Novel Tau-Informed Initialization for Maximum Likelihood Estimation of Copulas with Discrete Margins

Anna van Es Eva Cantoni

University of Geneva

anna.vanes@unige.ch

European Meeting of Statisticians
25 August, 2026

Presentation Overview

Motivation

- Goal: exact,computationally stable maximum likelihood (ML) for Gaussian copulas with discrete margins, with emphasis on low-count Poisson settings.
- Challenge: the joint probability mass function requires multivariate normal (MVN) **rectangle probabilities**, i.e. finite differences of a copula cumulative distribution function. When **many counts are zero**, the optimization procedure suffers both from identifiability and numerical stability issues, especially for the Gaussian copula parameters.
- Main idea: use Kendall's tau for discrete margins to construct **IFM-like stable starting values** for the correlation matrix, then run **exact ML** with an **unconstrained parametrization** and **analytic score** functions.

Key contributions

1. **Three tau-informed initializers** tailored to low-count discrete margins: exact, low-intensity approximation, and transformation-based.
2. **IFM-style start:** estimate margins first, then initialize dependence via discrete Kendall's tau logic.
3. **Exact ML with stability:** unconstrained reparameterization (log for λ, spherical-Cholesky for $\mathbf{P}$) and **analytic score functions**.

Method pipeline

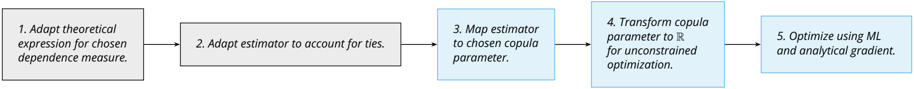


Figure: Schematic overview of the workflow. The contribution of the paper is highlighted in blue.

Steps 1 – 2 depend on the chosen dependence measure, Steps 4 – 5 depend on the copula parametrization and the numerical optimization routine, and Step 3 links the two. Hence, the **procedure can be adapted to various scenarios**, where one (or more) Steps would need to be adapted, depending on the choice of dependence measure and copula.

Discrete Gaussian copula with Poisson margins

Let $\mathbf{Y} = (Y_1, \ldots, Y_d)$ with Poisson margins $Y_j \sim \mathrm{Pois}(\lambda_j)$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_d)$ and dependence given by a Gaussian copula with correlation matrix $\boldsymbol{P}$. In this setting, the joint probability mass function (pmf) is expressed via rectangle probability differences as follows, where $u_{j\leftarrow}$ indicates the limit of u_j as it approaches from above [Joe, 2015]:

$$\begin{aligned} h(\mathbf{y}; \mathbf{P}, \boldsymbol{\lambda}) &= \mathbb{P}(Y_1 = y_1, \ldots, Y_d = y_d) = \mathbb{P}(y_{1\leftarrow} < Y_1 \leq y_1, \ldots, y_{d\leftarrow} < Y_d \leq y_d) \\ &= \sum_{t_1=0}^{1} \cdots \sum_{t_d=0}^{1} (-1)^{t_1+\ldots+t_d} C_{\mathbf{P}}(u_{1_{t_1}}, \ldots, u_{d_{t_d}}), \end{aligned} \tag{1}$$

with $u_{j0} = F_{Y_j}(y_j), u_{j1} = F_{Y_j}(y_{j\leftarrow})$, F_{Y_j} the Poisson cumulative distribution function with parameter λ_j. Note that since we focus on the integer-valued case here, $F(y_{j\leftarrow}) = F(y_j - 1)$.

Discrete Gaussian copula with Poisson margins

$$h(\mathbf{y}; \mathbf{P}, \boldsymbol{\lambda}) = \mathbb{P}(Y_1 = y_1, \ldots, Y_d = y_d) = \mathbb{P}(y_{1\leftarrow} < Y_1 \leq y_1, \ldots, y_{d\leftarrow} < Y_d \leq y_d)$$

$$= \sum_{t_1=0}^{1} \cdots \sum_{t_d=0}^{1} (-1)^{t_1+\ldots+t_d} C_{\mathbf{P}}(u_{1_{t_1}}, \ldots, u_{d_{t_d}}),$$

The equation above is computationally costly, however one can show it is feasible if:

- The initialization is done efficiently;
- The reparametrization allows for unconstrained optimization;
- The numerical procedure leverages the analytical expression of the gradient.

Step 3: Discrete Kendall's tau-informed initialization

Why Kendall's tau

For low counts, Pearson correlation of counts can be badly attenuated, while rank-based measures remain interpretable and often more stable under discretization.

Rank-based measures are more appropriate to describe dependence. Our proposal leverages on Kendall's tau.

Definition (Kendall's tau [Kendall, 1938])

Formally, for two independent copies $(X_1, Y_1), (X_2, Y_2)$ of (X, Y) with joint distribution function H, Kendall's tau is the difference between the probabilities of concordance and discordance, that is

$$\tau = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0).$$

For continuous margins this reduces to

$$\tau = 2P((X_1 - X_2)(Y_1 - Y_2) > 0) - 1.$$

Step 3: Discrete Kendall's tau-informed initialization

Population definitions of Kendall's tau for two discrete variables

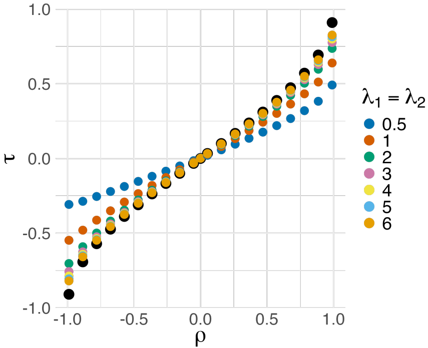
- τ_b of Kendall [Kendall, 1945] (ties correction)
- general τ for discrete margins [Nikoloulopoulos and Karlis, 2009]
- τ_H of [Pimentel, 2009] (zero inflation correction)
- τ_A of [Perrone et al., 2023] (ties and zero inflation correction)

Step 3: Discrete Kendall's tau-informed initialization

Initializer options:

- **Option 1 (IFM-type approach):** estimate the Poisson parameters, empirical Kendall's tau by using $\hat{\tau}_A$ from [Perrone et al., 2023], solve for $\hat{\rho}$ by minimizing a tau discrepancy implied by the Gaussian copula with Poisson margins.
- **Option 2 (Skellam approach):** use an alternative representation of Kendall's tau using Skellam distribution. This approach is theoretical rather than practical.
- **Option 3 (Arcsine approximation):** use a corrected relationship inspired by the continuous arcsine form, adjusted for discreteness. **Option 3a** and **3b** use different approximation functions:
  - Option 3a: logistic shrinkage
  - Option 3b: τ_b inspired scaling

Option 3: intuition



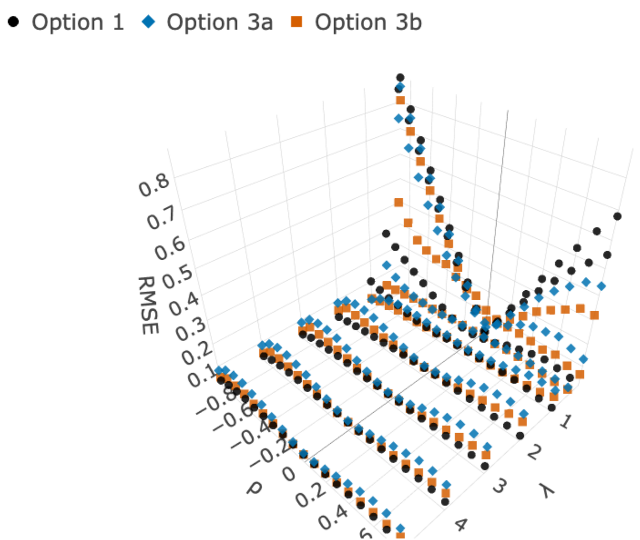
For a bivariate Gaussian copula with continuous margins, and correlation parameter ρ, we have the following expression for Kendall's tau (τ):

$$\tau = \frac{2}{\pi} \arcsin(\rho).$$

Using the properties of the transformation of the arcsin function, we assume for a Gaussian copula with Poisson margins,

$$\tau_{\text{Pois}} \approx \underbrace{a \frac{2}{\pi} \arcsin(b\,\rho)}_{a,\, b \text{ are functions of the Poisson parameters}} .$$

RMSE for options for tau-informed initialization



Steps 4,5: Unconstrained optimization with analytic score

Unconstrained parameters

1. $\eta_j = \log(\lambda_j) \in \mathbb{R}$,
2. reparametrize **P** via spherical-Cholesky parameterization.

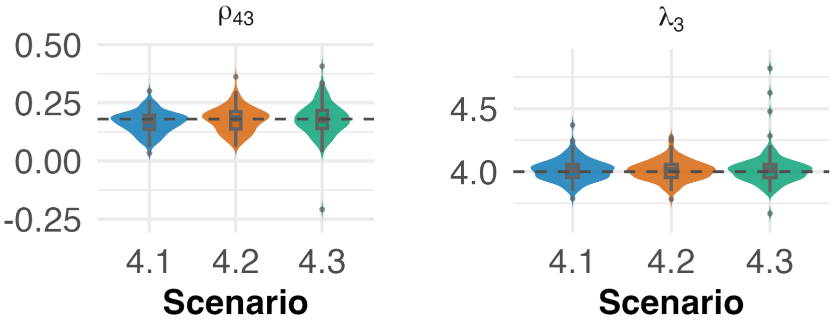
Reparametrization alongside with analytical gradient allows for **better curvature** and **scaling** for Newton-type optimization, as well as **reduction in numerical noise** and **improved convergence and computational speed**, specifically in higher dimensions.

Simulation highlights

- $n = 500$, 100 replications;
- $d = 2$; $\rho_{21} = -0.5$;
- $d = 4$: $\boldsymbol{\rho} = (\rho_{21}, \rho_{31}, \rho_{41}, \rho_{32}, \rho_{42}, \rho_{43}) = (-0.42, -0.23, 0.73, 0.21, -0.64, 0.18)$;
- Analytic vs Numeric gradient;
- Starting from Option 1 vs empirical correlation ("Corr").

Scenario	d	Lambdas	Gradient	Start	$\overline{\text{RMSE}}$	$\overline{\text{Bias}}(\times 10^3)$	Time
4.1	4	0.6, 2, 4, 0.8	Analytic	Corr	0.054	-2.233	$19H43M44.03S$
4.2	4	0.6, 2, 4, 0.8	Analytic	Option 1	0.050	-0.821	$17H8M15.31S$
4.3	4	0.6, 2, 4, 0.8	Numeric	Option 1	0.065	-2.139	$3d20H9M21.88S$

Simulation highlights



Conclusion and practical guidance

- **If counts are very small and zeroes are common:** prioritize the discrete tau-matching initializer (Option 1) for robust dependence starts.
- **If λ is moderate or large:** the approximation-based initializer can be competitive and very fast.
- **For optimization:** prefer the unconstrained parameterization and analytic score functions for stable Newton-type iterations.

Extensions and further work

- For sparse dependence, impose sparsity pattern in a modified Cholesky factor
- Extension possible to zero-inflated, hurdle or negative binomial settings
- Analytical gradients easily adaptable to different margins
- Automatic differentiation could be explored (however implementation for Gaussian cdf is lacking)
- Other dependence measures, e.g. Spearman's rho
- Other copulas, e.g. Archimedean (only positive dependence)

References

Joe, H. (2015).
Dependence Modeling with Copulas.
Chapman and Hall/CRC.

Kendall, M. G. (1938).
A New Measure of Rank Correlation.
Biometrika, 30(1/2):81–93.

Kendall, M. G. (1945).
The treatment of ties in ranking problems.
Biometrika, 33(3):239–251.

Nikoloulopoulos, A. K. and Karlis, D. (2009).
Modeling Multivariate Count Data Using Copulas.
Communications in Statistics - Simulation and Computation, 39(1):172–187.

Perrone, E., Van Den Heuvel, E. R., and Zhan, Z. (2023).
Kendall's tau estimator for bivariate zero-inflated count data.
Statistics & Probability Letters, 199:109858.

Pimentel, R. S. (2009).
Kendall's Tau and Spearman's Rho for Zero-Inflated Data.

Thank you!

Questions? Comments?

van Es, Anna and Cantoni, Eva. "Novel tau-informed initialization for maximum likelihood estimation of copulas with discrete margins" Dependence Modeling, vol. 14, no. 1, 2026, pp. 20250020. https://doi.org/10.1515/demo-2025-0020