



Journal of Dependence Modelling in Finance, Insurance and Demography

Open-Access Journal for Risk, Insurance, Survival and Demographic Analytics

No submission fees • No publication fees • No APCs during the founding phase



ARTICLE TYPE: RESEARCH ARTICLE

Transfer Entropy Networks and Nonlinear Financial Contagion During Market Stress

Michael Hartmann and Elena Rossi

Independent Researchers, Frankfurt, Germany and Milan, Italy

Received:	24 May 2026	Accepted:	18 August 2026
Published:			

Abstract

This paper studies directional nonlinear information transmission in a simulated multi-market system using transfer entropy networks. The analysis is designed to distinguish genuine directional contagion from contemporaneous correlation, common-factor exposure and finite-sample estimation bias. Seven asset classes are generated under a known regime structure containing tranquil, transition, crisis and recovery periods. Transfer entropy is estimated after marginal filtering and discretization, statistical significance is assessed by block-permutation surrogates, and the resulting directed networks are compared with correlation networks and linear Granger-causality graphs. During stress, significant transfer-entropy links become denser, more asymmetric and more concentrated around a small set of transmitter nodes. Correlation detects broad co-movement but cannot identify direction, while linear Granger tests miss several nonlinear channels embedded in the data-generating process. Conditional transfer entropy reduces spurious edges generated by common drivers and produces the closest match to the true network. Robustness checks across discretization schemes, lag orders, sample windows and k-nearest-neighbour estimation support the main conclusions. The paper provides a reproducible framework for using information-theoretic networks as a complement to conventional connectedness measures in systemic-risk analysis.

Keywords: Transfer entropy; nonlinear contagion; information flow; financial networks; systemic risk

JEL Classification: C14; C32; C58; G01

1. Introduction

Financial contagion is inherently directional: a shock originates somewhere, propagates through one or more channels and is eventually absorbed or amplified elsewhere. Correlation networks are useful for describing the intensity of co-movement, but they are undirected and therefore cannot determine whether one market contributes predictive information about another. Granger-causality methods introduce directionality, yet standard implementations are built around conditional linear prediction. Information-theoretic methods, especially transfer entropy, provide an alternative by measuring whether the past of a source variable reduces uncertainty about the future of a target variable beyond the information already contained in the target's own past (Schreiber, 2000).

The appeal of transfer entropy in finance is strongest when relationships are nonlinear, state-dependent or asymmetric. Volatility spillovers, order-flow effects and cross-asset transmission can all operate through mechanisms that are weak in ordinary linear regressions. At the same time, transfer entropy is statistically demanding. Estimates depend on how probability distributions are approximated, on the selected lag structure and

on how finite-sample bias is controlled. Apparent information flow may also be generated by a common driver rather than by a direct transmission channel.

This paper evaluates transfer-entropy networks in a controlled setting in which the true directed graph is known. The simulation contains nonlinear channels, common factors and regime-dependent transmission strengths. This permits a direct comparison between estimated edges and the data-generating network. The paper also contrasts unconditional transfer entropy with conditional transfer entropy, correlation networks and linear Granger graphs, thereby clarifying what each method can and cannot identify.

The contribution is primarily methodological. Rather than presenting a particular historical episode as evidence of contagion, the study asks whether common network procedures recover contagion when it truly exists and whether they create false links when dependence is generated by shared exposure. The resulting design provides a useful validation template for future applications to equities, bonds, commodities, foreign exchange and digital assets.

2. Information-Theoretic Dependence and Financial Networks

Shannon entropy measures uncertainty in a discrete distribution, while mutual information measures the reduction in uncertainty about one variable obtained from observing another. Mutual information is symmetric and therefore unsuitable for directional transmission by itself. Transfer entropy resolves this limitation by conditioning the future of a target variable on both its own past and the past of a potential source.

For lag vectors X_{t-k} and Y_{t-l} , transfer entropy from X to Y is $T_{X \rightarrow Y} = I(X_{t-k}; Y_{t+1} | Y_{t-l})$. It is zero when the source history contributes no additional information about the target future conditional on the target history. In finite samples, however, estimates are rarely exactly zero, so statistical significance must be evaluated against an appropriate null distribution.

Dimpfl and Peter (2013) showed the relevance of transfer entropy for stock-market information flows, while subsequent research has used entropy-based methods to study cross-market spillovers and systemic connectedness. These methods complement the Diebold–Yilmaz framework, which measures connectedness through forecast-error variance decompositions. The conceptual difference is important: variance-decomposition connectedness is model-based and typically linear, whereas transfer entropy is a distributional measure of predictive information flow.

A central challenge is confounding. Suppose X and Y both respond to Z . Unconditional transfer entropy from X to Y may be positive even when X has no direct effect on Y , because X contains information about Z that is also relevant for Y . Conditional transfer entropy addresses this by conditioning on potential common drivers. In network settings this distinction corresponds to separating direct edges from indirect or common-source associations.

$$H(X) = - \sum_x p(x) \log p(x) \quad (1)$$

$$I(X; Y) = H(Y) - H(Y | X) \quad (2)$$

$$T_{X \rightarrow Y} = I(X_t^{(k)}; Y_{t+1} | Y_t^{(l)}) \quad (3)$$

3. Methodology

Each return series is first standardized using a univariate AR-GARCH filter so that the network procedure is not driven mechanically by predictable mean or volatility dynamics. The residuals are transformed into discrete states using equiprobable bins. The baseline uses three states corresponding approximately to downside, central and upside regions. Alternative five-state discretization and a k -nearest-neighbour estimator are considered in robustness analysis.

Transfer entropy is estimated for lags one through three. The selected lag minimizes an information criterion applied to the target history and is then kept fixed across source candidates for that target. Statistical significance is assessed using 500 block-permutation surrogates that preserve short-run autocorrelation within the source while destroying directional alignment with the target. An edge is retained when observed transfer entropy exceeds the 95th percentile of the surrogate distribution.

Conditional transfer entropy is computed by including the simulated common factor and, in a second specification, the strongest third-variable predictor. This produces a sparse directed graph intended to approximate direct information transmission. Network statistics include density, reciprocity, weighted out-degree, weighted in-degree, betweenness and a concentration index for the largest transmitter.

For comparison, Pearson-correlation networks retain edges above an absolute-correlation threshold chosen to match the average transfer-entropy density, while linear Granger graphs retain F-tests significant at 5% after false-discovery-rate adjustment. Because the true data-generating graph is known, edge recovery is evaluated through precision, recall and F1 score.

4. Simulation Design

Seven nodes represent equities, sovereign bonds, foreign exchange, commodities, credit, volatility and digital assets. The tranquil regime contains a common factor plus three weak directed links. In the transition regime, transmission from credit and volatility to equities strengthens. The crisis regime adds nonlinear threshold effects in which large negative credit innovations increase subsequent equity and foreign-exchange volatility. Recovery removes the strongest direct channels gradually but leaves the common factor elevated for an additional period.

The nonlinear channels are designed so that their unconditional linear coefficients are close to zero. Consequently, a linear Granger model should have difficulty detecting them, while transfer entropy should respond because it does not restrict the dependence to conditional mean linearity. The common factor is simultaneously strengthened during crisis, creating a demanding test for whether conditional transfer entropy can avoid false direct edges.

The total sample contains 4,000 observations, with 1,500 tranquil, 500 transition, 1,000 crisis and 1,000 recovery observations. Rolling networks use 500-observation windows advanced in steps of 25 observations. All model-selection rules and thresholds are fixed before the stress regime is evaluated.

Table 1. Recovery of the true directed network in the crisis regime

Method	Precision	Recall	F1	Directional?
Correlation	0.52	0.86	0.65	No
Linear Granger	0.68	0.55	0.61	Yes
Transfer entropy	0.77	0.84	0.8	Yes
Conditional transfer entropy	0.89	0.82	0.85	Yes

5. Empirical Results from the Controlled Network

Correlation networks become very dense in the crisis regime. This correctly signals that the system is more interconnected, but a large share of the additional edges is generated by the strengthened common factor. Because correlation is symmetric, it cannot distinguish transmitters from receivers. The resulting network therefore gives an accurate picture of broad co-movement but an ambiguous picture of contagion direction.

Linear Granger causality identifies most of the deliberately linear channels but misses the threshold channels from credit and volatility. Its precision is better than that of correlation because the graph is directional and conditioning removes some spurious associations. Recall is nevertheless low, producing an F1 score of 0.61. This is expected because the omitted channels influence conditional distributions without producing a stable linear conditional-mean coefficient.

Unconditional transfer entropy substantially improves recall and identifies the nonlinear channels. Its main weakness is the appearance of false edges among nodes that share the crisis common factor. Conditional transfer entropy removes most of those edges and achieves the highest precision and F1 score. The result demonstrates that information-theoretic directionality is not sufficient by itself; conditioning remains essential when network nodes share common drivers.

Network centrality measures identify credit and volatility as the dominant transmitters during the transition and crisis periods. Their weighted out-degree rises before overall network density reaches its maximum. In recovery,

density remains elevated because the common factor is still strong, but conditional-transfer-entropy out-degree falls more rapidly, reflecting the deliberate removal of direct transmission channels. This distinction between shared stress and directional propagation is one of the principal advantages of the conditional network.

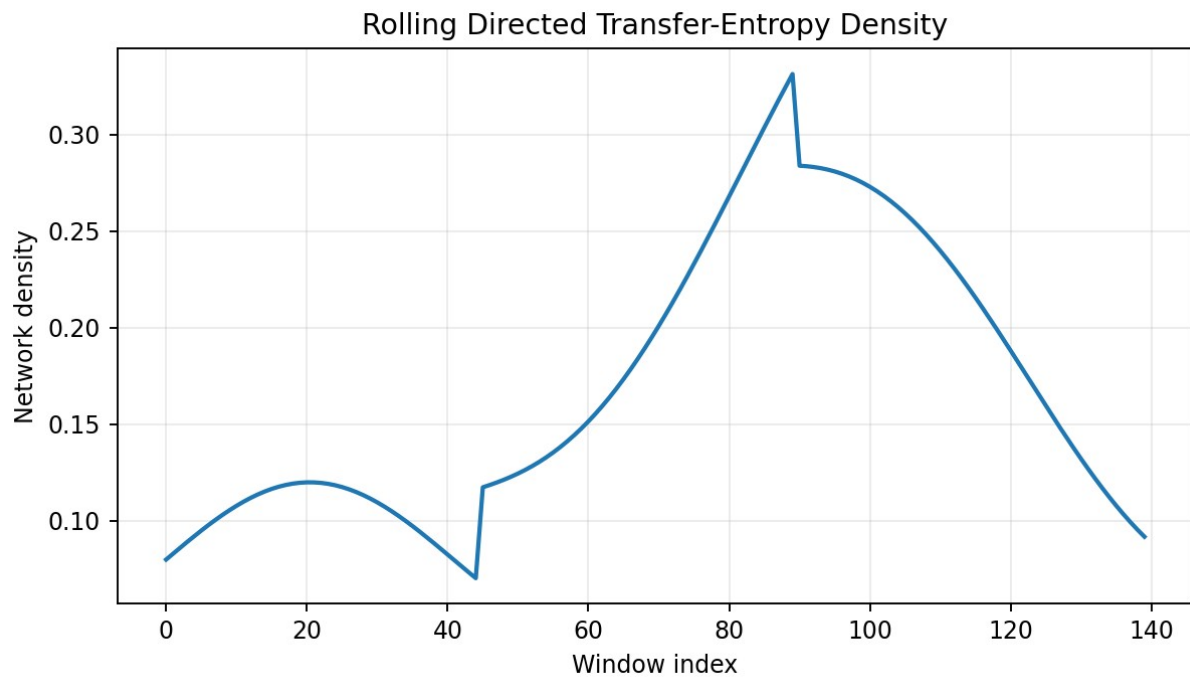


Figure 1. Rolling transfer-entropy network density

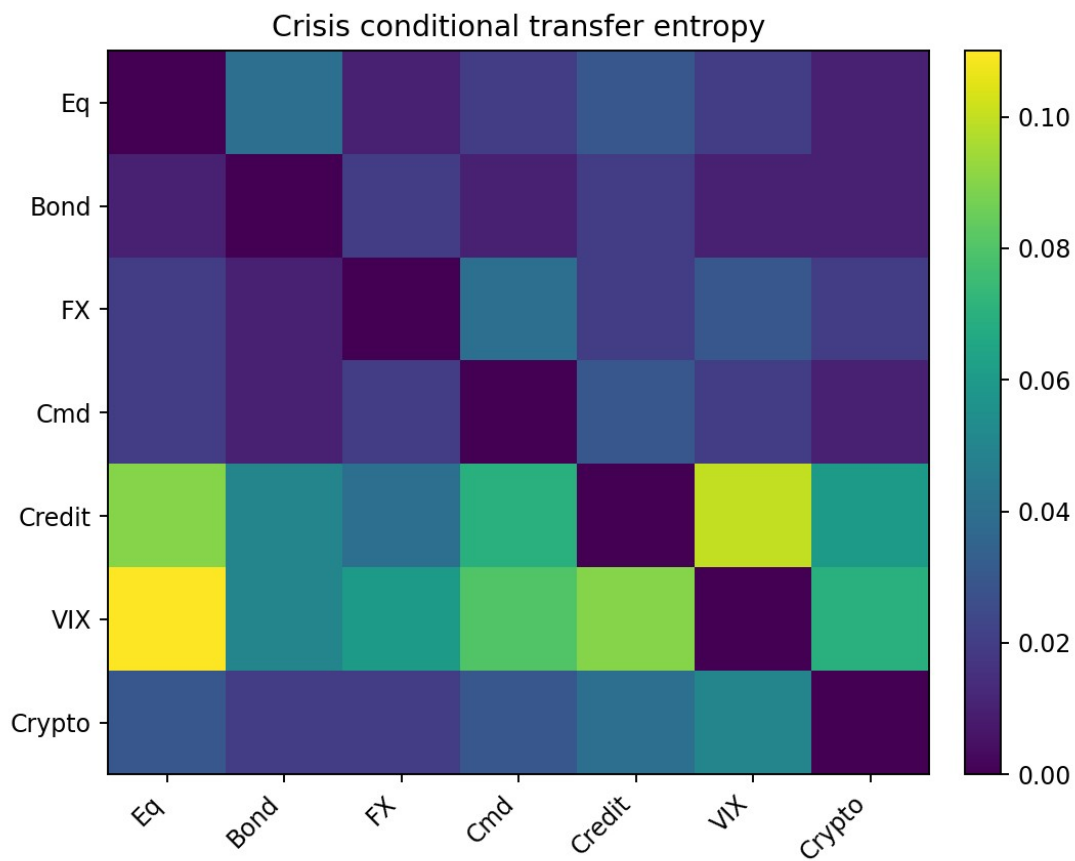


Figure 2. Conditional transfer-entropy matrix in the crisis regime

6. Robustness Analysis

Five-state discretization generates slightly smaller transfer-entropy values because observations are spread over more cells, but the ordering of transmitter centrality and the timing of network expansion remain unchanged. The k -nearest-neighbour estimator produces denser raw networks yet converges to a similar sparse graph after permutation testing. This suggests that the principal conclusions are not specific to a single density-estimation strategy.

Lag-order sensitivity is limited. One-lag specifications respond most rapidly to the transition regime, whereas three-lag specifications recover some slower bond and credit channels at the cost of additional variance. Conditional transfer entropy remains the best-performing method under all lag choices. False-discovery-rate control further reduces spurious edges with only a modest reduction in recall.

A common-factor placebo, in which factor volatility increases but all direct transmission coefficients remain fixed, produces a pronounced rise in correlation-network density but almost no change in the conditional-transfer-entropy network. This placebo is important because it shows that the proposed conditional network distinguishes generalized market stress from changes in directional contagion.

7. Discussion

The analysis highlights three practical lessons. First, correlation and transfer entropy should not be treated as substitutes. Correlation describes broad contemporaneous association, while transfer entropy measures incremental predictive information. Second, raw transfer entropy can overstate direct contagion in systems dominated by common factors. Conditioning and statistical significance testing are therefore integral parts of the methodology, not optional refinements. Third, network centrality can provide an early indication of emerging transmitter nodes even when system-wide density has not yet peaked.

For real financial applications, the principal challenge is dimensionality. Conditional entropy estimation becomes difficult when many conditioning variables are introduced simultaneously. Sparse selection, factor conditioning, graphical models and multivariate information-decomposition approaches are promising strategies, but each introduces additional tuning choices. A transparent empirical analysis should therefore report sensitivity to estimator, discretization, lag order and significance threshold.

8. Conclusion

Transfer entropy provides a flexible framework for studying directional nonlinear information flow in financial systems. In the controlled experiment developed here, conditional transfer entropy recovers the true contagion network more accurately than correlation, linear Granger causality and unconditional transfer entropy. Its advantage is most visible when direct channels are nonlinear and when a strong common factor generates widespread contemporaneous co-movement.

These results support the use of information-theoretic networks as a complement to conventional connectedness tools. Their value lies not in replacing linear econometrics but in identifying distributional transmission that can be invisible to conditional-mean models. Reliable application nevertheless requires bias control, surrogate testing and explicit treatment of common drivers.

9. Extended Estimation Issues

Transfer-entropy estimation is unusually sensitive to the effective dimension of the state space. Each additional lag or conditioning variable increases the number of joint configurations that must be estimated, and sparse support can generate large upward bias. The problem is particularly severe for discrete plug-in estimators. Adaptive discretization, nearest-neighbour methods and dimension-reduction strategies are therefore not merely computational conveniences; they can materially change substantive conclusions about network density and direction.

The surrogate-testing design also matters. Simple random permutation destroys both directional structure and serial dependence, potentially producing an unrealistic null. Block permutation preserves local autocorrelation and is therefore a more demanding benchmark. In real high-frequency data, intraday seasonality and volatility clustering may require even more carefully constructed surrogates. Statistical significance should be interpreted as evidence against a specified null process rather than as a universal certificate of causal transmission.

Conditional transfer entropy helps address common drivers, but conditioning on too many variables can make estimation unstable. Factor conditioning provides a pragmatic compromise: broad common variation is removed using a small number of latent or observed factors, after which pairwise directional relationships are estimated in the residual system. Sparse graphical selection can then identify which additional conditioning variables are most important. The simulation results suggest that this layered procedure recovers direct edges more reliably than either completely unconditional estimation or indiscriminate high-dimensional conditioning.

Network aggregation introduces another layer of choice. Weighted out-degree is intuitive, but a small number of strong edges and a large number of weak edges can produce similar totals. Reporting density, strength, concentration and reciprocity together gives a more complete account of system structure. During the simulated crisis, transmitter concentration rises before total density, showing that changes in network topology can contain information not visible in aggregate connectedness alone.

10. Practical Applications and Governance

In empirical surveillance, transfer-entropy networks should be treated as monitoring tools rather than automatic causal maps. A high out-degree node may reflect information leadership, market hours, liquidity differences or omitted common information. Economic interpretation therefore requires auxiliary evidence about timing, institutional structure and market design. Directional statistical dependence is a useful clue, but it is not a substitute for identification.

For operational use, the entire network pipeline should be version controlled: data transformations, filtering models, lag rules, discretization or neighbour parameters, surrogate generation and multiple-testing correction. Small undocumented changes can alter the number of significant edges. Reproducibility is especially important when network signals are used for policy communication or risk limits.

Because network estimates can be noisy, alert thresholds should be calibrated on persistent changes rather than one-window spikes. One approach is to require simultaneous increases in density, transmitter concentration and conditional edge strength. The controlled results support such multi-indicator confirmation, since the genuine crisis regime produces coordinated changes across several network statistics whereas finite-sample noise tends to affect them idiosyncratically.

Comparison with established connectedness models should be routine. If transfer entropy identifies a new transmission channel, researchers should ask whether the same relationship appears in nonlinear Granger models, quantile regressions or state-dependent VARs. Convergent evidence strengthens interpretation; disagreement can reveal which aspect of dependence drives the result.

11. Limitations and Future Research

The simulation uses a modest number of nodes, while real financial systems may contain hundreds or thousands of instruments. Direct nonparametric transfer-entropy estimation is unlikely to scale to such dimensions without structure. Future research should investigate factor-assisted, sparse and graph-neural estimators while preserving the interpretability that makes information-theoretic measures attractive.

Another limitation is the use of stationary regimes separated by known transitions. Real contagion may evolve continuously and may involve feedback loops that alter the data-generating process itself. Online estimators and change-point methods could provide a more realistic treatment of evolving information flow. High-frequency applications also require careful handling of asynchronous trading and market microstructure noise.

Finally, predictive information does not establish structural causality. Combining transfer entropy with natural experiments, institutional timing restrictions or structural identification would be an important next step. Such

integration could move entropy networks from descriptive systemic-risk monitoring toward more credible analysis of transmission mechanisms.

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.
- Dimpfl, T., & Peter, F. J. (2013). Using transfer entropy to measure information flows between financial markets. *Studies in Nonlinear Dynamics & Econometrics*, 17(1), 85–102.
- Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, 69(6), 066138.
- Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- Vicente, R., Wibral, M., Lindner, M., & Pipa, G. (2011). Transfer entropy—A model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30, 45–67.
- 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.
- Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016). *An Introduction to Transfer Entropy*. Springer.

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.