# The Geometry of Hidden-Variable Detectability: A Quartic Detection Theory

Tutorial Notes

Yuda Bi[1] and Vince D. Calhoun[1]

[1] *Georgia Institute of Technology, Atlanta, GA 30332, USA*

April 2026

**Abstract**

These notes provide a self-contained mathematical derivation of the quartic detection theory for hidden variables on statistical manifolds. We start from the Kullback–Leibler divergence and its local quadratic structure, derive the quartic detection law $D^{\min}_{\mathrm{KL,loc}}(\lambda) = C\lambda^4 + O(\lambda^6)$ as a geometric consequence of tangent-space absorption on reduced model manifolds, then specialize to two physical systems: the random-effects model (proving the pairing principle) and spectral inference under hidden persistent forcing (proving the existence of a spectrally dark regime at timescale coalescence). We then show how cross-spectral structure breaks the single-channel impossibility and certifies entropy production. Every derivation is presented step by step with full physical and mathematical commentary.

These notes accompany:

(i) *Why Single Probes Cannot Detect Hidden Forcing: A Quartic Detection Law* (PRL, under review)

(ii) *Timescale Coalescence Makes Hidden Persistent Forcing Spectrally Dark* (PRE, under review)

(iii) *Cross Spectra Break the Single-Channel Impossibility* (PRL, under review)

(iv) *Conditioning on a Volatility Proxy Compresses the Apparent Timescale of Collective Market Correlation* (PRE, under review)

## Contents

## Prerequisites and reading guide

**What this tutorial is.** These notes are a self-contained, book-style exposition of the quartic detection theory developed in the four companion papers listed above. Every theorem is proved in full (no "it can be shown that…"), and each derivation is accompanied by physical and mathematical commentary explaining *why* each step is taken.

**What background is assumed.** The reader should be comfortable with:

- *Linear algebra:* eigenvalues, matrix determinants, trace, rank-one updates.
- *Probability and statistics:* Gaussian distributions, expectation, variance, covariance matrices, maximum likelihood.
- *Calculus:* Taylor expansion, partial differentiation, integration.

If any of these topics feel rusty, **read Appendix A first**—it provides self-contained introductions to every mathematical tool used in the tutorial, with worked examples and pointers

to where each tool appears in the main text. More advanced prerequisites (contour integration, spectral densities, Lyapunov equations) are covered in Appendix A and are referenced inline when they first arise.

**How to read this tutorial.**

- *Sections 1–2* set up the problem and the mathematical language. Read these sequentially.
- *Section 3* derives the general quartic detection law—the central result. This is the theoretical core.
- *Sections 4–5* apply the general law to two concrete systems: the random-effects model (discrete, multi-sensor) and spectral inference (time series). These can be read in either order.
- *Section 6* shows how cross-spectral structure breaks the single-channel impossibility. It builds on Section 5.
- *Section 7* connects cross-spectral detection to thermodynamics. It requires Section 6.
- *Section 8* is a standalone classification result that can be read independently after Section 3.
- *Section 10* is a detailed worked case study applying the full framework to a financial scenario with concrete numbers. Read this after Sections 6–7 for a hands-on walkthrough.
- *Section 9* provides pointers to the broader literature for further reading.
- *Section 12* contains exercises for each major section—work through these to test your understanding.

## Notation

| Symbol | Meaning |
|---|---|
| $D_{\mathrm{KL}}(p\|q)$ | Kullback–Leibler divergence from $p$ to $q$ |
| $D^{\min}_{\mathrm{KL,loc}}(\lambda)$ | Minimum KL divergence from the true model (at coupling $\lambda$) to the best-fit null model |
| $\lambda$ | Coupling strength of the hidden variable |
| $\lambda_c^{\mathrm{pop}}(N)$ | Detection boundary: smallest $\lambda$ detectable with $N$ observations |
| $h$ | Perturbation function: the leading-order effect of the hidden variable on the observed distribution |
| $\mathcal{T}$ | Tangent space of the null model at the reference point |
| $\Pi_{\mathcal{T}}$ | Orthogonal projection onto $\mathcal{T}$ |
| $R$ | Normal residual: $R = (I - \Pi_{\mathcal{T}})\, h$, the unabsorbed part of the perturbation |
| $C$ | Quartic coefficient: $C = \frac{1}{2}\|R\|^2$ (density space) or $C = \frac{1}{4}\|R\|^2$ (Whittle spectral) |
| $G_{ij}$ | Fisher information matrix |
| $e_j$ | Score functions (tangent basis vectors) |
| $\mathcal{P}_c(\omega)$ | Pole polynomial: $\|1 - ce^{-i\omega}\|^2 = 1 + c^2 - 2c\cos\omega$ |
| $S_{\mathrm{null}}$, $S_{\mathrm{true}}$, $S_0$ | Null, true, and population-level power spectral densities |
| $\mathbf{1}$ | Vector of all ones |
| $n$ | Number of sensors (or observed channels) |
| $N$ | Number of independent observations (sample size) |

# 1 Background: why hidden variables matter

## 1.1 The attribution problem in complex systems

Slow collective observables—persistent fluctuations in brain connectivity, climate variability, financial correlation—are routinely interpreted as evidence for intrinsic memory, metastability, or critical slowing down. The same empirical appearance can, however, arise when a moderately relaxing observable tracks an equilibrium that is displaced by a persistent external or latent field. Distinguishing these two mechanisms is difficult whenever the field cannot be manipulated experimentally, which is precisely the situation in ecological synchrony, dynamic brain connectivity, and financial markets.

The first logical task is therefore *attribution*: before positing intrinsic memory, one should ask whether the apparent persistence is inherited from the motion of a driver. That question is sharper than a generic model-selection exercise. A slow observable can often be fit acceptably by several low-dimensional descriptions—bare mean reversion, multistability, hidden-variable models, and field-coupled dynamics—so good in-sample fit alone does not establish mechanism. What is needed is a sequence of discriminations that asks which parts of the observed persistence survive conditioning, placebo substitution, and model-free controls.

## 1.2 A concrete example: financial correlation dynamics

As a running motivating example, consider financial markets. One way to summarize how correlated 237 stocks are on a given day is to compute the largest eigenvalue of their correlation matrix—call it $\psi_1(t)$. (When $\psi_1$ is large, stocks move together; when it is small, they move independently.) Plotting $\psi_1(t)$ over 2004–2023 reveals slow fluctuations with apparent persistence. One interpretation is that the stock market has "intrinsic memory"—it takes a long time to relax after a shock. But an alternative interpretation is that $\psi_1$ is simply tracking the VIX volatility index, an external driver, and inheriting the VIX's own persistence.

Paper (iv) in this series tests this alternative: a VIX-coupled mean-reverting model reduces the effective relaxation time of $\psi_1$ from 298 to 61 trading days and improves the statistical fit by $\Delta\text{BIC} = 109$ (BIC is a model-comparison criterion that penalizes model complexity; see Eq. 2 below—a $\Delta$BIC of 109 is overwhelming evidence). Placebo drivers with matching autocorrelation but no physical relationship to the market fail ($\Delta\text{BIC}_{\max} = 2.7$). The broader point: the "slow dynamics" of the correlation observable may be largely inherited from the motion of VIX, not from intrinsic memory.

## 1.3 From attribution to detectability theory

The preceding example raises a more fundamental question. Even if a hidden field is present, *how much data is needed to detect it?* The naive statistical answer—$N \sim \lambda^{-2}$ observations (up to logarithmic factors) for a coupling of strength $\lambda$—turns out to be wrong by orders of magnitude. The remainder of these notes develops the exact theory of hidden-variable detectability, proving that the correct scaling is $N \sim \lambda^{-4}$ up to logarithmic corrections (the **quartic detection law**; see Eq. 11 for the precise form), that single probes are provably

blind (the **pairing principle**), and that cross-spectral structure provides a constructive escape from the impossibility.

## 1.4 Prior work and our contribution

The $n^{-1/4}$ detection rate for over-specified mixture models was first established by Chen [1]; the non-standard null distribution for boundary variance-component testing was derived by Self and Liang [2]. Both results treat specific statistical problems and provide asymptotic distributions or convergence rates.

What the present work adds is:

(a) *Exact coefficients $C$.* We compute the quartic coefficient $C$ in closed form for every system studied—random effects ($C = \frac{1}{2}\|R\|^2$ in density space), spectral and cross-spectral inference ($C = \frac{1}{4}\|R\|^2$ in the Whittle convention; see Remark 2.2), and Poisson count data—not just the rate but the prefactor.

(b) *Pairing principle combinatorics.* The sensor-count dependence $C = n(n-1)/(4\sigma^4) = \binom{n}{2}/(2\sigma^4)$ reveals that detection power grows with the number of sensor *pairs*, not individual sensors.

(c) *Darkness classification.* The tangent-space absorption criterion $h \in \mathcal{T}$ provides a geometric necessary and sufficient condition for exact undetectability, unifying the Gaussian single-probe impossibility, timescale coalescence, and the dimension-gap principle.

(d) *Cross-spectral escape.* The off-diagonal block of the spectral matrix is orthogonal to all diagonal tangent directions, providing a constructive escape from single-channel darkness.

(e) *Thermodynamic bridge.* The cross-spectral quartic coefficient is linked quantitatively to the entropy production rate under one-way coupling.

## 1.5 Scope of the quartic law

It is important to state the scope precisely. The standard detection scaling $\lambda_c \propto N^{-1/2}$ is correct for non-degenerate problems where the perturbation is not absorbed by model refitting. The quartic law applies to a specific but ubiquitous class of problems with two features: (i) the hidden coupling enters the observables at $O(\lambda^2)$ rather than $O(\lambda)$—the generic situation for power spectra, covariances, and correlations, which are quadratic in the dynamical variables; and (ii) the reduced (null) model has enough parameters to absorb the leading-order $O(\lambda^2)$ perturbation through reparametrization. When both conditions hold, the detectable signal drops from $O(\lambda^2)$ to $O(\lambda^4)$, and the data requirement increases accordingly. The theory does *not* claim that all detection problems are quartic.

# 2 Mathematical foundations

## 2.1 Kullback–Leibler divergence

**Definition 2.1.** For probability distributions $p$ and $q$ over the same sample space, the **Kullback–Leibler divergence** is

$$D_{\mathrm{KL}}(p\|q) = \int p(x) \log \frac{p(x)}{q(x)}\, dx. \tag{1}$$

The KL divergence satisfies $D_{\mathrm{KL}} \geq 0$ with equality if and only if $p = q$ almost everywhere (Gibbs' inequality). It is not symmetric. Physically, $D_{\mathrm{KL}}(p\|q)$ is the expected log-likelihood ratio when data come from $p$ but are evaluated under $q$. In detection theory it controls the exponent of the error probability: larger $D_{\mathrm{KL}}$ means the two distributions are easier to distinguish.

For model selection via the Bayesian Information Criterion (BIC), the model with the hidden variable is preferred when the log-likelihood gain exceeds $(\Delta k/2) \log N$, where $\Delta k$ is the parameter-count difference and $N$ is the sample size. For $N$ independent observations the expected log-likelihood gain is $N \cdot D_{\mathrm{KL}}$, so the detection boundary is determined by

$$N \cdot D^{\min}_{\mathrm{KL,loc}}(\lambda_c^{\mathrm{pop}}) = \frac{\Delta k}{2} \log N. \tag{2}$$

The BIC balance (2) is valid under standard regularity conditions: the observations are independent and identically distributed (or satisfy a mixing condition for time series), the null model is correctly specified, and the parameter space is compact. At the boundary $C = 0$, the BIC penalty may be replaced by the RLCT correction of Watanabe [3]; see Section 9 for further discussion.

## 2.2 Local quadratic approximation

Near a reference distribution $p_0$, write $p = p_0(1 + \epsilon)$ with $\int p_0\, \epsilon = 0$. Then

$$D_{\mathrm{KL}}(p_0(1+\epsilon_1) \,\|\, p_0(1+\epsilon_2)) = \frac{1}{2}\|\epsilon_1 - \epsilon_2\|^2_{L^2(p_0)} + O(\|\epsilon\|^3), \tag{3}$$

where $\|f\|^2_{L^2(p_0)} = \int f(x)^2\, p_0(x)\, dx$. Locally the KL divergence is a squared distance in the $L^2(p_0)$ Hilbert space. This is the key fact that converts the detection problem into a projection problem.

## 2.3 Whittle divergence for time series

In Sections 5–6 we will apply the quartic detection law to time series. For a stationary time series, the complete second-order statistical information is captured by the **power spectral density** $S(\omega)$, which decomposes the total variance of the process into contributions from each frequency $\omega$. (Think of it as the "frequency-domain fingerprint" of the time series.) Rather than working with the full time-domain likelihood, we use the Whittle approximation [4], which replaces the time-domain KL divergence with a frequency-domain integral

that depends only on the ratio of spectral densities. This is both computationally simpler and physically more transparent, because the hidden driver's effect on each frequency is directly visible.

For stationary Gaussian time series the frequency-domain KL divergence is

$$D_{\mathrm{KL}}^{\mathrm{Whittle}}(S_1 \| S_2) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[ \frac{S_1(\omega)}{S_2(\omega)} - \log \frac{S_1(\omega)}{S_2(\omega)} - 1 \right] d\omega, \tag{4}$$

where $S_1, S_2$ are power spectral densities [4, 5]. The induced inner product is $\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega)\, g(\omega)\, d\omega$.

*Remark* 2.2 (Whittle convention for the leading prefactor). Writing $S_1/S_2 = 1 + \delta$, the integrand $[(1+\delta) - \log(1+\delta) - 1]$ expands as $\delta^2/2 + O(\delta^3)$. Hence

$$D_{\mathrm{KL}}^{\mathrm{Whittle}} \approx \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\delta^2}{2}\, d\omega = \frac{1}{4} \|\delta\|_{L^2}^2,$$

where $\| \cdot \|_{L^2}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cdot)^2\, d\omega$. The leading prefactor is $1/4$, not $1/2$.

In the density-space expansion (3) the prefactor is $1/2$; the extra factor of $1/2$ in the Whittle setting arises because the spectral ratio is real-valued on $[-\pi, \pi]$. We track the convention explicitly throughout:

$$\text{density space: } C = \tfrac{1}{2}\|R\|^2; \qquad \text{spectral (Whittle): } C = \tfrac{1}{4}\|R\|_{L^2}^2.$$

# 3 The general quartic detection law

With the mathematical language in place—KL divergence as a measure of distinguishability, its local quadratic structure as a squared distance, and the Whittle divergence for time series—we now derive the central result: the quartic detection law. The key idea is that when a hidden variable perturbs an observed distribution, the null model can partially absorb this perturbation by refitting its parameters. Only the part of the perturbation that cannot be absorbed—the normal residual—contributes to detectability.

## 3.1 Setup: hidden perturbation of a parametric family

Let $\{p(\cdot|\boldsymbol{\eta}) : \boldsymbol{\eta} \in \Theta \subseteq \mathbb{R}^d\}$ be a parametric family of distributions—the "null" or "reduced" model. A hidden variable with coupling $\lambda$ perturbs the true distribution:

$$p_{\mathrm{true}}(x|\lambda) = p(x|\boldsymbol{\eta}_0)\big(1 + \lambda^2\, h(x) + O(\lambda^4)\big), \tag{5}$$

where $h \in L^2(p_0)$ with $\int p_0\, h = 0$ is the **perturbation function**. The entry at $O(\lambda^2)$ rather than $O(\lambda)$ is the generic case for spectra, covariances, and correlations, because these observables are quadratic in the dynamical variables. To see this concretely, consider the random-effects model of Section 4: the observation is $Y = \mu + \lambda u + \varepsilon$, where $u \sim N(0,1)$ is the hidden variable. The covariance of two sensors sharing $u$ is $\mathrm{Cov}(Y_1, Y_2) = \lambda^2$—quadratic in $\lambda$ because covariance is a second moment of the linear coupling $\lambda u$. More generally, any observable that depends on the joint second moments of a linearly coupled system will have its leading perturbation at $O(\lambda^2)$.

## 3.2 Tangent space and normal residual

The tangent space of the null model at $\boldsymbol{\eta}_0$ is

$$\mathcal{T} = \operatorname{span}\{e_1, \ldots, e_d\}, \qquad e_j(x) = \frac{\partial}{\partial \eta_j} \log p(x|\boldsymbol{\eta}) \Big|_{\boldsymbol{\eta}_0}. \tag{6}$$

These are the **score functions**—the directions in function space accessible by infinitesimal parameter changes. To build intuition: for a Gaussian $N(\mu, \sigma^2)$, the score with respect to $\mu$ is $e_\mu(x) = (x-\mu)/\sigma^2$, and the score with respect to $\sigma^2$ is $e_{\sigma^2}(x) = [(x-\mu)^2 - \sigma^2]/(2\sigma^4)$. Changing $\mu$ by $d\mu$ shifts the density as $p \to p(1 + d\mu \cdot e_\mu)$, so $e_\mu$ is the "direction" of that shift (see Appendix A.3 for details).

Let $G_{ij} = \langle e_i, e_j \rangle$ be the Fisher information matrix, and let $\Pi_{\mathcal{T}} : L^2(p_0) \to \mathcal{T}$ be the orthogonal projection:

$$\Pi_{\mathcal{T}} h = \sum_{i,j} \langle h, e_i \rangle \, (G^{-1})_{ij} \, e_j. \tag{7}$$

The **normal residual** is $R = (I - \Pi_{\mathcal{T}})\, h$—the component of the perturbation perpendicular to all parameter shifts. Physically, $\Pi_{\mathcal{T}} h$ is the part of the hidden effect that the null model can mimic by adjusting its parameters; $R$ is the part it cannot mimic, hence the only part that is detectable.

## 3.3 Derivation

**Theorem 3.1** (Quartic detection law)**.** *Assume the null family is regular: the Fisher information matrix $G$ is positive definite (identifiable model), and the parametrization is twice differentiable. Under the perturbation (5),*

$$D_{\mathrm{KL,loc}}^{\min}(\lambda) = \min_{\boldsymbol{\eta}} D_{\mathrm{KL}}\big(p_{\mathrm{true}}(\cdot|\lambda) \,\|\, p(\cdot|\boldsymbol{\eta})\big) = \frac{\lambda^4}{2} \|R\|^2_{L^2(p_0)} + O(\lambda^6). \tag{8}$$

*Proof.* Write $\boldsymbol{\eta} = \boldsymbol{\eta}_0 + \delta\boldsymbol{\eta}$. By the local quadratic approximation (3),

$$D_{\mathrm{KL}}\big(p_{\mathrm{true}} \,\|\, p(\cdot|\boldsymbol{\eta}_0 + \delta\boldsymbol{\eta})\big) = \frac{1}{2} \Big\| \lambda^2 h - \textstyle\sum_i \delta\eta_i \, e_i \Big\|^2 + O\big((|\delta\boldsymbol{\eta}| + \lambda^2)^3\big). \tag{9}$$

This is a quadratic function of $\delta\boldsymbol{\eta}$—it has the form $\frac{1}{2}\|v - A\delta\boldsymbol{\eta}\|^2$ with $v = \lambda^2 h$ and "$A\delta\boldsymbol{\eta}$" $= \sum_i \delta\eta_i \, e_i$. Setting $\partial/\partial(\delta\eta_k) = 0$ gives $\sum_j G_{kj} \, \delta\eta_j = \lambda^2 \langle h, e_k \rangle$ for each $k$, which is the standard least-squares normal equation (see Appendix A.2). Solving:

$$\delta\eta_i^* = \lambda^2 \sum_j (G^{-1})_{ij} \langle h, e_j \rangle + O(\lambda^4), \tag{10}$$

which is precisely $\lambda^2$ times the projection coefficients from Eq. (7). Substituting back,

$$\begin{aligned} D_{\mathrm{KL,loc}}^{\min}(\lambda) &= \tfrac{1}{2} \big\| \lambda^2 h - \lambda^2 \Pi_{\mathcal{T}} h \big\|^2 + O(\lambda^6) \\ &= \tfrac{\lambda^4}{2} \big\| (I - \Pi_{\mathcal{T}})\, h \big\|^2 + O(\lambda^6) \\ &= \tfrac{\lambda^4}{2} \|R\|^2 + O(\lambda^6). \end{aligned}$$

□

The result has a simple geometric reading. The true distribution deviates from the null manifold by $\lambda^2 h$. When we project this deviation onto the tangent space (i.e., refit the null model), the tangent component is absorbed and only the normal component $R$ survives. The squared length of $R$, times $\lambda^4/2$, is the minimum KL divergence.

*Remark* 3.2 (Connection to Efron's curvature). The normal residual $\|R\|^2$ is structurally identical to Efron's statistical curvature measure [6], but repurposed: Efron's curvature governs estimation efficiency, while our $\|R\|^2$ governs hidden-variable detectability.

## 3.4 Detection boundary

Writing $C = \frac{1}{2}\|R\|^2$ and combining Theorem 3.1 with the BIC balance (2):

$$\boxed{\lambda_c^{\text{pop}}(N) = \left[\frac{\Delta k \, \log N}{2\, C\, N}\right]^{1/4} \propto \left(\frac{\log N}{N}\right)^{1/4}.} \tag{11}$$

If the KL divergence were $O(\lambda^2)$ (as in standard, non-degenerate detection), the boundary would scale as $(\log N/N)^{1/2}$. The quartic law changes the exponent from $1/2$ to $1/4$. For a coupling of $\lambda/\sigma = 0.1$ the quartic penalty requires roughly $10^3$ times more data than the quadratic expectation—a qualitative change in what is experimentally feasible.

*Remark* 3.3 (Range of validity and finite-sample behavior). The quartic approximation $D_{\text{KL,loc}}^{\min} \approx C\lambda^4$ is accurate when the $O(\lambda^6)$ remainder is small relative to the leading term—roughly when $\lambda \lesssim 0.5$ in units where the null model parameters are $O(1)$. For larger couplings, higher-order terms become significant and the exact KL divergence (e.g., Eq. 14 for the random-effects model) should be used.

The BIC detection boundary (11) is asymptotic in $N$. At finite sample sizes ($N \lesssim 10^3$), the BIC penalty $\frac{\Delta k}{2}\log N$ may not accurately approximate the Bayes factor, and the actual detection threshold can differ from (11) by an $O(1)$ multiplicative factor. For precise finite-sample calibration, parametric bootstrap or Monte Carlo simulation of the likelihood ratio test is recommended.

## 3.5 The darkness criterion

**Corollary 3.4.** *The hidden variable is undetectable at all coupling strengths (at quartic order) if and only if the perturbation lies in the tangent space:*

$$C = 0 \quad \Longleftrightarrow \quad R = 0 \quad \Longleftrightarrow \quad h \in \mathcal{T}.$$

When this holds the system is in a **dark regime**: the hidden forcing is present and dynamically active, yet locally invisible to any statistical test based on the reduced model.

*Remark* 3.5 (Key takeaway from Section 3). The quartic detection law $D_{\text{KL,loc}}^{\min} = C\lambda^4$ follows from a single geometric fact: the null model absorbs the tangent component of the perturbation, leaving only the normal residual $R$. Everything reduces to computing $\|R\|^2$ for specific systems—which is what we do in the next two sections.

# 4 System I: random-effects model and the pairing principle

We now apply the general quartic law (Theorem 3.1) to our first concrete system: the Gaussian random-effects model, where $n$ sensors share a common hidden source. This system is tractable enough that we can compute everything in closed form—the exact KL divergence, the quartic coefficient, and the detection boundary—and the result reveals a fundamental structural principle: detection requires sensor *pairs*, not individual sensors.

## 4.1 Model

Consider $n$ sensors sharing a common hidden source. Following the convention in the PRL manuscript, we index observation windows by $i = 1, \ldots, N$ and sensors by $j = 1, \ldots, n$:

$$Y_{ij} = \mu + \lambda\, u_i + \varepsilon_{ij}, \qquad u_i \sim N(0,1), \;\; \varepsilon_{ij} \sim N(0, \sigma^2), \tag{12}$$

for windows $i = 1, \ldots, N$ and sensors $j = 1, \ldots, n$. The hidden source $u_i$ is shared across all $n$ sensors within window $i$, but varies independently across windows. All random variables $u_i$, $\varepsilon_{ij}$ are mutually independent. Integrating out $u_i$, the marginal for the $i$-th window vector (across sensors) takes the form of a spiked covariance model:

$$\boldsymbol{Y}_i \sim N\big(\mu\mathbf{1},\; \sigma^2 I_n + \lambda^2 \mathbf{1}\mathbf{1}^\top\big). \tag{13}$$

The null model ($\lambda = 0$) is $\boldsymbol{Y}_i \sim N(\tilde{\mu}\mathbf{1}, \tilde{\sigma}^2 I_n)$ with two parameters $(\tilde{\mu}, \tilde{\sigma}^2)$.

## 4.2 Pseudo-true parameters

The best-fit null parameters under the true model are $\tilde{\mu}^* = \mu$ (the mean is unaffected) and $\tilde{\sigma}^{2*} = \sigma^2 + \lambda^2$ (the null inflates the noise estimate to absorb the hidden variance). This is tangent-space absorption: the null explains away the marginal variance increase but cannot mimic the cross-sensor correlation $\lambda^2/(\sigma^2 + \lambda^2)$.

## 4.3 Exact KL divergence

**Lemma 4.1.** *For all $\lambda$,*

$$D^{\min}_{\mathrm{KL,loc}}(\lambda) = \tfrac{1}{2}\big[n \ln(1+t) - \ln(1+nt)\big], \qquad t = \lambda^2/\sigma^2. \tag{14}$$

*Proof.* The multivariate Gaussian KL formula gives

$$D_{\mathrm{KL}} = \tfrac{1}{2}\big[\mathrm{tr}(\Sigma_2^{-1}\Sigma_1) - n + \ln(\det \Sigma_2 / \det \Sigma_1)\big]$$

with $\Sigma_1 = \sigma^2 I_n + \lambda^2 \mathbf{1}\mathbf{1}^\top$ and $\Sigma_2 = (\sigma^2 + \lambda^2) I_n$. For the trace, we compute each factor explicitly:

$$\Sigma_2^{-1} = \frac{1}{\sigma^2 + \lambda^2}\, I_n,$$

so

$$\Sigma_2^{-1}\Sigma_1 = \frac{1}{\sigma^2+\lambda^2}\big(\sigma^2 I_n + \lambda^2 \mathbf{1}\mathbf{1}^\top\big).$$

Taking the trace:

$$\operatorname{tr}(\Sigma_2^{-1}\Sigma_1) = \frac{n\sigma^2 + \lambda^2\,\operatorname{tr}(\mathbf{1}\mathbf{1}^\top)}{\sigma^2+\lambda^2} = \frac{n\sigma^2 + n\lambda^2}{\sigma^2+\lambda^2} = n,$$

where we used $\operatorname{tr}(\mathbf{1}\mathbf{1}^\top) = \mathbf{1}^\top\mathbf{1} = n$. For the determinant ratio, the matrix determinant lemma gives $\det\Sigma_1 = \sigma^{2n}(1 + n\lambda^2/\sigma^2)$ and $\det\Sigma_2 = (\sigma^2+\lambda^2)^n$, so $\ln(\det\Sigma_2/\det\Sigma_1) = n\ln(1+t) - \ln(1+nt)$. Since the trace contribution is exactly $n - n = 0$, the result follows. □

## 4.4 The pairing principle

**Theorem 4.2** (Pairing principle)**.** *For $n = 1$: $D^{\min}_{\mathrm{KL,loc}}(\lambda) = 0$ for all $\lambda$. This is exact darkness—not a perturbative artifact.*

*For $n \geq 2$: the quartic expansion of (14) gives*

$$\boxed{C = \frac{n(n-1)}{4\sigma^4}.} \tag{15}$$

*The combinatorial factor $n(n-1) = 2\binom{n}{2}$ counts sensor pairs sharing the hidden effect.*

*Proof.* Taylor-expand $\ln(1+t) = t - t^2/2 + O(t^3)$:

$$\begin{aligned} D^{\min}_{\mathrm{KL,loc}} &= \tfrac{1}{2}\big[n(t - t^2/2) - (nt - n^2t^2/2) + O(t^3)\big] \\ &= \tfrac{n^2-n}{4}\,t^2 + O(t^3) = \frac{n(n-1)}{4\sigma^4}\,\lambda^4 + O(\lambda^6). \end{aligned} \tag{16}$$

At $n = 1$: $D^{\min}_{\mathrm{KL,loc}} = \frac{1}{2}[\ln(1+t) - \ln(1+t)] = 0$ identically. □

Why is $n = 1$ exactly dark? The marginal $Y \sim N(\mu, \sigma^2+\lambda^2)$ is *exactly in the null family* $\{N(\tilde\mu, \tilde\sigma^2)\}$. The perturbation has two components (shift in mean, shift in variance), and the null model has two free parameters—the tangent space is dimension-saturated. There is no residual, hence no detection, regardless of coupling strength or sample size.

For $n = 2$, the hidden source induces a correlation between the two sensors that the null model (diagonal covariance) cannot reproduce. This correlation is the normal residual. Each additional sensor adds $n-1$ new detectable pairs, but with diminishing returns: $\lambda_c \propto n^{-1/2}$ rather than $n^{-1}$.

**Example 4.3** (Experimental design—worked computation)**.** Suppose $\sigma^2 = 1$, $\Delta k = 1$ (one hidden parameter), and $N = 10^4$ observations. For $n = 2$ sensors, the quartic coefficient is $C = n(n-1)/(4\sigma^4) = 2\cdot 1/(4\cdot 1) = 1/2$. From (11):

$$\lambda_c = \left[\frac{\Delta k\,\log N}{2CN}\right]^{1/4} = \left[\frac{1\cdot\ln(10^4)}{2\cdot\frac{1}{2}\cdot 10^4}\right]^{1/4} = \left[\frac{9.21}{10^4}\right]^{1/4} \approx [9.21\times 10^{-4}]^{1/4} \approx 0.174.$$

Similarly: $C(n{=}3) = 3 \cdot 2/4 = 3/2 \Rightarrow \lambda_c \approx 0.131$; $C(n{=}4) = 4 \cdot 3/4 = 3 \Rightarrow \lambda_c \approx 0.110$.

To detect $\lambda/\sigma = 0.1$ with $n = 4$ sensors: solve $C\lambda^4 N = \frac{1}{2}\ln N$ for $N$, giving $N \geq 1.5\times10^5$ time windows. Compare with the naive $N \sim \lambda^{-2}$ expectation of $N \sim 100$: the quartic penalty increases the data requirement by a factor of $\sim 1500$.

*Remark* 4.4 (Key takeaway from Section 4). A single sensor is *exactly* dark (the KL divergence is identically zero for all $\lambda$ and all $N$). Detection requires at least two sensors, and the detection power grows as $\binom{n}{2}$—the number of sensor *pairs* sharing the hidden effect. This is the **pairing principle**.

# 5 System II: spectral inference and the dark regime

The random-effects model of the previous section operated in a "snapshot" setting: $N$ independent observations, each a vector of $n$ sensor readings. We now turn to a fundamentally different setting—time series—where the hidden variable drives the system *persistently* over time. The natural language for time-series detection is spectral: the power spectral density replaces the probability density, and the Whittle divergence replaces the ordinary KL divergence. This setting reveals a new phenomenon: the **spectrally dark regime**, where the hidden driver becomes undetectable not because of insufficient sensors, but because its timescale matches the observed system's timescale.

**Assumption 5.1.** Throughout Sections 5–7, all processes are jointly Gaussian and stationary.

## 5.1 The solvable benchmark

The driven AR(1) model:

$$X_{t+1} = a\,X_t + \lambda\,F_t + \varepsilon_t, \qquad F_{t+1} = b\,F_t + \eta_t, \tag{17}$$

with $|a| < 1$, $|b| < 1$, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $\eta_t \sim N(0, \sigma_\eta^2)$, $\varepsilon \perp \eta$. Only $\{X_t\}$ is observed.

Define the pole polynomial $\mathcal{P}_c(\omega) = |1 - ce^{-i\omega}|^2 = 1 + c^2 - 2c\cos\omega$. The null (one-pole) spectrum is

$$S_{\text{null}}(\omega; \tilde{a}, \tilde{\sigma}^2) = \frac{\tilde{\sigma}^2}{\mathcal{P}_{\tilde{a}}(\omega)}, \tag{18}$$

and the true observed spectrum is

$$S_{\text{true}}(\omega; \lambda) = \frac{\sigma_\varepsilon^2}{\mathcal{P}_a(\omega)} + \frac{\lambda^2\sigma_\eta^2}{\mathcal{P}_a(\omega)\,\mathcal{P}_b(\omega)} = S_0(\omega)\big(1 + \lambda^2\,h(\omega)\big), \tag{19}$$

where $S_0 = \sigma_\varepsilon^2/\mathcal{P}_a$ and the perturbation function is

$$h(\omega) = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2\,\mathcal{P}_b(\omega)}. \tag{20}$$

Physically, the hidden driver adds a second pole at $b$ to the observed spectrum.

## 5.2 Tangent space of the one-pole family

The null family has two parameters $(\tilde{a}, \tilde{\sigma}^2)$. Expanding the spectral ratio to first order about the null:

$$\frac{S_{\text{null}}}{S_0} = 1 + u\,\tilde{e}_1(\omega) + v\,\tilde{e}_2(\omega) + O(u^2 + v^2), \tag{21}$$

with $u = \delta\sigma^2/\sigma_\varepsilon^2$, $v = \delta a$, and tangent vectors

$$\tilde{e}_1(\omega) = 1, \qquad \tilde{e}_2(\omega) = \frac{2(\cos\omega - a)}{\mathcal{P}_a(\omega)}. \tag{22}$$

**Lemma 5.2** (Contour integral identities)**.** *For $|c| < 1$ and $|a| < 1$, $|b| < 1$ with $a \neq b$, the following identities hold (all integrals are over $[-\pi, \pi]$ with measure $d\omega/2\pi$):*

$$\langle 1, 1/\mathcal{P}_c \rangle = \frac{1}{1-c^2}, \tag{23}$$

$$\langle \cos\omega, 1/\mathcal{P}_c \rangle = \frac{c}{1-c^2}, \tag{24}$$

$$\langle 1/\mathcal{P}_c, 1/\mathcal{P}_c \rangle = \frac{1+c^2}{(1-c^2)^3}, \tag{25}$$

$$\langle 1/(\mathcal{P}_a\mathcal{P}_b), 1 \rangle = \frac{1}{(1-a^2)(1-b^2)} \quad \text{for } a \neq b. \tag{26}$$

*Each follows from the substitution $z = e^{i\omega}$ and evaluation of the residue at the pole $z = c$ (or poles $z = a, z = b$) inside the unit circle.*

*Proof.* For (23): write $\mathcal{P}_c(\omega) = |1 - ce^{-i\omega}|^2 = (1 - cz^{-1})(1 - cz)$ with $z = e^{i\omega}$. The substitution $dz = iz\,d\omega$ gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{d\omega}{\mathcal{P}_c} = \frac{1}{2\pi i}\oint \frac{dz}{z(1-cz^{-1})(1-cz)} = \frac{1}{2\pi i}\oint \frac{dz}{(z-c)(1-cz)}.$$

The pole at $z = c$ (inside $|z| = 1$ since $|c| < 1$) has residue $1/(1-c^2)$. The pole at $z = 1/c$ is outside.

For (24): replace $\cos\omega = (z + z^{-1})/2$ in the numerator. Using the same contour substitution as above:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos\omega}{\mathcal{P}_c}\,d\omega = \frac{1}{2\pi i}\oint \frac{(z+z^{-1})/2}{(z-c)(1-cz)}\,dz = \frac{1}{2\pi i}\oint \frac{z^2+1}{2z(z-c)(1-cz)}\,dz.$$

The integrand has poles at $z = 0$, $z = c$ (inside), and $z = 1/c$ (outside). The residue at $z = c$ is $(c^2+1)/[2c(1-c^2)]$. The residue at $z = 0$ is $-1/(2c)$. Summing:

$$\frac{c^2+1}{2c(1-c^2)} - \frac{1}{2c} = \frac{c^2+1-(1-c^2)}{2c(1-c^2)} = \frac{2c^2}{2c(1-c^2)} = \frac{c}{1-c^2}.$$

For (25): we need $\frac{1}{2\pi}\int 1/\mathcal{P}_c^2\,d\omega$. Differentiating (23) with respect to $c^2$, or by direct residue calculation with a double pole at $z = c$:

$$\frac{1}{2\pi i}\oint \frac{z\,dz}{(z-c)^2(1-cz)^2} = \frac{d}{dz}\left[\frac{z}{(1-cz)^2}\right]_{z=c} = \frac{1+c^2}{(1-c^2)^3}.$$

For (26): the integrand has simple poles at $z = a$ and $z = b$ inside the unit circle (assuming $a \neq b$), and the sum of residues gives $1/[(1-a^2)(1-b^2)]$. □

**Lemma 5.3** (Orthogonality of tangent basis)**.** *The tangent basis is orthogonal:* $\langle \tilde{e}_1, \tilde{e}_2 \rangle = 0$*, with* $\|\tilde{e}_1\|^2 = 1$ *and* $\|\tilde{e}_2\|^2 = 2/(1-a^2)$*.*

*Proof. Step 1:* $\|\tilde{e}_1\|^2 = 1$. Since $\tilde{e}_1 = 1$, we have $\|\tilde{e}_1\|^2 = \langle 1, 1 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\omega = 1$.

*Step 2:* $\langle \tilde{e}_1, \tilde{e}_2 \rangle = 0$. We compute

$$\langle 1, \tilde{e}_2 \rangle = \frac{2}{2\pi} \int_{-\pi}^{\pi} \frac{\cos\omega - a}{\mathcal{P}_a(\omega)}\, d\omega = 2\left[\frac{a}{1-a^2} - \frac{a}{1-a^2}\right] = 0,$$

using (24) for $\langle \cos\omega, 1/\mathcal{P}_a \rangle = a/(1-a^2)$ and (23) for $\langle 1, 1/\mathcal{P}_a \rangle = 1/(1-a^2)$.

*Step 3:* $\|\tilde{e}_2\|^2 = 2/(1-a^2)$. Jensen's formula establishes

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \mathcal{P}_a(\omega)\, d\omega = 0$$

for $|a| < 1$. Differentiating both sides twice with respect to $a$ yields the identity

$$\frac{4}{2\pi} \int_{-\pi}^{\pi} \frac{(a - \cos\omega)^2}{\mathcal{P}_a(\omega)^2}\, d\omega = \frac{2}{2\pi} \int_{-\pi}^{\pi} \frac{1}{\mathcal{P}_a(\omega)}\, d\omega = \frac{2}{1-a^2},$$

which gives $\|\tilde{e}_2\|^2 = 2/(1-a^2)$.

Alternatively, one can compute $\|\tilde{e}_2\|^2$ directly:

$$\|\tilde{e}_2\|^2 = 4\left[\langle \cos^2\omega/\mathcal{P}_a^2 \rangle - 2a\langle \cos\omega/\mathcal{P}_a^2 \rangle + a^2 \langle 1/\mathcal{P}_a^2 \rangle\right].$$

Using the contour integrals and the identity $\cos^2\omega = (1 + \cos 2\omega)/2$, all three terms can be evaluated by residues, and the result simplifies to $2/(1-a^2)$. □

## 5.3 Projection and the quartic coefficient

**Theorem 5.4** (Spectral quartic law)**.**

$$D_{\mathrm{KL,loc}}^{\min}(\lambda) = C(a, b, \sigma_\varepsilon, \sigma_\eta)\, \lambda^4 + O(\lambda^6), \tag{27}$$

*with the exact quartic coefficient*

$$\boxed{C = \frac{\sigma_\eta^4}{2\sigma_\varepsilon^4} \cdot \frac{b^2(a-b)^2}{(1-b^2)^3(1-ab)^2}.} \tag{28}$$

*Proof.* The quartic coefficient is $C = \frac{1}{4}\|R\|^2$ (Whittle convention). We compute:

$$\langle h, \tilde{e}_1 \rangle = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \cdot \frac{1}{1-b^2}, \tag{29}$$

$$\langle h, \tilde{e}_2 \rangle = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \cdot \frac{2b}{(1-ab)(1-b^2)}, \tag{30}$$

$$\|h\|^2 = \frac{\sigma_\eta^4}{\sigma_\varepsilon^4} \cdot \frac{1+b^2}{(1-b^2)^3}. \tag{31}$$

The residual norm is

$$\|R\|^2 = \|h\|^2 - \frac{\langle h, \tilde{e}_1\rangle^2}{\|\tilde{e}_1\|^2} - \frac{\langle h, \tilde{e}_2\rangle^2}{\|\tilde{e}_2\|^2} = \frac{\sigma_\eta^4}{\sigma_\varepsilon^4} \cdot \frac{2b^2(a-b)^2}{(1-b^2)^3(1-ab)^2}, \tag{32}$$

and $C = \frac{1}{4}\|R\|^2$ gives the stated result. □

Let us trace the algebra explicitly. Substituting (29)–(31):

$$\begin{aligned}\|R\|^2 &= \frac{\sigma_\eta^4}{\sigma_\varepsilon^4}\left[\frac{1+b^2}{(1-b^2)^3} - \frac{1}{(1-b^2)^2} - \frac{1-a^2}{2}\cdot\frac{4b^2}{(1-ab)^2(1-b^2)^2}\right] \\ &= \frac{2\sigma_\eta^4 b^2}{\sigma_\varepsilon^4(1-b^2)^3}\left[1 - \frac{(1-a^2)(1-b^2)}{(1-ab)^2}\right].\end{aligned} \tag{33}$$

The bracketed factor simplifies: $(1-ab)^2 - (1-a^2)(1-b^2) = a^2 - 2ab + b^2 = (a-b)^2$. Therefore

$$\|R\|^2 = \frac{2\sigma_\eta^4 b^2}{\sigma_\varepsilon^4(1-b^2)^3} \cdot \frac{(a-b)^2}{(1-ab)^2},$$

and $C = \frac{1}{4}\|R\|^2$ yields (28).

## 5.4 The spectrally dark regime

**Corollary 5.5.** *$C = 0$ if and only if $b = 0$ (trivial: white driver) or $a = b$ (timescale coalescence).*

At coalescence $a = b$, the perturbation becomes $h(\omega) = (\sigma_\eta^2/\sigma_\varepsilon^2)/\mathcal{P}_a(\omega)$, which is a linear combination of $\tilde{e}_1$ and $\tilde{e}_2$—i.e., it lies entirely in the tangent space $\mathcal{T}$. So $R = 0$.

Physically, when the hidden driver's relaxation timescale ($\tau_{\text{hidden}} = -1/\ln|b|$) matches the observed system's timescale ($\tau_{\text{obs}} = -1/\ln|a|$), the spectral perturbation is indistinguishable from a change in noise level and pole location. The data cost diverges:

$$\lambda_c^{\text{pop}} \propto |a-b|^{-1/2} \to \infty \quad \text{as} \quad b \to a. \tag{34}$$

## 5.5 Tangent absorption: derivation of the pseudo-true parameters

The pseudo-true parameters $(\tilde{a}^*, \tilde{\sigma}^{2*})$ minimize $D_{\text{KL}}^{\text{Whittle}}(S_{\text{true}}\|S_{\text{null}})$. We derive them from the Whittle first-order conditions.

*First-order condition for $\tilde{\sigma}^2$.* Write $r(\omega) = S_{\text{true}}(\omega)/S_{\text{null}}(\omega)$. The Whittle KL divergence is $\frac{1}{4\pi}\int[r - \log r - 1]\,d\omega$. Since $S_{\text{null}} = \tilde{\sigma}^2/\mathcal{P}_{\tilde{a}}$, we have $\partial_{\tilde{\sigma}^2}\log S_{\text{null}} = 1/\tilde{\sigma}^2$, so

$$\begin{aligned}\frac{\partial}{\partial\tilde{\sigma}^2}D_{\text{KL}}^{\text{Whittle}} &= \frac{1}{4\pi}\int\left[-\frac{1}{\tilde{\sigma}^2} + \frac{1}{\tilde{\sigma}^2\, r(\omega)}\right]d\omega = 0 \\ &\implies \quad \langle r^{-1}\rangle = 1 \quad \implies \quad \tilde{\sigma}^{2*} = \frac{\langle S_{\text{true}}\rangle_{\mathcal{P}}}{\langle S_{\text{true}}\cdot\mathcal{P}_{\tilde{a}}\rangle}.\end{aligned}$$

To extract the $O(\lambda^2)$ correction, write $\tilde{\sigma}^2 = \sigma_\varepsilon^2 + \delta\sigma^2$ and $r(\omega) = S_{\text{true}}/S_{\text{null}}$. At $\lambda = 0$, $r = 1$ and the condition is satisfied. At $O(\lambda^2)$: $S_{\text{true}} = S_0(1+\lambda^2 h)$ and $S_{\text{null}} = (\sigma_\varepsilon^2 + \delta\sigma^2)/\mathcal{P}_a$, so

$r \approx 1 + \lambda^2 h - \delta\sigma^2/\sigma_\varepsilon^2$. Thus $r^{-1} \approx 1 - \lambda^2 h + \delta\sigma^2/\sigma_\varepsilon^2$, and the condition $\langle r^{-1} \rangle = 1$ gives $\delta\sigma^2 = \lambda^2 \sigma_\varepsilon^2 \langle h \rangle = \lambda^2 \sigma_\eta^2/(1-b^2)$, using $\langle h \rangle = (\sigma_\eta^2/\sigma_\varepsilon^2)\langle 1/\mathcal{P}_b \rangle = \sigma_\eta^2/[\sigma_\varepsilon^2(1-b^2)]$ from (23). Therefore:

$$\tilde{\sigma}^{2*}(\lambda) = \sigma_\varepsilon^2 + \lambda^2 \frac{\sigma_\eta^2}{1-b^2} + O(\lambda^4). \tag{35}$$

*First-order condition for $\tilde{a}$.* Since $\partial_{\tilde{a}} \log S_{\text{null}} = -\partial_{\tilde{a}} \log \mathcal{P}_{\tilde{a}} = 2(\cos\omega - \tilde{a})/\mathcal{P}_{\tilde{a}}$, the condition $\partial_{\tilde{a}} D_{\text{KL}} = 0$ becomes $\langle r \cdot \tilde{e}_2 \rangle = \langle \tilde{e}_2 \rangle$, which at $O(\lambda^2)$ gives

$$\tilde{a}^*(\lambda) = a + \lambda^2 \frac{\sigma_\eta^2}{\sigma_\varepsilon^2} \cdot \frac{b(1-a^2)}{(1-ab)(1-b^2)} + O(\lambda^4). \tag{36}$$

The noise estimate absorbs the total hidden variance $\lambda^2\sigma_\eta^2/(1-b^2)$ (the $\tilde{e}_1$ component), and the pole shifts toward $b$ (the $\tilde{e}_2$ component). At coalescence $a = b$, the pole shift vanishes and all of the perturbation is absorbed into the noise level.

## 5.6 AR(2) driver: darkness is not generic

**Theorem 5.6.** *If the hidden driver is AR(2) with characteristic roots $z_1, z_2$ and $z_2 \neq 0$, then $C_{\text{AR(2)}} > 0$ for all parameter values, including $z_1 = a$. The coefficient interpolates smoothly: $C_{\text{AR(2)}} \sim C_0\, z_2^4$ as $z_2 \to 0$.*

*Proof.* The AR(2) hidden driver has characteristic polynomial $(1 - z_1 z^{-1})(1 - z_2 z^{-1})$ and spectral density $S_F(\omega) \propto 1/[\mathcal{P}_{z_1}(\omega)\mathcal{P}_{z_2}(\omega)]$. The perturbation function is therefore

$$h(\omega) = \frac{\alpha}{\mathcal{P}_{z_1}(\omega)\,\mathcal{P}_{z_2}(\omega)},$$

where $\alpha$ collects the noise variance ratio. Using partial fractions (valid when $z_1 \neq z_2$):

$$\frac{1}{\mathcal{P}_{z_1}\mathcal{P}_{z_2}} = \frac{A}{\mathcal{P}_{z_1}} + \frac{B}{\mathcal{P}_{z_2}} + (\text{cross terms}),$$

where $A$ and $B$ are determined by the partial-fraction decomposition in the variable $\cos\omega$.

The tangent space $\mathcal{T} = \text{span}\{\tilde{e}_1, \tilde{e}_2\} = \text{span}\{1,\, 2(\cos\omega - a)/\mathcal{P}_a\}$ can absorb any function of the form $c_1 + c_2/\mathcal{P}_a$. At coalescence $z_1 = a$, the $1/\mathcal{P}_a$ component of $h$ lies in $\mathcal{T}$ and is absorbed (as in the AR(1) dark regime). However, the residual contains $1/\mathcal{P}_{z_2}$, which is *not* a linear combination of 1 and $(\cos\omega - a)/\mathcal{P}_a$ unless $z_2 = 0$ (trivial) or $z_2 = a$ (which would make the AR(2) degenerate into AR(1)). To verify: projecting $1/\mathcal{P}_{z_2}$ onto $\tilde{e}_1$ and $\tilde{e}_2$ and computing the residual norm using the contour integrals of Lemma 5.2:

$$\|h - \Pi_{\mathcal{T}} h\|^2 \propto \frac{z_2^2(a - z_2)^2}{(1 - z_2^2)^3(1 - az_2)^2} > 0 \quad \text{whenever } z_2 \neq 0 \text{ and } z_2 \neq a.$$

This is the same functional form as the AR(1) coefficient (28) with $b$ replaced by $z_2$. Continuity in $z_2$ gives the smooth interpolation $C_{\text{AR(2)}} \sim C_0\, z_2^4$ as $z_2 \to 0$, recovering the AR(1) dark limit. □

The dark regime requires exact spectral shape matching—the hidden driver must have the same pole structure as the null family. Any additional spectral complexity (a second pole, a moving-average component) produces a residual that the two-dimensional tangent space cannot absorb. Darkness is structurally specific, not generic.

*Remark* 5.7 (Key takeaway from Section 5). Hidden persistent forcing from an AR(1) driver produces a quartic detection law $D^{\min}_{\mathrm{KL,loc}} = C\lambda^4$ with an exact coefficient $C$ that vanishes at timescale coalescence $a = b$. This **spectrally dark regime** means that even with unlimited data, a single observed time series cannot distinguish between an undriven system and one driven by a hidden source with a matching timescale. The culprit is tangent-space absorption: the spectral perturbation lies entirely within the null model's tangent space.

# 6 Breaking the impossibility: cross-spectral detection

We have established two impossibility results: a single Gaussian sensor cannot detect a hidden source (Section 4), and a single time series cannot detect a hidden driver at timescale coalescence (Section 5). Can we escape these impossibilities? The answer is yes: by observing *two* channels simultaneously and analyzing their *cross spectrum*—the frequency-domain correlation between the two channels. The key insight is that the cross spectrum lives in a subspace of the spectral matrix that is orthogonal to all single-channel (diagonal) tangent directions, so no amount of single-channel reparametrization can absorb it.

## 6.1 The single-channel impossibility

Lucente et al. [7] proved that no time-irreversibility measure can detect departure from equilibrium in a scalar Gaussian time series from a linear system. Our results strengthen this: at $n = 1$ the hidden source is not merely hard to detect—it is *exactly undetectable* for all coupling strengths.

## 6.2 Two-channel model

Two observed channels share the same hidden driver:

$$\begin{aligned} X^{(i)}_{t+1} &= a_i\, X^{(i)}_t + \lambda\, u_i\, F_t + \varepsilon^{(i)}_{t+1}, \quad i = 1, 2, \\ F_{t+1} &= b\, F_t + \eta_{t+1}, \qquad u_1^2 + u_2^2 = 1. \end{aligned} \tag{37}$$

The null hypothesis is a diagonal spectral matrix (independent channels, no cross-spectral power).

## 6.3 The cancellation identity

**Lemma 6.1** (Cancellation identity)**.** *The ratio* $|S_{12}(\omega)|^2/[S^0_{11}(\omega)\, S^0_{22}(\omega)]$ *is exactly independent of the observed-channel dynamics* $(a_1, a_2)$*:*

$$\frac{|S_{12}(\omega)|^2}{S^0_{11}\, S^0_{22}} = \frac{\lambda^4 u_1^2 u_2^2}{\sigma^2_{\varepsilon_1}\, \sigma^2_{\varepsilon_2}}\, S_F(\omega)^2, \tag{38}$$

*where $S_F = \sigma_\eta^2/\mathcal{P}_b$ is the hidden driver's spectral density.*

*Proof.* The cross spectrum is $|S_{12}|^2 = \lambda^4 u_1^2 u_2^2 \, |H_1|^2 |H_2|^2 \, S_F^2$, where $H_i(\omega) = 1/(1 - a_i e^{-i\omega})$ is the transfer function of channel $i$. The null diagonal spectra are $S_{ii}^0 = \sigma_{\varepsilon_i}^2 |H_i|^2$. In the ratio, $|H_1|^2$ and $|H_2|^2$ cancel identically. □

## 6.4 Multivariate Whittle KL and the auto–cross decomposition

Before stating the cross-spectral quartic law, we must derive the KL divergence for multivariate spectral matrices and show *why* it decomposes into auto-spectral and cross-spectral contributions.

**Lemma 6.2** (Multivariate Whittle KL divergence)**.** *For an $m$-channel stationary Gaussian process with true spectral matrix $\boldsymbol{S}(\omega)$ and null spectral matrix $\boldsymbol{S}^0(\omega)$, the Whittle KL divergence is*

$$D_{\mathrm{KL}}^{\mathrm{Whittle}}(\boldsymbol{S} \| \boldsymbol{S}^0) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[ \mathrm{tr}\big((\boldsymbol{S}^0)^{-1} \boldsymbol{S}\big) - \log\det\big((\boldsymbol{S}^0)^{-1} \boldsymbol{S}\big) - m \right] d\omega. \tag{39}$$

This is the matrix generalization of the scalar Whittle divergence (4): for $m = 1$, tr = id and log det = log, recovering (4).

Now consider the two-channel case ($m = 2$) with diagonal null: $\boldsymbol{S}^0 = \mathrm{diag}(S_{11}^0, S_{22}^0)$. The true spectral matrix has off-diagonal entries $S_{12}(\omega) = S_{21}(\omega)^*$ induced by the hidden driver. Writing $\boldsymbol{R} = (\boldsymbol{S}^0)^{-1} \boldsymbol{S}$:

$$\mathrm{tr}(\boldsymbol{R}) = \frac{S_{11}}{S_{11}^0} + \frac{S_{22}}{S_{22}^0}, \tag{40}$$

$$\det(\boldsymbol{R}) = \frac{S_{11} S_{22} - |S_{12}|^2}{S_{11}^0 S_{22}^0}. \tag{41}$$

Using $\log\det(\boldsymbol{R}) = \log(R_{11} R_{22}) + \log(1 - |S_{12}|^2/(S_{11} S_{22}))$, and expanding the second logarithm to leading order in $|S_{12}|^2 = O(\lambda^4)$:

$$\log\det(\boldsymbol{R}) = \log \frac{S_{11}}{S_{11}^0} + \log \frac{S_{22}}{S_{22}^0} + \frac{|S_{12}|^2}{S_{11} S_{22}} + O(\lambda^8).$$

Substituting into (39) and collecting terms, the KL divergence separates:

$$D_{\mathrm{KL}}^{\mathrm{Whittle}} = \underbrace{D_{\mathrm{auto}}^{(1)} + D_{\mathrm{auto}}^{(2)}}_{\text{scalar Whittle for each channel}} + \underbrace{\frac{1}{4\pi} \int \frac{|S_{12}|^2}{S_{11}^0 S_{22}^0} \, d\omega}_{D_{\mathrm{cross}}} + O(\lambda^6), \tag{42}$$

where we used $S_{ii} = S_{ii}^0 + O(\lambda^2)$ in the denominator of the cross term.

Under the diagonal null $\mathcal{M}_0 = \{(S_{11}^0, S_{22}^0, 0)\}$, the tangent space at any null point consists of perturbations that change the diagonal entries $(S_{11}, S_{22})$ while keeping $S_{12} = 0$. The cross spectrum $S_{12}$ lives in the *off-diagonal subspace* of the spectral matrix, which is orthogonal to all diagonal tangent directions. Therefore:

(a) The projection of the off-diagonal perturbation onto the diagonal tangent space is *zero*: no diagonal reparametrization can absorb cross-spectral power.

(b) Each auto-spectral term $D^{(i)}_{\text{auto}}$ inherits the scalar coalescence factor $(a_i - b)^2$ and vanishes at $a_i = b$, while $D_{\text{cross}}$ does not.

The cancellation identity (Lemma 6.1) then implies that $D_{\text{cross}}$ depends only on the hidden driver's spectral density $S_F$, not on the observed-channel filters.

## 6.5 Derivation of $C_{\text{cross}}$

From the decomposition above, the off-diagonal KL contribution at leading order is

$$D_{\text{cross}} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{|S_{12}(\omega)|^2}{S^0_{11}(\omega)\, S^0_{22}(\omega)}\, d\omega + O(\lambda^6). \tag{43}$$

Substituting the cancellation identity (38) and the AR(1) driver spectral density $S_F = \sigma_\eta^2/\mathcal{P}_b$:

$$\begin{aligned} D_{\text{cross}} &= \frac{\lambda^4 u_1^2 u_2^2}{\sigma^2_{\varepsilon_1}\, \sigma^2_{\varepsilon_2}} \cdot \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\sigma_\eta^4}{\mathcal{P}_b(\omega)^2}\, d\omega + O(\lambda^6) \\ &= \frac{\lambda^4 u_1^2 u_2^2\, \sigma_\eta^4}{\sigma^2_{\varepsilon_1}\, \sigma^2_{\varepsilon_2}} \cdot \frac{1}{2} \cdot \frac{1+b^2}{(1-b^2)^3} + O(\lambda^6), \end{aligned} \tag{44}$$

where we used the contour integral (25): $\langle 1/\mathcal{P}_b, 1/\mathcal{P}_b \rangle = (1+b^2)/(1-b^2)^3$. Therefore

$$C_{\text{cross}} = \frac{u_1^2 u_2^2\, \sigma_\eta^4 (1+b^2)}{2\, \sigma^2_{\varepsilon_1}\, \sigma^2_{\varepsilon_2}\, (1-b^2)^3}. \tag{45}$$

## 6.6 Cross-spectral quartic law

**Theorem 6.3** (Cross-spectral detectability)**.** *The cross-spectral KL contribution satisfies*

$$\boxed{D_{\text{cross}}(\lambda) = C_{\text{cross}}\, \lambda^4 + O(\lambda^6), \qquad C_{\text{cross}} = \frac{u_1^2 u_2^2}{\sigma^2_{\varepsilon_1} \sigma^2_{\varepsilon_2}}\, \mathcal{I}_F,} \tag{46}$$

*where* $\mathcal{I}_F = \frac{1}{4\pi} \int_{-\pi}^{\pi} S_F(\omega)^2\, d\omega$. *For an AR(1) hidden driver,*

$$C_{\text{cross}} = \frac{u_1^2 u_2^2\, \sigma_\eta^4 (1+b^2)}{2\, \sigma^2_{\varepsilon_1}\, \sigma^2_{\varepsilon_2}\, (1-b^2)^3}.$$

*Crucially,* $C_{\text{cross}} > 0$ *at exact timescale coalescence* $a_1 = a_2 = b$.

At coalescence, where all auto-spectral measures vanish:

$$D^{\min}_{\text{KL,loc}}(\lambda)\big|_{a_1 = a_2 = b} = C_{\text{cross}}\, \lambda^4 + O(\lambda^6) > 0. \tag{47}$$

*Remark* 6.4 (Key takeaway from Section 6). The cross spectrum provides a constructive escape from single-channel darkness. Because the cross-spectral power $|S_{12}(\omega)|^2$ lives in the off-diagonal block of the spectral matrix—orthogonal to all single-channel tangent directions—it cannot be absorbed by refitting the null model. Moreover, the cancellation identity shows that $D_{\text{cross}}$ depends only on the hidden driver's spectrum, not on the observed channels' dynamics. Two channels suffice to detect hidden forcing even at exact timescale coalescence.

# 7 The thermodynamic bridge

The cross-spectral detection result of the previous section has a striking physical implication: it provides a *quantitative* connection to thermodynamics. If a hidden driver forces two observed channels, the system is out of equilibrium (it dissipates energy), and the cross-spectral detection signal is directly proportional to the entropy production rate. This section makes that connection precise.

*Remark* 7.1 (Discrete-to-continuous correspondence). The cross-spectral quartic law of Section 6 was derived in discrete time with $\omega \in [-\pi, \pi]$. To connect with the continuous-time EPR below, we use the standard correspondence: a discrete-time AR(1) process $X_{t+1} = aX_t + \varepsilon_t$ with $a = e^{-\gamma \Delta t}$ approximates the continuous-time OU process $dX = -\gamma X\, dt + \sigma\, dW$ at sampling interval $\Delta t$. The discrete-time Whittle integral $\frac{1}{2\pi}\int_{-\pi}^{\pi}(\cdots)\, d\omega$ becomes $\frac{1}{2\pi}\int_{-\infty}^{\infty}(\cdots)\, d\omega$ in continuous time, with pole polynomials $\mathcal{P}_c(\omega) = |1 - ce^{-i\omega}|^2$ replaced by Lorentzian denominators $\gamma^2 + \omega^2$. The structural results—cancellation identity, orthogonality of cross-spectral block to diagonal tangent directions—carry over unchanged. The cross-spectral coefficient retains the same form $C_{\text{cross}} = u_1^2 u_2^2 \cdot \mathcal{I}_F/(\sigma_{\varepsilon_1}^2 \sigma_{\varepsilon_2}^2)$ with $\mathcal{I}_F$ now computed via Cauchy integrals in the upper half-plane. In particular, $C_{\text{cross}} > 0$ at coalescence, just as in discrete time.

In continuous time, the one-way coupled Ornstein–Uhlenbeck system

$$dX_i = -\gamma_i X_i\, dt + \lambda u_i F\, dt + \sqrt{2D_i}\, dW_i, \quad i = 1, 2, \tag{48}$$

$$dF = -\gamma_f F\, dt + \sqrt{2D_f}\, dW_f, \tag{49}$$

has the following exact entropy production rate.

**Theorem 7.2** (Exact EPR under one-way coupling)**.** *Assume one-way coupling: the hidden driver $F$ forces the observed channels $X_1, X_2$, but $X_1, X_2$ do not feed back into $F$. Under this condition,*

$$\dot{\Sigma}_{\text{total}} = \alpha_2\, \lambda^2 \quad \textit{exactly for all } \lambda, \tag{50}$$

*where $\alpha_2 = u_1^2 D_f/[D_1(\gamma_1 + \gamma_f)] + u_2^2 D_f/[D_2(\gamma_2 + \gamma_f)] > 0$.*

*If $X$ feeds back into $F$ (bidirectional coupling), the EPR acquires additional terms and is no longer a simple quadratic in $\lambda$.*

*Proof. Step 1: Lyapunov equation.* Under one-way coupling, the drift and diffusion matrices for $\boldsymbol{X} = (X_1, X_2, F)^\top$ are

$$A = \begin{pmatrix} -\gamma_1 & 0 & \lambda u_1 \\ 0 & -\gamma_2 & \lambda u_2 \\ 0 & 0 & -\gamma_f \end{pmatrix}, \qquad D = \begin{pmatrix} 2D_1 & 0 & 0 \\ 0 & 2D_2 & 0 \\ 0 & 0 & 2D_f \end{pmatrix}.$$

The steady-state covariance $\Sigma$ satisfies $A\Sigma + \Sigma A^\top + D = 0$ (see Appendix A.9).

*Step 2: Solve the Lyapunov equation.* The triangular structure of $A$ (the third row has no coupling from $X_1, X_2$ back to $F$) means the $(3,3)$ block decouples:

$$-2\gamma_f \Sigma_{FF} + 2D_f = 0 \quad \Longrightarrow \quad \Sigma_{FF} = D_f/\gamma_f.$$

For the off-diagonal $(i, 3)$ blocks, the Lyapunov equation gives

$$-(\gamma_i + \gamma_f)\Sigma_{X_iF} + \lambda u_i \Sigma_{FF} = 0 \quad \Longrightarrow \quad \Sigma_{X_iF} = \frac{\lambda u_i D_f}{\gamma_f(\gamma_i + \gamma_f)}.$$

Note that $\Sigma_{X_iF}$ is *linear* in $\lambda$—this is the key structural consequence of one-way coupling.

*Step 3: Decompose the drift into reversible and irreversible parts.* At equilibrium, detailed balance requires $D^{-1}A\Sigma$ to be symmetric. Decompose $A = A_{\text{rev}} + A_{\text{irr}}$ where $A_{\text{irr}}$ captures the antisymmetric (irreversible) part. Since $A$ is lower-triangular and $D$ is diagonal, the only irreversible contributions come from the off-diagonal coupling terms $\lambda u_i$ in positions $(1, 3)$ and $(2, 3)$ of $A$. Concretely:

$$(D^{-1}A)_{\text{antisym}} = \frac{1}{2}\big(D^{-1}A - A^\top D^{-1}\big),$$

whose only nonzero entries are in the $(i, 3)$ and $(3, i)$ positions.

*Step 4: Compute the EPR.* The entropy production rate for a linear Gaussian system is (see Appendix A.10):

$$\dot{\Sigma}_{\text{total}} = 2\,\text{tr}\big[(D^{-1}A\Sigma)^2_{\text{antisym}}\, D\big].$$

The antisymmetric matrix has entries proportional to $\lambda u_i \Sigma_{X_iF}/(2D_i) \propto \lambda^2$. Squaring and taking the trace:

$$\dot{\Sigma}_{\text{total}} = \sum_{i=1}^{2} \frac{\lambda^2 u_i^2 \Sigma_{FF}}{D_i} \cdot \frac{2D_f}{(\gamma_i + \gamma_f)} = \lambda^2 \sum_{i=1}^{2} \frac{u_i^2 D_f}{D_i(\gamma_i + \gamma_f)} = \alpha_2\, \lambda^2. \tag{51}$$

The result is *exact* for all $\lambda$ (not just a small-$\lambda$ approximation) because $\Sigma_{X_iF}$ is exactly linear in $\lambda$ and $\Sigma_{FF}$ is independent of $\lambda$—both consequences of one-way coupling. □

**Corollary 7.3** (EPR–detectability bridge)**.** *Under one-way coupling,*

$$\boxed{\dot{\Sigma}^2_{\text{total}} = \frac{\alpha_2^2}{C_{\text{cross}}}\, D_{\text{cross}} + O(\lambda^6).} \tag{52}$$

*This is a* quantitative *relationship: it not only certifies nonequilibrium ($D_{\text{cross}} > 0 \Rightarrow \dot{\Sigma} > 0$), but also* bounds *the entropy production rate from the observable cross-spectral divergence. Specifically, $\dot{\Sigma}_{\text{total}} = (\alpha_2/\sqrt{C_{\text{cross}}})\sqrt{D_{\text{cross}}} + O(\lambda^4)$, so a measurement of $D_{\text{cross}}$ translates directly into a lower bound on $\dot{\Sigma}$.*

A single probe can sit inside a system with arbitrarily large true entropy production and measure exactly zero irreversibility. Two probes, sharing the same hidden bath, break this thermodynamic blindness—and the bridge equation tells you quantitatively how much dissipation the cross-spectral signal certifies.

# 8 The dimension-gap principle

We now step back from specific systems and ask a general question: *when* is single-sensor darkness possible, and when is it not? All the darkness results so far involved Gaussian observations. Is this a coincidence, or is there a deeper structural principle? The answer is the **dimension-gap principle**: single-sensor darkness requires the null model to have enough parameters to absorb all the cumulants shifted by the hidden variable. When the number of affected cumulants exceeds the number of null parameters, detection is possible even from a single sensor.

Single-sensor darkness is not universal. It requires a structural condition relating the complexity of the null model to the complexity of the hidden perturbation.

**Theorem 8.1** (Dimension-gap principle)**.** *Let the null model be a minimal $d$-parameter exponential family with sufficient statistic $T(x) = (T_1(x), \dots, T_d(x)) \in \mathbb{R}^d$. Let $K$ be the number of cumulants of the marginal distribution that shift at $O(\lambda^2)$ when the hidden variable is activated:*

$$K = \big|\{r : \partial_{\lambda^2}\kappa_r\big|_{\lambda=0} \neq 0\}\big|.$$

*If $K > d$, then $C > 0$ (detectable at $n = 1$). If $K \leq d$, then $C$ may vanish.*

*Proof.* In a minimal exponential family, the tangent space at the null is $\mathcal{T} = \mathrm{span}\{T_k - \mu_k : k = 1, \dots, d\}$, where $\mu_k = E_0[T_k]$. This span has dimension exactly $d$ by minimality (the sufficient statistics are not redundant).

The perturbation $h$ shifts $K$ cumulants at $O(\lambda^2)$. The cumulant-shift vector $\delta\kappa = (\partial_{\lambda^2}\kappa_r|_{\lambda=0})_{r=1}^K \in \mathbb{R}^K$ determines $h$'s projection onto $\mathcal{T}$. The key fact linking cumulants to sufficient statistics is that in a minimal exponential family, the mean parameters $E_\theta[T_k] = \partial A/\partial\theta_k$ are in bijection with the natural parameters $\theta_k$ via the gradient map $\nabla A$ (which is a diffeomorphism by positive definiteness of $\nabla^2 A = G$). Therefore, shifting $d$ mean parameters—equivalently, the first $d$ cumulants associated with the sufficient statistics—is equivalent to shifting the $d$ natural parameters. Note that higher cumulants ($r > d$) are also functions of $\boldsymbol{\theta}$, but they are *determined* by the first $d$ and do not provide additional degrees of freedom. The tangent-space projection $\Pi_{\mathcal{T}} h$ absorbs the perturbation only if $h$ can be written as a linear combination of the centered sufficient statistics $T_k - \mu_k$.

If $K > d$: the image of $\mathcal{T}$ in cumulant space is a $d$-dimensional subspace of $\mathbb{R}^K$. The set of cumulant shifts that lie entirely in this $d$-dimensional image is a proper linear subspace of codimension $\geq K - d \geq 1$, hence has Lebesgue measure zero in $\mathbb{R}^K$. For the specific perturbations arising from hidden-variable models studied here, $\Pi_{\mathcal{T}}^{\perp} h \neq 0$ is verified directly, giving $C = \frac{1}{2}\|\Pi_{\mathcal{T}}^{\perp} h\|^2 > 0$.

If $K \leq d$: the tangent space can absorb all $K$ shifted cumulants. In particular, the $d$ centered sufficient statistics provide $d \geq K$ degrees of freedom to match $K$ cumulant constraints, and $C = 0$ is achievable. The Gaussian case ($K = d = 2$) is the canonical example. □ □

The principle is stated as:

$$\dim(\mathcal{T}) \geq K \implies \text{darkness possible at } n{=}1; \quad \dim(\mathcal{T}) < K \implies C > 0. \tag{53}$$

- *Gaussian sensors* ($\dim \mathcal{T} = 2$, affects mean + variance): saturated ⇒ exact darkness.
- *Poisson sensors* ($\dim \mathcal{T} = 1$, affects mean + variance): gap ⇒ detectable at $n = 1$.
- *Negative binomial sensors* ($\dim \mathcal{T} = 2$, affects mean + variance): saturated ⇒ darkness restored.

**Corollary 8.2** (Gaussian as worst case)**.** *Among exponential families with $d = K$ (saturated), the Gaussian maximizes the detection threshold because its tangent space exactly absorbs the perturbation. Formally: for any $d$-parameter exponential family with $K = d$ affected cumulants, the coefficient $C$ may vanish (as it does for the Gaussian), and the detection threshold diverges to $\lambda_c = \infty$ at $n = 1$. Any family with $d < K$ has $C > 0$ and a finite detection threshold. Therefore, among observation models with the same number of affected cumulants, Gaussian observations are the hardest case for single-probe detection.*

**Example 8.3** (Gaussian vs. Poisson)**.** *Gaussian:* $d = 2$ (mean and variance) and $K = 2$ (the hidden variable shifts both the first and second cumulants)—the tangent space is saturated, and $n = 1$ is exactly dark.

*Poisson with hidden multiplicative noise.* Consider $Y \mid Z \sim \mathrm{Pois}(\theta_0 Z)$, where $Z = 1+\lambda^2\xi$ with $E[\xi] = 0$, $E[\xi^2] = 1$. The marginal distribution is

$$p(y|\lambda) = E_Z\left[\frac{(\theta_0 Z)^y e^{-\theta_0 Z}}{y!}\right] = \frac{\theta_0^y e^{-\theta_0}}{y!}\big(1 + \lambda^2[\theta_0(y-\theta_0) - \theta_0 + \tfrac{1}{2}\theta_0^2(y-\theta_0)^2/\theta_0^2] + O(\lambda^4)\big).$$

Expanding $(\theta_0 Z)^y e^{-\theta_0 Z}$ to $O(\lambda^2)$ and collecting terms, the perturbation function is

$$h(y) = \frac{(y-\theta_0)^2 - \theta_0}{2}.$$

The null family is Poisson with $d = 1$ parameter $\theta$. The tangent space is one-dimensional: $\mathcal{T} = \mathrm{span}\{y - \theta_0\}$. Projecting $h$ onto $\mathcal{T}$:

$$\Pi_{\mathcal{T}} h = \frac{\langle h, y-\theta_0\rangle}{\|y-\theta_0\|^2}(y-\theta_0) = \frac{E[(y-\theta_0)^3 - \theta_0(y-\theta_0)]/(2)}{E[(y-\theta_0)^2]}(y-\theta_0) = \frac{\theta_0/2}{\theta_0}(y-\theta_0) = \tfrac{1}{2}(y-\theta_0),$$

where we used the Poisson cumulants $\kappa_2 = \kappa_3 = \theta_0$. The residual is $R(y) = h(y) - \Pi_{\mathcal{T}} h = [(y-\theta_0)^2 - \theta_0]/2 - (y-\theta_0)/2$, and

$$C^{\mathrm{Pois}} = \tfrac{1}{2}\|R\|^2 = \tfrac{\theta_0^2}{4} > 0.$$

The hidden variable is detectable from a single sensor because $K = 2$ (both mean $E[Y]$ and variance $\mathrm{Var}(Y)$ shift) but $d = 1$, producing a dimension gap.

Enlarging the null to negative binomial ($d = 2$) restores darkness at $n = 1$, confirming that it is the *gap* between $d$ and $K$ that governs detectability.

# 9 Further reading and connections to the literature

This section places the quartic detection theory in the context of the broader statistical and physical literature. Each subsection connects to a different intellectual tradition and provides pointers for readers who want to explore further. This material is not needed to understand the main results of the tutorial, but it provides valuable perspective on where the ideas come from and how they relate to other fields.

## 9.1 Coarse-grained inference and the Mori–Zwanzig formalism

The physical motivation for this work is the problem of detecting hidden slow modes from coarse observations. In climate physics, Hasselmann's stochastic climate model [8, 9] established that unresolved fast atmospheric forcing can mimic slow oceanic variability; Penland and Sardeshmukh [10] showed that the linear inverse model framework can extract predictable dynamics from coarse observations, but cannot distinguish intrinsic persistence from hidden forcing. In nonequilibrium thermodynamics, the Mori–Zwanzig formalism [11–14] projects out fast degrees of freedom, but this projection can erase signatures of dissipation: Mehl et al. [15] demonstrated that hidden slow variables can mask the fluctuation theorem, and Lucente et al. [7] proved that no time-irreversibility measure can detect departure from equilibrium in a scalar Gaussian time series from a linear system. The present work provides the geometric mechanism underlying these impossibilities: tangent-space absorption explains *why* certain hidden variables are invisible and *when* additional observations can restore detectability.

## 9.2 Le Cam's local asymptotic normality

Le Cam's theory of local asymptotic normality (LAN) [16, 17] provides the foundational framework for local hypothesis testing. Under LAN, the log-likelihood ratio for alternatives at distance $O(1/\sqrt{n})$ from the null is asymptotically normal with variance equal to the Fisher information. The population-level KL scaling $D_{\mathrm{KL}} \sim \delta\theta^\top G\, \delta\theta/2$ is the classical quadratic approximation. Our quartic law arises when the leading perturbation is absorbed by the tangent space, leaving only the next-order residual. This is a population-level phenomenon that produces the non-standard $n^{-1/4}$ detection rate—the regime studied by Chen [1] for over-specified mixtures and by Dacunha-Castelle and Gassiat [18] for testing the order of mixture and ARMA models via locally conic parametrizations.

## 9.3 Semiparametric efficiency (BKRW)

The semiparametric efficiency framework of Bickel, Klaassen, Ritov, and Wellner [19] decomposes the score function into an efficient component (orthogonal to the nuisance tangent space) and a nuisance component. When the efficient Fisher information vanishes ($I_{\mathrm{eff}} = 0$), $\sqrt{n}$-consistent estimation of the parameter of interest is impossible. Van der Vaart [20] develops the connection between least-favorable submodels and the projection geometry. Our darkness condition $h \in \mathcal{T}$ is mathematically equivalent to $I_{\mathrm{eff}} = 0$ in the semiparametric framework when the reduced model parameters are treated as nuisance. The distinction is in the physical question: semiparametric theory asks *how efficiently can we estimate* $\lambda$*?*; we ask *can we detect that* $\lambda \neq 0$ *at all?* The two questions coincide at the $n^{-1/2}$ rate but diverge at the $n^{-1/4}$ rate [1, 21].

## 9.4 Information geometry and Efron's statistical curvature

Amari's information geometry [22, 23] provides the natural language for the tangent-normal decomposition on statistical manifolds. The $m$-projection onto a submanifold and the asso-

ciated Pythagorean theorem are foundational; however, the Pythagorean theorem is exact only for $e$-flat or $m$-flat submanifolds [24] and degrades to an inequality for curved manifolds. Our result uses the local quadratic expansion instead, which holds under the regularity condition of Theorem 3.1.

Efron's statistical curvature [6] measures second-order efficiency loss in curved exponential families via the squared norm of the acceleration's normal component—structurally identical to $\|R\|^2$, but applied to a different problem (efficiency loss in estimation vs. detectability of hidden variables). The connection between curvature and model complexity was further developed by Kass and Vos [25].

## 9.5 Model misspecification and White's pseudo-true parameters

White [26] established that under model misspecification, the MLE converges to a *pseudo-true* parameter $\theta^*$ minimizing the KL divergence from the true distribution to the model family. Our $\tilde{\sigma}^{2*} = \sigma^2 + \lambda^2$ is the pseudo-true variance in the random-effects model. Vuong [27] developed likelihood-ratio tests for non-nested model comparison using KL divergence—the natural framework for our BIC-based detection criterion. The specific analysis of $D^{\min}_{\mathrm{KL,loc}}(\lambda)$ as a power series in a coupling parameter $\lambda$ around the null model is, to our knowledge, new: White and Vuong treat misspecification as fixed and study estimator asymptotics, whereas we study the *scaling* of the population KL minimum as the hidden perturbation is turned on.

## 9.6 The BBP spiked covariance transition

The marginal covariance $\sigma^2 I_n + \lambda^2 \mathbf{1}\mathbf{1}^\top$ of our random-effects model is a rank-one spiked covariance. In random matrix theory, the BBP transition [28] establishes a sharp threshold for eigenvalue separation in spiked Wishart matrices as the matrix dimension grows. Johnstone [29] initiated the study of the largest eigenvalue distribution in high-dimensional PCA. Our $n{=}1 \to n{=}2$ dark-to-detectable transition is a finite-dimensional analogue—both involve a rank-one perturbation becoming detectable as the observation dimension exceeds a critical value—though the mathematical mechanisms differ (our result is a KL projection at fixed $n$; BBP is a random-matrix limit as $n \to \infty$).

## 9.7 Singular learning theory and Watanabe's RLCT

At the dark boundary $C \to 0$, the Fisher information for the hidden variable degenerates, placing the problem in the singular regime studied by Watanabe [3]. The real log-canonical threshold (RLCT) replaces the BIC penalty term $\frac{d}{2}\log n$ with $\lambda_{\mathrm{RLCT}} \log n$, where $\lambda_{\mathrm{RLCT}} \leq d/2$ is a rational algebraic invariant. Drton [30] derived non-standard LRT distributions at singularities using the tangent-cone geometry, and Drton, Kubjas, and Weisman [31] recently computed RLCT bounds for factor analysis. The connection between our smooth-regime quartic law ($C > 0$) and the singular-regime RLCT (at the boundary $C = 0$) is an important open direction.

### 9.8 Partial-observation EPR and thermodynamic uncertainty relations

In nonequilibrium thermodynamics, partial observation systematically underestimates entropy production [32, 33]. Thermodynamic uncertainty relations [34, 35] provide lower bounds on dissipation from current fluctuations, but these bounds are loose under coarse graining. Dechant and Sasa [36] tightened dissipation bounds using multi-current covariances; Ohga, Ito, and Kolchinsky [37] bounded thermodynamic affinity from time-domain cross-correlation asymmetry. Our companion paper [38] shows that the cross-spectral coefficient provides a frequency-domain bridge to entropy production under one-way coupling, complementing these time-domain approaches.

## 10 Case study: detecting a hidden volatility factor in financial markets

This section applies the full quartic detection framework to a concrete financial scenario. We assign specific numerical values to every parameter, carry out every computation step by step, and trace the entire chain: single-channel darkness → spectral quartic law → coalescence singularity → cross-spectral escape → thermodynamic certification. The reader should be able to reproduce every number with a calculator.

### 10.1 The scenario

Suppose we observe daily log-returns of two sector indices:

- Channel 1: a large-cap technology index (think: Nasdaq-100),
- Channel 2: a financial-sector index (think: XLF).

Both indices are influenced by an unobserved market-wide volatility factor $F_t$ (think: a latent version of the VIX). We model each index's deviation from its mean as a mean-reverting process driven by both idiosyncratic noise and the hidden factor.

**The model** (discrete-time, daily observations). Each channel is driven by its own idiosyncratic noise and by a common hidden factor:

$$X_{t+1}^{(1)} = a_1\, X_t^{(1)} + \lambda_1\, F_t + \varepsilon_{t+1}^{(1)}, \tag{54}$$

$$X_{t+1}^{(2)} = a_2\, X_t^{(2)} + \lambda_2\, F_t + \varepsilon_{t+1}^{(2)}, \tag{55}$$

$$F_{t+1} = b\, F_t + \eta_{t+1}. \tag{56}$$

Here $\lambda_i$ is the coupling strength from the hidden factor to channel $i$. This is the model of Section 6 with $\lambda_i = \lambda u_i$; for this case study we set $\lambda_1 = \lambda_2 \equiv \lambda$ (equal coupling) for simplicity, corresponding to $u_1 = u_2 = 1/\sqrt{2}$ and overall coupling $\lambda_{\text{total}} = \lambda\sqrt{2}$.

**Parameter values.** We choose values motivated by empirical financial data:

| Parameter | Symbol | Value | Interpretation |
|---|---|---|---|
| Tech relaxation | $a_1$ | 0.90 | Half-life $\approx 6.6$ days |
| Finance relaxation | $a_2$ | 0.85 | Half-life $\approx 4.3$ days |
| Hidden factor relaxation | $b$ | 0.95 | Half-life $\approx 13.5$ days |
| Per-channel coupling | $\lambda$ | 0.15 | Same for both channels |
| Tech noise variance | $\sigma^2_{\varepsilon_1}$ | 1.0 | Normalized |
| Finance noise variance | $\sigma^2_{\varepsilon_2}$ | 1.0 | Normalized |
| Hidden factor noise variance | $\sigma^2_{\eta}$ | 1.0 | Normalized |
| Sample size | $N$ | 2500 | $\approx$10 years of daily data |

The half-life is computed as $\tau_{1/2} = -\ln 2/\ln|a|$; for $a_1 = 0.90$, $\tau_{1/2} = 0.693/0.105 \approx 6.6$ days.

## 10.2 Step 1: can we detect the hidden factor from a single channel?

Consider observing *only* Channel 1 (the tech index). The null model is a one-pole AR(1): $X_{t+1} = \tilde{a}\, X_t + \tilde{\varepsilon}_t$. From Theorem 5.4, the quartic coefficient is

$$C_1 = \frac{\sigma^4_\eta}{2\sigma^4_{\varepsilon_1}} \cdot \frac{b^2(a_1 - b)^2}{(1 - b^2)^3(1 - a_1 b)^2}.$$

**Substituting values.** We compute each factor:

$$\begin{aligned}
\sigma^4_\eta/\sigma^4_{\varepsilon_1} &= 1.0^4/1.0^4 = 1, \\
b^2 &= 0.95^2 = 0.9025, \\
(a_1 - b)^2 &= (0.90 - 0.95)^2 = (-0.05)^2 = 0.0025, \\
(1 - b^2)^3 &= (1 - 0.9025)^3 = 0.0975^3 = 9.27 \times 10^{-4}, \\
(1 - a_1 b)^2 &= (1 - 0.90 \cdot 0.95)^2 = (1 - 0.855)^2 = 0.145^2 = 0.02103.
\end{aligned}$$

Therefore:

$$\begin{aligned}
C_1 &= \frac{1}{2} \cdot \frac{0.9025 \times 0.0025}{9.27 \times 10^{-4} \times 0.02103} \\
&= \frac{1}{2} \cdot \frac{2.256 \times 10^{-3}}{1.949 \times 10^{-5}} \\
&= \frac{1}{2} \times 115.8 = \boxed{57.9}.
\end{aligned} \tag{57}$$

**Detection boundary.** From Eq. (11) with $\Delta k = 1$ (one hidden parameter: $\lambda$):

$$\begin{aligned}\lambda_c^{(1)} &= \left[\frac{\Delta k\, \ln N}{2\, C_1\, N}\right]^{1/4} = \left[\frac{1 \times \ln 2500}{2 \times 57.9 \times 2500}\right]^{1/4} \\ &= \left[\frac{7.824}{2.895 \times 10^5}\right]^{1/4} = \left[2.703 \times 10^{-5}\right]^{1/4} \\ &= \boxed{0.0721}. \end{aligned} \tag{58}$$

Since $\lambda = 0.15 > \lambda_c^{(1)} = 0.072$, the hidden factor is **detectable from Channel 1 alone**—but only barely (the ratio is about $2\times$).

**How much data would we need?** At the actual coupling $\lambda = 0.15$:

$$N_{\min} \approx \frac{\Delta k\, \ln N_{\min}}{2\, C_1\, \lambda^4} \approx \frac{\ln N_{\min}}{2 \times 57.9 \times 0.15^4} = \frac{\ln N_{\min}}{0.05858}.$$

Since $\ln N_{\min}$ varies slowly, use $\ln N_{\min} \approx 7$ (initial guess $N \sim 1000$):

$$N_{\min} \approx 7/0.05858 \approx 120 \text{ observations.}$$

With 10 years of data ($N = 2500$), we are well above this threshold.

## 10.3 Step 2: what happens near timescale coalescence?

Now suppose the hidden factor's timescale *matches* the tech index's timescale. Set $b = a_1 = 0.90$ (coalescence):

$$C_1\big|_{b=a_1} = \frac{\sigma_\eta^4}{2\sigma_{\varepsilon_1}^4} \cdot \frac{b^2\,(a_1 - b)^2}{(1-b^2)^3(1-a_1 b)^2}\bigg|_{b=0.90} = \frac{0.81 \times \mathbf{0}}{(\cdots)} = \boxed{0}.$$

At exact coalescence, $C_1 = 0$: the hidden factor is **exactly undetectable** from Channel 1, regardless of coupling strength or sample size. This is the spectrally dark regime of Corollary 5.5.

**Near coalescence.** The quartic coefficient scales as $C_1 \propto (a_1 - b)^2$. Setting $b = 0.90 + \delta$:

$$C_1(\delta) \approx \frac{0.81\,\delta^2}{2 \times (0.19)^3 \times (0.19)^2} \approx \frac{0.81}{2 \times 6.86 \times 10^{-3} \times 0.0361}\,\delta^2 \approx 1635\,\delta^2.$$

For $\delta = 0.05$ (our original $b = 0.95$): $C_1 \approx 1635 \times 0.0025 = 4.09$. Wait—this doesn't match our earlier $C_1 = 57.9$. The discrepancy arises because the approximation $C_1 \propto (a - b)^2$ holds only for small $\delta$; at $\delta = 0.05$, the factors $(1 - b^2)^3$ and $(1 - ab)^2$ change significantly between $b = 0.90$ and $b = 0.95$. This illustrates an important practical point: the coalescence singularity's influence extends only a few percent in parameter space.

To see the approach to darkness more precisely, we tabulate $C_1$ at several values of $b$:

| $b$ | $\lvert a_1 - b\rvert$ | $C_1$ | $\lambda_c^{(1)}$ ($N$=2500) |
|---|---|---|---|
| 0.95 | 0.05 | 57.9 | 0.072 |
| 0.92 | 0.02 | 1.6 | 0.177 |
| 0.91 | 0.01 | 0.25 | 0.282 |
| 0.905 | 0.005 | 0.05 | 0.420 |
| 0.900 | 0.000 | 0 | $\infty$ |

As $b \to a_1 = 0.90$, the detection boundary $\lambda_c$ diverges—the hidden factor becomes invisible. At $b = 0.91$ (just 1% from coalescence), the detection boundary is $\lambda_c = 0.28$, meaning our coupling $\lambda = 0.15$ is already *below* the threshold—undetectable with $N = 2500$. To detect it, we would need to solve $N_{\min} \cdot C \cdot \lambda^4 = \frac{1}{2} \ln N_{\min}$ with $C = 0.25$, $\lambda = 0.15$. Iterating from an initial guess gives $N_{\min} \approx 42{,}000$ observations (∼170 years of daily data).

**Physical interpretation.** When the hidden volatility factor relaxes at the same rate as the tech index, a one-pole null model can perfectly mimic the observed spectrum by simply adjusting its pole and noise level. The hidden factor's spectral signature is absorbed into the null model's parameters. This is tangent-space absorption in action.

## 10.4 Step 3: the null model absorbs the hidden factor—pseudo-true parameters

When we fit a one-pole AR(1) to Channel 1 data generated by the true two-pole model, what parameters do we get? From Eqs. (35)–(36) (with the per-channel coupling $\lambda$):

$$\begin{aligned} \tilde{\sigma}^{2*} &= \sigma_{\varepsilon_1}^2 + \lambda^2 \frac{\sigma_\eta^2}{1-b^2} = 1.0 + 0.15^2 \times \frac{1.0}{1-0.9025} \\ &= 1.0 + 0.0225 \times 10.26 = 1.0 + 0.231 = \boxed{1.231}, \end{aligned} \tag{59}$$

and

$$\begin{aligned} \tilde{a}^* &= a_1 + \lambda^2 \frac{\sigma_\eta^2}{\sigma_{\varepsilon_1}^2} \cdot \frac{b(1-a_1^2)}{(1-a_1 b)(1-b^2)} \\ &= 0.90 + 0.0225 \times \frac{0.95 \times 0.19}{0.145 \times 0.0975} \\ &= 0.90 + 0.0225 \times \frac{0.1805}{0.01414} \\ &= 0.90 + 0.0225 \times 12.77 = 0.90 + 0.287 = \boxed{1.187}. \end{aligned} \tag{60}$$

Wait—$\tilde{a}^* = 1.187 > 1$? This would be an unstable AR(1). The issue is that the $O(\lambda^2)$ perturbative formula (36) is only valid for small $\lambda$. With our moderate coupling $\lambda = 0.15$, the factor $b(1-a_1^2)/[(1-a_1 b)(1-b^2)] = 12.77$ is large, making the $O(\lambda^2)$ correction (0.287) comparable to the leading term (0.90). The perturbative expansion has broken down.

*Remark* 10.1 (Lesson: perturbative vs. exact). This is an important cautionary lesson for the reader. The perturbative pseudo-true formulas are derived under $\lambda \ll 1$ and are most reliable when the correction is small compared to the leading term. The amplification factor $b(1-a^2)/[(1-ab)(1-b^2)]$ diverges as $b \to 1$ (slow hidden factor) or $b \to a$ (coalescence), making the perturbative regime narrower. For moderate couplings, one should minimize the exact Whittle KL divergence numerically. For our case study, exact numerical minimization gives $\tilde{a}^* \approx 0.93$ and $\tilde{\sigma}^{2*} \approx 1.18$—the null AR(1) shifts its pole toward the hidden factor's pole $b = 0.95$ and inflates its noise estimate, but remains stable.

Despite the perturbative breakdown in the pseudo-true formulas, the quartic coefficient $C_1$ and the detection boundary $\lambda_c$ computed in Steps 1–2 remain valid: they depend on the *residual norm* $\|R\|^2$, which is computed exactly from the inner products in Theorem 5.4, not from the perturbative pseudo-true parameters.

The key qualitative message survives: the null model *partially absorbs* the hidden factor by shifting its parameters. The pseudo-true pole $\tilde{a}^* \approx 0.93$ lies between the true system pole $a_1 = 0.90$ and the hidden pole $b = 0.95$—a compromise that minimizes the spectral mismatch.

## 10.5 Step 4: cross-spectral escape from darkness

Now observe *both* channels. The cross spectrum $S_{12}(\omega)$ arises because both channels share the hidden driver $F$. From the cancellation identity (Lemma 6.1) with per-channel couplings $\lambda_1 = \lambda_2 = \lambda$:

$$\frac{|S_{12}(\omega)|^2}{S_{11}^0\, S_{22}^0} = \frac{\lambda^4\, \sigma_\eta^4}{\sigma_{\varepsilon_1}^2\, \sigma_{\varepsilon_2}^2\, \mathcal{P}_b(\omega)^2}.$$

Integrating over frequency (Section 6), the cross-spectral quartic coefficient is

$$C_{\text{cross}} = \frac{\sigma_\eta^4\,(1+b^2)}{2\,\sigma_{\varepsilon_1}^2\,\sigma_{\varepsilon_2}^2\,(1-b^2)^3}.$$

(This is the formula from Theorem 6.3 with the $u_i^2$ factors set to 1, since we defined $\lambda_i = \lambda$ directly.)

**Substituting values:**

$$1 + b^2 = 1 + 0.9025 = 1.9025,$$
$$(1-b^2)^3 = 0.0975^3 = 9.27 \times 10^{-4}.$$

Therefore:

$$C_{\text{cross}} = \frac{1.0 \times 1.9025}{2 \times 1.0 \times 1.0 \times 9.27 \times 10^{-4}} = \frac{1.9025}{1.854 \times 10^{-3}} = \boxed{1026}. \tag{61}$$

**Cross-spectral detection boundary:**

$$\lambda_c^{(\text{cross})} = \left[\frac{\ln N}{2\, C_{\text{cross}}\, N}\right]^{1/4} = \left[\frac{7.824}{2 \times 1026 \times 2500}\right]^{1/4}$$
$$= \left[\frac{7.824}{5.13 \times 10^6}\right]^{1/4} = \left[1.525 \times 10^{-6}\right]^{1/4} = \boxed{0.0351}. \tag{62}$$

Compare with the single-channel boundary $\lambda_c^{(1)} = 0.072$: the cross-spectral boundary is **51% lower**, meaning much weaker couplings become detectable. The cross spectrum is a far more sensitive probe.

**At coalescence** ($b = a_1 = 0.90$):

$$C_{\text{cross}}\big|_{b=0.90} = \frac{1 + 0.81}{2 \times (1-0.81)^3} = \frac{1.81}{2 \times 0.19^3} = \frac{1.81}{0.01372} = \boxed{131.9}.$$

So $\lambda_c^{(\text{cross})}|_{\text{coal}} = [7.824/(2 \times 131.9 \times 2500)]^{1/4} = [1.186 \times 10^{-5}]^{1/4} = 0.0587$.

The auto-spectral coefficient is $C_1 = 0$ at coalescence, but $C_{\text{cross}} = 131.9 > 0$. The cross spectrum **breaks through the darkness**: with $\lambda = 0.15 > 0.059 = \lambda_c^{(\text{cross})}$, the hidden factor is detectable even when single-channel methods are provably blind.

| **Method** | $C$ | $\boldsymbol{\lambda_c}$ ($N$=2500) | **Detects** $\lambda$=0.15**?** |
|---|---|---|---|
| $b = 0.95$ *(separated timescales)* | | | |
| Auto-spectral (Ch. 1) | 57.9 | 0.072 | Yes |
| Cross-spectral | 1026 | 0.035 | Yes |
| $b = 0.90$ *(timescale coalescence)* | | | |
| Auto-spectral (Ch. 1) | 0 | $\infty$ | **No** (dark) |
| Cross-spectral | 131.9 | 0.059 | **Yes** |

## 10.6 Step 5: the thermodynamic bridge—certifying nonequilibrium

The cross-spectral signal does more than detect the hidden factor—it certifies that the system is out of equilibrium and provides a quantitative bound on the dissipation rate.

**Computing the EPR.** From Theorem 7.2, the exact entropy production rate under one-way coupling is $\dot{\Sigma} = \alpha_2 \lambda^2$. We work in continuous time via the standard correspondence (Remark 7.1). The discrete-time parameters map to continuous-time rates via $\gamma = -\ln a/\Delta t$ (with $\Delta t = 1$ day):

$$\begin{aligned}\gamma_1 &= -\ln(0.90) = 0.1054 \text{ day}^{-1},\\ \gamma_2 &= -\ln(0.85) = 0.1625 \text{ day}^{-1},\\ \gamma_f &= -\ln(0.95) = 0.05129 \text{ day}^{-1}.\end{aligned}$$

With $D_1 = D_2 = D_f = 0.5$ (matching the discrete-time noise variances via $D = \sigma^2/(2\Delta t)$), and per-channel coupling $\lambda$, the EPR coefficient is:

$$\begin{aligned}\alpha_2 &= \frac{D_f}{D_1(\gamma_1 + \gamma_f)} + \frac{D_f}{D_2(\gamma_2 + \gamma_f)}\\ &= \frac{0.5}{0.5 \times 0.1567} + \frac{0.5}{0.5 \times 0.2138}\\ &= \frac{1}{0.1567} + \frac{1}{0.2138} = 6.38 + 4.68 = \boxed{11.06 \text{ day}^{-1}}.\end{aligned} \tag{63}$$

(Here each channel contributes independently to the EPR: the first term is Channel 1's contribution, the second is Channel 2's.)

The entropy production rate is:

$$\dot{\Sigma} = \alpha_2 \lambda^2 = 11.06 \times 0.15^2 = 11.06 \times 0.0225 = \boxed{0.249 \text{ day}^{-1}}. \tag{64}$$

**The bridge equation.** From Corollary 7.3:

$$\dot{\Sigma}^2 = \frac{\alpha_2^2}{C_{\text{cross}}} D_{\text{cross}}.$$

Rearranging: $D_{\text{cross}} = C_{\text{cross}}\lambda^4 = 1026 \times 0.15^4 = 1026 \times 5.063 \times 10^{-4} = 0.519$. Then:

$$\dot{\Sigma} = \frac{\alpha_2}{\sqrt{C_{\text{cross}}}}\sqrt{D_{\text{cross}}} = \frac{11.06}{\sqrt{1026}}\sqrt{0.519} = \frac{11.06}{32.03} \times 0.720 = 0.345 \times 0.720 = 0.249 \text{ day}^{-1}.$$

The two calculations agree, confirming the bridge equation. The observable cross-spectral divergence $D_{\text{cross}} = 0.519$ translates, via the bridge, into a certified entropy production rate of $\dot{\Sigma} = 0.249$ day$^{-1}$.

**Physical interpretation.** The hidden volatility factor continuously injects energy into the two-channel system: it pushes both channels in the same direction, creating correlations that would not exist in equilibrium. A single channel sees only its own marginal dynamics, which *look* like equilibrium (the auto-spectrum is consistent with an AR(1)—detailed balance appears satisfied). But the cross-channel correlation reveals the hidden forcing: the system is dissipating entropy at rate 0.249 day$^{-1}$, and the bridge equation certifies this from the observable spectral structure alone.

## 10.7 Step 6: summary of the case study

We now collect all the results in a single narrative:

**(1) Model setup.** Two sector indices are driven by a hidden volatility factor with relaxation rate $b = 0.95$ and coupling $\lambda = 0.15$.

**(2) Single-channel detection.** From the tech index alone, the quartic coefficient is $C_1 = 57.9$, giving a detection boundary $\lambda_c = 0.072$. Since $\lambda = 0.15 > 0.072$, detection is possible with $N = 2500$ daily observations—but the margin is modest.

**(3) Approach to darkness.** As the hidden factor's timescale approaches the tech index's timescale ($b \to a_1$), the quartic coefficient $C_1 \to 0$ and the detection boundary $\lambda_c \to \infty$. At exact coalescence ($b = a_1 = 0.90$), the hidden factor is **exactly undetectable** from any single channel.

**(4) Cross-spectral escape.** Observing both channels, the cross-spectral coefficient is $C_{\text{cross}} = 1026$ (separated timescales) or $C_{\text{cross}} = 131.9$ (coalescence). In both cases, $\lambda_c^{(\text{cross})} < 0.15$, so the hidden factor is detectable via the cross spectrum. The cross spectrum breaks through the single-channel impossibility.

**(5) Thermodynamic certification.** The cross-spectral signal certifies that the system is out of equilibrium, with an entropy production rate $\dot{\Sigma} = 0.249$ day$^{-1}$. A single channel would measure zero irreversibility.

**(6) Practical implication.** In financial markets, the apparent "intrinsic persistence" of a single index may be inherited from a hidden common factor. Detecting this requires either (a) a known proxy for the factor (as in Paper (iv)), or (b) cross-spectral analysis of multiple indices sharing the factor. The quartic detection law quantifies exactly how much data is needed, and the cross-spectral bridge provides a thermodynamic interpretation.

# 11 Summary

The results assemble into a layered architecture:

| Layer | Result | Key object |
|---|---|---|
| Universal law | $D^{\min}_{\mathrm{KL,loc}} = C\lambda^4 + O(\lambda^6)$ | Normal residual $R$ |
| Structural impossibility | Single probe blind ($C = 0$ at $n$=1) | Pairing principle |
| Spectral darkness | $C \to 0$ at timescale coalescence | Tangent alignment |
| Cross-spectral escape | $C_{\mathrm{cross}} > 0$ at coalescence | Off-diagonal orthogonality |
| Thermodynamic bridge | $D_{\mathrm{cross}} > 0 \Rightarrow \dot{\Sigma} > 0$ (quantitative) | EPR certification |
| Dimension gap | Dark iff $d \geq K$ | Poisson counterexample |
| Gaussian worst case | Gaussian is the hardest family at $n$=1 | Dimension saturation |

The last row deserves emphasis. Among all exponential families with the same sufficient statistics, the Gaussian family is the *hardest* for single-probe detection because its two-parameter tangent space exactly saturates the two cumulants affected by the hidden variable. Any family with fewer parameters per cumulant (e.g., Poisson) has a dimension gap and allows single-probe detection. This "Gaussian-as-worst-case" result is a core conclusion of the general theory.

The hierarchy is:

1. A single observation channel is exactly dark for Gaussian observations—hidden forcing is indistinguishable from equilibrium at all coupling strengths and all sample sizes.

2. Two channels sharing the hidden driver restore quartic detectability via the cross spectrum, even at the coalescence singularity where auto-spectral methods fail.

3. The cross-spectral witness certifies that the full system is out of equilibrium, and the bridge equation (52) provides a quantitative bound linking the observable spectral structure to the thermodynamic arrow of time.

# 12 Exercises

These exercises are designed to reinforce the main ideas of the tutorial. They range from straightforward verification to more challenging extensions.

### General theory (Section 3)

**E1. Verify the quadratic approximation.** For $p_0 = N(0,1)$ and $p_\epsilon = N(0, 1+\epsilon)$, compute $D_{\mathrm{KL}}(p_\epsilon \| p_0)$ exactly and verify that it equals $\frac{1}{2}\epsilon^2/(2) + O(\epsilon^3)$ for small $\epsilon$.

**E2. One-parameter family.** Let the null model be $p(x|\mu) = N(\mu, 1)$ and let the true model be $p_{\mathrm{true}} = N(0, 1+\lambda^2)$. Compute $h(x)$ in the perturbation expansion, the tangent space $\mathcal{T}$, the projection $\Pi_{\mathcal{T}} h$, and the residual $R$. What is the quartic coefficient $C$? (Hint: the mean shift is fully absorbed; check whether the variance shift is also absorbed.)

**E3. Detection boundary.** Using $C = 1/2$, $\Delta k = 1$, compute the detection boundary $\lambda_c$ for $N = 10^3, 10^4, 10^5, 10^6$. Plot $\lambda_c$ vs. $N$ on a log-log scale and verify the $N^{-1/4}$ slope.

### Random-effects model (Section 4)

**E4. Verify exact darkness at $n = 1$.** Substitute $n = 1$ into $D^{\min}_{\mathrm{KL,loc}} = \frac{1}{2}[\ln(1+t) - \ln(1+t)]$ and confirm $D^{\min}_{\mathrm{KL,loc}} = 0$ identically. Explain in words why a single sensor cannot detect a hidden mean shift.

**E5. Three sensors.** For $n = 3$, $\sigma^2 = 2$, compute $C$ and the detection boundary for $N = 5000$. How many observation windows would you need to detect $\lambda = 0.2$?

### Spectral inference (Section 5)

**E6. Coalescence limit.** Show that $C(a, b) \to 0$ as $b \to a$ by substituting $b = a + \delta$ into Eq. (28) and expanding to leading order in $\delta$. What is the rate: $C \sim |\delta|^\alpha$ for what $\alpha$?

**E7. White driver.** Verify that $C = 0$ when $b = 0$ (white noise driver) by substituting into (28). Explain physically: why can a one-pole null model perfectly absorb a white-noise perturbation?

### Cross-spectral detection (Section 6)

**E8. Cancellation identity.** Verify the cancellation identity (Lemma 6.1) by writing out $|S_{12}|^2$, $S^0_{11}$, $S^0_{22}$ in terms of transfer functions and checking that $|H_1|^2$ and $|H_2|^2$ cancel.

**E9. Optimal coupling allocation.** For fixed total coupling $u_1^2 + u_2^2 = 1$, what allocation $(u_1, u_2)$ maximizes $C_{\mathrm{cross}}$? (Answer: $u_1 = u_2 = 1/\sqrt{2}$.)

# A Mathematical prerequisites

This appendix provides self-contained introductions to the mathematical tools used throughout the tutorial. Each topic includes definitions, physical intuition, worked examples, and pointers to where the tool appears in the main text. Readers comfortable with graduate-level probability and linear algebra may skip directly to the sections they need.

## A.1 Inner product spaces and $L^2$ spaces

An **inner product space** is a vector space $V$ equipped with a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ (or $\mathbb{C}$) that is linear in its first argument, symmetric ($\langle f, g \rangle = \langle g, f \rangle$), and positive definite ($\langle f, f \rangle > 0$ for $f \neq 0$). The induced **norm** is $\|f\| = \sqrt{\langle f, f \rangle}$, and the **distance** between two elements is $d(f, g) = \|f - g\|$.

The space $L^2(p_0)$ consists of all functions $f$ satisfying $\int f(x)^2 \, p_0(x) \, dx < \infty$, with inner product

$$\langle f, g \rangle_{L^2(p_0)} = \int f(x) \, g(x) \, p_0(x) \, dx.$$

When $p_0$ is a probability density, this is the expectation $E_{p_0}[f \cdot g]$.

*Where it appears.* The entire quartic law (Theorem 3.1) is derived by computing distances in $L^2(p_0)$. The perturbation function $h$ and the score functions $e_j$ live in this space, and the quartic coefficient $C = \frac{1}{2}\|R\|^2$ is a squared norm.

*Example.* Let $p_0 = N(0, 1)$ and $f(x) = x$, $g(x) = x^2 - 1$. Then $\langle f, g \rangle = E[x(x^2 - 1)] = E[x^3] - E[x] = 0$ (since odd moments of the standard normal vanish). So $f$ and $g$ are orthogonal in $L^2(p_0)$.

In the spectral (Whittle) setting, the relevant $L^2$ space has inner product

$$\langle f, g \rangle_{L^2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) \, g(\omega) \, d\omega,$$

which is the standard $L^2$ inner product on the circle $[-\pi, \pi]$ with normalized Lebesgue measure.

## A.2 Orthogonal projection in Hilbert spaces

Given a closed subspace $\mathcal{T} \subset V$ and a vector $h \in V$, the **orthogonal projection** of $h$ onto $\mathcal{T}$ is the unique element $\Pi_{\mathcal{T}} h \in \mathcal{T}$ that minimizes $\|h - v\|$ over all $v \in \mathcal{T}$. Equivalently, $\Pi_{\mathcal{T}} h$ is characterized by the condition

$$\langle h - \Pi_{\mathcal{T}} h, \, v \rangle = 0 \qquad \text{for all } v \in \mathcal{T}.$$

The **residual** $R = h - \Pi_{\mathcal{T}} h$ is perpendicular to the entire subspace. The Pythagorean theorem gives $\|h\|^2 = \|\Pi_{\mathcal{T}} h\|^2 + \|R\|^2$.

When $\mathcal{T}$ has a finite orthonormal basis $\{e_1, \ldots, e_d\}$, the projection is simply

$$\Pi_{\mathcal{T}} h = \sum_{j=1}^{d} \langle h, e_j \rangle \, e_j.$$

If the basis is orthogonal but not normalized ($\langle e_i, e_j \rangle = 0$ for $i \neq j$ but $\|e_j\| \neq 1$), the formula becomes

$$\Pi_{\mathcal{T}} h = \sum_{j=1}^{d} \frac{\langle h, e_j \rangle}{\|e_j\|^2} \, e_j.$$

For a general (non-orthogonal) basis, the projection involves the Gram matrix $G_{ij} = \langle e_i, e_j \rangle$:

$$\Pi_{\mathcal{T}} h = \sum_{i,j} \langle h, e_i \rangle \, (G^{-1})_{ij} \, e_j.$$

This is Eq. (7) in the main text.

*Where it appears.* The quartic law arises because $\Pi_{\mathcal{T}} h$ (the tangent-space projection) is absorbed by model refitting, and only $R = h - \Pi_{\mathcal{T}} h$ (the normal residual) contributes to the KL divergence. In Sections 4 and 5, we compute these projections explicitly for the Gaussian random-effects and spectral models.

## A.3 Fisher information and score functions

For a parametric family $\{p(\cdot|\theta) : \theta \in \Theta\}$, the **score function** at $\theta_0$ is

$$s(x) = \frac{\partial}{\partial \theta} \log p(x|\theta) \Big|_{\theta=\theta_0}.$$

It has zero mean: $E_{\theta_0}[s] = 0$. The **Fisher information** is its variance:

$$I(\theta_0) = E_{\theta_0}[s(x)^2] = \mathrm{Var}_{\theta_0}[s(x)] = \langle s, s \rangle_{L^2(p_0)}.$$

For multivariate parameters $\boldsymbol{\theta} \in \mathbb{R}^d$, the Fisher information matrix is $G_{ij} = E[s_i \, s_j]$ where $s_j = \partial_{\theta_j} \log p$.

*Intuition.* The score function tells you how the log-likelihood changes when you wiggle the parameter. Large Fisher information means the data is sensitive to parameter changes—good for estimation, good for detection. The Fisher information matrix is the metric tensor of the statistical manifold (in Amari's information geometry).

*Where it appears.* The tangent space $\mathcal{T} = \mathrm{span}\{e_1, \ldots, e_d\}$ in Section 3 is exactly the span of the score functions. The Gram matrix of the projection (7) is the Fisher information matrix.

*Example.* For $p(x|\mu, \sigma^2) = N(\mu, \sigma^2)$, the scores are $s_\mu = (x - \mu)/\sigma^2$ and $s_{\sigma^2} = [(x-\mu)^2 - \sigma^2]/(2\sigma^4)$. The Fisher information matrix is $G = \mathrm{diag}(1/\sigma^2, \, 1/(2\sigma^4))$.

## A.4 Contour integration on the unit circle

Many spectral integrals in this tutorial have the form $\frac{1}{2\pi} \int_{-\pi}^{\pi} R(e^{i\omega}) \, d\omega$, where $R$ is a rational function. These are evaluated by the substitution $z = e^{i\omega}$, $d\omega = dz/(iz)$, converting the integral to a contour integral over the unit circle $|z| = 1$.

*Method.* The pole polynomial $\mathcal{P}_c(\omega) = 1 + c^2 - 2c \cos \omega$ factors as $\mathcal{P}_c = (z - c)(z^{-1} - c)$ on the unit circle. For $|c| < 1$, the pole at $z = c$ is inside the circle and the pole at $z = 1/c$ is outside. By the residue theorem:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) \, d\omega = \frac{1}{2\pi i} \oint_{|z|=1} f\left( \frac{z + z^{-1}}{2} \right) \frac{dz}{iz} = \sum_{\text{poles inside}} \mathrm{Res}(\cdots).$$

*Key identities* (all require $|c| < 1$):

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{d\omega}{\mathcal{P}_c(\omega)} = \frac{1}{1-c^2}, \tag{65}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\cos\omega}{\mathcal{P}_c(\omega)}\,d\omega = \frac{c}{1-c^2}, \tag{66}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{d\omega}{\mathcal{P}_c(\omega)^2} = \frac{1+c^2}{(1-c^2)^3}. \tag{67}$$

*Derivation of* (65). Write $\mathcal{P}_c = (z-c)(1-cz)/z$ on $|z| = 1$. The integrand becomes $z/[(z-c)(1-cz)]$. Pole inside: $z = c$ with residue $c/[(1-c^2)] \cdot 1/1 = 1/(1-c^2)$... Actually, the full calculation: the integral is $\frac{1}{2\pi i}\oint \frac{dz}{(z-c)(1-cz)} = \mathrm{Res}_{z=c}\frac{1}{(z-c)(1-cz)} = \frac{1}{1-c^2}$.

*Where it appears.* Every spectral quartic coefficient computation in Sections 5–6 relies on these integrals.

## A.5 Taylor expansion and order notation

For a smooth function $f(\lambda)$, the **Taylor expansion** around $\lambda = 0$ is

$$f(\lambda) = f(0) + f'(0)\lambda + \frac{f''(0)}{2}\lambda^2 + \cdots + \frac{f^{(n)}(0)}{n!}\lambda^n + O(\lambda^{n+1}).$$

The notation $O(\lambda^k)$ means "a quantity bounded by $C|\lambda|^k$ for some constant $C$ and all $|\lambda|$ sufficiently small."

*Key expansions used in this tutorial:*

$$\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \cdots, \qquad |t| < 1, \tag{68}$$

$$\frac{1}{1+t} = 1 - t + t^2 - \cdots, \qquad |t| < 1, \tag{69}$$

$$(1+t)^{-1/2} = 1 - \frac{t}{2} + \frac{3t^2}{8} - \cdots. \tag{70}$$

*Where it appears.* The pairing principle (Theorem 4.2) uses (68) to extract the quartic term from $\frac{1}{2}[n\ln(1+t) - \ln(1+nt)]$. The detection boundary (11) uses the quartic scaling $D_{\mathrm{KL,loc}}^{\min} \sim C\lambda^4$ to solve for $\lambda_c^{\mathrm{pop}}$.

## A.6 Matrix determinant lemma and Woodbury identity

The **matrix determinant lemma** states that for an invertible $n \times n$ matrix $A$ and vectors $u, v \in \mathbb{R}^n$:

$$\det(A + uv^\top) = (1 + v^\top A^{-1}u)\det A. \tag{71}$$

*Special case.* For $A = \sigma^2 I_n$ and $u = v = \lambda\mathbf{1}$:

$$\det(\sigma^2 I_n + \lambda^2 \mathbf{1}\mathbf{1}^\top) = \sigma^{2n}(1 + n\lambda^2/\sigma^2).$$

The related **Woodbury matrix identity** gives the inverse:

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}. \tag{72}$$

For the rank-one case: $(\sigma^2 I + \lambda^2 \mathbf{1}\mathbf{1}^\top)^{-1} = \frac{1}{\sigma^2} I - \frac{\lambda^2}{\sigma^2(\sigma^2+n\lambda^2)}\mathbf{1}\mathbf{1}^\top$.

*Where it appears.* The exact KL divergence for the random-effects model (Lemma 4.1) uses the matrix determinant lemma to compute $\det \Sigma_1$.

## A.7 Properties of Kullback–Leibler divergence

Beyond the definition (Eq. 1), several properties of KL divergence are used:

(i) **Non-negativity**: $D_{\mathrm{KL}}(p\|q) \geq 0$, with equality iff $p = q$ a.e.

(ii) **Asymmetry**: $D_{\mathrm{KL}}(p\|q) \neq D_{\mathrm{KL}}(q\|p)$ in general.

(iii) **Multivariate Gaussian formula**: For $p_i = N(\mu_i, \Sigma_i)$:

$$D_{\mathrm{KL}}(p_1\|p_2) = \frac{1}{2}\Big[\mathrm{tr}(\Sigma_2^{-1}\Sigma_1) - n + \ln\frac{\det\Sigma_2}{\det\Sigma_1} + (\mu_2-\mu_1)^\top\Sigma_2^{-1}(\mu_2-\mu_1)\Big]. \tag{73}$$

(iv) **Additivity for independent variables**: If $p = p_1 \otimes p_2$ and $q = q_1 \otimes q_2$, then $D_{\mathrm{KL}}(p\|q) = D_{\mathrm{KL}}(p_1\|q_1) + D_{\mathrm{KL}}(p_2\|q_2)$.

(v) **Chi-squared approximation**: For $p = p_0(1+\epsilon)$ with $\|\epsilon\|$ small:

$$D_{\mathrm{KL}}(p\|p_0) \approx \frac{1}{2}\int \epsilon(x)^2\, p_0(x)\, dx = \frac{1}{2}\chi^2(p\|p_0).$$

*Where it appears.* Property (iii) is used in the random-effects model (Lemma 4.1). Property (v) is the local quadratic approximation (3) that underpins the entire quartic law.

## A.8 Exponential families and sufficient statistics

An **exponential family** is a parametric family of the form

$$p(x|\boldsymbol{\theta}) = h(x)\exp\big(\boldsymbol{\theta}^\top T(x) - A(\boldsymbol{\theta})\big),$$

where $\boldsymbol{\theta} \in \mathbb{R}^d$ is the **natural parameter**, $T(x) \in \mathbb{R}^d$ is the **sufficient statistic**, $A(\boldsymbol{\theta}) = \ln\int h(x)e^{\boldsymbol{\theta}^\top T(x)}\,dx$ is the **log-partition function**, and $h(x)$ is the base measure.

Key properties:

- The mean of the sufficient statistic is $E_\theta[T(x)] = \nabla A(\boldsymbol{\theta})$.
- The covariance is $\mathrm{Cov}_\theta[T(x)] = \nabla^2 A(\boldsymbol{\theta})$, which equals the Fisher information matrix in the natural parametrization.
- The family is **minimal** if the components of $T$ are affinely independent.
- The **cumulants** of the distribution are derivatives of $A$: $\kappa_r = \partial^r A/\partial\theta^r$.

*Examples.*

| Family | $\boldsymbol{\theta}$ | $T(x)$ | $d$ |
|---|---|---|---|
| Gaussian $N(\mu, \sigma^2)$ | $(\mu/\sigma^2,\ -1/(2\sigma^2))$ | $(x,\ x^2)$ | 2 |
| Poisson $\mathrm{Pois}(\lambda)$ | $\ln \lambda$ | $x$ | 1 |
| Negative Binomial | $(\ln p,\ \ln(1-p))$ | $(x,\ r)$ | 2 |

*Where it appears.* The dimension-gap principle (Theorem 8.1) classifies darkness by comparing $d = \dim(\mathcal{T})$ (the number of free parameters) with $K$ (the number of independently affected cumulants). The tangent space $\mathcal{T} = \mathrm{span}\{T_k - \mu_k\}$ is spanned by the centered sufficient statistics.

## A.9 The Lyapunov equation

For a linear stochastic system $d\boldsymbol{X} = A\boldsymbol{X}\,dt + \sqrt{D}\,d\boldsymbol{W}$, the steady-state covariance $\Sigma = \lim_{t\to\infty} \mathrm{Cov}[\boldsymbol{X}(t)]$ satisfies the **continuous Lyapunov equation**:

$$A\Sigma + \Sigma A^\top + D = 0, \tag{74}$$

provided all eigenvalues of $A$ have negative real parts (the system is stable).

*Intuition.* The Lyapunov equation expresses a balance: the drift matrix $A$ dissipates variance (the terms $A\Sigma + \Sigma A^\top$ have negative trace for stable $A$), while the diffusion matrix $D$ injects variance. At steady state, injection equals dissipation.

For a $2 \times 2$ system with $A = \mathrm{diag}(-\gamma_1, -\gamma_2)$ and $D = \mathrm{diag}(2D_1, 2D_2)$:

$$\Sigma = \mathrm{diag}\left(\frac{D_1}{\gamma_1},\ \frac{D_2}{\gamma_2}\right).$$

Off-diagonal coupling in $A$ produces off-diagonal elements in $\Sigma$.

*Where it appears.* The EPR proof sketch in Section 7 uses the Lyapunov equation to compute the steady-state covariance of the 3D Ornstein–Uhlenbeck system $(X_1, X_2, F)$.

## A.10 Entropy production rate in linear systems

For a linear Gaussian system $d\boldsymbol{X} = A\boldsymbol{X}\,dt + \sqrt{D}\,d\boldsymbol{W}$ at steady state, the **entropy production rate** (EPR) quantifies the irreversibility—the rate at which the system generates thermodynamic entropy due to nonequilibrium driving.

At equilibrium, detailed balance holds: $A\Sigma = \Sigma A^\top$ (equivalently, $D^{-1}A\Sigma$ is symmetric). When this condition is violated, the EPR is

$$\dot{\Sigma} = -\operatorname{tr}\big(A\,D^{-1}A\,\Sigma\big) + \operatorname{tr}\big(D^{-1}A\,\Sigma\,A^\top D^{-1}D\big) = \operatorname{tr}\big[(D^{-1}A\Sigma)^2_{\mathrm{antisym}} \cdot D\big] \geq 0, \tag{75}$$

where the antisymmetric part captures the broken detailed balance.

*Intuition.* EPR measures the "restless current" in the system: probability flowing in loops rather than relaxing to equilibrium. A system with hidden forcing has nonzero EPR even if each observed channel *appears* to be in equilibrium. The cross-spectral bridge (Corollary 7.3) shows that cross-channel correlations can detect this hidden irreversibility.

*Key fact for one-way coupling.* If $A$ is block-triangular (the hidden driver $F$ does not receive feedback from the observed channels), then the steady-state of $F$ is independent of the coupling strength $\lambda$, and the EPR simplifies to $\dot{\Sigma} = \alpha_2 \lambda^2$ (exact for all $\lambda$).

*Where it appears.* The thermodynamic bridge in Section 7 connects $\dot{\Sigma}$ to the cross-spectral detectability $D_{\text{cross}}$, providing a physical interpretation of the quartic detection law.

### A.11 Jensen's formula and the logarithmic integral

**Jensen's formula** relates the integral of the logarithm of an analytic function over the unit circle to its zeros inside the circle. For our purposes, the key consequence is:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \mathcal{P}_a(\omega)\, d\omega = 0 \qquad \text{for } |a| < 1, \tag{76}$$

where $\mathcal{P}_a(\omega) = |1 - ae^{-i\omega}|^2$. This follows because $\ln|1 - ae^{-i\omega}|$ has no singularity inside the unit circle when $|a| < 1$, and the mean value property of harmonic functions gives $\frac{1}{2\pi} \int \ln|1 - ae^{-i\omega}|\, d\omega = \ln|1| = 0$.

*Why it matters.* Differentiating (76) with respect to $a$ gives $\frac{1}{2\pi} \int \frac{2(a - \cos\omega)}{\mathcal{P}_a}\, d\omega = 0$, confirming $\langle \tilde{e}_1, \tilde{e}_2 \rangle = 0$ (Lemma 5.3). Differentiating *twice*:

$$\frac{4}{2\pi} \int \frac{(\cos\omega - a)^2}{\mathcal{P}_a^2}\, d\omega = \frac{2}{2\pi} \int \frac{d\omega}{\mathcal{P}_a} = \frac{2}{1 - a^2},$$

which gives $\|\tilde{e}_2\|^2 = 2/(1 - a^2)$.

*Where it appears.* The orthogonality proof in Section 5 and the tangent-norm calculation both rely on Jensen's formula and its derivatives.

## References


[1] J. Chen, Ann. Stat. **23**, 221 (1995).

[2] S. G. Self and K.-Y. Liang, J. Am. Stat. Assoc. **82**, 605 (1987).

[3] S. Watanabe, *Algebraic Geometry and Statistical Learning Theory* (Cambridge University Press, 2009).

[4] P. Whittle, Ark. Mat. **2**, 423 (1953).

[5] M. B. Priestley, *Spectral Analysis and Time Series* (Academic Press, 1981).

[6] B. Efron, Ann. Stat. **3**, 1189 (1975).

[7] D. Lucente, A. Baldassarri, A. Puglisi, A. Vulpiani, and M. Viale, Phys. Rev. Res. **4**, 043103 (2022).

[8] K. Hasselmann, Tellus **28**, 473 (1976).

[9] C. Frankignoul and K. Hasselmann, Tellus **29**, 289 (1977).

[10] C. Penland and P. D. Sardeshmukh, J. Climate **8**, 1999 (1995).

[11] R. Zwanzig, Phys. Rev. **124**, 983 (1961).

[12] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).

[13] A. J. Chorin, O. H. Hald, and R. Kupferman, Proc. Natl. Acad. Sci. USA **97**, 2968 (2000).

[14] D. Givon, R. Kupferman, and A. Stuart, Nonlinearity **17**, R55 (2004).

[15] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Phys. Rev. Lett. **108**, 220601 (2012).

[16] L. Le Cam, Univ. California Publ. Stat. **3**, 37 (1960).

[17] L. Le Cam, *Asymptotic Methods in Statistical Decision Theory* (Springer, 1986).

[18] D. Dacunha-Castelle and E. Gassiat, Ann. Stat. **27**, 1178 (1999).

[19] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, and J. A. Wellner, *Efficient and Adaptive Estimation for Semiparametric Models* (Springer, 1998) originally Johns Hopkins UP, 1993.

[20] A. W. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, 1998).

[21] J. C. Escanciano, Econometric Theory **38**, 301 (2022).

[22] S.-i. Amari and H. Nagaoka, *Methods of Information Geometry* (American Mathematical Society, 2000).

[23] S.-i. Amari, *Information Geometry and Its Applications* (Springer, 2016).

[24] I. Csiszár and F. Matúš, IEEE Trans. Inform. Theory **49**, 1474 (2003).

[25] R. E. Kass and P. W. Vos, *Geometrical Foundations of Asymptotic Inference* (Wiley, 1997).

[26] H. White, Econometrica **50**, 1 (1982).

[27] Q. H. Vuong, Econometrica **57**, 307 (1989).

[28] J. Baik, G. Ben Arous, and S. Péché, Ann. Probab. **33**, 1643 (2005).

[29] I. M. Johnstone, Ann. Stat. **29**, 295 (2001).

[30] M. Drton, Ann. Stat. **37**, 979 (2009).

[31] M. Drton, K. Kubjas, and R. Weisman, arXiv:2511.15419 (2025).

[32] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[33] M. Esposito, Phys. Rev. E **85**, 041125 (2012).

[34] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).

[35] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).

[36] A. Dechant and S.-i. Sasa, Phys. Rev. Res. **3**, L042012 (2021).

[37] N. Ohga, S. Ito, and A. Kolchinsky, Phys. Rev. Lett. **131**, 077101 (2023).

[38] Y. Bi and V. D. Calhoun, Phys. Rev. Lett. (2026), companion paper.