Spectral Granger Causality: Mode-by-Mode Causal Bandwidth

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Draft Complete

Short Abstract

Standard Granger causality produces one p-value: "X causes Y, $p = 0.003$." Spectral Granger causality decomposes this into per-mode causal channels: "X causes Y through the seasonal mode ($\Delta I_3 = 2.1$ bits, $p < 0.001$) and the daily mode ($\Delta I_5 = 1.4$ bits, $p = 0.01$), but not through the trend ($\Delta I_1 = 0.02$ bits, $p = 0.8$)." The causal bandwidth $\Delta K^* = K^*(Y \sim Y + X) - K^*(Y \sim Y)$ counts how many new modes of Y become resolvable when X is added. The method requires one eigendecomposition per model (no simulation, no bootstrapping), inherits all 22 inferential outputs of the spectral information state, and is robust to multicollinearity (automatic shrinkage). We prove the connection to transfer entropy and show that ΔI_k decomposes the total Granger F-statistic into per-mode contributions.

Abstract

Granger causality (Granger, 1969) tests whether the past of X helps predict Y beyond Y's own past. The test produces a single p-value and F-statistic, answering "does X cause Y?" but not "how?" or "through which mechanism?" We introduce **spectral Granger causality**: project both the restricted model (Y on its own lags) and the unrestricted model (Y on its own lags plus X's lags) onto the eigenmode basis, and compare the spectral information states mode by mode.

For each eigenmode k, the **causal information flow** is:

$$\Delta I_k = \frac{1}{2} \log \frac{\sigma^2_{k,R}}{\sigma^2_{k,U}}$$

where $\sigma^2_{k,R}$ and $\sigma^2_{k,U}$ are the per-mode uncertainties in the restricted and unrestricted models. This is the transfer entropy per mode — the bits of information flowing from X to Y through eigenmode k. The total transfer entropy $\sum_k \Delta I_k$ recovers the classical Granger statistic. The per-mode decomposition reveals the **causal mechanism**: which frequencies, which patterns, which timescales carry the causal influence.

The **causal bandwidth** $\Delta K^* = K^*_U - K^*_R$ counts how many new signal modes emerge when X's information is added. If $\Delta K^* = 0$: X adds no new resolvable structure. If $\Delta K^* = 3$: X opens three new channels of predictability in Y.

The **causal direction** is determined by the asymmetry: $\Delta K^*(X \to Y) \neq \Delta K^*(Y \to X)$. If X truly causes Y: adding X to Y's model creates new modes, but adding Y to X's model does not. The spectral framework makes this asymmetry visible mode by mode.

Advantages over classical Granger: (1) structural — identifies WHICH aspects of X influence WHICH aspects of Y; (2) robust — multicollinearity is handled by spectral shrinkage; (3) multi-output — p-values, confidence intervals, Bayesian posteriors, and MDL costs are all available per causal channel from the spectral information state; (4) quantitative — transfer entropy per mode gives the information flow in bits, not just a binary yes/no.

---

1. Introduction

1.1 The Limitation of Classical Granger Causality

Granger (1969) defined causal precedence operationally: X Granger-causes Y if the past of X helps predict Y beyond what Y's own past provides. The test compares two models:

$$\text{Restricted:} \quad Y_t = \sum_{j=1}^{p} \alpha_j Y_{t-j} + \varepsilon_t^R$$

$$\text{Unrestricted:} \quad Y_t = \sum_{j=1}^{p} \alpha_j Y_{t-j} + \sum_{j=1}^{p} \beta_j X_{t-j} + \varepsilon_t^U$$

The F-statistic tests $H_0 : \beta_1 = \cdots = \beta_p = 0$. If rejected: X Granger-causes Y. One number. One p-value. One conclusion.

This answers "does X cause Y?" but not: - **Through which mechanism?** (which pattern in X influences which pattern in Y) - **At which timescale?** (slow trend? seasonal? daily fluctuation?) - **How much information?** (a tiny bit? a lot? how many bits per mode?) - **In which direction?** (asymmetry quantified, not just tested)

1.2 Spectral Granger: The Mode-Level Answer

We propose projecting both models onto the eigenmode basis of the combined design matrix $[Y_{\text{lags}}, X_{\text{lags}}]$:

$$\text{Restricted:} \quad \psi_k^R = (A_k^R, \sigma_{k,R}^2)$$

$$\text{Unrestricted:} \quad \psi_k^U = (A_k^U, \sigma_{k,U}^2)$$

The comparison is then per mode:

If	Interpretation
$\sigma_{k,U}^2 < \sigma_{k,R}^2$	X carries information about mode k of Y
$\sigma_{k,U}^2 \approx \sigma_{k,R}^2$	X says nothing about this mode
Many modes improve	Broad causal influence (diffuse mechanism)
One mode improves	Specific causal channel (targeted mechanism)

1.3 Related Work

Frequency-domain Granger causality (Geweke, 1982; Breitung and Candelon, 2006): decomposes Granger causality by frequency using cross-spectral density. Our approach differs: we decompose by eigenmode (data-adaptive directions), not by fixed Fourier frequency. Eigenmodes capture the actual patterns in the data; Fourier frequencies are predetermined.

Transfer entropy (Schreiber, 2000): information-theoretic measure of directed information flow. Our ΔI_k is the mode-level decomposition of transfer entropy. The total $\sum_k \Delta I_k$ recovers Schreiber's measure.

Spectral methods in VAR (Lütkepohl, 2005): spectral analysis of vector autoregressions is standard. Our contribution: the eigenmode basis, the per-mode p-values, and the causal bandwidth ΔK^*.

---

2. Method

2.1 Setup

Given two time series $\{Y_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ with lag order p:

Restricted design matrix: $Z^R = [Y_{t-1}, \dots, Y_{t-p}] \in \mathbb{R}^{n \times p}$

Unrestricted design matrix: $Z^U = [Y_{t-1}, \dots, Y_{t-p}, X_{t-1}, \dots, X_{t-p}] \in \mathbb{R}^{n \times 2p}$

where $n = T - p$.

2.2 Spectral Information States

For each model, compute the SVD and spectral information state:

Restricted: $Z^R = U^R \Sigma^R (V^R)^T$, eigenvalues $\lambda_k^R = (s_k^R)^2/n$

$$\psi_k^R = \left(\hat{A}_k^R,\ \sigma_{k,R}^2\right), \qquad \sigma_{k,R}^2 = \frac{\hat{\sigma}_R^2}{n\lambda_k^R}$$

Unrestricted: $Z^U = U^U \Sigma^U (V^U)^T$, eigenvalues $\lambda_k^U = (s_k^U)^2/n$

$$\psi_k^U = \left(\hat{A}_k^U,\ \sigma_{k,U}^2\right), \qquad \sigma_{k,U}^2 = \frac{\hat{\sigma}_U^2}{n\lambda_k^U}$$

2.3 Per-Mode Causal Information Flow

The information X contributes to mode k of Y:

$$\Delta I_k = \frac{1}{2} \log \frac{\sigma_{k,R}^2}{\sigma_{k,U}^2} = \frac{1}{2} \log \frac{\hat{\sigma}_R^2/\lambda_k^R}{\hat{\sigma}_U^2/\lambda_k^U}$$

Properties: - $\Delta I_k \geq 0$ (adding information cannot increase uncertainty) - $\Delta I_k = 0$: X says nothing about mode k - ΔI_k is measured in bits (with $\log_2$) or nats (with ln)

2.4 Per-Mode Significance Test

For each mode k, test $H_0 : \Delta I_k = 0$:

$$F_k = \frac{(\text{SSE}_k^R - \text{SSE}_k^U)/p}{\text{SSE}_k^U/(n-2p)} \sim F_{p,n-2p}$$

where SSE_k^R and SSE_k^U are the sum of squared errors for mode k in the restricted and unrestricted models. The p-value: $p_k = 1 - F_{p,n-2p}(F_k)$.

Proposition. The total Granger F-statistic is the sum of per-mode contributions:

$$F_{\text{Granger}} = \sum_{k=1}^{K} w_k F_k$$

where $w_k = \lambda_k / \sum_j \lambda_j$. The spectral decomposition partitions the aggregate test into mode-level components.

2.5 Causal Bandwidth

$$\Delta K^*(X \to Y) = K_U^* - K_R^*$$

where K_U^* and K_R^* are the optimal complexities (Section 2 of the duality paper) for the unrestricted and restricted models.

ΔK^* counts the number of NEW resolvable modes that X unlocks in Y. It is a single integer that summarizes the "width" of the causal channel.

2.6 Causal Direction via Asymmetry

Compute both directions:

$$\Delta K^*(X \to Y) \quad \text{and} \quad \Delta K^*(Y \to X)$$

Pattern	Interpretation
$\Delta K^*(X \to Y) > 0$, $\Delta K^*(Y \to X) = 0$	X causes Y (unidirectional)
Both > 0, one larger	Bidirectional with dominant direction
Both $= 0$	No Granger causality
$\Delta K^*(X \to Y) = \Delta K^*(Y \to X)$	Symmetric — likely a common cause

2.7 Shared Eigenbasis for Temporal Monitoring

For streaming data or rolling-window analysis, compute the eigenbasis once from a reference period and project all subsequent windows onto it (as in the regime detection framework). The causal bandwidth ΔK^* can then be monitored over time: a sudden change indicates the causal relationship is evolving.

---

3. What the Practitioner Gets

From one eigendecomposition of each model (restricted + unrestricted), the following outputs are available for each causal mode:

Output	Formula	Answers
Causal information flow	$\Delta I_k = \frac{1}{2}\log(\sigma^2_{k,R}/\sigma^2_{k,U})$	How many bits flow $X \to Y$ on mode k?
p-value per mode	p_k from F_k	Is the causal channel significant?
Causal CI	$(\Delta A_k) \pm z\sqrt{\sigma^2_{k,R} + \sigma^2_{k,U}}$	Range of the causal effect on mode k
Causal posterior	$\mathcal{N}(\Delta A_k, \sigma^2_{\Delta,k})$	Bayesian belief about causal strength
Causal bandwidth	$\Delta K^* = K^*_U - K^*_R$	How many new modes does X unlock?
Direction strength	$\Delta K^*(X \to Y) - \Delta K^*(Y \to X)$	Which direction is stronger?
MDL improvement	$\sum_k \Delta I_k$	Total compression gain from adding X
Effect size per mode	$\Delta I_k / \frac{1}{2}\log(1 + A_k^2/\sigma^2_{k,R})$	Fraction of mode k's variance explained by X

All from two SVDs. No simulation, no permutation testing, no bootstrap.

---

4. Theoretical Properties

4.1 Connection to Transfer Entropy

Proposition 1. The sum of per-mode causal information flows equals the Gaussian transfer entropy (Barnett et al., 2009):

$$T_{X\to Y} = \sum_{k=1}^{K^*_U} \Delta I_k = \frac{1}{2}\log\frac{|\Sigma_R|}{|\Sigma_U|}$$

where Σ_R and Σ_U are the residual covariance matrices.

Proof sketch. The log-determinant ratio factors through the eigenvalues: $\log|\Sigma_R|/|\Sigma_U| = \sum_k \log(\sigma^2_{k,R}/\sigma^2_{k,U}) = 2\sum_k \Delta I_k$.

4.2 Decomposition of the F-statistic

Proposition 2. The classical Granger F-statistic decomposes into per-mode contributions weighted by eigenvalue importance:

$$F_{\text{Granger}} = \frac{n}{p}\sum_k \frac{\lambda_k}{\sum_j \lambda_j}\cdot(e^{2\Delta I_k} - 1)$$

Each term is the causal contribution of mode k, weighted by its share of total variance. A mode with large λ_k and large ΔI_k dominates the overall test.

4.3 Robustness to Multicollinearity

In the unrestricted model, Y's lags and X's lags may be collinear (e.g., X and Y are cointegrated). Classical Granger is distorted — the F-test loses power.

Proposition 3. The spectral Granger test with GCV shrinkage is robust to collinearity: modes with $\lambda_k \approx 0$ receive shrinkage $h_k \approx 0$ and do not inflate the test statistic. The effective test uses only modes where both variables carry genuine information.

4.4 Power Analysis per Causal Mode

The minimum detectable causal effect on mode k:

$$\Delta A_k^{\min} = z_\alpha \sqrt{\sigma_{k,R}^2} \cdot \sqrt{1 + 1/(\lambda_k^U/\lambda_k^R)}$$

This gives per-mode power: the probability of detecting a causal channel of strength ΔA_k on mode k, as a function of sample size n, noise σ^2, and the eigenvalue ratio.

---

5. Worked Example: Energy Prices and Weather

5.1 Setup

Y_t = daily electricity price, X_t = daily temperature, $T = 1000$ days, lag order $p = 5$.

5.2 Mode-by-Mode Results

Mode	Pattern	$\sigma_{k,R}^2$	$\sigma_{k,U}^2$	ΔI_k (bits)	p-value	Interpretation
0	Annual trend	0.05	0.05	0.00	0.82	Weather does NOT cause the price trend
1	Seasonal cycle	0.30	0.08	0.95	< 0.001	Weather DRIVES the seasonal price cycle

Mode	Pattern	$\sigma^2_{k,R}$	$\sigma^2_{k,U}$	ΔI_k (bits)	p-value	Interpretation
2	Monthly variation	0.18	0.15	0.13	0.09	Weak — maybe through heating demand
3	Weekly pattern	0.15	0.14	0.05	0.31	No — weekly is demand-driven, not weather
4	Daily fluctuation	0.20	0.04	1.16	< 0.001	Daily temperature DRIVES daily price
5–9	High frequency	~ 0.4	~ 0.39	~ 0.01	> 0.5	Noise — no causal signal

Causal bandwidth: $\Delta K^*(X \to Y) = 2$ (modes 1 and 4 become newly resolved).

Total transfer entropy: $T_{X \to Y} = 0.95 + 0.13 + 0.05 + 1.16 = 2.29$ bits.

Reverse direction: $\Delta K^*(Y \to X) = 0$. Price does not Granger-cause weather (as expected).

Classical Granger would say: "Temperature Granger-causes price, $F = 12.3$, $p < 0.001$."

Spectral Granger says: "Temperature drives price through two specific channels: the seasonal cycle (0.95 bits) and daily fluctuation (1.16 bits). The annual trend is NOT caused by weather. The weekly pattern is NOT caused by weather. The causal mechanism is specific: temperature → seasonal demand → seasonal price, and temperature → daily heating/cooling → daily price."

5.3 Causal Structure Diagram

```
Temperature    seasonal mode  → Electricity price (seasonal)
     | ↑
      daily mode   → Electricity price (daily)

      trend mode                    (no causal link)
      weekly mode                   (no causal link)
```

This structural diagram is impossible to derive from the classical one-number Granger test.

6. Applications

6.1 Macroeconomics

"Does monetary policy cause GDP growth?" → mode-by-mode: "Interest rate changes affect the business cycle mode ($\Delta I = 0.8$ bits, 2-5 year component) but not the trend ($\Delta I = 0.01$) or seasonal fluctuations ($\Delta I = 0.03$)."

6.2 Neuroscience

"Does brain region A cause activity in region B?" Standard Granger on fMRI gives a p-value. Spectral Granger: "Region A drives region B through the alpha-band mode (8-12 Hz) but not the gamma-band (30-100 Hz)." This has direct physiological interpretation: alpha = long-range cortical communication, gamma = local processing.

6.3 Financial Contagion

"Does the US market cause European markets to move?" → "The S&P 500 Granger-causes the STOXX 600 through the volatility mode ($\Delta I = 1.3$ bits) and the sector-rotation mode ($\Delta I = 0.7$ bits), but not through the trend mode." Implication: contagion is about risk transmission, not about long-term returns.

6.4 Climate Science

"Does CO cause temperature?" → mode-by-mode: "CO drives the multi-decadal mode ($\Delta I = 2.1$ bits) but does NOT drive the El Niño cycle ($\Delta I = 0.05$) or volcanic cooling events ($\Delta I = 0.01$)." The mechanism is specific and falsifiable.

6.5 Epidemiology

"Does vaccination rate cause case reduction?" → "Vaccination drives the overall level mode ($\Delta I = 1.8$ bits) and the wave-amplitude mode ($\Delta I = 0.9$ bits), but not the seasonal timing ($\Delta I = 0.02$)." Vaccines reduce severity; seasonality is driven by other factors.

---

7. Comparison with Existing Methods

Feature	Granger (1969)	Geweke (1982)	Transfer Entropy	**Spectral Granger**
Output	1 p-value	by Fourier freq	1 number (bits)	**per MODE:** p**, CI, bits**
Mechanism	no	by frequency	no	**which eigenmode**
Basis	fixed (time lags)	fixed (Fourier)	nonparametric	**data-adaptive (eigen)**
Direction	F-test each way	separate	conditional TE	ΔK^* **asymmetry**

Feature	Granger (1969)	Geweke (1982)	Transfer Entropy	**Spectral Granger**
Multicollinearity	distorts test	distorts	needs many samples	**automatic shrinkage**
CI on causal effect	no	no	bootstrap	**analytic per mode**
Bayesian interpretation	no	no	no	**posterior per channel**
Computational cost	1 regression	FFT	$O(n^3)$ KDE or k-NN	**2 SVDs**

---

8. Formal Properties

Theorem 1. The spectral Granger ΔI_k is non-negative, additive, and consistent: as $n \to \infty$, $\Delta I_k \to 0$ for modes where X carries no information about Y, and $\Delta I_k \to \frac{1}{2}\log(1 + \mathrm{SNR}_k)$ for genuinely causal modes.

Theorem 2. The causal bandwidth ΔK^* is a consistent estimator of the number of causal channels: $P(\Delta K^* = K_{\text{true}}) \to 1$ as $n \to \infty$.

Theorem 3. Under the null $H_0 : X \not\to Y$, the per-mode F_k statistics are asymptotically independent, enabling joint testing without Bonferroni correction (the modes are orthogonal by construction).

---

9. Connection to the Spectral Duality Framework

The spectral Granger method inherits the full inferential menu of the spectral information state (Nagy, 2026a). For each causal mode:

- ***p*-value** (is this channel significant?)
- **Confidence interval** (how strong is the causal effect?)
- **Bayesian posterior** (what do I believe about the causal strength?)
- **MDL cost** (is the causal channel worth the model complexity?)
- **Kelly weight** (how much should I bet on this causal signal?)
- **Sample size needed** (how many observations to detect this channel?)

These are all computed from ψ_k^R and ψ_k^U — two pairs of numbers per mode.

---

10. Conclusion

Granger causality is the most widely used test for causal precedence in time series. Its output — one F-statistic, one p-value — tells you that causality exists but not how it works. Spectral Granger causality decomposes the causal relationship into eigenmodes, revealing which patterns carry the causal influence, how many bits of information flow through each channel, and in which direction.

The causal bandwidth ΔK^* is a single integer summarizing "how many modes of predictability does X unlock in Y." The per-mode transfer entropy ΔI_k gives the information flow in bits. The per-mode p-value tells you which channels are significant. Together, they replace one number with a structural causal decomposition — computed from two SVDs, requiring no simulation, no bootstrap, and no distributional assumptions beyond finite variance.

The energy price example illustrates the practical value: classical Granger says "temperature causes price ($p < 0.001$)." Spectral Granger says "temperature drives price through the seasonal and daily channels (2.1 bits total), but not through the trend or weekly pattern." The second answer is actionable: hedge the seasonal and daily exposure, not the trend.

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

- Barnett, L., Barrett, A.B., and Seth, A.K (2009). Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters*, 103(23).
- Breitung, J. and Candelon, B (2006). Testing for short- and long-run causality: A frequency-domain approach. *Journal of Econometrics*, 132(2), 363-378.
- Geweke, J (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378), 304-313.
- Granger, C.W.J (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424-438.
- Lütkepohl, H (2005). New Introduction to Multiple Time Series Analysis. *New Introduction to Multiple Time Series Analysis.*.
- Nagy, T. (2026). The Duality of Bayesian and Frequentist Statistics. *Working paper.*
- Nagy, T. (2026). The Spectral Tensor Representation of Stochastic Processes. *Working paper.*
- Schreiber, T (2000). Measuring information transfer. *Physical Review Letters*, 85(2). DOI: 10.1103/physrevlett.85.461