

Learning notes: Gaussian measures, equivalence, and the aquifer prior

Raj Mahat

Started 14 October 2026 – compiled October 14, 2026

These notes go slowly through the mathematics behind the `gw_demarcation` experiment. Every figure is produced by a Jupyter notebook in this folder, so each statement can be checked by running code. Chapter 1 pairs with `01_feldman_hajek.ipynb`.

Contents

1	Feldman–Hájek: when are two Gaussian priors “the same kind of thing”?	1
1.1	Equivalent and singular measures	1
1.2	Infinitely many independent coordinates	2
1.3	Mean shifts and the Cameron–Martin space	4
1.4	The Feldman–Hájek theorem	6
1.5	Random functions: range versus amplitude	7
1.6	Back to the aquifer, in one dimension	8
1.7	Take-aways	10

1 Feldman–Hájek: when are two Gaussian priors “the same kind of thing”?

Why we care. In the project we have a reference prior P , the stationary SPDE $(\kappa_0^2 - \Delta)u = \xi$, and alternatives Q in which wells, a leakage bump, or a conductivity lens are added. Whenever we compare, learn, or sample, we implicitly use the density ratio dQ/dP : likelihood ratios, KL divergences, MCMC acceptance ratios, “correct the reference by a few directions”. On a fixed mesh this ratio always exists. The question is whether it *survives mesh refinement*. If it does not, every one of those quantities blows up as the mesh gets finer, and no fixed-size correction can repair the reference. The Feldman–Hájek theorem tells us exactly when this happens.

Intuition. Two Gaussian measures are equivalent if and only if a single, infinitely resolved sample cannot tell you which of the two produced it. If one perfect sample *can* tell them apart, they are mutually singular.

Everything below builds toward this sentence and then checks it with numbers.

1.1 Equivalent and singular measures

Definition 1.1. Let P, Q be probability measures on the same space.

- Q is *absolutely continuous* w.r.t. P ($Q \ll P$) if $P(E) = 0 \Rightarrow Q(E) = 0$ for every event E .

- P and Q are *equivalent* ($P \sim Q$) if $Q \ll P$ and $P \ll Q$: they have the same impossible events.
- P and Q are *mutually singular* ($P \perp Q$) if there is an event E with $P(E) = 1$ and $Q(E) = 0$: they live on disjoint sets.

By the Radon–Nikodym theorem, $Q \ll P$ exactly when there is a density $\frac{dQ}{dP}$ with $Q(E) = \int_E \frac{dQ}{dP} dP$. For equivalent measures the density is finite and positive (P -almost surely).

Finite dimensions: equivalence is the default. $P = \mathcal{N}(0, 1)$ and $Q = \mathcal{N}(0, 1.5^2)$ have

$$\frac{dQ}{dP}(x) = \frac{1}{1.5} \exp\left(\frac{x^2}{2}(1 - 1.5^{-2})\right),$$

which is finite and positive for every x (Fig. 1b). The only way two Gaussians on $\mathbb{R}^n$ can be singular is if one is degenerate and lives on a lower-dimensional subspace (Fig. 1c). In general, two *non-degenerate* Gaussians on $\mathbb{R}^n$ are always equivalent, however different their means and covariances are.

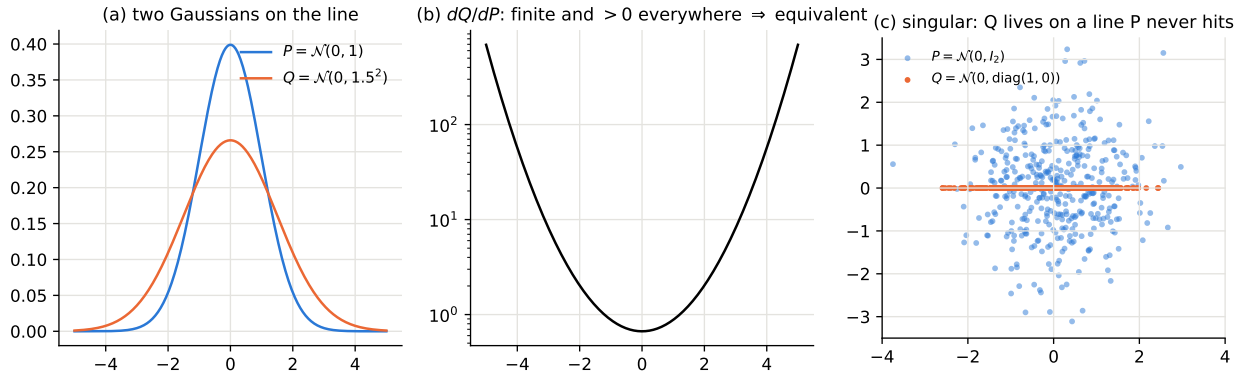


Figure 1: Finite dimensions. (a,b) Two non-degenerate Gaussians are equivalent: the density ratio is finite and positive everywhere. (c) A degenerate Gaussian sits on a line, which has probability 0 under $\mathcal{N}(0, I_2)$, so the two measures are singular.

So singularity is an *infinite-dimensional* phenomenon. On a function space, as on an ever-finer mesh, it arises without any degeneracy.

1.2 Infinitely many independent coordinates

Take $x = (x_1, x_2, \dots)$ with

$$P = \mathcal{N}(0, I), \quad Q = \mathcal{N}(0, \text{diag}(1 + t_j)), \quad t_j > -1.$$

t_j is the relative change of variance in coordinate j . Restricted to the first n coordinates, both measures are non-degenerate Gaussians and the log density ratio of one sample is a sum of independent terms:

$$\ell_n(x) = \log \frac{dQ}{dP}(x) = \sum_{j=1}^n \left[-\frac{1}{2} \log(1 + t_j) + \frac{1}{2} \frac{t_j}{1 + t_j} x_j^2 \right]. \quad (1)$$

Mean and spread of the one-sample statistic. Using $\mathbb{E}x_j^2 = 1$ under P , $\mathbb{E}x_j^2 = 1 + t_j$ under Q and $\text{Var}(x_j^2) = 2(\mathbb{E}x_j^2)^2$:

$$\begin{aligned}\mathbb{E}_Q \ell_n &= \frac{1}{2} \sum_j (t_j - \log(1 + t_j)) = \text{KL}(Q\|P) \approx \frac{1}{4} \sum_j t_j^2, \\ \mathbb{E}_P \ell_n &= -\frac{1}{2} \sum_j \left(\log(1 + t_j) - \frac{t_j}{1+t_j} \right) = -\text{KL}(P\|Q) \approx -\frac{1}{4} \sum_j t_j^2, \\ \text{Var}_P \ell_n &= \frac{1}{2} \sum_j \left(\frac{t_j}{1+t_j} \right)^2 \approx \frac{1}{2} \sum_j t_j^2,\end{aligned}$$

where $\approx$ is for small t_j . The two histograms of ℓ_n (one for samples from P , one for samples from Q) are therefore separated by about $\frac{1}{2} \sum t_j^2$, while each has width about $\sqrt{\frac{1}{2} \sum t_j^2}$. Their separation in units of width is

$$\frac{\text{gap}}{\text{width}} \approx \sqrt{\frac{1}{2} \sum_{j \leq n} t_j^2}.$$

Proposition 1.2 (Diagonal Feldman–Hájek). *If $\sum_j t_j^2 = \infty$, one sample identifies the measure with certainty as $n \rightarrow \infty$, so $P \perp Q$. If $\sum_j t_j^2 < \infty$, the histograms keep a fixed overlap and $P \sim Q$.*

The rigorous version is **Kakutani’s theorem** (1948). Products of equivalent measures $\bigotimes P_j$ and $\bigotimes Q_j$ are equivalent iff the product of Hellinger affinities $\prod_j \rho_j$ is positive, and otherwise singular. For one Gaussian coordinate

$$\rho_j = \int \sqrt{p_j q_j} dx = \left(\frac{2\sqrt{1+t_j}}{2+t_j} \right)^{1/2}, \quad -\log \rho_j = \frac{t_j^2}{16} + O(t_j^3),$$

so $\prod \rho_j > 0 \iff \sum t_j^2 < \infty$. Note the *dichotomy*: there is no in-between. Gaussian measures are either equivalent or singular.

i.i.d. case. With $t_j = 0.2$ for all j , every coordinate is a harmless 20% change, yet $\sum t_j^2 = \infty$. Fig. 2 shows the best one-sample test, “say Q iff $\ell_n > 0$ ”, going from coin-flipping at $n = 1$ to never erring at $n = 10^4$. The reason is the law of large numbers: $\frac{1}{n} \sum x_j^2 \rightarrow 1$ under P and $\rightarrow 1.2$ under Q . These two events are disjoint and carry probability one under their own measure.

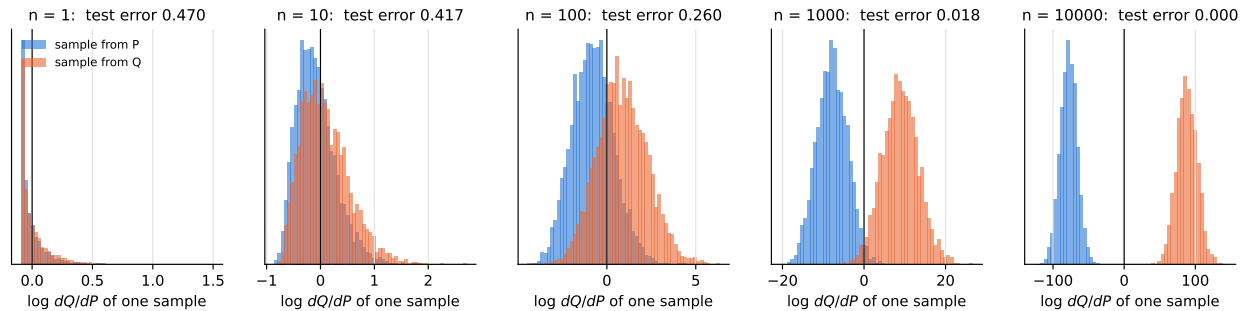


Figure 2: $t_j = 0.2$ in every coordinate. Histograms of the one-sample log likelihood ratio (1) under P (blue) and Q (orange). The title gives the error of the best test. As the number of coordinates grows the histograms separate completely: singular.

Decaying perturbations. Fig. 3 compares three sequences: $t_j = 0.2$ ($\sum t_j^2 = \infty$, fast), $t_j = j^{-1/4}$ ($\sum t_j^2 = \sum j^{-1/2} = \infty$, slow), and $t_j = 1/j$ ($\sum t_j^2 = \pi^2/6$). For $t_j = 1/j$ the affinity, the test error and the histograms all *freeze* as n grows: extra coordinates bring no new evidence. That is what equivalence looks like.

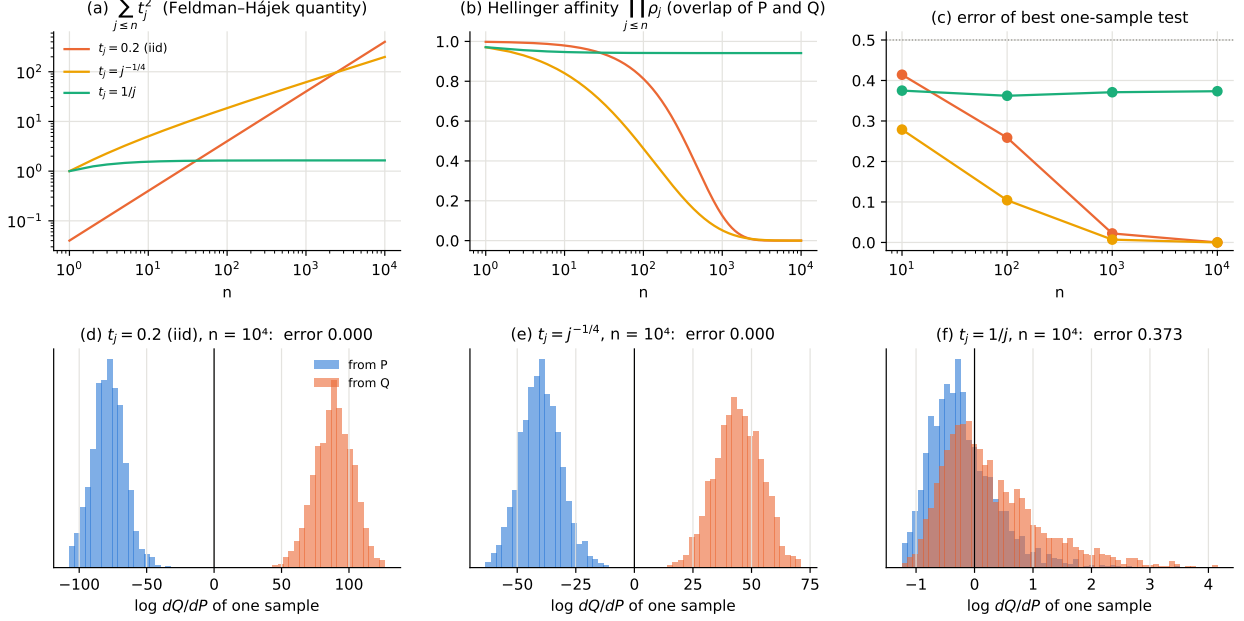


Figure 3: (a) The Feldman–Hájek quantity $\sum_{j \leq n} t_j^2$. (b) The overlap of P and Q (Hellinger affinity). (c) The best one-sample test error. (d–f) Histograms at $n = 10^4$. Only $t_j = 1/j$, with a finite sum, keeps a permanent overlap.

1.3 Mean shifts and the Cameron–Martin space

So far P and Q had the same mean and different variances. Now keep the covariance and move the mean:

$$P = \mathcal{N}(0, I_n), \quad Q = \mathcal{N}(m, I_n), \quad m = (m_1, \dots, m_n).$$

Every x_j still has variance 1; under Q it is centred at m_j instead of 0.

The one-sample statistic. Writing out the two Gaussian densities, the x_j^2 terms cancel and

$$\ell_n(x) = \log \frac{dQ}{dP}(x) = \sum_{j=1}^n \left(m_j x_j - \frac{1}{2} m_j^2 \right) = \langle m, x \rangle - \frac{1}{2} \|m\|^2.$$

This is a *matched filter*: project the sample onto the direction m and compare with the midpoint. It is linear in x , so it is exactly Gaussian:

$$\begin{aligned} x \sim P : \quad \langle m, x \rangle &\sim \mathcal{N}(0, \|m\|^2) & \Rightarrow \quad \ell_n &\sim \mathcal{N}\left(-\frac{1}{2}\|m\|^2, \|m\|^2\right), \\ x \sim Q : \quad \langle m, x \rangle &= \|m\|^2 + \langle m, z \rangle & \Rightarrow \quad \ell_n &\sim \mathcal{N}\left(+\frac{1}{2}\|m\|^2, \|m\|^2\right). \end{aligned}$$

The two histograms are separated by $\|m\|^2$ and each has width $\|m\|$. Gap over width is $\|m\|$, and the best test (say Q iff $\ell_n > 0$) has error $\Phi(-\|m\|/2)$.

Geometric picture. Along the direction $m/\|m\|$ the noise has standard deviation 1 (white noise has the same spread in every direction), while the signal has length $\|m\|$. The signal-to-noise ratio is $\|m\|$. Each coordinate contributes m_j^2 to $\|m\|^2$, so m_j^2 is the evidence coordinate j supplies, just as t_j^2 was for variances.

$$P \sim Q \iff \|m\|^2 = \sum_j m_j^2 < \infty.$$

Examples (Fig. 4).

- $m_j = 1/j$: $\|m\|^2 \rightarrow \pi^2/6 \approx 1.64$, $\|m\| \approx 1.28$. The error freezes at $\Phi(-0.64) \approx 0.26$ however many coordinates are added: **equivalent**.
- $m_j = 1/\sqrt{j}$: $\|m\|^2 = \sum 1/j \approx \log n + 0.58 \rightarrow \infty$. The error decays to 0, but only logarithmically: at $n = 10^4$, $\|m\| \approx 3.1$ and the error is still ≈ 0.06 . Even so, it is **singular**. Each individual shift $m_j \rightarrow 0$, but not fast enough.

When the coordinates are not white: the Cameron–Martin space. For a general prior $P = \mathcal{N}(0, C)$ with eigenpairs (λ_j, ϕ_j) , coordinate j has noise standard deviation $\sqrt{\lambda_j}$. A shift $m_j = \langle m, \phi_j \rangle$ must be measured in units of that noise: whitening gives $\tilde{m}_j = m_j/\sqrt{\lambda_j}$, and

$$P \sim Q \iff \sum_j \frac{m_j^2}{\lambda_j} = \|C^{-1/2}m\|^2 < \infty \iff m \in \mathcal{H}_P := C^{1/2}H.$$

$\mathcal{H}_P$ is the *Cameron–Martin space* of P . Because $\lambda_j \rightarrow 0$ quickly for a smooth prior, fine-scale coordinates are expensive. A shift that has any appreciable fine-scale content is fatal, because the prior has almost no noise there to hide it in.

For our SPDE prior. $(\kappa^2 - \Delta)u = \xi$ has $\lambda_j = (\kappa^2 + \mu_j)^{-2}$, where the μ_j are the Dirichlet Laplacian eigenvalues. Then

$$\|C^{-1/2}m\|^2 = \sum_j (\kappa^2 + \mu_j)^2 m_j^2 = \|(\kappa^2 - \Delta)m\|_{L^2}^2,$$

so $\mathcal{H}_P = \{m : (\kappa^2 - \Delta)m \in L^2, m = 0 \text{ on } \partial\Omega\}$, the Sobolev space H^2 with boundary conditions. Two consequences follow.

- **A typical sample is *not* in its own Cameron–Martin space.** For a sample u , $(\kappa^2 - \Delta)u = \xi$ is white noise, and $\|\xi\|_{L^2}^2 = \sum \xi_j^2 = \infty$. So shifting the prior by one of its own typical draws gives a singular measure. The Cameron–Martin space is much smoother than the samples: it has probability zero under P .
- **Deterministic wells.** A known pumping $q = \xi + a g$ shifts the mean by $m = a(\kappa^2 - \Delta)^{-1}g$. Then $\|C^{-1/2}m\| = |a| \|g\|_{L^2}$, so the shifted prior is equivalent iff $g \in L^2$. A smooth or extended well is fine. A *point* well $g = \delta_{x_0}$ is not in L^2 (in any dimension), so its mean shift is singular. On a mesh, δ becomes a spike of height $1/h^d$ at one node, and $\|g\|^2 \approx h^{-d} \rightarrow \infty$.

The *random* wells of the project ($a \sim \mathcal{N}(0, c)$) give the covariance version of the same condition. In whitened coordinates the added eigenvalue is $t = c \|g\|_{L^2}^2$, finite iff $g \in L^2$ (§1.6: $t = c h g^\top g$).

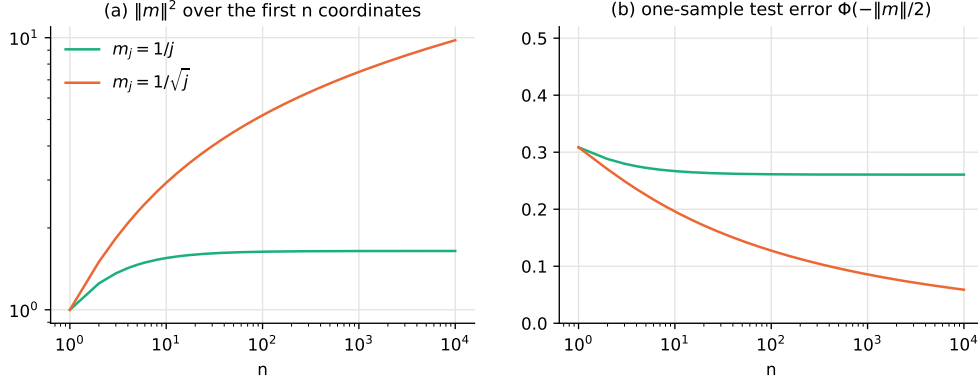


Figure 4: Mean shifts. $m_j = 1/j$ has finite $\|m\|$, so the measures are equivalent. $m_j = 1/\sqrt{j}$ has infinite $\|m\|$, so they are singular.

1.4 The Feldman–Hájek theorem

Remove the diagonal assumption. Let H be a separable Hilbert space ($L^2(\Omega)$ for us), and $P = \mathcal{N}(m_0, C_0)$, $Q = \mathcal{N}(m_1, C_1)$ Gaussian measures with injective trace-class covariances.

Theorem 1.3 (Feldman 1958, Hájek 1958). *P and Q are either equivalent or mutually singular. They are equivalent if and only if*

- (i) $C_0^{1/2}H = C_1^{1/2}H$ (same Cameron–Martin space);
- (ii) $m_1 - m_0 \in C_0^{1/2}H$;
- (iii) $T := C_0^{-1/2}C_1C_0^{-1/2} - I$ is a Hilbert–Schmidt operator, i.e. $\|T\|_{HS}^2 = \sum_j t_j^2 < \infty$, where t_j are the eigenvalues of T .

See Da Prato & Zabczyk (2014, Thm. 2.25) or Bogachev (1998, Thm. 2.7.2).

How to read it: whitening. Let $z = C_0^{-1/2}(u - m_0)$. Under P , z is white noise, $\mathcal{N}(0, I)$. Under Q (taking $m_1 = m_0$), $z \sim \mathcal{N}(0, I + T)$. In the eigenbasis of T this is exactly the diagonal problem of the previous subsection, with the t_j now the eigenvalues of T . Condition (iii) is $\sum t_j^2 < \infty$. Condition (ii) is $\sum m_j^2 < \infty$ in whitened coordinates. Condition (i) is the technical requirement that T is well defined and $I + T$ is invertible, so no t_j reaches -1 .

KL divergence. When $P \sim Q$ (zero means),

$$\text{KL}(Q\|P) = \frac{1}{2} \sum_j (t_j - \log(1 + t_j)) = \frac{1}{2} (\text{tr } T - \log \det(I + T)), \quad (2)$$

which is finite exactly when $\sum t_j^2 < \infty$, since $t - \log(1 + t) \approx t^2/2$. If $P \perp Q$, then $\text{KL} = \infty$.

The mesh-refinement test. On an N -node mesh everything is finite. Compute the $N \times N$ matrix T_N and watch $\|T_N\|_F$ and KL_N as N grows. A plateau means $P \sim Q$. Growth without bound means $P \perp Q$. This is the diagnostic used in the main project (figures F1, F2).

1.5 Random functions: range versus amplitude

Take the 1D version of the reference prior on $[0, 1]$ with $u(0) = u(1) = 0$:

$$(\kappa^2 - \partial_x^2) u = \tau \xi.$$

The sine functions $\phi_j(x) = \sqrt{2} \sin(j\pi x)$ are eigenfunctions of $-\partial_x^2$ with $\lambda_j = (j\pi)^2$. Hence

$$u = \sum_j c_j \phi_j, \quad c_j \sim \mathcal{N}\left(0, \frac{\tau^2}{(\kappa^2 + \lambda_j)^2}\right) \text{ independent.}$$

Two such priors are diagonal in the same basis, with $1 + t_j = \text{var}_Q(c_j)/\text{var}_P(c_j)$.

- **Range change** $\kappa_0 \rightarrow \kappa_1$ (with τ fixed): $t_j = \left(\frac{\kappa_0^2 + \lambda_j}{\kappa_1^2 + \lambda_j}\right)^2 - 1 \approx \frac{2(\kappa_0^2 - \kappa_1^2)}{\pi^2 j^2}$. So $\sum t_j^2 < \infty$: **equivalent**.
- **Amplitude change** $\tau_0 \rightarrow \tau_1$: $t_j = \tau_1^2/\tau_0^2 - 1$ for every j . This is the i.i.d. case: **singular**.

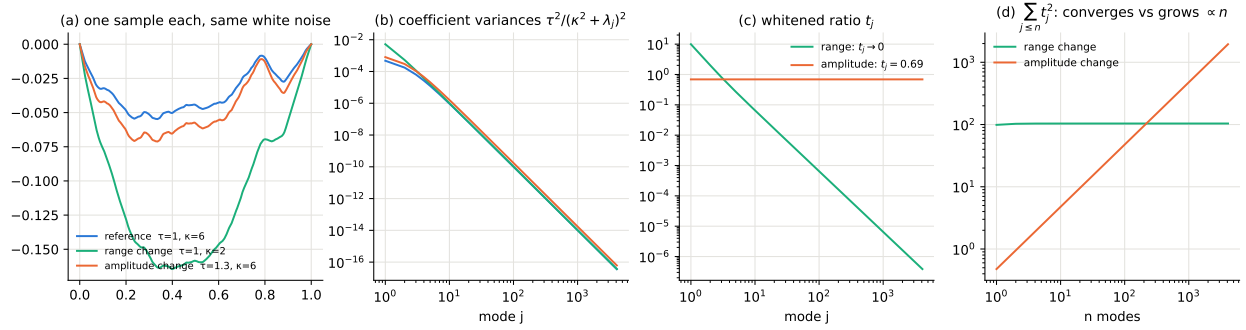


Figure 5: (a) One sample from the reference ($\tau = 1, \kappa = 6$), a range change ($\kappa = 2$) and an amplitude change ($\tau = 1.3$), all driven by the same white noise. (b) Coefficient variances: the range change merges with the reference at high modes, while the amplitude change stays parallel forever. (c) t_j . (d) $\sum_{j=1}^n t_j^2$ converges (to ≈ 104 , dominated by the first mode) for the range change and grows linearly for the amplitude change.

The surprise (Fig. 5a). The range change looks the most different to the eye, with bigger and smoother excursions, yet it is the *equivalent* one. The amplitude change looks almost like the reference, yet it is *singular*. Equivalence is decided by the **fine scales**, because there are infinitely many of them. A large-scale difference lives in finitely many modes, however dramatic it looks, and a single sample can never confirm it with certainty.

Higher dimensions. In d dimensions the number of modes with $\lambda \leq \Lambda$ grows like $\Lambda^{d/2}$, and for a range change $t \sim 1/\lambda$. Then $\sum t^2 \sim \int_0^\infty \Lambda^{d/2-1} \Lambda^{-2} d\Lambda < \infty$ iff $d < 4$. Range changes are harmless in $d \leq 3$, which is Bolin & Kirchner's Cor. 3.3.

The flip side: what one sample can teach you (microergodicity). Fit (τ, κ) by maximum likelihood to *one* sample resolved to N modes. The Fisher information is

$$I_\tau(N) = \frac{2N}{\tau^2} \rightarrow \infty, \quad I_\kappa(N) = \frac{1}{2} \sum_{j \leq N} \left(\frac{4\kappa}{\kappa^2 + \lambda_j} \right)^2 \rightarrow \text{const} < \infty.$$

$\hat{\tau}$ becomes exact as the sample is refined, while $\hat{\kappa}$ never settles (Fig. 6). In the notebook the spread of $\hat{\tau}$ falls from 0.26 to 0.01 between $N = 16$ and 4096, while the spread of $\hat{\kappa}$ stays near 2.5. This is the same fact as before, seen from the other side: parameters that change the measure's class are learnable from one sample, and parameters that keep it equivalent are not. τ (more precisely $\sigma^2 \kappa^{2\nu}$ in Matérn form) is the *microergodic* parameter (Zhang 2004).

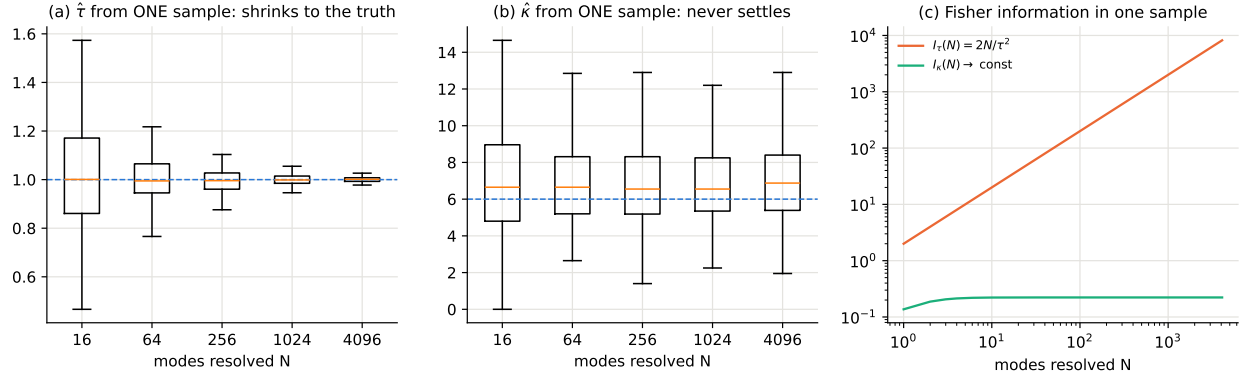


Figure 6: Maximum-likelihood estimates from a single sample (300 replicates, truth $\tau = 1, \kappa = 6$). (a) $\hat{\tau}$ converges as the resolution grows. (b) $\hat{\kappa}$ does not. (c) Fisher information: unbounded for τ , bounded for κ .

Link to the project. Because κ costs almost nothing in likelihood, it is the knob a stationary fit turns to absorb unmodelled large-scale variance. With the wells present, M1 slid to $\hat{\kappa} \approx 0$.

1.6 Back to the aquifer, in one dimension

Discretize $-(Ku')' + \kappa^2(x)u = q$ on N interior nodes with spacing $h = 1/(N+1)$, with K evaluated at cell faces. This gives a tridiagonal A with $Au = q$. With discrete white noise $\xi_i \sim \mathcal{N}(0, 1/h)$,

$$\Sigma = h^{-1} A^{-1} A^{-\top}, \quad \text{precision } Q = h A^\top A.$$

Choosing the square root $\Sigma_0^{1/2} = h^{-1/2} A_0^{-1}$ gives

$$I + T = \Sigma_0^{-1/2} \Sigma_1 \Sigma_0^{-\top/2} = BB^\top, \quad B = A_0 A_1^{-1},$$

so the eigenvalues of T are the squared singular values of $A_0 A_1^{-1}$ minus one. The reference is $K \equiv 1$, $\kappa^2 \equiv 36$. The three truths are:

truth	change	place in the equation
well	$q = \xi + a g$, $a \sim \mathcal{N}(0, 50)$, g a smooth bump with $\int g^2 = 1$	source
leakage	$\kappa^2(x) = 36 - 30 e^{-((x-0.75)/0.06)^2}$	zeroth order
lens	$K(x) = e^{b(x)}$, $b \approx 1$ on $[0.4, 0.6]$	principal part

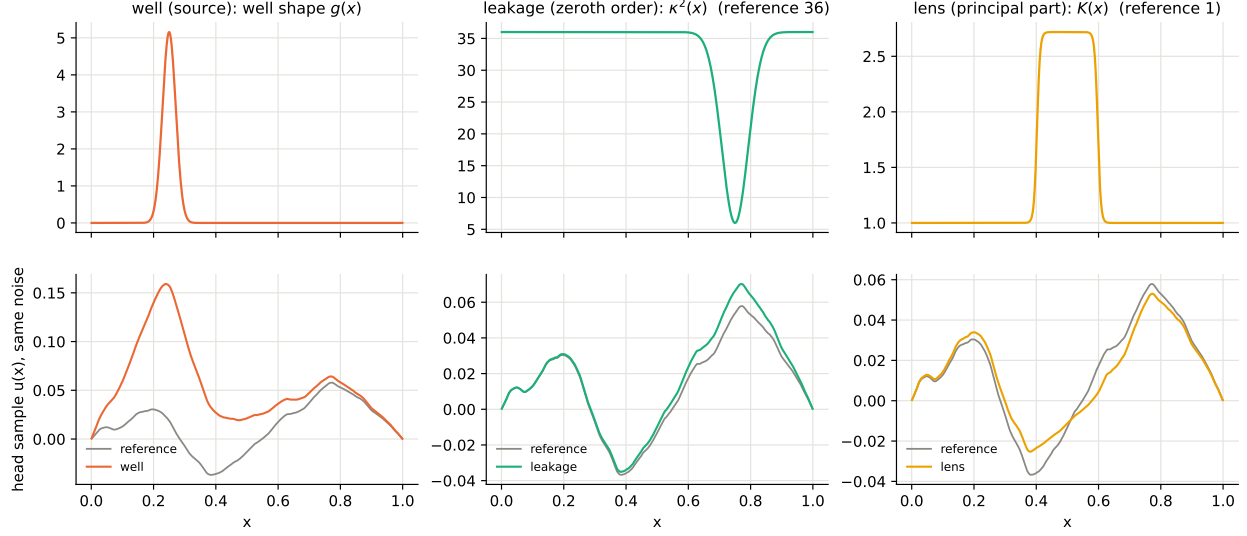


Figure 7: Top: the three departures. Bottom: one head sample under the reference (grey) and under each truth (colour), all from the same white noise.

What to expect.

- The **well** adds the rank-one term $c A_0^{-1} g g^\top A_0^{-\top}$. In whitened coordinates this is a single eigenvalue $t = c h g^\top g \approx c \int g^2 = 50$ and zeros elsewhere. That is trivially Hilbert–Schmidt: equivalent.
- The **leakage** change perturbs the operator by a bounded multiplication, $\delta \kappa^2$, which is lower order than $-\partial_x^2$. Whitened, it behaves like $(\kappa^2 - \partial_x^2)^{-1}$ times something bounded, so $t_j \sim 1/\lambda_j$: Hilbert–Schmidt in $d \leq 3$. Equivalent.
- The **lens** changes the highest-order term. At fine scales $A_1 \approx -\partial_x K \partial_x$ and $A_0 \approx -\partial_x^2$, so $BB^\top$ acts locally like multiplication by $K(x)^{-2}$, and T like multiplication by $K^{-2} - 1$. A multiplication operator is not compact. A fixed *fraction* of all N eigenvalues (the fraction of the domain covered by the lens) sits near $e^{-2} - 1 \approx -0.86$, so $\|T\|_F^2 \propto N$. Singular.

N	well	leakage	lens
32	50.00	0.638	2.34
128	50.00	0.635	4.46
512	50.00	0.635	8.80
2048	50.00	0.635	17.52

$\|T_N\|_F$ from the notebook. The lens grows by $\sqrt{2}$ per doubling of N .

In 2D, $N = n^2$, so $\|T\|_F \propto n$ and $\text{KL} \propto n^2$, which is exactly what the main project measured (F1). Note the scale. The well has a *much larger* $\|T\|_F$ (50) than the lens at any mesh shown here, yet it is the lens that is singular. As with the range change, what matters is not the size of the departure but whether it reaches the fine scales.

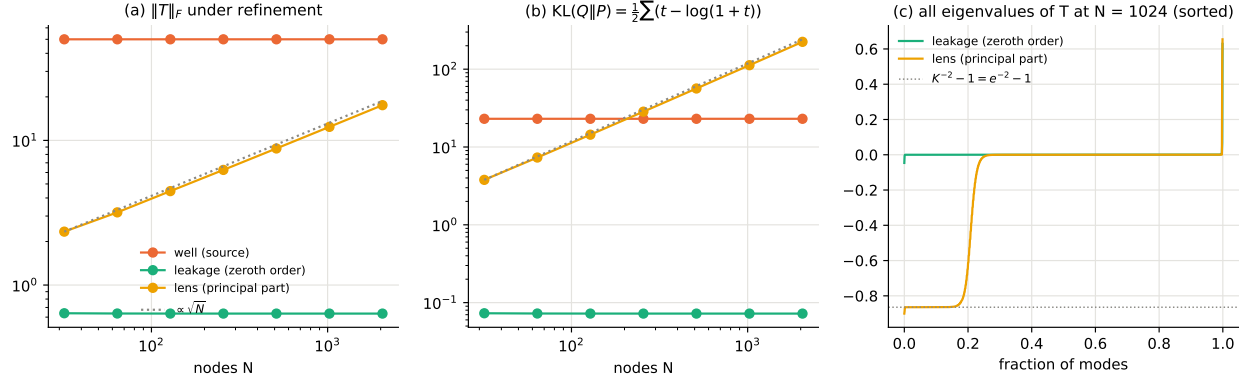


Figure 8: (a) $\|T\|_F$ and (b) KL under mesh refinement. The well and the leakage are flat; the lens grows like $\sqrt{N}$ and N . (c) All eigenvalues of T at $N = 1024$: about 20% of them sit at the symbol value $K^{-2} - 1$.

The one-sample test, for heads. Repeat the experiment of Fig. 2 with real head fields, using

$$\log \frac{d\mu_1}{d\mu_0}(u) = \log |A_1| - \log |A_0| - \frac{h}{2} (\|A_1 u\|^2 - \|A_0 u\|^2),$$

and a Woodbury form for the well. Fig. 9 shows the result. For the well and the leakage the error of the best test does not change with N : a finer mesh gives no new evidence. For the lens the error drops to zero. A finely resolved head field carries the lens’s fingerprint in every small-scale wiggle.

1.7 Take-aways

1. Equivalence means the same null sets, which means a density dQ/dP exists, which means one perfect sample cannot tell the measures apart.
2. Diagonal case: equivalent iff $\sum t_j^2 < \infty$ (variances) and $\sum m_j^2 < \infty$ (means). Feldman–Hájek is the same statement without a basis: T is Hilbert–Schmidt and the mean lies in the Cameron–Martin space.
3. The decision is made at the **fine scales**. Large-scale changes (range κ , wells, leakage) are harmless however big they look. Changes that persist down to the smallest scales (amplitude τ , conductivity K) are fatal however small they look.
4. Practically: what a fine sample determines exactly (τ, K) cannot be “corrected” with finitely many directions, and what it never determines (κ) is a cheap knob a fit can slide along.

Check yourself. 1. Is $t_j = 1/\sqrt{j}$ equivalent or singular? Compute $\sum t_j^2$.

2. A constant conductivity $K \equiv 2$: equivalent or singular? (Hint: $-\nabla \cdot (2\nabla u) + \kappa^2 u = \xi$ is $2(\frac{\kappa^2}{2} - \Delta)u = \xi$. Which parameter of §1.5 changed?)
3. The well’s source shape g must be in L^2 , so that its head response $(\kappa^2 - \Delta)^{-1}g$ is in H^2 . What happens to $t = ch g^\top g$ when g becomes a delta at the grid scale ($g_i = 1/h$ at one node)?

Next chapter (planned). Whitening and the operator T in 2D; why the lens’s T is “multiplication by $K^{-2} - 1$ ” (symbols of differential operators); then how the finite-rank model M3 sees T through sample covariances (Marchenko–Pastur, BBP).

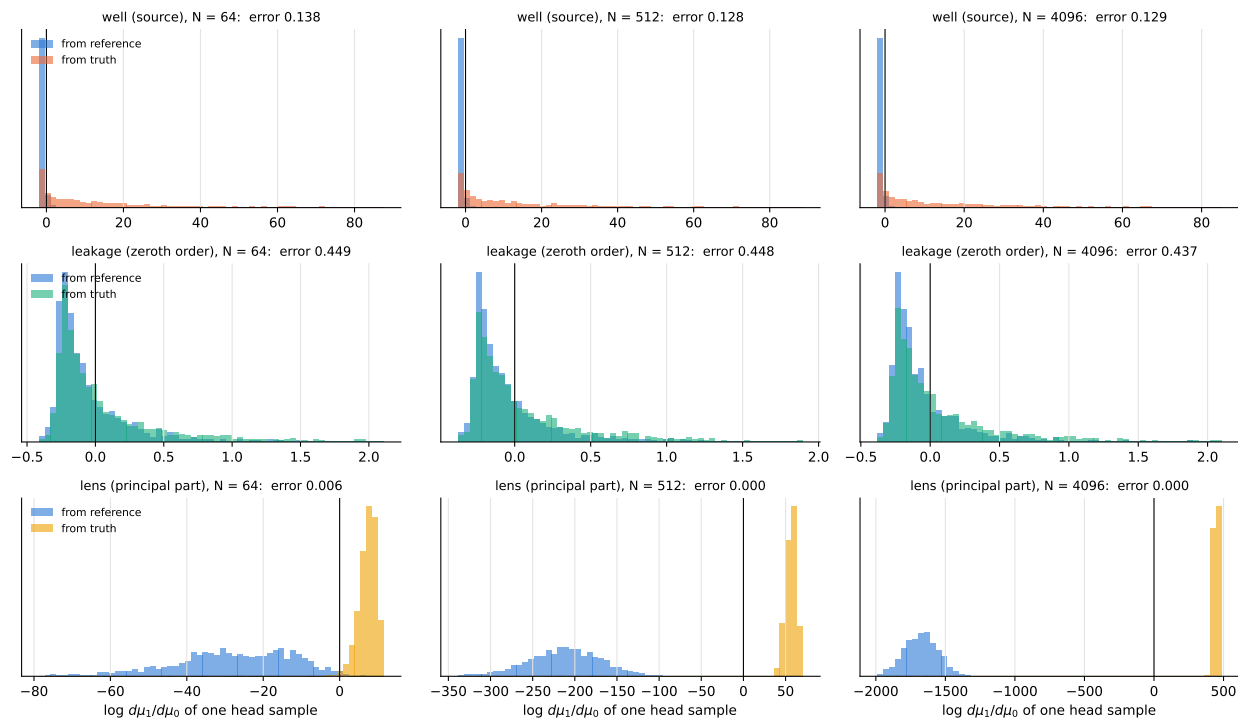


Figure 9: One-sample log likelihood ratio, reference versus truth, at $N = 64, 512, 4096$. The well (error ≈ 0.13) and the leakage (error ≈ 0.44) are stable: equivalent. The lens error is $0.006 \rightarrow 0.000$: singular.

References

- V. I. Bogachev, *Gaussian Measures*, AMS, 1998 (Thm. 2.7.2).
- G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, 2nd ed., CUP, 2014 (Thm. 2.25).
- J. Feldman, Equivalence and perpendicularity of Gaussian processes, *Pacific J. Math.* 8 (1958).
- J. Hájek, On a property of normal distributions of any stochastic process, *Czech. Math. J.* 8 (1958).
- S. Kakutani, On equivalence of infinite product measures, *Ann. Math.* 49 (1948).
- H. Zhang, Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics, *JASA* 99 (2004).
- D. Bolin, K. Kirchner, Equivalence of measures and asymptotically optimal linear prediction for Gaussian random fields with fractional-order covariance operators, *Bernoulli* 29(2) (2023).