Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semi-parametric location mixtures from ranked set samples: what the sampling design and the ranks contribute

Seyed Jalaleddin Moniri

jjmoniri@gmail.com

Independent Researcher https://orcid.org/0009-0005-2346-9711

Research Article

Keywords: imperfect ranking, concomitant variable, EM-type algorithm, misplacement probabilities, kernel density estimation, Monte Carlo simulation

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11863149/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Semi-parametric location mixtures from ranked set samples: what the sampling design and the ranks contribute

Seyed Jalaleddin Moniri
Independent researcher, Vancouver, BC, Canada
Corresponding author: Seyed Jalaleddin Moniri, jjmoniri@gmail.com

Abstract

Ranked set sampling (RSS) is often proposed for estimating finite mixtures when units can be ranked cheaply on a concomitant variable. For a two-component location mixture of an unknown symmetric density, we ask how much of the gain from RSS comes from the sampling design and how much from the ranks. We restate an EM-type algorithm with latent true ranks, correct the normalising constant of its rank-aware mixing-proportion update, and note that its modified location step is an unbiased estimating equation that does not use the ranks. A simulation study of 420 configurations with 1,000 replicates each, covering three component laws and seven ranking scenarios, gives three main findings. The gain over simple random sampling comes from the design, and fitting the pooled sample while ignoring the ranks captures it. The rank-aware estimator does not improve on that fit and is clearly worse with heavy-tailed components. When ranking uses a noisy concomitant, the constant-weight misplacement model is misspecified; even with the component law known, the resulting bias does not shrink over the sample sizes studied, so the rank-aware fit falls further behind as the sample grows. For heavy-tailed components, a likelihood-type location update improves the rank-ignoring fit.

Keywords: imperfect ranking; concomitant variable; EM-type algorithm; misplacement probabilities; kernel density estimation; Monte Carlo simulation.

1 Introduction

Ranked set sampling (RSS) is used when measuring a unit is expensive but a small set of units can be ordered cheaply, by eye or through an inexpensive concomitant variable (McIntyre 1952; Dell and Clutter 1972; Chen et al. 2004). In each set of H units drawn from the population, only the unit judged to hold a prescribed rank is measured. Repeating this over n cycles yields $N = nH$ measured units that are independent but not identically distributed, and that are spread more evenly across the population than a simple random sample (SRS) of the same size.

Finite mixture models are a natural target for RSS. In osteoporosis research, for example, bone mineral density is measured by dual X-ray absorptiometry while age, weight or body-mass index are available at no cost, and the population is often modelled as a mixture of healthy and affected individuals (Omidvar et al. 2018; Alvandi et al. 2024). When the component shape is unknown, a widely used semi-parametric model is the location mixture of a single symmetric density,

$$g(x) = \sum_{j=1}^{M} \pi_j f(x - \mu_j). \tag{1}$$

This paper asks how much ranked set sampling improves estimation in model (1), and where the improvement comes from. Two sources must be separated. The *design* spreads the measured

units over the population; any consistent estimator applied to the pooled sample benefits from this without using the ranks. The *ranks* carry further information about where each measured unit lies in the population distribution, which only a likelihood that conditions on the ranks can use. Using the ranks requires a model for ranking errors, and the model matters: a rank-aware estimator built on a wrong ranking model can do worse than one that ignores the ranks.

Our starting point is the EM-type algorithm for model (1) under RSS developed in the author's MSc thesis (Moniri 2021), which extends the latent-variable construction of Hatefi et al. (2014, 2015) from parametric to semi-parametric components. The contributions of this paper are as follows.

1. We restate the algorithm, correct the normalising constant of its rank-aware update for the mixing proportions, and note, using a component-wise version of the RSS case of a lemma of Hatefi et al. (2015), that its modified location step is an unbiased estimating equation from which the ranks drop out (Section 4).

2. We document by direct simulation that the constant-weight misplacement model is misspecified when ranking is done on a noisy concomitant, and we quantify the consequences for mixture estimation, including a bias that does not shrink over the sample sizes studied (Sections 3 and 6.5).

3. We give a parametric benchmark for how much the ranks could add beyond the design (Section 4.6).

4. In a simulation study covering three component laws, four combinations of set size and number of cycles, and seven ranking scenarios, we separate the contribution of the design from that of the ranks by comparing every estimator with a rank-ignoring fit on the same ranked set sample, and we report failure rates by cause for every method (Sections 5 and 6).

The results lead to practical recommendations (Section 8): collect data by RSS when ranking is cheap, fit the pooled sample while ignoring the ranks, choose the location update according to tail weight, and avoid constant-weight rank-aware likelihoods when ranking relies on a concomitant.

2 Literature review and research gap

Ranked set sampling and imperfect ranking. RSS was introduced by McIntyre (1952); Dell and Clutter (1972) gave its order-statistic theory and the concomitant ("perceptual error") model of imperfect ranking used here, in which units are ranked on the response plus independent normal noise. A general account is Chen et al. (2004). Several models for imperfect rankings exist (see, e.g., Frey 2007). For rank-ignoring procedures, Presnell and Bohn (1999) showed that random-sample U-statistics applied to ranked set sample data are asymptotically at least as efficient as on an SRS, whatever the accuracy of the ranking. For parametric estimation of a single distribution, Taconeli and Rodrigues de Lara (2024) found that maximum likelihood under the perfect-ranking assumption performed best under perfect ranking but often produced highly biased estimators when ranking errors were present, and that maximum likelihood under SRS assumptions, among other alternatives, was preferable under imperfect ranking.

Finite mixtures from ranked set samples. Hatefi et al. (2014) developed maximum likelihood estimation and an EM algorithm for finite mixtures from ranked set samples, treating the data as independent order statistics, that is, assuming perfect ranking; they introduced the latent

component labels and counts used below. Hatefi et al. (2015) extended the approach to partially rank-ordered set samples with ranking errors between judgment subsets. They modelled the errors by a doubly stochastic misplacement-probability matrix, which for subsets of size one gives the constant-weight form (2) below. They also proposed a modified EM algorithm in which the hazard-rate terms of the location score are replaced by their expectations, so that the location update coincides with the SRS update. In their simulation study, imperfect samples were generated with the concomitant model of Dell and Clutter (1972), misplacement probabilities estimated in a first simulation stage were treated as the true values, and the estimators were found to be slightly biased. Their efficiency comparisons were made between designs (PROS, RSS and SRS), and failed replicates were replaced by new data sets. Alvandi et al. (2024) developed a Bayesian version for imperfect RSS with the same misplacement-probability model, again estimating the misplacement matrix in a first stage from concomitant ranking and comparing RSS-based with SRS-based estimators (design details as described in the preprint version, Alvandi et al. 2023). Fisher-information comparisons (Hatefi and Jafari Jozani 2013) and judgment post-stratification (Omidvar et al. 2018) have also been studied. The PhD thesis of Hatefi (2014) uses the same constant-weight form for the judgment-rank density. The present paper builds on the semi-parametric extension in Moniri (2021).

Semi-parametric symmetric location mixtures. Model (1) is identifiable for $M = 2$ under mild conditions (Bordes et al. 2006; Hunter et al. 2007). It can be fitted by EM-like algorithms that alternate posterior weights with a symmetrised kernel estimate of f (Bordes et al. 2007; Benaglia et al. 2009a), implemented in the `mixtools` package (Benaglia et al. 2009b), or by distance-based and smoothed-likelihood methods (Butucea and Vandekerkhove 2014; Wu et al. 2017; Seo 2017; Levine et al. 2011).

Research gap. In the three mixture studies above, the rank-aware estimators are compared across designs, chiefly with estimators based on an SRS of the same size. That comparison does not separate the gain due to the design, which any estimator applied to the pooled ranked set sample would share, from the gain due to the ranks. None of the three studies compares the rank-aware estimator with a rank-ignoring estimator applied to the same ranked set sample, and none examines whether the constant-weight misplacement model describes the concomitant ranking used to generate the data. Both questions are central for semi-parametric mixtures, where the component density must also be estimated. To our knowledge they have not been addressed for this model; we have not, however, made an exhaustive search of the wider RSS literature, and similar points may have been made for other estimation problems.

3 Model and sampling design

3.1 Location mixture

Let f be a density symmetric about zero with distribution function F, and let X follow (1) with mixing proportions $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)$ and locations $\mu_1 < \cdots < \mu_M$. Write $G(x) = \sum_j \pi_j F(x - \mu_j)$ and $\bar{G} = 1 - G$. We take $M = 2$ throughout.

3.2 Balanced ranked set sampling and ranking error

In cycle $i = 1, \ldots, n$ and for each judgment rank $r = 1, \ldots, H$, a set of H units is drawn independently from (1), the units are ranked by judgment, and the unit given rank r is measured, yielding

Table 1: Constant-weight model against concomitant ranking. Normal components, $\boldsymbol{\pi} = (0.4, 0.6)$, $\boldsymbol{\mu} = (0, 3)$, $\sigma = 0.5$; 4×10^5 sets per setting. "Model" is the mixture (2) with $\boldsymbol{\alpha}$ estimated from the same sets; KS is the Kolmogorov distance between the actual and model distributions.

H	ρ	Judged rank	Actual mean	Model mean	Difference (SE)	KS
3	0.9	1	0.553	0.819	−0.267 (0.002)	0.100
3	0.9	3	2.932	2.686	0.246 (0.001)	0.089
3	0.5	1	1.109	1.308	−0.199 (0.002)	0.071
3	0.5	3	2.456	2.269	0.187 (0.002)	0.067
5	0.9	1	0.096	0.361	−0.265 (0.002)	0.105
5	0.9	5	3.183	2.970	0.213 (0.001)	0.077
5	0.5	1	0.837	1.035	−0.198 (0.002)	0.075
5	0.5	5	2.655	2.479	0.176 (0.002)	0.066

$X_{[r]i}$. Let α_{rh} be the probability that the unit judged to have rank r is in fact the h-th smallest of its set. Because each set yields exactly one unit per judgment rank and one unit per true rank, $\boldsymbol{\alpha} = (\alpha_{rh})$ is doubly stochastic. The constant-weight misplacement model, used for mixtures by Hatefi et al. (2015) and Alvandi et al. (2024), assumes

$$f_{[r]}(x) = \sum_{h=1}^{H} \alpha_{rh}\, g_{(h:H)}(x), \qquad g_{(h:H)}(x) = g(x)\, b_h\{G(x)\}, \tag{2}$$

where $b_h(u) = H\binom{H-1}{h-1} u^{h-1}(1-u)^{H-h}$ is the Beta$(h, H-h+1)$ density. Model (2) is exact under perfect ranking ($\boldsymbol{\alpha} = I$). Otherwise it is an assumption: a misranked unit is taken to be distributed like a typical order statistic of its true rank, whatever its value.

In our simulations ranking follows the concomitant model of Dell and Clutter (1972): units are ranked on $U = X + \tau\varepsilon$ with $\varepsilon \sim N(0,1)$ independent of X, and $\tau = \mathrm{sd}(X)\sqrt{(1-\rho^2)/\rho^2}$, so that $\mathrm{corr}(X, U) = \rho$. Under this mechanism two units swap ranks mostly when their values are close, so a misranked unit is *not* a typical draw from the order statistic of its true rank.

Table 1 checks this directly. For each setting we drew 4×10^5 sets, recorded the misplacement frequencies $\hat{\boldsymbol{\alpha}}$, and compared the actual distribution of the unit judged to have rank r with the mixture (2) evaluated with the same $\hat{\boldsymbol{\alpha}}$. The two disagree by many standard errors at every $\rho < 1$ tested: the unit judged lowest is on average 0.20–0.27 lower than the constant-weight model predicts, and the unit judged highest correspondingly higher. An estimator that uses the ranks through (2) therefore targets the wrong distribution under concomitant ranking.

3.3 Rank-ignoring estimation is valid under any ranking

Proposition 1. *In a balanced ranked set sample, whatever the ranking mechanism,* $H^{-1}\sum_{r=1}^{H} f_{[r]} = g$.

Proof. Consider a single set of H independent draws from g, ranked by any mechanism, and let Y_r denote the unit given judgment rank r. Because every unit of the set receives exactly one judgment rank, $(Y_1, \ldots, Y_H)$ is a permutation of the set, so for any bounded function ψ, $\sum_r \psi(Y_r)$ equals the sum of ψ over the set and $\sum_r \mathbb{E}\,\psi(Y_r) = H\,\mathbb{E}_g\psi(X)$. In a balanced design the measured unit $X_{[r]}$ has the distribution of Y_r, hence $H^{-1}\sum_r \mathbb{E}\,\psi(X_{[r]}) = \mathbb{E}_g\psi(X)$ for all such ψ, which is the claim. □

Under (2) the identity also follows from $\sum_r \alpha_{rh} = 1$ and $\sum_h b_h \equiv H$. Proposition 1 implies that the pooled ranked set sample has the population marginal on average, so estimating equations that are unbiased for i.i.d. data from g remain unbiased when applied to it, under any ranking quality; this is the mechanism behind results such as those of Presnell and Bohn (1999).

4 Estimation and theoretical results

4.1 Likelihood and latent structure

Under (2) the observed-data log-likelihood is

$$\ell(\boldsymbol{\pi}, \boldsymbol{\mu}, f, \boldsymbol{\alpha}) = \sum_{i=1}^{n} \sum_{r=1}^{H} \Big[\log g(x_{[r]i}) + \log \sum_{h=1}^{H} \alpha_{rh}\, b_h\{G(x_{[r]i})\} \Big]. \tag{3}$$

Following Hatefi et al. (2014) for perfect ranking and Hatefi et al. (2015) for imperfect ranking, introduce for each measured unit a latent true rank $\boldsymbol{\Delta}$ (multinomial with probabilities $\boldsymbol{\alpha}_{r\cdot}$), its component label $\mathbf{Z}$, and the component counts $\mathbf{W}$ and $\mathbf{V}$ of the $h-1$ smaller and $H-h$ larger units in its set. Given $\Delta = h$, these are independent multinomials with 1, $h-1$ and $H-h$ trials and probabilities $\boldsymbol{\pi}$, and the complete-data log-likelihood separates into terms in $\boldsymbol{\alpha}$, in $\boldsymbol{\pi}$, and in $(\boldsymbol{\mu}, f)$. Summing the complete-data density over $(\mathbf{Z}, \mathbf{W}, \mathbf{V})$ recovers (2); we checked this identity numerically by exact enumeration at 64 points (largest relative difference 4×10^{-15}).

4.2 E-step

With $\phi_{rh}(x) \propto \alpha_{rh} b_h\{G(x)\}$ the posterior probability of true rank h, and

$$\tau_j(x) = \frac{\pi_j f(x-\mu_j)}{g(x)}, \quad \beta_j(x) = \frac{\pi_j F(x-\mu_j)}{G(x)}, \quad \gamma_j(x) = \frac{\pi_j \bar{F}(x-\mu_j)}{\bar{G}(x)},$$

the conditional expectations of Z_j, W_j and V_j given x and $\Delta = h$ are τ_j, $(h-1)\beta_j$ and $(H-h)\gamma_j$.

4.3 M-step for the mixing proportions and the misplacement matrix

Maximising the expected complete-data log-likelihood in $\boldsymbol{\pi}$ gives

$$\pi_j^{\text{new}} = \frac{1}{nH^2} \sum_{i,r} \Big[\tau_j + (\bar{h}-1)\beta_j + (H-\bar{h})\gamma_j \Big](x_{[r]i}), \qquad \bar{h} = \sum_h h\, \phi_{rh}, \tag{4}$$

the expected share of component j among all nH^2 units in the sampled sets. Each measured unit accounts for H set members, because $\sum_j \{\tau_j + (h-1)\beta_j + (H-h)\gamma_j\} = H$ for every h, so the normalising constant is nH^2; the thesis (Moniri 2021) prints nH^3. The update of $\boldsymbol{\alpha}$ maximises $\sum_{r,h} \Phi_{rh} \log \alpha_{rh}$, with $\Phi_{rh} = \sum_i \phi_{rh}(x_{[r]i})$, over doubly stochastic (here also symmetric) matrices; the stationarity conditions give $\alpha_{rh} = \Phi_{rh}/(\lambda_r + \nu_h)$ with multipliers found by Newton's method.

4.4 Locations: the modified step is an estimating equation

In the expected complete-data log-likelihood, the score for μ_j has a density term, $\sum \tau_j\, \partial_{\mu_j} \log f$, and two distribution-function terms weighted by β_j and γ_j. Following the modified EM algorithm of Hatefi et al. (2015, Section 5), the algorithm of Moniri (2021) replaces the two distribution-function terms by their expectations. The following lemma specialises Lemma 3 of Hatefi et al.

(2015), stated for partially rank-ordered sets and summed over the components, to balanced RSS (subsets of size one) and to each component separately; we include a short proof in our notation.

Lemma 1. *Let $(X_{[1]}, \dots, X_{[H]})$ be one cycle of a balanced ranked set sample satisfying* (2) *with doubly stochastic $\boldsymbol{\alpha}$, and let Δ_{rh}, W_{rj}, V_{rj} be the latent true-rank indicator and counts attached to $X_{[r]}$. For any measurable S with the expectations below finite,*

$$\sum_{r=1}^{H}\sum_{h=1}^{H} \mathbb{E}\big[\Delta_{rh} W_{rj}\, S(X_{[r]})\big] = H(H-1)\,\pi_j\, \mathbb{E}\big[S(X)F(X-\mu_j)\big],$$
$$\sum_{r=1}^{H}\sum_{h=1}^{H} \mathbb{E}\big[\Delta_{rh} V_{rj}\, S(X_{[r]})\big] = H(H-1)\,\pi_j\, \mathbb{E}\big[S(X)\bar{F}(X-\mu_j)\big],$$

where the expectations on the right are under (1).

Proof. Conditioning on $X_{[r]} = x$ and $\Delta_{rh} = 1$, $\mathbb{E}[W_{rj} \mid \cdot] = (h-1)\pi_j F(x-\mu_j)/G(x)$, and the joint density of $(X_{[r]}, \Delta_{rh} = 1)$ is $\alpha_{rh}\, g(x)\, b_h\{G(x)\}$. Hence the left side of the first identity equals

$$\pi_j \int S(x)F(x-\mu_j)\frac{g(x)}{G(x)} \sum_h (h-1)\, b_h\{G(x)\} \sum_r \alpha_{rh}\, dx.$$

Column sums of $\boldsymbol{\alpha}$ are one, and $\sum_{h=1}^{H}(h-1)b_h(u) = H\,\mathbb{E}[B]$ with $B \sim \mathrm{Bin}(H-1, u)$, that is, $H(H-1)u$. With $u = G(x)$ the factor $G(x)$ cancels, giving $H(H-1)\pi_j \int S(x)F(x-\mu_j)g(x)\,dx$. The second identity follows in the same way from $\sum_h (H-h)b_h(u) = H(H-1)(1-u)$. □

In the location score the distribution-function terms are

$$\sum \Delta_{rh} W_{rj}\, \partial_{\mu_j} \log F(x-\mu_j) \quad \text{and} \quad \sum \Delta_{rh} V_{rj}\, \partial_{\mu_j} \log \bar{F}(x-\mu_j).$$

Taking $S = -f/F$ in the first identity and $S = f/\bar{F}$ in the second, their expectations are $-H(H-1)\pi_j\mathbb{E}[f(X-\mu_j)]$ and $+H(H-1)\pi_j\mathbb{E}[f(X-\mu_j)]$. They cancel. Removing both terms therefore leaves an estimating equation whose expectation at the true parameter is that of the density term alone, which is zero; and since τ_j does not depend on the true rank, nothing rank-dependent remains. Solving the remaining equation in the moment form used by Moniri (2021) gives

$$\mu_j^{\text{new}} = \frac{\sum_{i,r} \tau_j(x_{[r]i})\, x_{[r]i}}{\sum_{i,r} \tau_j(x_{[r]i})}, \tag{5}$$

which is exactly the SRS update, as Hatefi et al. (2015) observed for their modified algorithm. The modified algorithm is therefore not a maximisation of the expected complete-data log-likelihood, and the ranks enter only through (4). We confirmed both identities and the cancellation by numerical integration ($H = 4$, symmetric doubly stochastic $\boldsymbol{\alpha}$; agreement to 8 digits, and to 4×10^{-16} for the cancellation).

4.5 Component density

Given posterior weights and locations, f is estimated by the symmetrised weighted Gaussian kernel estimate of the centred data,

$$\hat{f}(u) = \frac{1}{2N}\sum_{k=1}^{N}\sum_{j=1}^{M} \tau_j(x_k)\big\{K_h(u - x_k + \mu_j) + K_h(u + x_k - \mu_j)\big\},$$

and F by its exact integral, which is a proper distribution function. The bandwidth is the rule of thumb of Silverman (1986), $h = 0.9\,\min(s, \mathrm{IQR}/1.34)N^{-1/5}$, applied to the weighted, symmetrised centred residuals (the *centred* rule). A rule computed from the raw pooled data (the *pooled* rule) includes the separation between the components in its scale estimate and oversmooths f; Section 6.1 shows the consequence.

4.6 How much can the ranks add? A parametric benchmark

To separate the contribution of the design from that of the ranks, consider a case favourable to the ranks: f and $\boldsymbol{\alpha}$ known and $\boldsymbol{\theta} = (\boldsymbol{\pi}, \boldsymbol{\mu})$ unknown. Three estimators of $\boldsymbol{\theta}$ are compared, per measured unit, as $n \to \infty$ with H fixed, under standard regularity conditions for parametric likelihoods. The full RSS likelihood (3) gives asymptotic variance $\mathcal{I}_R^{-1}$ with

$$\mathcal{I}_R = \frac{1}{H}\sum_{r=1}^{H} \mathbb{E}_r\big[s_r(X)s_r(X)^\top\big], \qquad s_r(x) = \nabla_{\boldsymbol{\theta}}\Big[\log g(x) + \log \sum_h \alpha_{rh} b_h\{G(x)\}\Big],$$

where $\mathbb{E}_r$ is expectation under $f_{[r]}$. The rank-ignoring likelihood $\sum \log g(x_{[r]i})$ treats the data as i.i.d.; by Proposition 1 its score has mean zero, and its sandwich variance is $A^{-1}BA^{-1}$, where $A = \mathbb{E}_g[ss^\top]$, $s = \nabla_{\boldsymbol{\theta}} \log g$, and

$$B = \frac{1}{H}\sum_r \mathrm{Var}_r\{s(X)\} = A - \frac{1}{H}\sum_r m_r m_r^\top, \qquad m_r = \mathbb{E}_r[s(X)].$$

The reduction $A - B$ is the design effect: stratification by judgment rank removes the between-stratum variation of the score. The same likelihood on an SRS of equal size has variance A^{-1}.

Table 2 gives the ratios of diagonal elements, computed by one-dimensional numerical integration for normal components with $\pi_1 = 0.4$. Under perfect ranking the ranks reduce the asymptotic variance by a further 4–18% beyond the design, while the design alone reduces the variance of $\hat{\pi}_1$ by 18–49%. For $\rho < 1$ the calculation assumes the constant-weight model (2); it describes data generated from that model, not concomitant ranking, and we report two such rows only for orientation. These ratios are benchmarks for a parametric model with f known, not bounds for semi-parametric estimators, which must also estimate f.

Table 2: Asymptotic variance ratios with f and $\boldsymbol{\alpha}$ known. Normal components, $\sigma = 0.5$, $\pi_1 = 0.4$, $\mu_1 = 0$. "Rank / ignore": full RSS likelihood against the rank-ignoring likelihood on the same RSS. "RSS / SRS": rank-ignoring likelihood on RSS against SRS (the design effect). Values below one favour the first estimator. The last two rows assume the constant-weight model (CW).

			Rank / ignore			RSS / SRS	
H	Separation	Ranking	π_1	μ_1	μ_2	π_1	μ_2
3	6σ	perfect	0.958	0.889	0.889	0.639	0.847
5	6σ	perfect	0.929	0.862	0.870	0.506	0.718
3	3σ	perfect	0.909	0.869	0.870	0.772	0.883
5	3σ	perfect	0.862	0.826	0.842	0.662	0.757
5	2σ	perfect	0.829	0.817	0.827	0.819	0.825
5	6σ	CW, $\alpha_{11} = 0.69$	0.944	0.936	0.929	0.704	0.912
5	6σ	CW, $\alpha_{11} = 0.42$	0.976	0.984	0.981	0.918	0.980

5 Simulation design

5.1 Data-generating models

The data are two-component location mixtures with $\mu_1 = 0$, $\mu_2 = \delta\sigma$ and component standard deviation $\sigma = 0.5$, where f is normal, Student t_5 or Laplace, each scaled to standard deviation σ, so that the separation δ is measured in standard deviations for every law. The factors are: set size and cycles $(H, n) \in \{(3, 30), (5, 18), (5, 30)\}$, giving $N = 90, 90, 150$, and $(5, 60)$ ($N = 300$, normal components only); separation $\delta \in \{6, 4, 3\}$; $\pi_1 \in \{0.4, 0.2\}$; and seven ranking scenarios: perfect ranking; concomitant ranking with $\rho \in \{0.95, 0.9, 0.7, 0.5\}$ (labelled C ρ); and data generated from the constant-weight model itself (CW ρ), with $\boldsymbol{\alpha}$ equal to the concomitant misplacement matrix at $\rho = 0.9$ and 0.7, so that the rank-aware methods are also assessed where their ranking model holds. For CW data, the true rank of each measured unit is drawn from the corresponding row of $\boldsymbol{\alpha}$ and the unit is the order statistic of that rank in a fresh set. The true $\boldsymbol{\alpha}$ for each concomitant setting is the symmetrised matrix of misplacement frequencies in 2×10^5 simulated sets.

The grid has 420 cells, each with 1,000 replicates. A further 12 sensitivity cells ($H = 5$, $n = 18$, $\pi_1 = 0.4$, normal, $\delta \in \{6, 4, 3\}$, perfect ranking and C 0.9) use the pooled bandwidth rule or a stochastic kernel step with averaging after burn-in (Bordes et al. 2007). Replicate i of cell k uses the seed $20261120 + 10^4 k + i$, so the results do not depend on the number of parallel workers or on the order in which cells are run.

5.2 Estimators

All methods use the same deterministic starting values: one-dimensional k-means started from the lower and upper quartiles, with the first $\hat{f}$ built from posterior weights under a normal working model. They iterate until the largest change in $(\boldsymbol{\pi}, \boldsymbol{\mu})$ is below 10^{-6} (and in $\boldsymbol{\alpha}$ below 10^{-4}), or until an iteration cap of 1,000 (200 for the likelihood-step methods RI-L, P-RI and P-RA). In every replicate all methods are fitted to the same ranked set sample.

RI-WM Rank-ignoring fit with the weighted-mean location update: E-step with τ_j, $\pi_j^{\text{new}} = N^{-1} \sum \tau_j$, update (5), kernel step. It is of the same type as the deterministic EM-like algorithms of Benaglia et al. (2009a), and with matched settings it reproduces the `mixtools`

implementation (Section 6.8).

RI-L Rank-ignoring fit with a likelihood-type location update: at each iteration $(\boldsymbol{\pi}, \boldsymbol{\mu})$ maximise $\sum \log \hat{g}(x_k)$ with the current $\hat{f}$ held fixed (BFGS with analytic gradient), followed by the kernel step; a plug-in version of the smoothed-likelihood idea (Levine et al. 2011; Seo 2017). Fitted on a sub-grid of 36 cells ($H = 5$; $N = 90, 150$; $\delta = 6, 4$; $\pi_1 = 0.4$; perfect ranking, C 0.9 and C 0.7; all three laws).

CDA($\hat{\boldsymbol{\alpha}}$) The rank-aware algorithm of Section 4, with update (4), update (5), one $\boldsymbol{\alpha}$ step per iteration (symmetric doubly stochastic, started at diagonal 0.6), and the kernel step.

CDA(I) The same algorithm with $\boldsymbol{\alpha}$ fixed at the identity, i.e. assuming perfect ranking.

SRS RI-WM fitted to an independent SRS of the same size.

P-RI, P-RA Parametric benchmarks with f known: the i.i.d. likelihood ignoring the ranks (P-RI), and the full RSS likelihood (3) with the true $\boldsymbol{\alpha}$ (P-RA).

We call the rank-aware algorithm CDA ("corrected derived algorithm") because it departs from the version in Moniri (2021) in four documented ways: the constant nH^2 in (4), the exact integral for $\hat{F}$, the centred bandwidth rule, and a deterministic kernel step run to convergence instead of a stochastic one. The pooled rule and the stochastic version are examined in the sensitivity cells.

5.3 Failures and performance measures

Each fit is classified as successful or failed. Failures have four mutually exclusive causes: an error raised by the fitting routine (no estimates); a numerical stop (non-finite likelihood or update); the iteration cap; or convergence to a degenerate solution, either merged locations (closer than 0.05 sample standard deviations) or a mixing proportion below $1/N$. The first two are numerical failures, the third a convergence failure, and the last two statistical degeneracy (collapse). Failed fits are never dropped: every replicate enters the bias and root mean squared error (RMSE) summaries, and failures are reported by cause.

For each method and parameter we report the ratio of its mean squared error to that of RI-WM on the same data sets. Our primary measure uses the replicates in which neither fit failed ("both-ok"); it compares accuracy conditional on success. At 4σ and 6σ failures are rare and the two versions agree closely, except in a few cells where a handful of collapsed fits dominate the all-replicate ratio; at 3σ the all-replicate version mixes accuracy with collapse. SRS comparisons use all replicates, since the samples are independent; their per-cell standard errors are larger (median 0.10, 10th to 90th percentile 0.07–0.15), so single cells should not be over-read. Monte Carlo standard errors of ratios use the delta method on the paired squared errors, and those of RMSEs the delta method on the squared errors. A method is called better or worse in a cell when its ratio differs from one by more than two standard errors; with 126 cells per law, about three cells on each side would pass this threshold by chance alone. Laws are compared on the 126 cells per law with $N \leq 150$. Per-cell bias, RMSE with standard errors, and failure counts by cause for every method are in Online Resource 1.

5.4 Software and use of generative AI

All computations were carried out in R (version 4.5.2) with the package `semirss`, written for this study in base R. Every reported number is an output of that code, run with the seeds given

above, and can be reproduced from the code and saved results, which will be deposited in a public repository (see Statements and Declarations).

The author used AI (Claude, Anthropic) to help with the writing. The author checked all of it and takes full responsibility for the manuscript.

6 Simulation results

6.1 Failures

No fit raised an error and none stopped on a numerical problem anywhere in the study (0 errors and 0 numerical stops among all fits). Table 3 and Figure 1 give failure rates by cause at 3σ. At 4σ and 6σ failures are rare: in any single cell they affect at most 1.8% of replicates for RI-WM and the SRS fit, 3.0% for CDA(I) and 5.7% for CDA($\hat{\boldsymbol{\alpha}}$), the largest rates occurring for Laplace components with $\pi_1 = 0.2$ at 4σ, mostly as small-proportion solutions. The number of iterations also differs between methods: the median over cells of the per-cell median number of iterations is 11, 18 and 35 for RI-WM at 6σ, 4σ and 3σ, similar for CDA(I) (14, 20, 38), and 167, 157 and 168 for CDA($\hat{\boldsymbol{\alpha}}$). Failures concentrate at 3σ, where the mixture density is barely bimodal, and are mostly merged locations; between 2% and 13% of failures, depending on method and law, stop at the iteration cap. For normal components the failure rate of RI-WM at 3σ falls from 11.7% at $N = 90$ to 3.8% and 0.2% at $N = 150$ and 300, and CDA($\hat{\boldsymbol{\alpha}}$) fails more often (16.8%, 7.2%, 0.6%), with more iteration-cap stops and more small-proportion solutions. The SRS fit fails about as often as RI-WM, so collapse is a property of the model at this separation and sample size rather than of the design. At 3σ, heavy-tailed components collapse less often than normal ones at the same standard-deviation separation.

In the sensitivity cells, the pooled bandwidth rule raised the failure rate of RI-WM at 3σ from 10.8–14.0% (matching main cells) to 75.5–76.5%, and at 4σ from 0.0–0.1% to 4.5–5.5%. The stochastic kernel step with averaging gave 17.1–19.2% at 3σ and 0.0–0.2% at 4σ. The sensitivity cells use their own seeds, so these are comparisons between independent samples.

6.2 Absolute performance

Table 4 shows bias and RMSE in two representative cells ($H = 5$, $N = 90$, $\pi_1 = 0.4$, 4σ), for normal and Laplace components under perfect ranking and C 0.9. Biases are small relative to RMSE for all semi-parametric methods; the differences between methods are differences in variance, and they are of the order of a few Monte Carlo standard errors per cell, which is why the paired ratios below are needed to resolve them.

6.3 The rank-aware algorithm against ignoring the ranks

Table 5 counts, for each law and parameter, the cells in which the rank-aware algorithm is better than, not distinguishable from, or worse than RI-WM; Table 6 and Figures 2 and 3 give the ratios. With $\boldsymbol{\alpha}$ estimated, CDA is worse than RI-WM for μ_2 in 62 of 126 cells for normal components, 112 for t_5 and 117 for Laplace, and better in at most 2. The picture for π_1 is similar. With heavy-tailed components the loss is systematic: median μ_2 ratios are 1.02–1.27 for t_5 and 1.03–1.30 for Laplace, largest under perfect ranking. Only for the smaller-mean component μ_1 with heavy tails is CDA better more often than worse, with median ratios down to 0.95 at 6σ.

The loss persists where the ranking model is correct: on the CW cells CDA is never better than RI-WM for μ_2. With normal components and $N = 300$, the ratios approach one at 4σ and 6σ.

Table 3: Failure rates at separation 3σ: total % of replicates, with the split (iteration cap / merged locations / proportion below $1/N$) in parentheses; mean over the cells of each row. No fit raised an error or stopped on a numerical problem.

Law	N	RI-WM (RSS)	RI-WM (SRS)	CDA($\hat{\boldsymbol{\alpha}}$)	CDA(I)
Normal	90	11.7 (0.4/10.9/0.4)	12.0 (0.3/11.2/0.5)	16.8 (1.0/14.0/1.8)	12.9 (0.2/12.1/0.6)
Normal	150	3.8 (0.3/3.4/0.2)	3.6 (0.2/3.2/0.1)	7.2 (1.0/4.9/1.2)	4.5 (0.2/4.0/0.3)
Normal	300	0.2 (0.1/0.1/0.0)	0.3 (0.1/0.2/0.0)	0.6 (0.3/0.2/0.1)	0.3 (0.1/0.3/0.0)
t_5	90	5.3 (0.3/4.2/0.7)	5.3 (0.3/4.3/0.7)	8.5 (0.7/5.7/2.1)	4.8 (0.3/3.8/0.7)
t_5	150	1.2 (0.1/0.9/0.1)	1.4 (0.1/1.1/0.1)	3.0 (0.5/1.5/1.1)	1.1 (0.1/0.7/0.3)
Laplace	90	4.9 (0.3/3.6/0.9)	5.1 (0.3/3.8/1.0)	8.1 (0.9/4.5/2.7)	4.0 (0.3/2.8/0.9)
Laplace	150	1.5 (0.1/1.1/0.3)	1.7 (0.1/1.2/0.4)	3.6 (0.8/1.4/1.4)	1.3 (0.2/0.6/0.5)

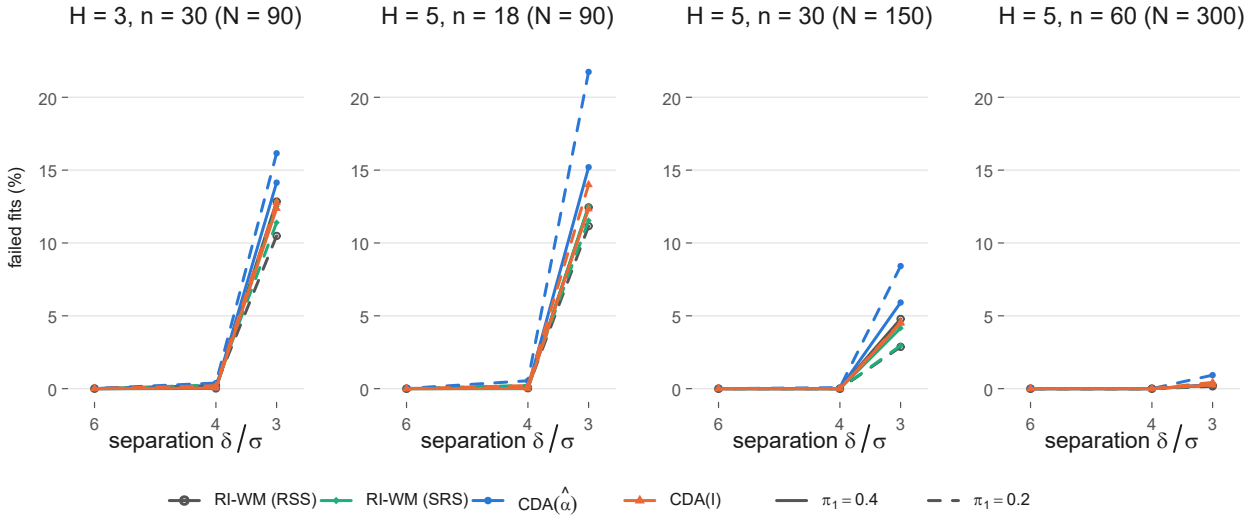


Figure 1: Failure rates (%) for normal components by separation, for each combination of set size and number of cycles; mean over the seven ranking scenarios, solid lines $\pi_1 = 0.4$, dashed lines $\pi_1 = 0.2$.

There, CDA is better for μ_2 in 8 of 42 cells and worse in 16: 4 of the better cells are at 4σ, with gains of 0.5–1.4%, and 4 are at 3σ with $\rho \leq 0.7$, with gains of 2.5–7.3%.

6.4 The design gain

Table 7 compares RI-WM on an SRS with RI-WM on the ranked set sample. At 4σ and 6σ and under perfect ranking, the MSE of $\hat{\pi}_1$ is 1.56–1.72 times larger with SRS; at $\rho = 0.5$ the factor falls to 1.10–1.16. The gain for μ_2 is smaller and has largely disappeared at $\rho = 0.5$. The pattern is the same for the three laws. For μ_2, CW data give gains similar to concomitant ranking at the matched ρ; for π_1 the gain on CW data is clearly smaller than under concomitant ranking at the same ρ, consistent with concomitant ranking errors mostly swapping units with close values. This gain needs no rank-aware estimator.

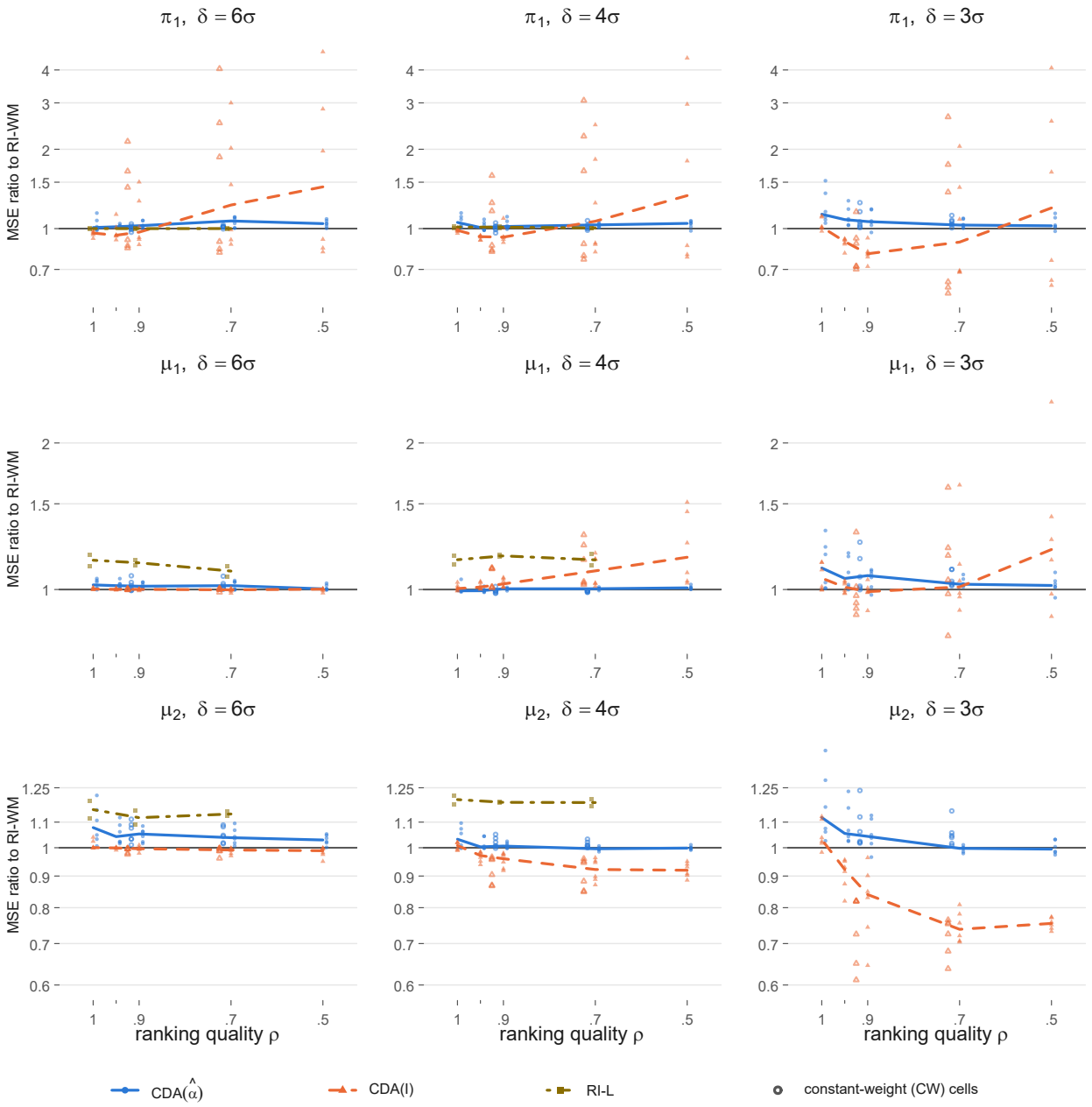


Figure 2: Normal components: MSE ratio against RI-WM by ranking quality, separation and parameter (both-ok replicates, $N \in \{90, 150\}$, log scale). Small symbols are individual cells (three size combinations by two values of π_1); lines join the medians over the concomitant and perfect-ranking cells; hollow symbols are constant-weight (CW) cells, drawn just right of their matched ρ. RI-L was fitted on the sub-grid only.

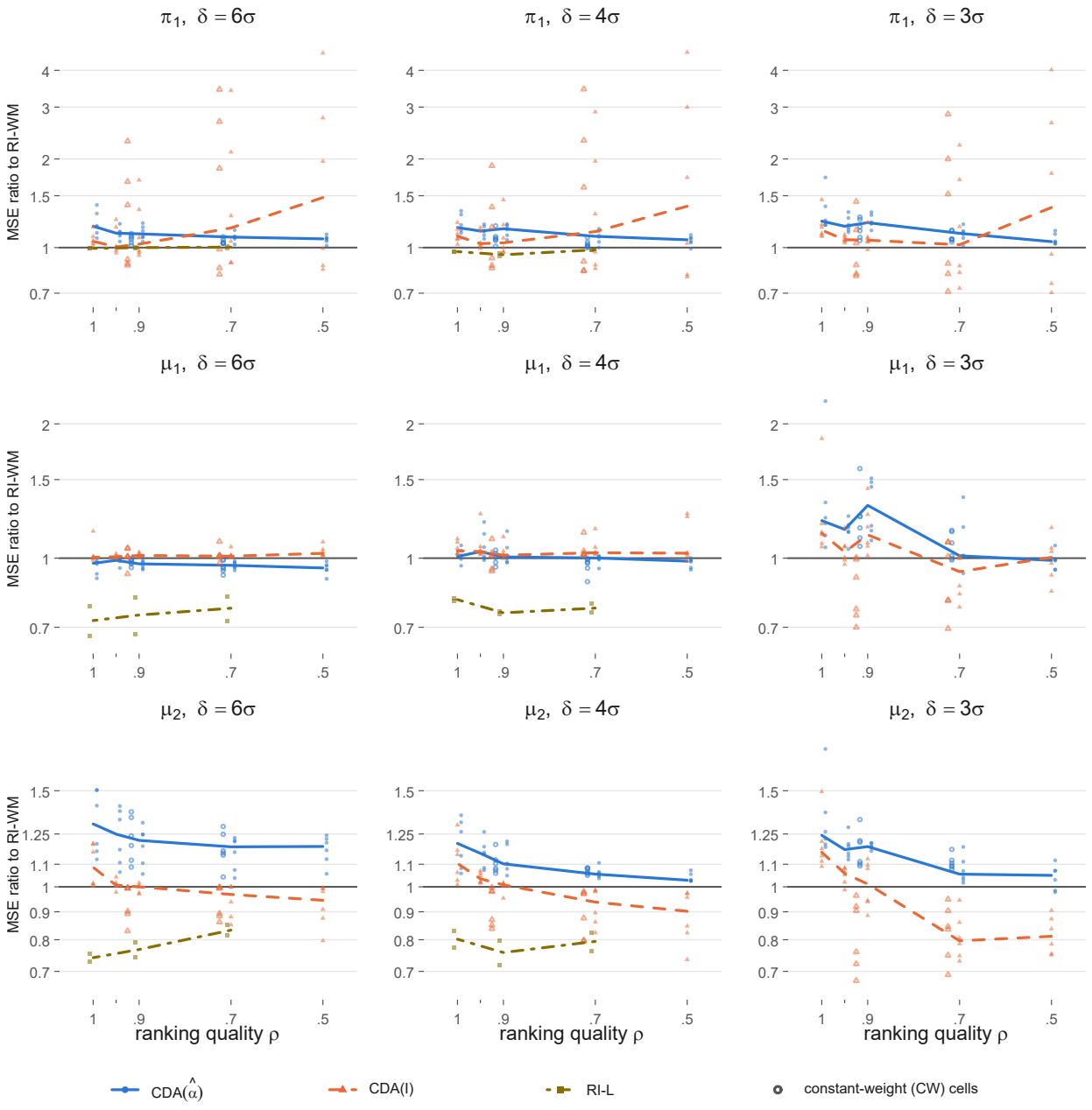


Figure 3: As Figure 2, for Laplace components. The corresponding figure for t_5 components is in Online Resource 1.

Table 4: Bias and RMSE (Monte Carlo SE) in representative cells: $H = 5$, $n = 18$ ($N = 90$), $\pi_1 = 0.4$, separation 4σ; all replicates; units of the parameter ($\sigma = 0.5$). Failed: % of replicates.

				π_1		μ_1		μ_2	
Law	Ranking	Method	Failed %	bias	RMSE (SE)	bias	RMSE (SE)	bias	RMSE (SE)
Normal	perfect	RI-WM	0.0	+0.0017	0.0420 (0.0010)	+0.0017	0.0905 (0.0022)	-0.0062	0.0692 (0.0016)
		RI-L	0.0	+0.0014	0.0421 (0.0010)	+0.0057	0.0961 (0.0022)	-0.0121	0.0751 (0.0019)
		SRS	0.0	-0.0006	0.0542 (0.0012)	+0.0044	0.0956 (0.0021)	-0.0065	0.0789 (0.0018)
		CDA($\hat{\boldsymbol{\alpha}}$)	0.0	-0.0014	0.0422 (0.0010)	+0.0008	0.0902 (0.0022)	-0.0073	0.0695 (0.0016)
		CDA(I)	0.0	+0.0007	0.0411 (0.0010)	+0.0014	0.0902 (0.0022)	-0.0066	0.0691 (0.0016)
		P-RI	0.0	+0.0035	0.0409 (0.0009)	-0.0015	0.0881 (0.0020)	+0.0020	0.0664 (0.0016)
		P-RA	0.0	+0.0034	0.0383 (0.0009)	+0.0015	0.0798 (0.0019)	+0.0007	0.0617 (0.0014)
Normal	C 0.9	RI-WM	0.1	-0.0016	0.0446 (0.0011)	+0.0086	0.1028 (0.0073)	-0.0090	0.0786 (0.0042)
		RI-L	0.2	-0.0017	0.0456 (0.0016)	+0.0137	0.1167 (0.0093)	-0.0135	0.0880 (0.0052)
		SRS	0.0	-0.0040	0.0539 (0.0012)	+0.0053	0.1011 (0.0024)	-0.0103	0.0751 (0.0017)
		CDA($\hat{\boldsymbol{\alpha}}$)	0.1	-0.0005	0.0450 (0.0011)	+0.0089	0.1030 (0.0073)	-0.0089	0.0785 (0.0042)
		CDA(I)	0.0	+0.0054	0.0415 (0.0010)	+0.0101	0.0964 (0.0022)	-0.0067	0.0740 (0.0017)
		P-RI	0.0	+0.0008	0.0428 (0.0010)	+0.0037	0.0924 (0.0022)	+0.0013	0.0700 (0.0015)
		P-RA	0.0	+0.0012	0.0428 (0.0010)	+0.0317	0.0959 (0.0022)	-0.0187	0.0720 (0.0015)
Laplace	perfect	RI-WM	0.1	+0.0008	0.0400 (0.0013)	+0.0018	0.0903 (0.0080)	-0.0018	0.0644 (0.0052)
		RI-L	0.0	+0.0012	0.0384 (0.0009)	+0.0079	0.0741 (0.0020)	-0.0051	0.0538 (0.0013)
		SRS	0.0	+0.0025	0.0564 (0.0012)	+0.0047	0.0859 (0.0020)	-0.0034	0.0687 (0.0016)
		CDA($\hat{\boldsymbol{\alpha}}$)	0.2	-0.0017	0.0438 (0.0021)	+0.0021	0.0901 (0.0080)	-0.0046	0.0680 (0.0050)
		CDA(I)	0.2	-0.0001	0.0410 (0.0011)	+0.0017	0.0911 (0.0079)	-0.0027	0.0660 (0.0051)
		P-RI	0.0	+0.0017	0.0376 (0.0008)	-0.0011	0.0654 (0.0016)	-0.0005	0.0485 (0.0012)
		P-RA	0.0	+0.0010	0.0367 (0.0008)	+0.0004	0.0593 (0.0014)	-0.0016	0.0440 (0.0010)
Laplace	C 0.9	RI-WM	0.0	-0.0020	0.0431 (0.0010)	-0.0052	0.0843 (0.0020)	-0.0033	0.0674 (0.0017)
		RI-L	0.0	-0.0018	0.0417 (0.0010)	+0.0026	0.0735 (0.0018)	-0.0070	0.0602 (0.0014)
		SRS	0.0	+0.0006	0.0547 (0.0012)	-0.0088	0.0894 (0.0021)	-0.0063	0.0688 (0.0016)
		CDA($\hat{\boldsymbol{\alpha}}$)	0.1	-0.0017	0.0463 (0.0011)	-0.0027	0.0835 (0.0020)	-0.0062	0.0707 (0.0018)
		CDA(I)	0.0	+0.0053	0.0431 (0.0010)	-0.0031	0.0850 (0.0020)	-0.0017	0.0679 (0.0017)
		P-RI	0.0	-0.0011	0.0413 (0.0009)	-0.0032	0.0698 (0.0017)	-0.0005	0.0551 (0.0014)
		P-RA	0.0	-0.0008	0.0413 (0.0010)	+0.0179	0.0705 (0.0017)	-0.0153	0.0569 (0.0014)

6.5 What the ranks add when f is known

Table 8 compares the two parametric benchmarks. Where the constant-weight model holds, the median MSE ratio for $\hat{\mu}_2$ is 0.82–0.87 under perfect ranking and 0.91–0.97 on CW data, for all three laws; for normal components this agrees with the asymptotic benchmark of Table 2. Under concomitant ranking the same likelihood is misspecified, and its ratios for μ_2 exceed one at $\rho = 0.9$.

The harm grows with the sample size. Table 9 and Figure 4 follow the normal cells with $H = 5$ at 6σ over $N = 90, 150, 300$. Under concomitant ranking the bias of $\hat{\mu}_2$ from P-RA does not shrink (at C 0.9 it is between -0.036σ and -0.043σ) while its standard deviation falls, so the MSE ratio rises: at C 0.9 it is 1.11, 1.14 and 1.40. On CW data and under perfect ranking the bias is negligible and the ratio stable. A bias that does not shrink over $N = 90$–300 is what an inconsistent estimator would show, but three sample sizes cannot establish inconsistency.

6.6 Assuming perfect ranking

CDA(I), which assumes perfect ranking, behaves differently from CDA($\hat{\boldsymbol{\alpha}}$) (Tables 5, 10 and 11). Under imperfect ranking it often improves $\hat{\mu}_2$: for normal components the median gain at 4σ is 3–9%, and it is better than RI-WM in 93 of the 108 normal cells with imperfect ranking. It

Table 5: Cells in which the rank-aware algorithm is better / not distinguishable / worse than RI-WM (paired MSE ratio, both-ok replicates, more than two Monte Carlo SE from one). $N \in \{90, 150\}$: 126 cells per law; 36 CW cells and 18 perfect-ranking cells per law in the restricted rows.

Method	Law	μ_1	μ_2	π_1
CDA($\hat{\boldsymbol{\alpha}}$)	Normal	8/76/42	2/62/62	2/67/57
	t_5	27/87/12	0/14/112	0/28/98
	Laplace	23/88/15	0/9/117	0/14/112
CDA(I)	Normal	5/73/48	95/26/5	70/17/39
	t_5	3/79/44	62/47/17	44/32/50
	Laplace	8/63/55	58/43/25	33/29/64
CDA($\hat{\boldsymbol{\alpha}}$), CW data only	Normal	7/19/10	0/17/19	2/24/10
	t_5	11/23/2	0/3/33	0/18/18
	Laplace	9/24/3	0/2/34	0/7/29
CDA(I), perfect ranking only	Normal	2/13/3	2/11/5	6/10/2
	t_5	0/13/5	0/3/15	0/10/8
	Laplace	0/8/10	0/1/17	0/2/16

does so while biasing the other two parameters. At 4σ and 6σ its median bias under imperfect ranking reaches 0.035 for $\hat{\pi}_1$ and 0.016σ for $\hat{\mu}_1$, against at most 0.004 and 0.006σ under perfect ranking, while its bias for $\hat{\mu}_2$ is at most 0.011σ under any ranking, similar to that of RI-WM (at most 0.010σ; Table 11). The $\hat{\pi}_1$ bias does not shrink over $N = 90$–300 (up to 0.042 for normal components, $H = 5$, 6σ; Table 9), so the MSE ratio for $\hat{\pi}_1$ grows with N; at C 0.7 it is 1.46, 1.99 and 3.57. Where its assumption is true, under perfect ranking, CDA(I) improves $\hat{\pi}_1$ slightly for normal components and is worse for μ_2 with heavy tails. A small-sample gain in one parameter is therefore no evidence that the ranks carry usable information; this parallels the finding of Taconeli and Rodrigues de Lara (2024) for single distributions that perfect-ranking likelihoods become biased under ranking error. We have not examined the mechanism behind the μ_2 gain.

6.7 The location update

Table 12 compares the two rank-ignoring fits on the sub-grid. The likelihood-type update RI-L is better for Laplace components in every cell (cell ratios 0.72–0.85 for μ_2), about neutral for t_5 (0.92–1.06) and worse for normal components (1.09–1.22). It barely changes $\hat{\pi}_1$. The size of the heavy-tail gain matches what is at stake: with f known, ignoring the ranks, the MSE of $\hat{\mu}_2$ is 0.57 times that of RI-WM for Laplace components at 4σ and 6σ, against 0.94 for normal components, so the weighted mean wastes much more information under heavy tails.

6.8 Comparison with an independent implementation

We compared RI-WM with `spEMsymloc` in the `mixtools` package (version 2.0.0.1; Benaglia et al. 2009b), on the data of the main study. With matched settings (the same starting centres, the same fixed bandwidth, tolerance 10^{-8} and a cap of 1,000 iterations; 8 cells with $H = 5$, $N = 90$, 200 replicates each), both implementations converged in all 1600 replicates, and wherever the fitted locations did not merge, their estimates agreed to within 10^{-8}, the convergence tolerance. With its default settings (a bandwidth computed once from the pooled data, and a cap of 100 iterations) on the 36 sub-grid cells, `spEMsymloc` matched RI-WM at 6σ (MSE ratios for $\hat{\mu}_2$ of 1.00–1.13 on both-ok replicates) but was clearly worse at 4σ: up to 9.6% of its fits failed in a cell, against

Table 6: Median MSE ratio of CDA($\hat{\boldsymbol{\alpha}}$) to RI-WM (both-ok replicates; $N \in \{90, 150\}$; median over the three size combinations and two values of π_1). C: concomitant ranking at the given ρ; CW: data generated from the constant-weight model.

Law	Sep.	perfect	C 0.95	C 0.9	C 0.7	C 0.5	CW 0.9	CW 0.7
(a) μ_2								
Normal	6	1.08	1.04	1.05	1.04	1.03	1.05	1.05
	4	1.03	1.00	1.01	1.00	1.00	1.01	1.01
	3	1.12	1.05	1.04	1.00	0.99	1.04	1.04
t_5	6	1.27	1.20	1.14	1.16	1.13	1.18	1.16
	4	1.12	1.11	1.07	1.06	1.06	1.08	1.06
	3	1.22	1.15	1.10	1.03	1.02	1.16	1.10
Laplace	6	1.30	1.25	1.22	1.18	1.19	1.22	1.16
	4	1.20	1.15	1.10	1.06	1.03	1.09	1.07
	3	1.24	1.17	1.19	1.06	1.05	1.16	1.10
(b) μ_1								
Normal	6	1.02	1.02	1.02	1.02	1.00	1.01	1.01
	4	0.99	0.99	1.00	1.00	1.01	0.99	0.99
	3	1.11	1.05	1.07	1.03	1.02	1.04	1.03
t_5	6	0.98	0.97	0.98	0.98	0.99	0.98	0.98
	4	0.98	0.99	0.99	0.96	0.99	0.97	0.98
	3	1.16	1.26	1.11	1.03	0.99	1.20	1.09
Laplace	6	0.97	0.99	0.97	0.96	0.95	0.98	0.95
	4	1.01	1.03	1.01	1.00	0.98	1.01	0.98
	3	1.21	1.16	1.31	1.01	0.99	1.20	1.05
(c) π_1								
Normal	6	1.01	1.02	1.03	1.07	1.04	1.00	1.02
	4	1.06	1.01	1.02	1.03	1.05	1.01	1.01
	3	1.13	1.08	1.06	1.03	1.02	1.05	1.05
t_5	6	1.12	1.09	1.08	1.08	1.06	1.04	1.01
	4	1.11	1.09	1.10	1.07	1.05	1.04	1.02
	3	1.22	1.15	1.10	1.09	1.06	1.14	1.08
Laplace	6	1.18	1.12	1.11	1.09	1.07	1.07	1.04
	4	1.17	1.14	1.16	1.09	1.06	1.07	1.02
	3	1.23	1.18	1.22	1.12	1.05	1.16	1.07

at most 0.1% for RI-WM, and its MSE for $\hat{\mu}_2$ was 1.41–2.19 times that of RI-WM on both-ok replicates (1.42–7.89 on all replicates). Since the implementations agree when their settings agree, these differences come from the default settings rather than the code, consistent with the effect of the pooled bandwidth rule in the sensitivity cells. Details are in Online Resource 1.

6.9 The misplacement matrix

The misplacement matrix is recovered only where the constant-weight model holds. For CDA($\hat{\boldsymbol{\alpha}}$), the bias of $\hat{\alpha}_{11}$, averaged over the cells of each law and ranking level, is within 0.004 of zero on CW data; under concomitant ranking it exceeds the averaged misplacement matrix by 0.08–0.10, and under perfect ranking it falls short of one by 0.04–0.05.

Table 7: Design gain: median ratio MSE(SRS)/MSE(RSS) for RI-WM over the cells with separations 4σ and 6σ ($N \in \{90, 150\}$; all replicates; independent samples). Values above one favour RSS.

Par.	Law	perfect	C 0.95	C 0.9	C 0.7	C 0.5	CW 0.9	CW 0.7
π_1	Normal	1.56	1.64	1.47	1.34	1.16	1.25	1.18
	t_5	1.72	1.63	1.52	1.34	1.12	1.23	1.09
	Laplace	1.66	1.61	1.59	1.31	1.10	1.23	1.18
μ_1	Normal	1.11	1.08	1.04	0.99	0.99	1.03	1.01
	t_5	1.19	1.04	1.04	0.98	1.05	1.06	1.04
	Laplace	0.98	1.06	1.07	1.03	0.98	1.06	1.07
μ_2	Normal	1.35	1.25	1.13	1.08	1.01	1.08	1.06
	t_5	1.29	1.18	1.12	1.05	1.04	1.10	1.04
	Laplace	1.25	1.11	1.11	1.03	1.03	1.16	1.06

Table 8: f known: median MSE ratio of the full RSS likelihood with the true $\boldsymbol{\alpha}$ (P-RA) to the rank-ignoring likelihood (P-RI), $N \in \{90, 150\}$, all replicates. Only the perfect and CW columns describe a correctly specified model.

Par.	Law	perfect	C 0.95	C 0.9	C 0.7	C 0.5	CW 0.9	CW 0.7
μ_2	Normal	0.87	1.01	1.04	1.02	1.00	0.94	0.97
	t_5	0.87	1.03	1.06	1.02	1.01	0.91	0.96
	Laplace	0.82	0.99	1.04	1.03	1.00	0.91	0.94
μ_1	Normal	0.87	1.00	1.04	1.01	1.00	0.96	0.98
	t_5	0.87	1.02	1.04	1.02	1.00	0.95	0.98
	Laplace	0.83	1.01	1.05	1.03	0.98	0.92	0.97
π_1	Normal	0.90	0.96	0.98	1.01	0.99	0.94	0.96
	t_5	0.92	0.97	0.97	0.98	0.99	0.93	0.96
	Laplace	0.91	0.97	0.97	0.99	0.99	0.95	0.96

Table 9: Trend in N: normal components, $H = 5$, separation 6σ (median over $\pi_1 \in \{0.4, 0.2\}$). Left: P-RA, bias and standard deviation of $\hat{\mu}_2$ in units of σ, and its MSE ratio to P-RI. Right: bias of $\hat{\pi}_1$ from CDA(I) and its MSE ratio to RI-WM (both-ok).

	f known, true $\boldsymbol{\alpha}$: $\hat{\mu}_2$ bias, SD (units of σ), MSE ratio									CDA, $\boldsymbol{\alpha} = I$: $\hat{\pi}_1$ bias; ratio		
	$N = 90$			$N = 150$			$N = 300$			90	150	300
Ranking	bias	SD	ratio	bias	SD	ratio	bias	SD	ratio			
perfect	-0.002	0.098	0.89	+0.001	0.075	0.86	+0.002	0.052	0.87	-0.003; 0.97	-0.001; 0.96	-0.000; 0.94
C 0.95	-0.033	0.115	1.06	-0.037	0.088	1.13	-0.037	0.063	1.29	+0.007; 0.96	+0.009; 1.03	+0.008; 1.18
C 0.9	-0.039	0.118	1.11	-0.036	0.094	1.14	-0.043	0.066	1.40	+0.012; 1.07	+0.014; 1.21	+0.014; 1.59
C 0.7	-0.028	0.126	1.05	-0.031	0.098	1.10	-0.031	0.068	1.21	+0.028; 1.46	+0.029; 1.99	+0.029; 3.57
C 0.5	-0.017	0.128	1.03	-0.015	0.099	1.01	-0.016	0.070	1.05	+0.039; 1.84	+0.041; 2.80	+0.042; 4.95
CW 0.9	-0.000	0.115	0.92	-0.002	0.093	0.94	+0.001	0.063	0.94	+0.021; 1.25	+0.022; 1.52	+0.023; 2.36
CW 0.7	-0.003	0.121	0.96	+0.003	0.094	0.96	-0.001	0.066	0.96	+0.034; 1.66	+0.034; 2.47	+0.035; 3.90

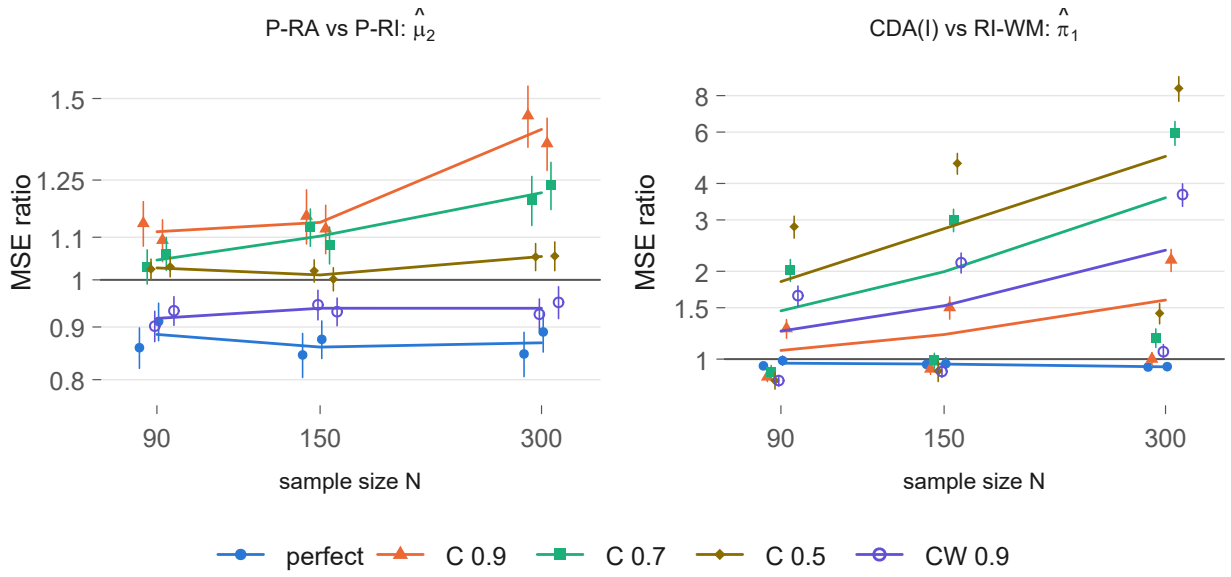


Figure 4: Trend in N (normal components, $H = 5$, 6σ, log scales). Each symbol is one cell ($\pi_1 = 0.4$ left of each pair, 0.2 right), with ± 2 Monte Carlo standard errors; lines join the medians over the two cells. Left: P-RA against P-RI for μ_2. Right: CDA(I) against RI-WM for π_1 (both-ok).

Table 10: Median MSE ratio of CDA(I) to RI-WM (both-ok replicates; $N \in \{90, 150\}$).

Law	Sep.	perfect	C 0.95	C 0.9	C 0.7	C 0.5	CW 0.9	CW 0.7
(a) μ_2								
Normal	6	1.00	1.00	1.00	0.99	0.99	1.00	0.99
	4	1.01	0.97	0.96	0.92	0.92	0.93	0.91
	3	1.03	0.92	0.84	0.74	0.75	0.77	0.74
t_5	6	1.07	1.00	0.99	0.97	0.95	0.97	0.94
	4	1.05	1.02	1.00	0.95	0.93	0.94	0.94
	3	1.08	0.99	0.93	0.79	0.80	0.82	0.80
Laplace	6	1.09	1.01	1.00	0.97	0.94	0.95	0.94
	4	1.10	1.03	1.01	0.94	0.90	0.93	0.92
	3	1.16	1.06	1.01	0.80	0.81	0.83	0.84
(b) μ_1								
Normal	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	4	1.00	1.01	1.03	1.09	1.16	1.06	1.09
	3	1.06	1.01	0.99	1.01	1.21	0.95	1.11
t_5	6	1.00	1.00	1.00	1.01	1.01	1.01	1.01
	4	1.01	1.01	1.01	1.02	1.06	1.01	1.04
	3	1.10	1.15	1.01	1.05	1.02	0.89	1.03
Laplace	6	1.00	1.01	1.01	1.01	1.03	1.01	1.01
	4	1.04	1.03	1.02	1.03	1.03	1.01	1.02
	3	1.14	1.03	1.13	0.93	1.00	0.84	0.91
(c) π_1								
Normal	6	0.96	0.94	0.97	1.23	1.44	1.17	1.38
	4	0.99	0.93	0.93	1.07	1.33	1.02	1.27
	3	1.01	0.89	0.80	0.89	1.20	0.80	1.00
t_5	6	1.01	0.99	1.05	1.25	1.47	1.09	1.35
	4	1.02	0.99	1.01	1.19	1.33	1.06	1.28
	3	1.09	0.99	0.95	1.01	1.34	0.93	1.26
Laplace	6	1.05	1.01	1.03	1.17	1.48	1.15	1.42
	4	1.09	1.03	1.04	1.13	1.38	1.05	1.25
	3	1.15	1.06	1.06	1.02	1.37	0.95	1.20

Table 11: Median bias of CDA(I) and RI-WM over the cells with separations 4σ and 6σ ($N \in \{90, 150\}$; all replicates); locations in units of σ.

		π_1		μ_1 (units of σ)		μ_2 (units of σ)	
Law	Ranking	CDA(I)	RI-WM	CDA(I)	RI-WM	CDA(I)	RI-WM
Normal	perfect	-0.002	-0.001	-0.005	-0.005	-0.007	-0.006
	C 0.95	+0.004	-0.001	+0.002	-0.001	-0.005	-0.006
	C 0.9	+0.008	-0.001	+0.004	-0.003	-0.004	-0.008
	C 0.7	+0.022	-0.001	+0.008	-0.002	-0.002	-0.009
	C 0.5	+0.032	-0.002	+0.013	+0.001	+0.004	-0.009
	CW 0.9	+0.017	-0.001	+0.005	-0.004	-0.002	-0.004
	CW 0.7	+0.030	-0.001	+0.011	-0.003	+0.001	-0.004
t_5	perfect	-0.004	-0.001	+0.006	-0.000	-0.011	-0.006
	C 0.95	+0.004	-0.002	-0.001	-0.003	-0.004	-0.005
	C 0.9	+0.010	-0.000	+0.005	+0.001	-0.001	-0.003
	C 0.7	+0.024	-0.001	+0.004	-0.006	+0.005	-0.005
	C 0.5	+0.033	-0.001	+0.016	+0.003	+0.007	-0.006
	CW 0.9	+0.018	-0.002	+0.002	-0.002	+0.001	-0.005
	CW 0.7	+0.031	-0.001	+0.012	+0.003	+0.006	-0.003
Laplace	perfect	-0.004	-0.001	+0.000	+0.001	-0.008	-0.005
	C 0.95	+0.005	-0.001	-0.001	-0.002	-0.003	-0.006
	C 0.9	+0.009	-0.001	+0.001	-0.002	-0.002	-0.007
	C 0.7	+0.022	-0.001	+0.010	+0.001	+0.007	-0.007
	C 0.5	+0.035	-0.001	+0.014	+0.002	+0.008	-0.010
	CW 0.9	+0.018	-0.001	+0.009	+0.003	-0.001	-0.008
	CW 0.7	+0.031	-0.001	+0.010	-0.000	+0.005	-0.008

Table 12: Rank-ignoring fits: MSE ratio of the likelihood-type update (RI-L) to the weighted-mean update (RI-WM); median [range] over the six sub-grid cells per row (both-ok replicates).

Law	Sep.	μ_1	μ_2	π_1
Normal	6	1.12 [1.06, 1.18]	1.13 [1.09, 1.19]	1.00 [1.00, 1.00]
	4	1.17 [1.12, 1.18]	1.18 [1.17, 1.22]	1.01 [1.00, 1.02]
t_5	6	0.99 [0.97, 1.02]	0.97 [0.95, 0.99]	1.00 [0.99, 1.00]
	4	0.98 [0.94, 1.03]	0.97 [0.92, 1.06]	0.99 [0.99, 1.00]
Laplace	6	0.75 [0.67, 0.82]	0.77 [0.73, 0.85]	0.99 [0.99, 1.01]
	4	0.78 [0.75, 0.81]	0.79 [0.72, 0.83]	0.97 [0.93, 0.98]

7 Real-data considerations

The thesis (Moniri 2021) illustrated the method by resampling from bone mineral density (BMD) measurements of women aged 50 and older from the third US National Health and Nutrition Examination Survey (NHANES III). We do not include a real-data illustration in this paper, for three reasons.

First, the provenance of the data file could not be fully reconstructed. The thesis reports 3,299 women; the file available to us has 3,240 records, of which 3,228 remain after removing records with missing-value codes, and the procedure used to extract it from the public NHANES III files is not documented. The thesis text ranks on body weight (correlation 0.535 with BMD in the file), while the accompanying code ranks on age (correlation -0.49), the concomitant also used in the bone-density example of Alvandi et al. (2023).

Second, the population is poorly suited to the model. The BMD distribution is right-skewed (sample skewness 0.48), and the fitted normal components are only about 2σ apart, so the mixture density is unimodal; the symmetric location mixture is then a weakly identified description rather than a reflection of two well-separated groups.

Third, and decisively, the fit is unstable at realistic sample sizes. In a feasibility check with 200 replicates per size (sets of five ranked on weight; target the same fit on all 3,228 records), the rank-ignoring fit failed in 69%, 54% and 37% of RSS samples of size 150, 300 and 600, and in 70%, 56% and 33% of SRS samples of the same sizes. A single-sample illustration would then depend heavily on which sample was drawn.

A defensible illustration would need data with clearly separated groups and a cheap, documented ranking variable, together with an extraction script from the original public files; we leave this to future work.

8 Discussion and limitations

For semi-parametric location mixtures, the benefit of ranked set sampling comes from the design. Fitting the pooled ranked set sample as if it were a simple random sample captures that benefit under any ranking mechanism (Proposition 1), and the gain is large for the mixing proportion when ranking is good (Table 7). The rank-aware EM-type algorithm with $\boldsymbol{\alpha}$ estimated does not improve on the rank-ignoring fit overall in our study; at $N = 300$ it is better in a few cells, by small margins (Section 6.3). Its location step does not use the ranks (Lemma 1), its rank-aware mixing-proportion update with estimated $\boldsymbol{\alpha}$ does not improve on the rank-ignoring one, and with heavy-tailed components it is clearly worse. With $\boldsymbol{\alpha}$ fixed at the identity, the same update helps $\hat{\pi}_1$ slightly when ranking is in fact perfect and the components are normal, but not otherwise.

Two findings deserve emphasis. First, the constant-weight misplacement model is not a description of ranking on a noisy concomitant (Table 1). A likelihood built on it is biased under concomitant ranking even with the component law and the true average misplacement matrix known, and over the sample sizes studied the bias does not shrink, so the rank-aware fit falls further behind the rank-ignoring fit as N grows (Table 9). This is consistent with the small biases reported by Hatefi et al. (2015) for parametric estimators fitted to data generated with the same concomitant model; our results suggest that such biases are a systematic consequence of the ranking model rather than a small-sample effect. Second, assuming perfect ranking when it is not can produce an apparent small-sample gain in one parameter alongside biases in the others that do not shrink with N (Section 6.6).

Even in the best case for the ranks, with f and $\boldsymbol{\alpha}$ known and the ranking model correct, the

median MSE ratio for $\hat{\mu}_2$ is 0.82–0.87 under perfect ranking and 0.91–0.97 under constant-weight ranking error (Tables 2 and 8). A semi-parametric method must also estimate f and, in practice, $\boldsymbol{\alpha}$, and $\hat{\boldsymbol{\alpha}}$ is biased whenever the constant-weight model fails (Section 6.9).

Recommendations.

1. Use RSS when ranking is cheap and reasonably accurate; the gain for the mixing proportion is substantial.

2. Fit the pooled sample with a rank-ignoring semi-parametric method. With components that may be heavy-tailed, prefer a likelihood-type location update; with near-normal components, the weighted-mean update is more efficient. Reporting both is a reasonable default.

3. Do not use rank-aware likelihoods built on the constant-weight misplacement model when ranking relies on a concomitant, and do not assume perfect ranking in the likelihood.

4. Estimate the kernel bandwidth from the centred residuals, not from the pooled data, and report the rate of degenerate fits; at about 3σ separation and $N \leq 150$, collapse is common for every method.

Limitations. The study covers two components with equal shape, three laws, set sizes 3 and 5, and N up to 150 for heavy-tailed laws and 300 for normal components. The likelihood-type update was evaluated on a sub-grid only, and we did not study a data-driven choice between the two location updates. The concomitant model is one model of imperfect ranking; other ranking mechanisms may interact differently with the constant-weight likelihood. The asymptotic benchmark of Table 2 covers normal components only. We give no real-data illustration (Section 7).

9 Conclusion and future research

For semi-parametric location mixtures, ranked set sampling pays off through its design, and the payoff is available to any estimator applied to the pooled sample. The ranks themselves add at most a modest further gain, and only when the ranking model is correct; under the concomitant ranking that motivates RSS in practice, a rank-aware likelihood built on constant misplacement weights is misspecified, and in our study (normal components, $H = 5$, N up to 300) its disadvantage grew with the sample size. Practitioners should therefore fit the pooled sample without the ranks, choosing the location update by tail weight.

Three directions follow. First, a ranking model that describes concomitant ranking, for example by modelling the concomitant jointly with the response, could make the ranks usable without the bias documented here. Second, a data-driven choice between the weighted-mean and likelihood-type location updates, or a robust update that performs well for all tail weights, would remove the need to know the tail behaviour in advance. Third, a real-data illustration with clearly separated groups and a documented ranking variable would show the practical size of the design gain.

Statements and Declarations

Funding. No funding was received for conducting this study.

Competing interests. The author has no relevant financial or non-financial interests to disclose.

Ethics approval. Not applicable. The study uses simulated data and, in Section 7, publicly available de-identified survey data.

Author contributions. Seyed Jalaleddin Moniri carried out the MSc thesis research on which this work builds (Moniri 2021), with the research question proposed by, and methodological and programming advice from, the thesis supervisor. For the present paper, the author framed the research question, designed and carried out the study, wrote the software, and wrote the manuscript. CRediT roles: conceptualisation; methodology; project administration; software; validation; writing – original draft; writing – review and editing.

Use of generative AI. The author used AI (Claude, Anthropic) to help with the writing. The author checked all of it and takes full responsibility for the manuscript.

Data and code availability. The simulation study uses simulated data only. The feasibility check in Section 7 uses bone mineral density records from the third US National Health and Nutrition Examination Survey (NHANES III), whose public-use files are available from the US National Center for Health Statistics (`https://www.cdc.gov/nchs/nhanes/`). The R package `semirss` (base R only), the scripts that run the simulation study and produce every table and figure, and the saved per-cell results will be deposited in a public repository under the MIT licence on publication. The original thesis code is not included.

Supplementary information. Online Resource 1 (ESM_1.pdf) contains an additional figure for t_5 components, the sensitivity cells, the details of the comparison with `mixtools::spEMsymloc`, and the results for every cell and method (bias, root mean squared error with its Monte Carlo standard error, median number of iterations, and failures by cause).

References

Amirhossein Alvandi, Sedigheh Omidvar, Armin Hatefi, Mohammad Jafari Jozani, Omer Ozturk, and Nader Nematollahi. Bayesian mixture modelling with ranked set samples. arXiv preprint arXiv:2306.08092, 2023. URL `https://doi.org/10.48550/arXiv.2306.08092`.

Amirhossein Alvandi, Sedigheh Omidvar, Armin Hatefi, Mohammad Jafari Jozani, Omer Ozturk, and Nader Nematollahi. Bayesian mixture modelling with ranked set samples. *Statistics in Medicine*, 43(19):3723–3741, 2024. `https://doi.org/10.1002/sim.10144`.

Tatiana Benaglia, Didier Chauveau, and David R. Hunter. An EM-like algorithm for semi- and nonparametric estimation in multivariate mixtures. *Journal of Computational and Graphical Statistics*, 18(2):505–526, 2009a. `https://doi.org/10.1198/jcgs.2009.07175`.

Tatiana Benaglia, Didier Chauveau, David R. Hunter, and Derek Young. mixtools: An R package for analyzing finite mixture models. *Journal of Statistical Software*, 32(6), 2009b. `https://doi.org/10.18637/jss.v032.i06`.

Laurent Bordes, Stéphane Mottelet, and Pierre Vandekerkhove. Semiparametric estimation of a two-component mixture model. *The Annals of Statistics*, 34(3):1204–1232, 2006. `https://doi.org/10.1214/009053606000000353`.

Laurent Bordes, Didier Chauveau, and Pierre Vandekerkhove. A stochastic EM algorithm for a semiparametric mixture model. *Computational Statistics & Data Analysis*, 51(11):5429–5443, 2007. https://doi.org/10.1016/j.csda.2006.08.015.

Cristina Butucea and Pierre Vandekerkhove. Semiparametric mixtures of symmetric distributions. *Scandinavian Journal of Statistics*, 41(1):227–239, 2014. https://doi.org/10.1111/sjos.12015.

Zehua Chen, Zhidong Bai, and Bimal K. Sinha. *Ranked Set Sampling: Theory and Applications*, volume 176 of *Lecture Notes in Statistics*. Springer, New York, 2004. https://doi.org/10.1007/978-0-387-21664-5.

T. R. Dell and J. L. Clutter. Ranked set sampling theory with order statistics background. *Biometrics*, 28(2):545–555, 1972. https://doi.org/10.2307/2556166.

Jesse C. Frey. New imperfect rankings models for ranked set sampling. *Journal of Statistical Planning and Inference*, 137(4):1433–1445, 2007. https://doi.org/10.1016/j.jspi.2006.02.013.

Armin Hatefi. *Mixture Model Analysis with Rank-Based Samples*. PhD thesis, University of Manitoba, Winnipeg, Canada, 2014.

Armin Hatefi and Mohammad Jafari Jozani. Fisher information in different types of perfect and imperfect ranked set samples from finite mixture models. *Journal of Multivariate Analysis*, 119: 16–31, 2013. https://doi.org/10.1016/j.jmva.2013.03.018.

Armin Hatefi, Mohammad Jafari Jozani, and Djemel Ziou. Estimation and classification for finite mixture models under ranked set sampling. *Statistica Sinica*, 24(2):675–698, 2014. https://doi.org/10.5705/ss.2012.178.

Armin Hatefi, Mohammad Jafari Jozani, and Omer Ozturk. Mixture model analysis of partially rank-ordered set samples: age groups of fish from length-frequency data. *Scandinavian Journal of Statistics*, 42(3):848–871, 2015. https://doi.org/10.1111/sjos.12140.

David R. Hunter, Shaoli Wang, and Thomas P. Hettmansperger. Inference for mixtures of symmetric distributions. *The Annals of Statistics*, 35(1):224–251, 2007. https://doi.org/10.1214/009053606000001118.

M. Levine, D. R. Hunter, and D. Chauveau. Maximum smoothed likelihood for multivariate mixtures. *Biometrika*, 98(2):403–416, 2011. https://doi.org/10.1093/biomet/asq079.

G. A. McIntyre. A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, 3:385–390, 1952. https://doi.org/10.1071/ar9520385.

Seyed Jalaleddin Moniri. Semi-parametric mixture models with ranked set samples. Master's thesis, Memorial University of Newfoundland, St. John's, Canada, 2021. URL https://doi.org/10.48336/M2W2-N809.

Sedigheh Omidvar, Mohammad Jafari Jozani, and Nader Nematollahi. Judgment post-stratification in finite mixture modeling: an example in estimating the prevalence of osteoporosis. *Statistics in Medicine*, 37(30):4823–4836, 2018. https://doi.org/10.1002/sim.7984.

Brett Presnell and Lora L. Bohn. U-statistics and imperfect ranking in ranked set sampling. *Journal of Nonparametric Statistics*, 10(2):111–126, 1999. https://doi.org/10.1080/10485259908832756.

Byungtae Seo. The doubly smoothed maximum likelihood estimation for location-shifted semiparametric mixtures. *Computational Statistics & Data Analysis*, 108:27–39, 2017. https://doi.org/10.1016/j.csda.2016.11.003.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.

Cesar Augusto Taconeli and Idemauro Antonio Rodrigues de Lara. Evaluation of maximum likelihood procedures in ranked set sampling under imperfect ranking. *Communications in Statistics – Simulation and Computation*, 53(12):5719–5741, 2024. https://doi.org/10.1080/03610918.2023.2196749.

Jingjing Wu, Weixin Yao, and Sijia Xiang. Computation of an efficient and robust estimator in a semiparametric mixture model. *Journal of Statistical Computation and Simulation*, 87(11): 2128–2137, 2017. https://doi.org/10.1080/00949655.2017.1318136.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf