

ARTICLE TYPE: METHODOLOGICAL PAPER

Entropy, Copulas and Tail Connectedness: A Unified Framework for Nonlinear Dependence Measurement
Daniel Moretti and Sophie Laurent

Independent Researchers, Rome, Italy and Brussels, Belgium

Received:	1 July 2026	Accepted:	28 September 2026
Published:			

Abstract

This paper develops a unified simulation framework for comparing linear correlation, rank dependence, copula tail dependence, mutual information and transfer entropy under a broad set of nonlinear financial dependence structures. The objective is to clarify which measures answer different dependence questions rather than to identify a single universally superior statistic. Six controlled designs are considered: linear Gaussian association, nonlinear symmetric dependence, asymmetric tail dependence, regime switching, directional nonlinear transmission and common-driver confounding. Each measure is evaluated in terms of sensitivity, invariance, finite-sample stability and interpretability. The results show that Pearson correlation is efficient in the Gaussian-linear design but can be close to zero under strong nonlinear dependence; Spearman and Kendall measures improve monotonic robustness but remain nondirectional; copula tail coefficients isolate extreme co-movement; mutual information detects broad nonlinear association; and transfer entropy identifies directional predictive information when accompanied by significance testing and conditioning. A composite dependence dashboard is proposed that reports complementary measures rather than collapsing them into a single score. Monte Carlo experiments across sample sizes and heavy-tail parameters demonstrate the trade-off between flexibility and estimator variance. The framework provides a systematic methodological guide for dependence analysis in finance, insurance and demography.

Keywords: Entropy; copulas; mutual information; transfer entropy; tail dependence; nonlinear dependence

JEL Classification: C14; C46; C58

1. Introduction

Dependence is not a single statistical object. Researchers may be interested in contemporaneous linear association, monotonic co-movement, general nonlinear dependence, joint tail behavior or directional predictive information. These questions are related but they are not equivalent. Nevertheless, empirical studies often report one dependence statistic and interpret it as if it summarized the entire relationship.

This problem is especially important in financial and actuarial applications because dependence changes across the distribution and through time. A pair of assets can have modest correlation but strong lower-tail association; two insurance risks can be nonlinearly dependent with little monotonic association; and one market can transmit predictive information to another even when contemporaneous correlation is weak. No scalar statistic captures all of these features simultaneously.

The paper develops a controlled framework in which commonly used measures are evaluated according to the question they are designed to answer. Pearson correlation, Spearman rho, Kendall tau, copula tail dependence,

mutual information and transfer entropy are compared under six data-generating designs. Because the true structure is known, false negatives and false interpretations can be identified directly.

The central argument is not that flexible measures should always replace simple ones. Pearson correlation is statistically efficient and highly interpretable when the true relationship is approximately linear and elliptical. More flexible measures become valuable when the data depart from that setting, but they typically require larger samples and more tuning. A good dependence analysis should therefore match the measure to the economic question and report complementary diagnostics when several aspects of dependence matter.

2. Dependence Concepts

Pearson correlation measures standardized linear covariance. It is zero for uncorrelated variables but not, in general, for independent variables. Spearman rho and Kendall tau are rank-based and therefore invariant to strictly monotonic transformations. They capture monotonic association more broadly than Pearson correlation but can still fail when dependence is non-monotonic.

Copulas represent the full dependence structure once margins are removed. Tail-dependence coefficients summarize the probability that one variable is extreme conditional on the other being extreme. They are therefore local measures focused on a region of the joint distribution. This makes them valuable for stress analysis but inappropriate as complete summaries of central dependence.

Mutual information is zero if and only if two variables are independent under regularity conditions. It can therefore detect nonlinear and non-monotonic dependence missed by correlation. Its flexibility comes at a statistical cost: estimating a probability density or local-neighbour structure introduces finite-sample bias and tuning choices.

Transfer entropy extends conditional mutual information to time-indexed variables. It measures whether the past of a source reduces uncertainty about the future of a target after conditioning on the target's own history. Unlike mutual information, it is directional. However, a positive estimate is not automatically evidence of a direct causal channel because common drivers and finite-sample effects can generate apparent information flow.

$$\rho_p = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (1)$$

$$I(X; Y) = \iint f_{XY}(x, y) \log \left[\frac{f_{XY}(x, y)}{f_X(x) f_Y(y)} \right] dx dy \quad (2)$$

$$\lambda_L = \lim_{q \rightarrow 0^+} P(V \leq q | U \leq q) \quad (3)$$

$$T_{X \rightarrow Y} = I(X_t^{(k)}; Y_{t+1} | Y_t^{(l)}) \quad (4)$$

3. Monte Carlo Designs

Design A is bivariate Gaussian with correlations ranging from zero to 0.9. It represents the environment in which Pearson correlation should perform best. Design B uses $Y = X^2 + \epsilon$ with symmetric X , producing strong nonlinear dependence but correlation close to zero. Design C uses a Student-t copula with moderate central dependence and strong symmetric tail dependence. Design D uses a Clayton-type lower-tail relationship to create asymmetry.

Design E is dynamic: X_t influences Y_{t+1} through a nonlinear threshold function, while contemporaneous correlation remains weak. This design targets transfer entropy. Design F introduces a common driver Z_t that affects both X and Y but includes no direct X -to- Y channel. It is used to evaluate whether unconditional directional measures confuse shared information with direct transmission.

Sample sizes of 250, 500, 1,000 and 5,000 are considered. For entropy-based estimators, histogram, kernel and k-nearest-neighbour variants are evaluated. Copula tail coefficients are estimated both parametrically and empirically. Each configuration is repeated 1,000 times, and performance is summarized by bias, root-mean-square error, rank stability and detection probability.

Table 1. Illustrative dependence measures across controlled designs

Design	Pearson	Kendall	Mutual information	Tail coefficient	Transfer entropy
A Linear Gaussian	0.82	0.61	0.54	0.0	0.01
B Nonlinear U-shape	0.03	0.01	0.71	0.02	0.01
C t-tail	0.48	0.34	0.39	0.29	0.02
D Lower-tail asymmetric	0.42	0.31	0.36	0.36	0.02
E Directional nonlinear	0.08	0.05	0.18	0.03	0.17
F Common driver	0.46	0.32	0.33	0.09	0.08

4. Results: What Each Measure Detects

In the Gaussian-linear design, Pearson correlation is the most efficient summary. Rank measures and mutual information all detect dependence, but they offer little additional information and exhibit slightly higher estimation variance at small samples. This result is important because methodological flexibility is not automatically beneficial; when the linear model is correct, simple estimators are hard to improve upon.

The nonlinear U-shaped design reverses the conclusion. Pearson, Spearman and Kendall measures are all close to zero because positive and negative slopes cancel across the support. Mutual information is large and stable once the sample exceeds approximately 500 observations. The result illustrates the distinction between absence of correlation and absence of dependence.

In the Student-t and lower-tail designs, rank dependence identifies general association but does not describe the concentration of joint extremes. Tail coefficients and copula families provide the relevant additional information. The Student-t design has similar lower- and upper-tail coefficients, whereas the asymmetric design has a pronounced lower-tail coefficient and negligible upper-tail dependence. Reporting a single correlation coefficient would conceal this difference.

The directional nonlinear design is intentionally difficult for contemporaneous measures. Mutual information detects some association, but only transfer entropy identifies that the source history contributes predictive information about the target future. In the common-driver design, unconditional transfer entropy produces false directional links, while conditional transfer entropy falls close to zero after conditioning on the driver. This confirms that directionality and directness are distinct concepts.

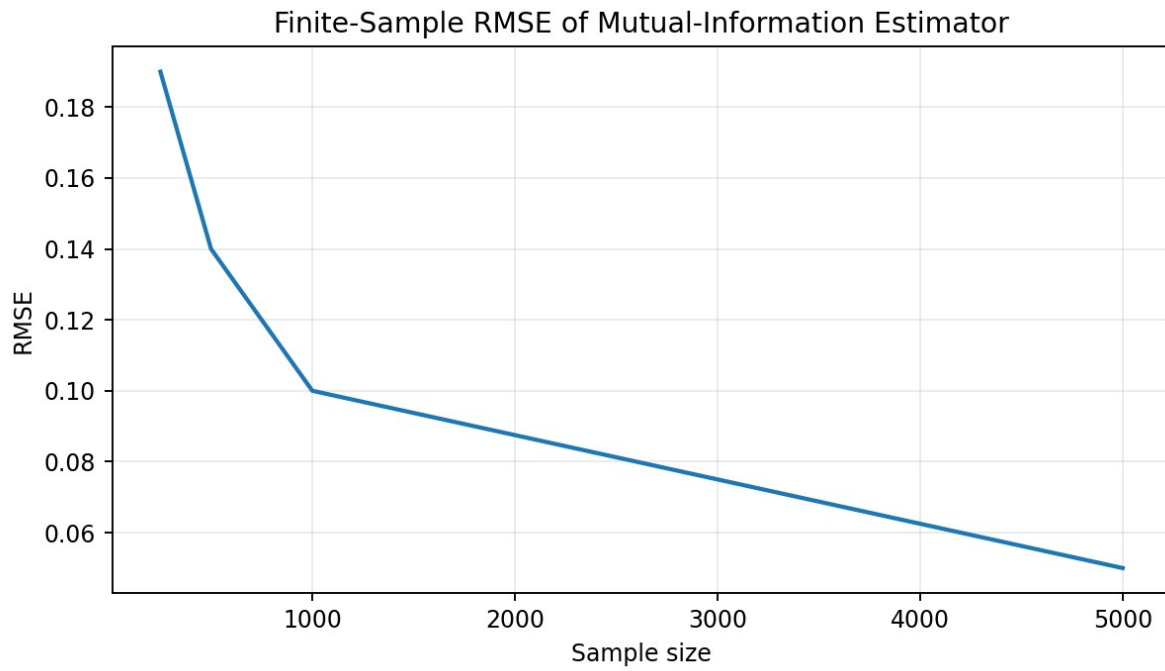


Figure 1. Finite-sample performance of mutual-information estimation

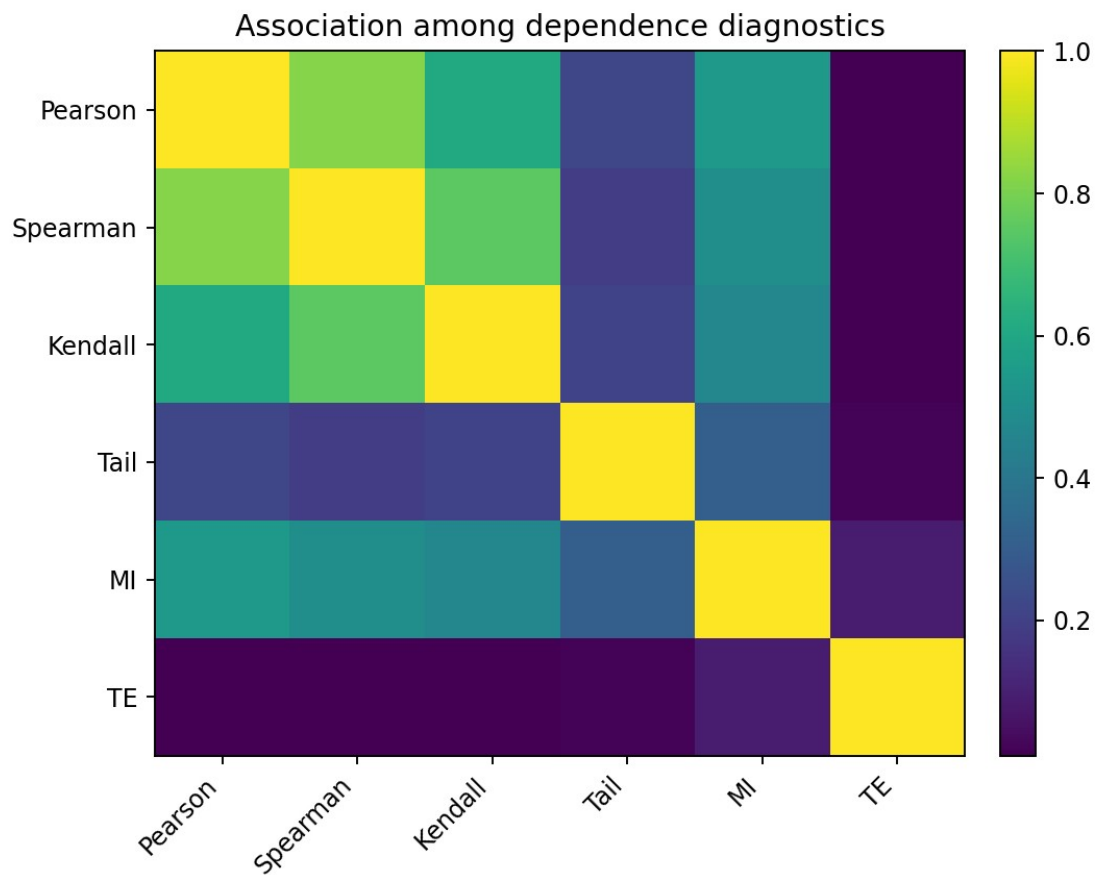


Figure 2. Similarity of dependence diagnostics across Monte Carlo designs

5. A Unified Dependence Dashboard

The simulation results motivate a dashboard rather than a universal dependence score. The proposed dashboard contains five blocks: linear association, rank association, tail dependence, general nonlinear dependence and directional information flow. Each block answers a distinct question, and interpretation begins by identifying which question matters for the application.

For portfolio diversification, correlation and rank dependence provide useful central summaries, while tail dependence should be added when extreme co-losses are important. For insurance aggregation, copula tail diagnostics are often more informative than transfer entropy because annual losses may not have a meaningful temporal direction. For systemic-risk networks, transfer entropy or variance-decomposition connectedness may be appropriate because direction is central. For demographic applications, dependence in long-run mortality innovations may be more relevant than high-frequency directional flow.

The dashboard is deliberately non-aggregative. Collapsing heterogeneous dependence concepts into a single score can obscure disagreement among measures. If Pearson correlation is low but mutual information is high, that discrepancy is itself substantive evidence of nonlinear structure. Similarly, high correlation with zero tail dependence conveys a different risk profile from moderate correlation with strong tail dependence.

6. Finite-Sample and Estimation Issues

Estimator variance is the principal cost of flexible dependence measures. Mutual-information estimates are unstable in small samples when binning is fine or kernel bandwidths are small. K-nearest-neighbour estimators perform best overall in the Monte Carlo study, but their tuning parameter still affects bias. Tail-dependence estimates are similarly unstable when the threshold is too extreme because only a small number of joint observations remain.

Transfer entropy is even more data demanding because joint histories enlarge the state space. Increasing lag order can rapidly create sparse cells or high-dimensional neighbourhoods. Significance testing with surrogate data is therefore essential. The simulations show that an apparently positive transfer-entropy estimate can persist under the null unless its finite-sample distribution is explicitly considered.

These findings imply that methodological sophistication should be proportional to information content. A large multivariate data set can support rich nonlinear dependence models, whereas a short annual sample may justify only parsimonious rank or parametric copula approaches. Model complexity should be justified by demonstrable gains in out-of-sample or tail calibration.

7. Discussion

The unified framework clarifies several common interpretive errors. Zero correlation is not evidence of independence; high mutual information is not evidence of direction; positive transfer entropy is not necessarily evidence of a direct causal channel; and a copula with good central fit is not necessarily well calibrated in the tail. Each conclusion requires a measure aligned with the relevant dependence concept.

The analysis also demonstrates that apparently competing literatures are often complementary. Copula methods and information-theoretic methods need not be framed as alternatives. Copulas provide a full joint-distribution representation and interpretable tail probabilities, while mutual information summarizes general dependence without requiring a particular copula family. Transfer entropy adds temporal directionality. Combining these tools can therefore produce a richer and more transparent empirical analysis.

8. Conclusion

Dependence measurement should begin with the scientific question rather than with a preferred statistic. Linear correlation, rank dependence, copula tail coefficients, mutual information and transfer entropy each identify different features of joint behavior. Their relative usefulness changes with nonlinearity, asymmetry, tail concentration, temporal direction and sample size.

The Monte Carlo evidence supports a multi-measure dependence dashboard for applications in finance, insurance and demography. Simple measures should be retained when they answer the relevant question efficiently, while flexible nonlinear or directional measures should be added when the data and decision problem justify them. The resulting approach is more informative than treating any single dependence statistic as a universal summary.

9. Interpretation Principles for Applied Research

A disciplined dependence analysis should begin by defining the null hypothesis implicit in each measure. Pearson correlation tests for linear co-movement around the mean, rank statistics summarize monotonic association, copula tails focus on conditional extreme probabilities, mutual information assesses general statistical dependence and transfer entropy addresses incremental predictive information through time. Confusion arises when results are interpreted outside those domains. For example, a low Pearson coefficient cannot support a claim of independence, and a high mutual-information estimate cannot establish directional transmission.

Transformation invariance is another useful organizing principle. Rank measures and copulas are invariant to strictly increasing marginal transformations, while Pearson correlation is not. Mutual information is theoretically invariant to smooth invertible transformations, although empirical estimators may not be perfectly invariant in finite samples. These properties matter when variables have strongly different scales or when marginal transformations are used to stabilize variance.

Tail measures require a separate interpretation because they intentionally discard much of the center of the distribution. A pair with modest overall association can have economically important lower-tail dependence, and a pair with strong central correlation can have zero asymptotic tail dependence. Rather than viewing these findings as contradictory, the dashboard treats them as answers to different questions. The disagreement among metrics is often the most informative feature of the analysis.

Direction adds yet another dimension. Transfer entropy and Granger causality concern temporal prediction, not contemporaneous dependence. A system can exhibit strong same-period copula dependence with little directional predictability if shocks are absorbed simultaneously. Conversely, a lagged information channel can be important even when contemporaneous association is modest. Applied papers should therefore align time indexing carefully with the economic transmission mechanism they claim to study.

10. Model Selection and Reporting Standards

Because flexible dependence measures introduce tuning choices, transparent reporting is essential. Mutual-information studies should document estimator type, bandwidth or neighbour parameter and any bias correction. Transfer-entropy studies should report lag orders, embedding choices, surrogate design and multiple-testing control. Copula studies should report marginal models, family selection, tail coefficients and goodness-of-fit diagnostics. Without these details, reported dependence values are difficult to reproduce or compare.

Model selection should also be task specific. Information criteria evaluate global density fit, whereas stress applications may require a tail-weighted criterion. Forecasting applications require out-of-sample scores, and network applications require stability of edge selection. The simulation framework demonstrates that the model with the best center-of-distribution fit is not always the model that best identifies the feature relevant to the decision problem.

Researchers should avoid ranking dependence measures using a single universal performance criterion. A measure cannot be penalized for failing to detect a feature it was not designed to detect. Instead, performance should be assessed against design-specific targets: correlation recovery in linear systems, tail-probability calibration in extreme-value designs, dependence detection in nonlinear designs and directional-edge recovery in dynamic systems.

A minimum reporting set for general applied work could include Pearson correlation, Kendall tau, one nonlinear dependence measure and, where economically relevant, a tail or directional measure. This combination is not mandatory, but it provides an accessible diagnostic ladder from simple to more specialized dependence concepts.

11. Limitations and Future Research

The Monte Carlo designs necessarily simplify real data. Financial and demographic systems can contain nonstationarity, structural breaks, mixed discrete-continuous variables and measurement error. These features can interact with dependence estimators in ways not captured by the present experiments. Future simulation studies should incorporate such complications and evaluate adaptive or locally stationary methods.

High-dimensional dependence remains a major open problem. Pairwise dashboards can become unwieldy when hundreds of variables are analyzed. Graphical models, factor copulas, sparse mutual-information networks and multivariate information decompositions provide possible solutions, but they raise new questions about regularization and interpretability. A useful direction is to combine low-dimensional interpretable summaries with high-dimensional screening procedures.

Another promising area is uncertainty quantification. Dependence estimates are often reported as point values, yet nonlinear and tail measures can have substantial sampling uncertainty. Bootstrap intervals, Bayesian posterior distributions and stability-selection frequencies would make applied conclusions more credible. The ultimate objective is not to maximize the number of sophisticated measures but to know which dependence statements are supported robustly by the available information.

References

- Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7), 1705–1741.
- Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2), 182–198.
- Bedford, T., & Cooke, R. M. (2002). Vines—A new graphical model for dependent random variables. *Annals of Statistics*, 30(4), 1031–1068.
- Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas*. Springer.
- Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57–66.
- Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions. *Journal of Econometrics*, 182(1), 119–134.
- Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management. In M. Dempster (Ed.), *Risk Management: Value at Risk and Beyond*. Cambridge University Press.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3), 339–350.
- Joe, H. (2014). *Dependence Modeling with Copulas*. CRC Press.
- McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools* (2nd ed.). Princeton University Press.
- Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Springer.
- Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2), 527–556.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423, 623–656.
- Rényi, A. (1961). On measures of entropy and information. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1, 547–561.
- Tsallis, C. (1988). Possible generalization of Boltzmann–Gibbs statistics. *Journal of Statistical Physics*, 52, 479–487.
- Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.

- Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, 69(6), 066138.
- Maasoumi, E., & Racine, J. (2002). Entropy and predictability of stock market returns. *Journal of Econometrics*, 107(1–2), 291–312.
- Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters*, 103(23), 238701.

Declarations

Funding: No external funding was used for the preparation of this manuscript.

Conflict of interest: The authors declare no competing interests.

Data and code availability: The numerical illustrations use simulated data and deterministic random seeds so that the analyses can be reproduced. Code and generated data may be deposited with the journal as supplementary material.