + o(1), \quad \text{as } s \rightarrow \pm\infty. \end{aligned}$$

The function v is continuous and bounded between 0 and $\text{Var}X$ (Proposition A.1.1). Let $\varepsilon > 0$. Then there exists $s_0 > 0$ such that $\text{Var}X - \varepsilon \leq v(s) \leq \text{Var}X \forall |s| > s_0$. The restriction of v to $[-s_0, s_0]$ is then continuous on a bounded interval, so it attains its bounds by the Extreme Value Theorem.

A.2 Proof of the affine invariance of $\phi^*(X)$ and $\phi^{**}(X)$

Proposition A.2.1 *The pivots s^* and s^{**} are equivariant under affine transformations. The corresponding folding ratios $\phi^*(X)$ and $\phi^{**}(X)$ are invariant under affine transformations.*

Proof Let us consider an affine transformation of the random variable X , $\tilde{X} = aX + b$, for real numbers a, b such that $a \neq 0$.

(i) We have:

$$\begin{aligned} s^* = \arg \min_{s \in \mathbb{R}} \text{Var}|X - s| &\iff \text{Var}|X - s^*| \leq \text{Var}|X - s|, \quad \forall s \in \mathbb{R}, \\ &\iff \text{Var}|aX + b - as^* - b| \leq \text{Var}|aX + b - as - b|, \\ &\quad \forall s \in \mathbb{R}. \end{aligned}$$

Let $\tilde{s} = as + b$ and $\tilde{s}^* = as^* + b$.

Then, $\text{Var}|\tilde{X} - \tilde{s}^*| \leq \text{Var}|\tilde{X} - \tilde{s}|, \forall \tilde{s} \in \mathbb{R} \iff \tilde{s}^* = \arg \min_{\tilde{s} \in \mathbb{R}} \text{Var}|\tilde{X} - \tilde{s}|$.

Similarly, for s^{**} , we obtain $\tilde{s}^{**} = as^{**} + b = \arg \min_{\tilde{s} \in \mathbb{R}} \text{Var}(\tilde{X} - \tilde{s})^2$.

(ii) We also have:

$$\phi^*(\tilde{X}) = \frac{\text{Var}|\tilde{X} - \tilde{s}^*|}{\text{Var}\tilde{X}} = \frac{\text{Var}|X - s^*|}{\text{Var}X} = \phi^*(X) \quad \text{and} \quad \phi^{**}(\tilde{X}) = \phi^{**}(X).$$

A.3 Proof of $\phi^*(U) = \phi^{**}(U) = 1/4$

Proposition A.3.1 The folding ratios $\phi^*(U)$ and $\phi^{**}(U)$ are equal to $1/4$, where $U \sim \mathcal{U}([l, u])$ on any bounded interval $[l, u]$.

Proof By the affine invariance of $\phi^*(X)$ and $\phi^{**}(X)$, it is sufficient to consider $U \sim \mathcal{U}([-R, R])$ with $R > 0$. For $\phi^*(U)$: Let $U \sim \mathcal{U}([-R, R])$ for $R \in \mathbb{R}_+^*$. We know that $\mathbb{E}[U] = 0$ and $\text{Var}[U] = \frac{R^2}{3}$.

$$\phi^*(U) = \frac{\text{Var}|U - s^*|}{\text{Var}U},$$

where $s^* = \arg \min_{s \in [-R, R]} \text{Var}|U - s|$. For $s \in [-R, R]$, $\text{Var}|U - s| = \mathbb{E}[(U - s)^2] - \mathbb{E}^2[|U - s|]$ where $\mathbb{E}[(U - s)^2] = s^2 + \frac{R^2}{3}$ and $\mathbb{E}[|U - s|] = \frac{1}{2R}(R^2 + s^2)$. We obtain:

$$\text{Var}|U - s| = \frac{1}{12} \left(R^2 + 6s^2 - \frac{3s^4}{R^2} \right), \quad \phi^*(U) = \frac{1}{4R^2} \left(R^2 + 6s^2 - \frac{3s^4}{R^2} \right).$$

There are three solutions to $\frac{\partial \text{Var}|U - s|}{\partial s} = s(1 - \frac{s^2}{R^2}) = 0$: $s = 0$ or $s = R$ or $s = -R$. The only solution for which $\frac{\partial^2 \text{Var}|U - s|}{\partial s^2} > 0$ is $s = 0$. Thus, $s^* = 0$, and the associated folding ratio $\phi^*(U) = 1/4$. For $\phi^{**}(U)$: the proof is available in Siffer [2019].

A.4 Proof of the affine invariance of $\Phi^*(X)$ and $\Phi^{**}(X)$

Proposition A.4.1 The SFR $\Phi^*(X)$ and $\Phi^{**}(X)$ are invariant by affine transformations.

Proof Let us consider an affine transformation of the random variable X , $\tilde{X} = aX + b$, where $a, b \in \mathbb{R}$ and $a \neq 0$. From Proposition A.3.1, we know that $\phi^*(U) = \phi^{**}(U) = 1/4$, where $U \sim \mathcal{U}([l, u])$ on any bounded interval $[l, u]$. Using the affine invariance of $\phi^*(X)$ and $\phi^{**}(X)$ (Proposition A.2.1), we have

$$\Phi^*(\tilde{X}) = \frac{\phi^*(\tilde{X})}{\phi^*(U)} = \frac{\phi^*(X)}{\phi^*(U)} = \Phi^*(X) \quad \text{and similarly} \quad \Phi^{**}(\tilde{X}) = \Phi^{**}(X).$$

A.5 Proof of Proposition 1

In this paragraph, we provide the proof of the explicit expression for $\Phi^*(X)$ for the k -Dirac mixture model (2), and establish the conditions under which the ratio leads to false indications of unimodality. X is standardized ($\mathbb{E}[X] = 0$ and $\text{Var}[X] = 1$) with ordered locations $\mu_1 < \dots < \mu_k$. Let us suppose that for any $g \in \{1, \dots, k-1\}$, we define $\alpha_g = \sum_{i=1}^g \varepsilon_i \mu_i$ and $\eta_g = \sum_{i=1}^g \varepsilon_i$.

In order to prove Proposition 1, we will need the following lemmas.

Lemma A.5.1 *This lemma is used in the proof of Proposition 1.*

The following results give explicit expressions for $\text{Var}|X - s|$.

1. $\text{Var}|X - s| = 1 + s^2 - \mathbb{E}^2[|X - s|]$.
2. For all $g \in \{1, \dots, k-1\}$ and $s \in [\mu_g, \mu_{g+1}]$, the following three equivalent expressions hold:

$$\begin{aligned} \text{Var}|X - s| &= 1 - 4\eta_g(1 - \eta_g) \left(\frac{\alpha_g}{\eta_g} - s \right) \left(\frac{\alpha_g}{1 - \eta_g} + s \right) \\ &= 1 - 4(\alpha_g - \eta_g s)(\alpha_g + (1 - \eta_g)s) \\ &= 4\eta_g(1 - \eta_g)s^2 + 4(2\eta_g - 1)\alpha_g s + 1 - 4\alpha_g^2. \end{aligned}$$

Proof 1. The expression is obtained by expanding:

$$\begin{aligned} \text{Var}|X - s| &= \mathbb{E}[(X - s)^2] - \mathbb{E}^2[|X - s|] \\ &= \mathbb{E}[X^2] - 2s\mathbb{E}[X] + s^2 - \mathbb{E}^2[|X - s|]. \end{aligned}$$

Since $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] = 1$, we obtain the result.

2. For all $g \in \{1, \dots, k-1\}$ and $s \in [\mu_g, \mu_{g+1}]$, results are obtained using Lemma A.5.1(1.) and the following expression:

$$\begin{aligned} \mathbb{E}^2[|X - s|] &= (-2\alpha_g + s(2\eta_g - 1))^2 \\ &= 4(\alpha_g - s\eta_g)(\alpha_g + s(1 - \eta_g)) + s^2. \end{aligned}$$

Lemma A.5.2 *This lemma is used in the proof of Proposition 1.*

The function $s \mapsto \text{Var}|X - s|$ is continuous and consists of piecewise parabolic segments defined on each interval $[\mu_g, \mu_{g+1}]$, for $g \in \{1, \dots, k-1\}$.

Proof Lemma A.5.1 states that the function is piecewise parabolic and we just need to prove that it is continuous. To do so, we use the last expression of $\text{Var}|X - s|$ from Lemma A.5.1(2.). To verify the continuity of the function $s \mapsto \text{Var}|X - s|$ at each point $s = \mu_g$, we evaluate the function from the left-hand limit ($s = \mu_g^-$) and the right-hand limit ($s = \mu_g^+$), and confirm that they are equal. Using the fact that $\alpha_g = \alpha_{g-1} + \varepsilon_g \mu_g$, and $\eta_g = \eta_{g-1} + \varepsilon_g$, we obtain:

$$\text{Var}|X - \mu_g^+| - \text{Var}|X - \mu_g^-| = 0.$$

Lemma A.5.3 *This lemma is used in the proof of Proposition 1.*

For each $g \in \{1, \dots, k-1\}$, define $s_g = \frac{\alpha_g}{2} \left(\frac{1}{\eta_g} - \frac{1}{1 - \eta_g} \right)$. Then, the minimizer of $\text{Var}|X - s|$ over the interval $[\mu_g, \mu_{g+1}]$ is given by

$$s_g^\dagger = \underset{s \in [\mu_g, \mu_{g+1}]}{\text{argmin}} \text{Var}|X - s| = s_g - \min(0, s_g - \mu_g) - \max(0, s_g - \mu_{g+1}),$$

which is equivalent to

$$s_g^\dagger = \begin{cases} \mu_g, & \text{if } s_g < \mu_g, \\ s_g, & \text{if } s_g \in [\mu_g, \mu_{g+1}], \\ \mu_{g+1}, & \text{if } s_g > \mu_{g+1}. \end{cases}$$

An exact pivot can be chosen as any solution of the minimization problem

$$s^* = \underset{s \in \mathbb{R}}{\text{argmin}} \text{Var}|X - s| = s_{g^*}^\dagger, \quad \text{where } g^* = \underset{g \in \{1, \dots, k-1\}}{\text{argmin}} \text{Var}|X - s_g^\dagger|.$$

Proof From Lemma A.5.1(1.), we have

$$\text{Var}|X - s| = 1 + s^2 - \left(\sum_{i=1}^k \varepsilon_i |\mu_i - s| \right)^2,$$

and its derivative satisfies

$$\frac{d}{ds} \text{Var}|X - s| = 0 \iff -s = \left(\sum_{i=1}^k \varepsilon_i \text{sign}(\mu_i - s) \right) \left(\sum_{i=1}^k \varepsilon_i |\mu_i - s| \right).$$

Since $s \in [\mu_1, \mu_k]$, let us consider $\mu_1 < \dots < \mu_g < s < \mu_{g+1} < \dots < \mu_k$, for $g \in \{1, \dots, k-1\}$. Then, $\text{sign}(\mu_i - s) = -1, \forall i \in \{1, \dots, g\}$, and $\text{sign}(\mu_i - s) = 1, \forall i \in \{g+1, \dots, k\}$. It yields

$$s_g = \frac{\alpha_g}{2} \left(\frac{1}{\eta_g} - \frac{1}{1 - \eta_g} \right).$$

Lemma A.5.2 states that the function $s \mapsto \text{Var}|X - s|$ is piecewise parabolic. On each interval $[\mu_g, \mu_{g+1}]$, $g \in \{1, \dots, k-1\}$, the point where the derivative vanishes corresponds to a minimum.

Moreover, $s^* \in [\mu_1, \mu_k]$. Otherwise, $\text{Var}|X - s| \geq 1$ for all $s \in [\mu_1, \mu_k]$, which contradicts Proposition A.1.1. Hence,

$$s^* = \underset{s \in \mathbb{R}}{\text{argmin}} \text{Var}|X - s| = s_{g^*}^\dagger, \quad \text{where} \quad g^* = \underset{g \in \{1, \dots, k-1\}}{\text{argmin}} \text{Var}|X - s_g^\dagger|.$$

Lemma A.5.4 This lemma is used in the proof of Proposition 1.

For all $g \in \{1, \dots, k-2\}$, the sequence $(s_g)_g$ is strictly increasing, i.e., $s_g < s_{g+1}$.

Proof We have $s_g = \frac{\alpha_g}{2} \left(\frac{1}{\eta_g} - \frac{1}{1 - \eta_g} \right)$ for $g \in \{1, \dots, k-1\}$. Using $\alpha_{g+1} = \alpha_g + \mu_{g+1}\varepsilon_{g+1}$ and $\eta_{g+1} = \eta_g + \varepsilon_{g+1}$, for $g \in \{1, \dots, k-2\}$, we obtain

$$s_{g+1} - s_g = \frac{\varepsilon_{g+1}}{2(1 - \eta_g)(1 - \eta_{g+1})} \left[-\alpha_g - \mu_{g+1}(1 - \eta_g) \right] + \frac{\varepsilon_{g+1}}{2\eta_g\eta_{g+1}} (-\alpha_g + \mu_{g+1}\eta_g),$$

where both $\frac{\varepsilon_{g+1}}{2(1 - \eta_g)(1 - \eta_{g+1})} > 0$ and $\frac{\varepsilon_{g+1}}{2\eta_g\eta_{g+1}} > 0$. As the locations are ordered in model (2):

- $\mu_i \leq \mu_g < \mu_{g+1} \quad \forall i \leq g$. This implies $-\alpha_g + \mu_{g+1}\eta_g > 0$.
- $\mu_i \geq \mu_{g+1} \quad \forall i > g$. As $\mathbb{E}(X) = 0$, we have $\sum_{i=g+1}^k \varepsilon_i \mu_i = -\alpha_g$. It results in $-\alpha_g - \mu_{g+1}(1 - \eta_g) \geq 0$.

Thus, $s_{g+1} - s_g > 0$.

Proof Proof of Proposition 1

1. We want to prove that $s^* = s_{g^*}$, where $g^* = \underset{g \in \{1, \dots, k-1\}}{\text{argmin}} \min_{s \in [\mu_g, \mu_{g+1}]} \text{Var}|X - s|$ and $s_g = \frac{\alpha_g}{2} \left(\frac{1}{\eta_g} - \frac{1}{1 - \eta_g} \right)$.

Let us prove the result by contradiction. By Lemma A.5.3, if $s^* \neq s_{g^*}$, then $s^* = \mu_{g+1}$ where $g \in \{1, \dots, k-2\}$. In this case we would have $s_g^\dagger = \mu_{g+1}$ meaning that $s_g \geq \mu_{g+1}$, and $s_{g+1}^\dagger = \mu_{g+1}$ meaning that $s_{g+1} \leq \mu_{g+1}$. It implies that $s_{g+1} \leq s_g$ which contradicts Lemma A.5.4. It results that $s^* = s_{g^*}$.

2. The expression of $\Phi^*(X)$ is then obtained by substituting s^* from the previous result into $\Phi^*(X) = 4 \text{Var}|X - s^*|$.

3. Solving the equation $\Phi^*(X) = 1$ yields $\frac{\alpha_{g^*}^2}{(\sqrt{3}/4)^2} + \frac{(\eta_{g^*} - \frac{1}{2})^2}{(1/2)^2} = 1$, which is the canonical equation of an ellipse in the variables $(\alpha_{g^*}, \eta_{g^*})$, centered at $(0, 1/2)$, with semi-minor axis $\sqrt{3}/4$ and semi-major axis $1/2$.

4. Solving $\Phi^*(X) \geq 1$, we obtain

$$\frac{\alpha_{g^*}^2}{(\sqrt{3}/4)^2} + \frac{(\eta_{g^*} - \frac{1}{2})^2}{(1/2)^2} \leq 1 \iff \alpha_{g^*}^2 \leq \frac{3}{4} \eta_{g^*} (1 - \eta_{g^*}).$$

Since $\alpha_{g^*} \leq 0$, this reduces to $-\frac{\sqrt{3}}{2} \sqrt{\eta_{g^*}(1 - \eta_{g^*})} \leq \alpha_{g^*} \leq 0$ which proves the result.

A.6 Proof of Proposition 2

Here, we provide the proof of the explicit expression for $\Phi^{**}(X)$ for the k -Dirac mixture model, and the conditions under which the ratio falsely indicates unimodality.

Proof *Proof of Proposition 2*

1. From Siffer [2019] (Theorem 6.1), and using that $\mathbb{E}[X] = 0$ and $\text{Var}[X] = 1$, we obtain

$$s^{**} = \frac{\text{Cov}(X, X^2)}{2\text{Var}X} = \frac{\mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2)}{2\text{Var}X} = \frac{\gamma}{2}, \text{ with } \gamma = \mathbb{E}[X^3].$$

2. The expression of $\Phi^{**}(X)$ is then obtained by substituting s^{**} from the previous result into $\Phi^{**}(X) = 4\text{Var}|X - s^{**}|$.

3. Solving $\Phi^{**}(X) = 1$, we find

$$\left(4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma - \sqrt{\gamma^2 + 3}\right) \left(4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma + \sqrt{\gamma^2 + 3}\right) = 0.$$

Since $0 < \eta_{g^{**}} < 1$, $-(2\eta_{g^{**}} - 1)\gamma \leq |\gamma|$. Since $\alpha_{g^{**}} < 0$ and $\sqrt{\gamma^2 + 3} > |\gamma|$, $4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma - \sqrt{\gamma^2 + 3} < 0$. Thus, there is a unique solution, which defines the line

$$4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma + \sqrt{\gamma^2 + 3} = 0.$$

4. Solving $\Phi^{**}(X) \geq 1$, we obtain $4\alpha_{g^{**}} - (2\eta_{g^{**}} - 1)\gamma + \sqrt{\gamma^2 + 3} \geq 0$.

A.7 Proof of Proposition 3

We consider a standardized random variable X following a 3-Dirac mixture with means μ_1 , μ_2 and μ_3 , sorted in ascending order. When the minimizer of $\text{Var}|X - s|$ is not unique, we take s^* to be the smallest minimizer.

Proof *Proof of Proposition 3*

From the constraint $\sum_{i=1}^3 \varepsilon_i = 1$, we deduce $\varepsilon_2 = 1 - \varepsilon_1 - \varepsilon_3$. Using $\mathbb{E}[X] = 0$, $\text{Var}[X] = 1$, and the ordering $\mu_2 < \mu_3$, we obtain

$$\begin{aligned} \mu_2 &= -\frac{\varepsilon_1}{1 - \varepsilon_1} \mu_1 - \frac{1}{1 - \varepsilon_1} \sqrt{\frac{\varepsilon_3}{1 - \varepsilon_1 - \varepsilon_3}} \sqrt{1 - \varepsilon_1(1 + \mu_1^2)} \\ \mu_3 &= -\frac{1}{\varepsilon_3} (\varepsilon_1 \mu_1 + \varepsilon_2 \mu_2). \end{aligned}$$

From the conditions $1 - \varepsilon_1(1 + \mu_1^2) > 0$ and $\mu_2 > \mu_1$, together with the fact that $\mu_1 < 0$ in the model, we deduce

$$\mu_1 \in \left[-\sqrt{\frac{1 - \varepsilon_1}{\varepsilon_1}}, -\sqrt{\frac{\varepsilon_3}{1 - \varepsilon_3}} \right]. \quad (5)$$

1. We seek conditions on μ_1 such that $s^* \in [\mu_1, \mu_2]$. By definition,

$$s^* \in [\mu_1, \mu_2] \iff \text{Var}[X - s_1^\dagger] \leq \text{Var}[X - s_2^\dagger].$$

By Proposition 1.1, this is equivalent to

$$\text{Var}[X - s_1] \leq \text{Var}[X - s_2^\dagger],$$

where, by Lemma A.5.3,

$$s_2^\dagger = \begin{cases} \mu_2, & \text{if } s_2 < \mu_2, \\ s_2, & \text{if } s_2 \geq \mu_2. \end{cases}$$

Consequently,

$$s^* \in [\mu_1, \mu_2] \iff \begin{cases} \text{Var}[X - s_1] \leq \text{Var}[X - \mu_2], & \text{if } s_2 < \mu_2, \\ \text{Var}[X - s_1] \leq \text{Var}[X - s_2], & \text{if } s_2 \geq \mu_2. \end{cases}$$

Combining the cases $s_2 < \mu_2$ (see Lemma A.7.1 below) and $s_2 \geq \mu_2$ (see Lemma A.7.2 below), we conclude that

$$s^* \in [\mu_1, \mu_2] \iff \mu_1 \leq \max\{U_1, U_2\}.$$

2. When $s^* \in [\mu_1, \mu_2]$, we have $g^* = 1$. Applying Lemma 1.4 with $g^* = 1$, we have $\Phi^*(X) \geq 1$ if and only if

$$\mu_1 \geq -\frac{\sqrt{3}}{2} \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} \geq -\sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}},$$

which directly gives the expression of $F(\varepsilon_1)$.

Lemma A.7.1 This lemma is used in the proof of Proposition 3.

$$s^* \in [\mu_1, \mu_2] \text{ and } s_2 < \mu_2 \iff \mu_1 < U_1,$$

where

$$U_1 = -\frac{2\varepsilon_2 + 4\varepsilon_1\varepsilon_3}{\sqrt{4\varepsilon_1\varepsilon_3(\varepsilon_2 + 4\varepsilon_1\varepsilon_3)}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}}.$$

Proof When $s^* \in [\mu_1, \mu_2]$, s_1 minimizes $\text{Var}[X - s]$ over $s \in [\mu_1, \mu_2]$, so the inequality

$$\text{Var}[X - s_1] \leq \text{Var}[X - \mu_2]$$

is always satisfied. It therefore remains to characterize the conditions under which $s_2 < \mu_2$. From Lemma A.5.3,

$$s_2 = -\frac{\mu_3}{2} \frac{2\varepsilon_3 - 1}{1 - \varepsilon_3}.$$

Expressing μ_1 and μ_2 as functions of μ_3 results in

$$\begin{aligned} \mu_2 &= -\frac{\varepsilon_3}{1-\varepsilon_3} \mu_3 + \frac{1}{1-\varepsilon_3} \sqrt{\frac{\varepsilon_1}{\varepsilon_2}} \sqrt{1-\varepsilon_3(1+\mu_3^2)}, \\ \mu_1 &= -\frac{1}{1-\varepsilon_3} \left(\varepsilon_3 \mu_3 + \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} \sqrt{1-\varepsilon_3(1+\mu_3^2)} \right). \end{aligned}$$

Imposing $\mu_2 < \mu_3$ and $1 - \varepsilon_3(1 + \mu_3^2) \geq 0$ yields

$$\mu_3 \in \left[\sqrt{\frac{\varepsilon_1}{1-\varepsilon_1}}, \sqrt{\frac{1-\varepsilon_3}{\varepsilon_3}} \right]. \quad (6)$$

Solving $s_2 < \mu_2$, we obtain

$$\mu_3 < 2 \sqrt{\frac{\varepsilon_1(1-\varepsilon_3)}{\varepsilon_2 + 4\varepsilon_1\varepsilon_3}}.$$

and the upper bound lies within the interval from (6). On this domain, μ_1 is nondecreasing in μ_3 , so that $s_2 < \mu_2 \iff \mu_1 < U_1$.

Lemma A.7.2 This lemma is used in the proof of Proposition 3.

$$s^* \in [\mu_1, \mu_2] \text{ and } s_2 \geq \mu_2 \iff \mu_1 \in [U_1, U_2],$$

where

$$U_2 = -\frac{\sqrt{2}}{2} \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} + \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}}.$$

Proof Following the line of reasoning of the proof of Lemma A.7.1,

$$s_2 \geq \mu_2 \iff \mu_1 \geq U_1.$$

Moreover, solving the inequality

$$\text{Var}[X - s_1] \leq \text{Var}[X - s_2]$$

using the first variance expression from Lemma A.5.1(2.) together with the explicit expressions of s_1 and s_2 from Lemma A.5.3 yields

$$\mu_1 \leq U_2.$$

U_2 lies within the interval from (5). Therefore,

$$s^* \in [\mu_1, \mu_2] \text{ and } s_2 \geq \mu_2 \iff \mu_1 \in [U_1, U_2].$$

B Additional Gaussian mixture illustrations

In this section, we examine several 3-Gaussian mixtures and give supplementary figures, extending the analysis beyond Section 2.2 of the main paper. This additional study aims to confirm that the Gaussian mixture model reproduces the results obtained for the Dirac mixture model in cases where the SNR is sufficiently large.

Fig. 2 illustrates a 3-Gaussian mixture $X \sim \varepsilon_1 \cdot \mathcal{N}(\mu_1, \sigma^2) + \varepsilon_2 \cdot \mathcal{N}(\mu_2, \sigma^2) + \varepsilon_3 \cdot \mathcal{N}(\mu_3, \sigma^2)$, with $\sigma^2 \in \{0.0225, 0.0625, 0.25\}$. The three configurations are as follows:

- In Fig. 2a: $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (1/3, 1/3, 1/3)$ and $(\mu_1, \mu_2, \mu_3) = (-2, 0, 2)$. It is an example where both $\Phi^*(X)$ and $\Phi^{**}(X)$ fail in the Dirac case ($\sigma^2 = 0$).
- In Fig. 2b: $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.2, 0.4, 0.4)$ and $(\mu_1, \mu_2, \mu_3) = (-2, 0, 2)$. It corresponds to Example 1.a in the main paper, where $\Phi^{**}(X)$ fails but $\Phi^*(X)$ does not.
- In Fig. 2c: $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.5, 0.2, 0.3)$ and $(\mu_1, \mu_2, \mu_3) = (-2, 0, 2.5)$. It illustrates a case where neither pivot fail in the Dirac case.

When both pivots fail (Fig. 2a) or neither fails (Fig. 2c) in the Dirac case, increasing σ^2 does not alter the conclusions of $\Phi^*(X)$ and $\Phi^{**}(X)$. When only the approximate pivot fails (Fig. 2b), a sufficiently small σ^2 reproduces the Dirac results ($\Phi^*(X) < 1$ and $\Phi^{**}(X) > 1$). However, as σ^2 increases, $\Phi^*(X)$ exceeds one, leading to an incorrect unimodal conclusion.

In all illustrated cases, both the DFTU and the dip test correctly detect multimodality, even when FTU* and FTU** fail to do so (Figs. 2a and 2b). These findings are consistent with the simulation results presented in Section 4 of the main paper.

C Details about the double-folding procedure

C.1 Application to the “pathological” example

This section analyzes the application of the double-folding procedure to the “pathological” example, explains the use of the approximate pivot in the second step, and shows that the procedure correctly detects multimodality, as noted in Section 3.1 of the main paper.

The model is the following balanced and symmetric 3-Dirac mixture:

$$X \sim \frac{1}{3} \delta_{-a} + \frac{1}{3} \delta_0 + \frac{1}{3} \delta_a.$$

The first step of the procedure corresponds to Example 8.1 from Siffer [2019] with the exact pivot approach, yielding $s_1^* = \pm \frac{a}{4}$ and $\Phi_1^{**}(X) = 1$. The first-step approximate pivot and associated SFR are $s_1^{**} = 0$ and $\Phi_1^{**}(X) = \frac{4}{3} > 1$, as in Example 8.1.

Folding with an exact pivot yields the following distribution:

$$\tilde{X}^* \sim \frac{1}{3} \delta_{\frac{1}{4}a} + \frac{1}{3} \delta_{\frac{3}{4}a} + \frac{1}{3} \delta_{\frac{5}{4}a}.$$

Since $\tilde{X}^* = \frac{1}{2}X + \frac{3}{4}a$ and the SFR are affine invariant, the unimodal conclusion of the first step persists in the second step.

However, folding with the approximate pivot, as done in the double-folding procedure, yields

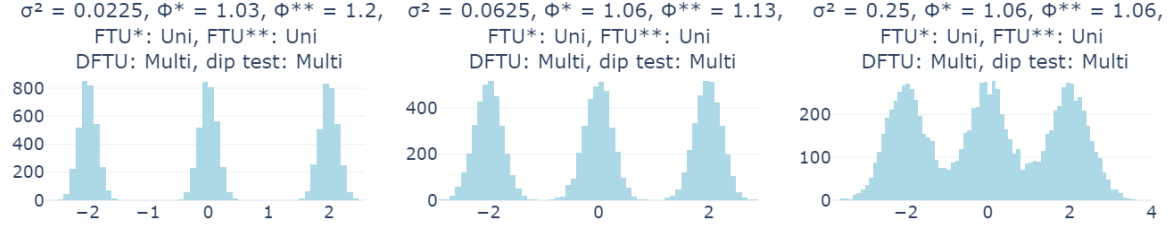
$$\tilde{X}^{**} \sim \frac{1}{3} \delta_0 + \frac{2}{3} \delta_a.$$

The folded distribution is no longer balanced or symmetric, leading to $s_2^* = s_2^{**} = \frac{1}{4}$ and $\Phi_2^*(X) = \Phi_2^{**}(X) = 0$, which indicates multimodality.

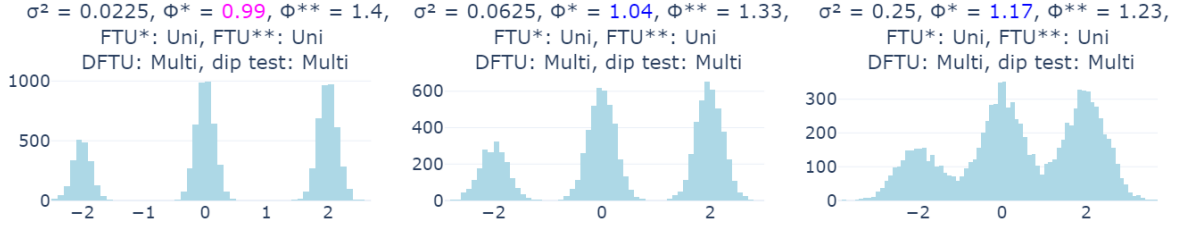
C.2 Details about the 3-Dirac mixture analysis

In Section 3.2 of the main paper, we apply the double-folding procedure to a 3-Dirac mixture. This paragraph details the analysis, in particular the optimization procedure involved.

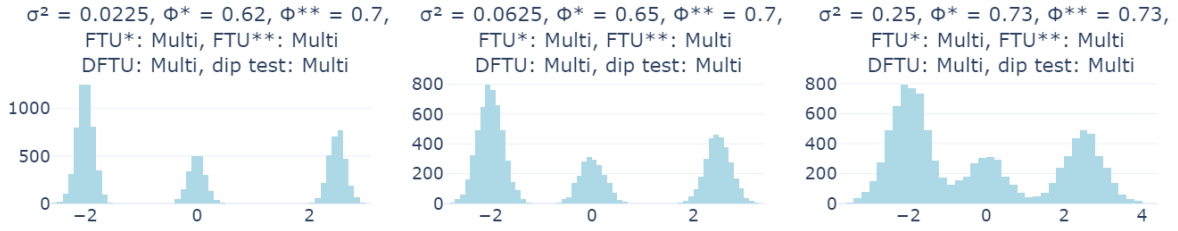
We consider a standardized random variable X following a 3-Dirac mixture with locations $\mu_1 < \mu_2 < \mu_3$. In this case, when the minimizer of $\text{Var}[X - s]$ is not unique, we take s^* to be the smallest minimizer. First, we analyze the



(a) Both pivots fail in the Dirac case



(b) Only the approximate pivot fails in the Dirac case



(c) Neither pivot fails in the Dirac case

Figure 2: 3-Gaussian mixtures under different configurations, together with the corresponding exact and approximate SFR (Φ^* and Φ^{**}) and the results of unimodality tests. “Uni” and “Multi” denote unimodal and multimodal, respectively. Each column corresponds to a value of the mixture within variance σ^2 . Each row corresponds to an example obtained in the Dirac case ($\sigma^2 = 0$): (a) both pivots fail, (b) only the approximate pivot fails, and (c) neither pivot fails.

first-step SFR. If the multimodality of the mixture is detected at this step, the procedure stops and returns the correct conclusion. We are therefore interested in the case where the first-step SFR fails, which can be fully characterized in terms of μ_1 , ε_1 , and ε_3 . According to Proposition 3.1, the existence of $s_1^* \in [\mu_1, \mu_2]$ is equivalent to

$$\mu_1 \leq U(\varepsilon_1, \varepsilon_3), \quad U(\varepsilon_1, \varepsilon_3) := \max\{U_1, U_2\}$$

where

$$U_1 = -\frac{\sqrt{2}}{2} \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} + \sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}} \quad \text{and} \quad U_2 = -\frac{2\varepsilon_2 + 4\varepsilon_1\varepsilon_3}{\sqrt{4\varepsilon_1\varepsilon_3(\varepsilon_2 + 4\varepsilon_1\varepsilon_3)}} \sqrt{\frac{\varepsilon_3}{1-\varepsilon_3}}.$$

Proposition 3.2 then states that, when $s_1^* \in [\mu_1, \mu_2]$, the first-step SFR fails if and only if

$$\mu_1 \geq F(\varepsilon_1).$$

Recall that the case $s_1^* \in [\mu_2, \mu_3]$ need not be analyzed separately. Indeed, by the affine invariance of the SFR, applying the transformation $X \mapsto -X$ maps this case to $s_1^* \in [\mu_1, \mu_2]$ while leaving the SFR unchanged. Therefore, it suffices to consider the case $s_1^* \in [\mu_1, \mu_2]$ without loss of generality.

To study the case $\mu_1 \in [F(\varepsilon_1), U(\varepsilon_1, \varepsilon_3)]$, we rely on numerical optimization using SciPy's SLSQP solver [Virtanen et al., 2020]. The optimization variables are $\mu_1 \in [-10, 0]$, $\varepsilon_1 \in [0.01, 0.99]$, and $\varepsilon_3 \in [0.01, 0.99]$, from which ε_2 , μ_2 , and μ_3 can be deduced, as detailed in the proof of Proposition 3. We impose the following constraints:

- $1 - \varepsilon_1 - \varepsilon_3 > 0$,
- $F(\varepsilon_1) \leq \mu_1 \leq U(\varepsilon_1, \varepsilon_3)$.

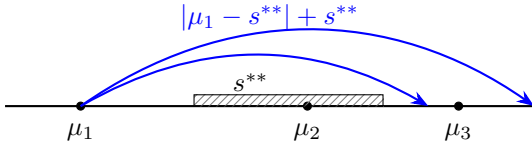
The first two optimization problems

$$\min [\gamma - \mu_1 - \mu_2], \quad \text{and} \quad \min [\mu_2 + \mu_3 - \gamma]$$

with $\gamma = \mathbb{E}[X^3]$, both yield positive values when solved numerically. This implies that

$$s^{**} = \frac{\gamma}{2} \in \left(\frac{\mu_1 + \mu_2}{2}, \frac{\mu_2 + \mu_3}{2} \right).$$

Let $\tilde{\mu}_1$, $\tilde{\mu}_2$, and $\tilde{\mu}_3$ denote the locations after the transformation $X \mapsto |X - s^{**}|$. The preceding result implies that $\tilde{\mu}_1 > \tilde{\mu}_2$ and $\tilde{\mu}_3 > \tilde{\mu}_2$, so $\tilde{\mu}_2$ is the smallest of the transformed locations.



To apply Proposition 3.2 in the second step of the DFTU, we standardize and order the locations $\tilde{\mu}_1$, $\tilde{\mu}_2$, and $\tilde{\mu}_3$. The resulting locations are denoted by $\mu_{(1)}^{(2)}$, $\mu_{(2)}^{(2)}$, and $\mu_{(3)}^{(2)}$, with $\mu_{(1)}^{(2)} < \mu_{(2)}^{(2)} < \mu_{(3)}^{(2)}$, and corresponding proportions $\varepsilon_{(1)}^{(2)}$, $\varepsilon_{(2)}^{(2)}$, and $\varepsilon_{(3)}^{(2)}$.

To apply Proposition 3.2, we need to verify that $s_2^* \in [\mu_{(1)}^{(2)}, \mu_{(2)}^{(2)}]$, which, according to Proposition 3.1, is equivalent to $\mu_{(1)}^{(2)} \leq U(\varepsilon_{(1)}^{(2)}, \varepsilon_{(3)}^{(2)})$. This leads to the third optimization problem

$$\min [U(\varepsilon_{(1)}^{(2)}, \varepsilon_{(3)}^{(2)}) - \mu_{(1)}^{(2)}]$$

whose numerical solution is positive, thereby ensuring that $s_2^* \in [\mu_{(1)}^{(2)}, \mu_{(2)}^{(2)}]$.

The final optimization problem

$$\min [\mu_{(1)}^{(2)} - F(\varepsilon_{(1)}^{(2)})]$$

attains a positive value using a numerical solver, ensuring that $\mu_{(1)}^{(2)} < F(\varepsilon_{(1)}^{(2)})$. According to Proposition 3.2, this implies that the second-step SFR cannot fail when the first-step one has failed.

C.3 DFTU algorithm

See Algorithm C.3.1.

D Details about the simulation study

In this section, we provide details about the simulation study. The objective is to compare the results of different unimodality tests across various mixture models. All computations were done in Python 3.13. Replication files are available from: <https://github.com/cbecquart/DFTU-Replication>.

For each mixture with k components, we generate 100 datasets, each containing 1,000 points. Specifically, we draw 1,000 ε_i points from the distribution of the i -th component of the mixture model, for $i \in \{1, \dots, k\}$. For each dataset, we apply the following tests at a 5% significance level:

- the FTU with the exact pivot, denoted FTU*, partially based on an implementation provided by A. Siffer (personal communication);

Algorithm C.3.1 Algorithm of the DFTU**Input:** Data matrix X , critical values q_1 and q_2 **Output:** Unimodal or Multimodal

```

 $s_1^* \leftarrow \operatorname{argmin}_{s \in \mathbb{R}} \operatorname{Var}|X - s|$  ▷ Exact pivot on  $X$ 
 $\Phi_1^*(X) \leftarrow 4 \operatorname{Var}|X - s_1^*|/\operatorname{Var}X$ 
if  $\Phi_1^*(X) < q_1$  then
  return Multimodal
else
   $s_1^{**} \leftarrow \operatorname{argmin}_{s \in \mathbb{R}} \operatorname{Var}(X - s)^2$  ▷ Approximated pivot on  $X$ 
   $\tilde{X} \leftarrow |X - s_1^{**}|$  ▷ Folding
   $s_2^* \leftarrow \operatorname{argmin}_{s \in \mathbb{R}} \operatorname{Var}|\tilde{X} - s|$  ▷ Exact pivot on  $\tilde{X}$ 
   $\Phi_2^*(X) \leftarrow 4 \operatorname{Var}|\tilde{X} - s_2^*|/\operatorname{Var}\tilde{X}$ 
  if  $\Phi_2^*(X) < q_2$  then
    return Multimodal
  else
    return Unimodal
  end if
end if

```

- the FTU with the approximate pivot, denoted FTU**, implemented as a Python translation of the R code available in the `Rfolding` package [Siffer, 2018];
- the DFTU, implemented according to Algorithm C.3.1, with parameter α_1 set to 3% and α_2 deduced from α and α_1 ;
- Hartigan’s dip test, implementation from the `dip.test` library [Urlus, 2022].

The critical values for all three folding-based tests are obtained via Monte Carlo simulation, as described in Section 3.1 of the main paper. For each test, we record the number of times unimodality is concluded across the 100 datasets. Table D.1 reports these counts. Given the strong polarization of the results, values close to 100 are reported as “Uni” (for unimodal) and those close to 0 as “Multi” (for multimodal) in Table 1 of the main paper. Table D.2 provides details about the distributions used in the simulation study. See also the comments in the main paper.

Distribution	FTU*	FTU**	DFTU	dip test
$\mathcal{N}(0, 1)$	100	100	100	100
$0.5 \cdot (\mathcal{N}_0 + \mathcal{N}_1)$	100	100	100	100
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([4, 8])$	0	0	0	0
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([1, 4])$	1	1	1	100
$1/3 \cdot \delta_{-2} + 1/3 \cdot \delta_0 + 1/3 \cdot \delta_2$	100	100	0	0
$1/3 \cdot \mathcal{N}_{-2} + 1/3 \cdot \mathcal{N}_0 + 1/3 \cdot \mathcal{N}_2$	100	100	4	0
$0.2 \cdot (\delta_{-3} + \delta_{-1.5} + \delta_{2.5} + \delta_4 + \delta_{11})$	100	100	0	0
$0.2 \cdot (\mathcal{N}_{-3} + \mathcal{N}_{-1.5} + \mathcal{N}_{2.5} + \mathcal{N}_4 + \mathcal{N}_{11})$	100	100	0	0

Table D.1: Simulation results. Counts of unimodality across the 100 datasets for the FTU with exact pivot (FTU*), the FTU with approximate pivot (FTU**), the DFTU, and the dip test. $\mathcal{N}_a = \mathcal{N}(a, 0.5)$.

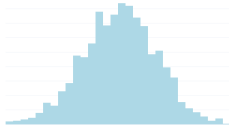
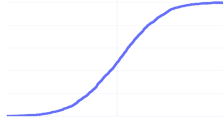
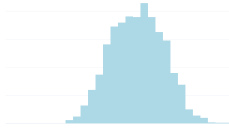
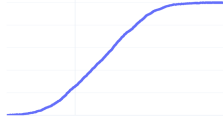
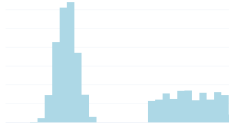
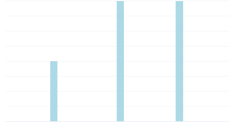
Distribution	Histogram	ECDF	Description
$\mathcal{N}(0, 1)$			Simple case of unimodality
$0.5 \cdot (\mathcal{N}_0 + \mathcal{N}_1)$			2-Gaussian mixture, not well separated
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([4, 8])$			Mixture of Gaussian and Uniform distributions, well separated
$0.6 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{U}([1, 4])$			Mixture of Gaussian and Uniform distributions, less separated
$0.2 \cdot \delta_{-2} + 0.4 \cdot \delta_0 + 0.4 \cdot \delta_2$			3-Dirac mixture
$0.2 \cdot \mathcal{N}_{-2} + 0.4 \cdot \mathcal{N}_0 + 0.4 \cdot \mathcal{N}_2$			3-Gaussian mixture, well separated

Table D.2: Details about the distributions used in the simulations.

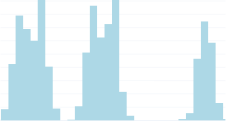
Distribution	Histogram	ECDF	Description
$0.2 \cdot (\delta_{-3} + \delta_{-1.5} + \delta_{2.5} + \delta_4 + \delta_{11})$			5-Dirac mixture
$0.2 \cdot (\mathcal{N}_{-3} + \mathcal{N}_{-1.5} + \mathcal{N}_{2.5} + \mathcal{N}_4 + \mathcal{N}_{11})$			5-Gaussian mixture, well separated

Table D.2: Details about the distributions used in the simulations (Continued).

Acknowledgments

We thank Alban Siffer for stimulating discussions and code about the FTU. A.M.R. and C.B. acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d’Avenir) program, grant ANR-17-EURE-0010. Part of C.B.’s work was done during an Erasmus+ visit funded by the University of Toulouse at the University of Helsinki.

References

- Argyris Kalogeratos and Aristidis Likas. Dip-means: an incremental clustering method for estimating the number of clusters. *Advances in neural information processing systems*, 25, 2012.
- Mahdi Shahbaba and Soosan Beheshti. Efficient unimodality test in clustering by signature testing. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8282–8286. IEEE, 2014.
- Andreas Adolfsson, Margareta Ackerman, and Naomi C Brownstein. To cluster, or not to cluster: An analysis of clusterability methods. *Pattern Recognition*, 88:13–26, 2019.
- Andreas Krause and Volkmar Liebscher. Multimodal projection pursuit using the dip statistic. 2005.
- P. Arabie and L. J. Hubert. Cluster Analysis in Marketing Research. In R. P. Bagozzi, editor, *Advanced methods in marketing research*, pages 160–189. Blackwell, Oxford, 1994.
- Andreas Alfons, Aurore Archimbaud, Klaus Nordhausen, and Anne Ruiz-Gazen. Tandem clustering with invariant coordinate selection. *Econometrics and Statistics*, In press, March 2024. ISSN 2452-3062. doi:10.1016/j.ecosta.2024.03.002. URL <https://www.sciencedirect.com/science/article/pii/S2452306224000121>.
- A. Ya Khintchine. On unimodal distributions. *Izvestiya Nauchno-Issledovatel’skogo Instituta Matematiki i Mekhaniki*, 2(2):1–7, 1938.
- John A Hartigan and Pamela M Hartigan. The dip test of unimodality. *The annals of Statistics*, pages 70–84, 1985.
- B. W. Silverman. Using Kernel Density Estimates to Investigate Multimodality. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 43(1):97–99, 1981.
- Alban Siffer, Pierre-Alain Fouque, Alexandre Termier, and Christine Largouët. Are your data gathered? In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2210–2218, London United Kingdom, 2018. ACM.
- Alban Siffer. *New statistical methods for data mining, contributions to anomaly detection and unimodality testing*. Doctoral dissertation, Rennes 1, 2019.
- Paraskevi Chasani and Aristidis Likas. The UU-test for statistical modeling of unimodal data. *Pattern Recognition*, 122:108272, 2022.
- Prodromos Kolyvakis and Aristidis Likas. A Multivariate Unimodality Test Harnessing the Dip Statistic of Mahalanobis Distances Over Random Projections. *arXiv preprint arXiv:2311.16614*, 2023.
- Fred C Leone, Lloyd S Nelson, and RB Nottingham. The folded normal distribution. *Technometrics*, 3(4):543–550, 1961.
- Peter J. Rousseeuw and Christophe Croux. Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.
- Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:10.1038/s41592-019-0686-2.
- Alban Siffer. Rfolding: The folding test of unimodality. <https://cran.r-project.org/package=Rfolding>, 2018.
- Ralph Urlus. diptest: A Python/C(++) implementation of Hartigan & Hartigan’s dip test for unimodality. <https://pypi.org/project/diptest/>, 2022.