

The Entropy of Finite Response

Causal kernels, closure, and spread

J. Landers

Draft note

Abstract

Idealized models often describe a response as instantaneous. This note separates the source from the physical mediator: the observed response is the source convolved with a passive causal kernel. Such kernels have two elementary operations, convex mixing for parallel channels and convolution for sequential stages. The resulting moment laws show that hidden passive stages cannot sharpen an already mediated response; sharpness lives only on a singular boundary of the kernel space. Replacing variance by entropy power makes the statement strict: by the entropy power inequality and the support constraint imposed by causality, every nontrivial sequential stage increases entropy power by more than simple additivity would predict. The causal interior is therefore pinned between two unreachable walls—the delta locus $N = 0$ and the Gaussian locus $N = \sigma^2$ —with real finite-response dynamics living between them.

1 From instantaneous onset to mediated response

Let an idealized force or source field be written as

$$F_{\text{ideal}}(x, t),$$

where $x \in \mathbb{R}^d$ is position and t is time. A physically observed response may be modeled as

$$F_{\text{actual}}(x, t) = \int_{\mathbb{R}^d} \int_{-\infty}^t G(x - x', t - s) F_{\text{ideal}}(x', s) ds dx'. \quad (1)$$

Here G is the mediation kernel. The source F_{ideal} is what is being mediated; the kernel G describes how mediation occurs. In compact notation,

$$F_{\text{actual}} = G * F_{\text{ideal}}.$$

This source–kernel separation is the organizing move of the note; the convolution algebra is standard, but here it is used to isolate what finite passive response does to sharpness.

This formalism captures the distinction between an idealized onset and a finite physical response. If $m\ddot{x} = F$ and one writes an ideal step force

$$F(t) = F_0 H(t - t_0),$$

then

$$a(t) = \frac{F_0}{m} H(t - t_0), \quad j(t) = \dot{a}(t) = \frac{F_0}{m} \delta(t - t_0).$$

Thus an instantaneous onset of acceleration corresponds, in the distributional sense, to an impulse in jerk. A causal temporal kernel K replaces this singularity by finite response:

$$F_{\text{actual}}(t) = \int_{-\infty}^t K(t - s) F_{\text{ideal}}(s) ds, \quad K(t) = 0 \text{ for } t < 0. \quad (2)$$

For the step source,

$$F_{\text{actual}}(t) = F_0 \int_0^{t-t_0} K(u) du, \quad j(t) = \frac{F_0}{m} K(t-t_0).$$

So the singular idealization

$$\delta(t-t_0)$$

is replaced by the finite kernel

$$K(t-t_0).$$

2 Two canonical geometries

The kernel formalism separates the source from the transport geometry. Two limiting cases are especially useful: sharp transport on an arrival surface and diffusive transport through a broadening profile.

Ballistic mediation. In three spatial dimensions, a sharp finite-speed wavefront is represented by

$$G_{\text{ballistic}}(x-x', t-s) = \frac{1}{4\pi r} \delta\left((t-s) - \frac{r}{c}\right) H(t-s), \quad r = \|x-x'\|. \quad (3)$$

This kernel has response concentrated at the arrival time $t-s = r/c$. It is not instantaneous; it is a delayed sharp arrival.

Diffusive mediation. In d dimensions, a diffusion-like response is represented by

$$G_{\text{diffusive}}(x-x', t-s) = \frac{H(t-s)}{(4\pi D(t-s))^{d/2}} \exp\left(-\frac{\|x-x'\|^2}{4D(t-s)}\right), \quad (4)$$

where $D > 0$ is the diffusion coefficient. This kernel spreads influence over a length scale

$$\ell(t-s) \sim \sqrt{D(t-s)}.$$

The ballistic kernel lives on a sharp arrival surface; the diffusive kernel fills space-time with a broadening cloud. Both are choices of G in Eq. (1), and the diffusive case later connects this finite-response picture to the entropy–Fisher-information identities used in information theory [4].

3 Kernel space: mixtures and compositions

There are two elementary ways to build a complicated mediation law from simpler ones.

First, distinct channels may act in parallel:

$$G_\lambda = (1-\lambda)G_1 + \lambda G_2, \quad 0 \leq \lambda \leq 1. \quad (5)$$

For example, a direct ballistic channel may coexist with a scattered diffusive tail:

$$G_\lambda = (1-\lambda)G_{\text{ballistic}} + \lambda G_{\text{diffusive}}.$$

Second, mediation stages may occur sequentially. If the source first passes through G_1 and then through G_2 , the total kernel is

$$G_{\text{total}} = G_2 * G_1. \quad (6)$$

Thus addition represents parallel physical channels, while convolution represents cascaded physical stages.

This suggests that physically admissible kernels form not a single family but a structured space: a constrained set closed under convex mixing and, under suitable assumptions, closed under composition. If one requires composition at all subdivision scales, this is the language of convolution semigroups familiar from Levy-process theory [1]; here only the finite two-stage version is needed. The following proposition makes the simplest temporal version precise.

4 A small closure and spread result

Definition 1 (Passive causal temporal kernel). *A kernel K is passive causal if*

$$K(t) = 0 \text{ for } t < 0, \quad K(t) \geq 0, \quad \int_0^\infty K(t) dt = 1,$$

and if its first two moments are finite. Define

$$\mu(K) = \int_0^\infty tK(t) dt, \quad \sigma^2(K) = \int_0^\infty (t - \mu(K))^2 K(t) dt.$$

Proposition 1 (Causal mediation spread). *Let K_1 and K_2 be passive causal temporal kernels.*

Parallel channels. For

$$K_\lambda = (1 - \lambda)K_1 + \lambda K_2, \quad 0 \leq \lambda \leq 1,$$

K_λ is passive causal, and

$$\mu(K_\lambda) = (1 - \lambda)\mu(K_1) + \lambda\mu(K_2), \tag{7}$$

$$\sigma^2(K_\lambda) = (1 - \lambda)\sigma^2(K_1) + \lambda\sigma^2(K_2) + \lambda(1 - \lambda)(\mu(K_2) - \mu(K_1))^2. \tag{8}$$

Sequential stages. For

$$K_{\text{total}} = K_2 * K_1,$$

K_{total} is passive causal, and

$$\mu(K_2 * K_1) = \mu(K_1) + \mu(K_2), \tag{9}$$

$$\sigma^2(K_2 * K_1) = \sigma^2(K_1) + \sigma^2(K_2). \tag{10}$$

Proof. Since nonnegative causal normalized functions are probability densities on $[0, \infty)$, interpret K_i as the delay distribution of a nonnegative random variable T_i . A convex mixture first chooses channel 1 with probability $1 - \lambda$ and channel 2 with probability λ . Equations (7) and (8) are the standard laws of total expectation and total variance.

For sequential stages, the total delay is

$$T_{\text{total}} = T_1 + T_2,$$

with independent stages. The density of $T_1 + T_2$ is $K_2 * K_1$. Since $T_1, T_2 \geq 0$, the convolution remains causal; normalization follows by Fubini's theorem. The expectation and variance of the sum give (9) and (10), as illustrated by the exponential cascade in Fig. 1. \square

The proposition formalizes a simple physical intuition: hidden passive stages accumulate delay and spread. A delta delay, $K(t) = \delta(t - t_0)$, has zero variance and lies on the sharp boundary of the space. Any nontrivial stage with positive variance makes the response less instantaneous. Parallel mixtures may also broaden a response even if each channel is individually sharp, as shown in Fig. 2, because different arrival times produce the additional term in Eq. (8).

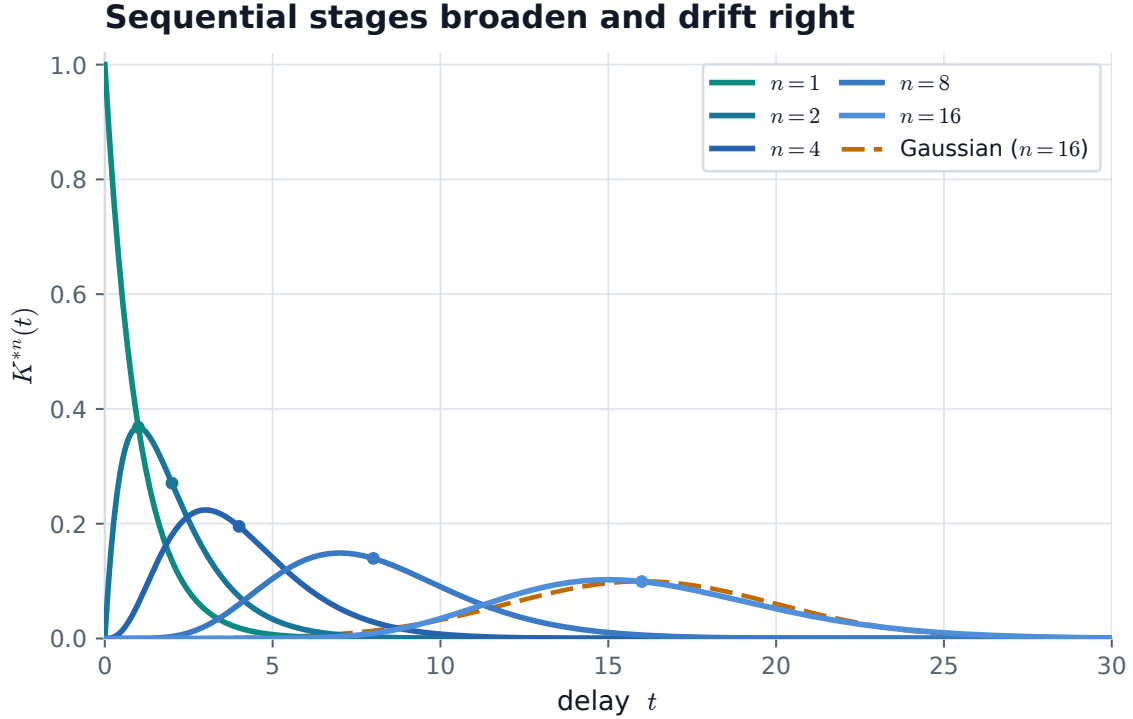


Figure 1: Sequential stages broaden and drift. Self-convolution of an exponential, $\Gamma(1, \theta)^{*n} = \Gamma(n, \theta)$, realizes the additive laws of Proposition 1 ($\mu = n, \sigma^2 = n$ at $\theta = 1$). The shape approaches a Gaussian in the usual convolution limit [2] but stays causal; the matched Gaussian for $n = 16$ (dashed) lies strictly outside on $t < 0$.

5 The sharp side of spread: an entropy-power refinement

Proposition 1 controls the second moment, and the variance law (10) is an exact equality. That is its virtue and its limitation: variance sees only the scale of a kernel, not the shape of how it disperses. Two stages of identical variance can spread response very differently. To see the shape, replace variance by the *entropy power*, the Shannon information-theoretic scale associated with differential entropy [5, 4].

For a passive causal kernel K with finite differential entropy

$$h(K) = - \int_0^{\infty} K(t) \log K(t) dt,$$

define the entropy power

$$N(K) = \frac{1}{2\pi e} e^{2h(K)}. \quad (11)$$

Two facts fix its meaning against the variance already in play. First, among all densities of a given variance the Gaussian maximizes entropy, so

$$N(K) \leq \sigma^2(K), \quad (12)$$

with equality if and only if K is Gaussian. Second, as K collapses to a delta, $h(K) \rightarrow -\infty$ and $N(K) \rightarrow 0$. Entropy power and variance vanish together on the sharp boundary, but N is the stricter

Parallel channels broaden without sharpening

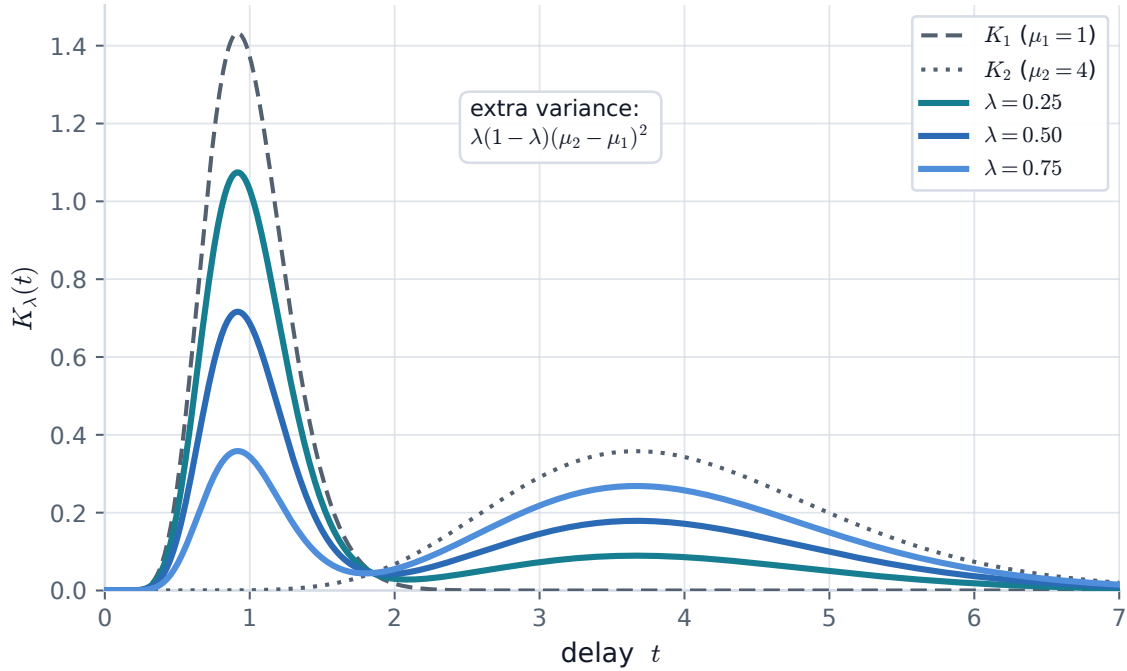


Figure 2: Parallel channels broaden without sharpening. Convex mixtures $K_\lambda = (1 - \lambda)K_1 + \lambda K_2$ of two sharp channels ($\mu_1 = 1, \mu_2 = 4$). Variance carries the extra term $\lambda(1 - \lambda)(\mu_2 - \mu_1)^2$ of Eq. (8); spread peaks near $\lambda = 0.5$ with neither channel altered.

of the two: bounded above by σ^2 , and sensitive to shape where σ^2 is not, as shown in the kernel slice in Fig. 3.

The composition law for N is not an equality.

Theorem 1 (Strict entropy-power spreading). *Let K_1, K_2 be passive causal temporal kernels with finite differential entropy, each carrying a density rather than a point mass. Then*

$$N(K_2 * K_1) \geq N(K_1) + N(K_2), \quad (13)$$

and the inequality is strict,

$$N(K_2 * K_1) > N(K_1) + N(K_2).$$

Proof. By Proposition 1, each K_i is the law of an independent nonnegative delay T_i , and $K_2 * K_1$ is the law of $T_1 + T_2$. Inequality (13) is the Shannon–Stam entropy power inequality [6, 3]. Equality there holds if and only if T_1 and T_2 are Gaussian. But a random variable whose density is supported on $[0, \infty)$ cannot be Gaussian, since the Gaussian law charges the whole real line. Equality is therefore unattainable for passive causal kernels, and the inequality is strict. The same support argument applied to (12) makes that bound strict as well. \square

The contrast with Proposition 1 is the whole point. Variance adds exactly, so the second-moment statement of irreversibility is neutral: it forbids sharpening but is indifferent to the act of composing. Entropy power instead *strictly increases beyond additivity* at every nontrivial stage. Cascaded causal

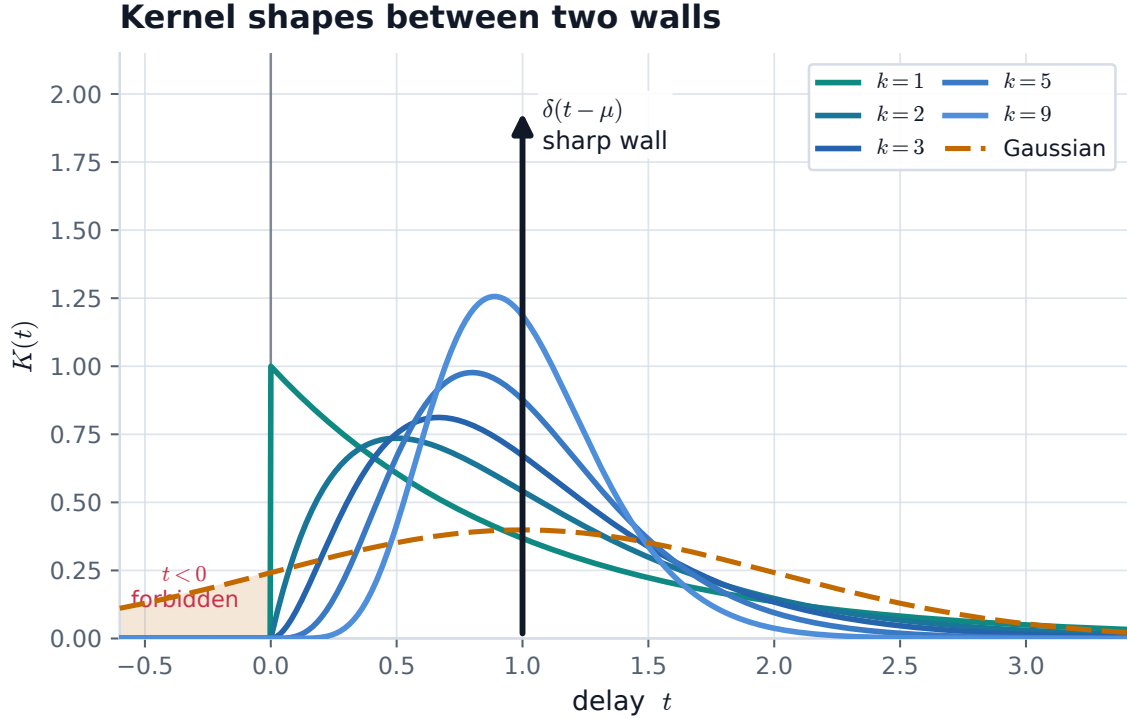


Figure 3: Passive causal kernel shapes. Gamma kernels of fixed mean $\mu = 1$ and increasing shape k interpolate from skewed toward bell-shaped while staying supported on $[0, \infty)$. The delta $\delta(t - \mu)$ marks the sharp wall ($N = 0$); the variance-matched Gaussian for the broadest case (dashed) marks the efficient wall ($N = \sigma^2$), unreachable as its mass leaks to $t < 0$ (shaded). Shown as a one-parameter slice through the kernel space, not a literal path between the walls.

mediation does not merely fail to sharpen the response; it produces strictly more dispersion than the sum of its parts, and it does so precisely because causality forbids the one shape—the Gaussian—for which composition would be efficient.

This sharpens the boundary picture into two boundaries. The interior of the passive causal kernel space is pinned below by the delta locus $N = 0$ (infinitely sharp, reached only in the limit) and, through (12), above by the Gaussian locus $N = \sigma^2$ (maximally efficient spread, where the entropy power inequality would close). Causal kernels live strictly between the two: never perfectly sharp, never perfectly efficient. Both ideals sit on the boundary, and the physics happens in between.

The strict gap is order-unity, not marginal: a single exponential stage has $N/\sigma^2 = e/(2\pi) \approx 0.43$, so the first composition already adds about 57% of a stage’s variance beyond additivity (Fig. 4).

The diffusive kernel of §2 is the flow that carries a response into that interior. Writing the profile of (4) as $p(\cdot, \tau)$, it solves the heat equation $\partial_\tau p = D \Delta p$, and de Bruijn’s identity gives the entropy production rate exactly [4]:

$$\frac{d}{d\tau} h(p(\cdot, \tau)) = D J(p(\cdot, \tau)), \quad J(p) = \int \frac{\|\nabla p\|^2}{p} d\xi, \quad (14)$$

with J the Fisher information of the current profile. Since $J \geq 0$, diffusion increases entropy monotonically, at a rate set by J . The discrete statement—each passive stage strictly adds entropy power—and

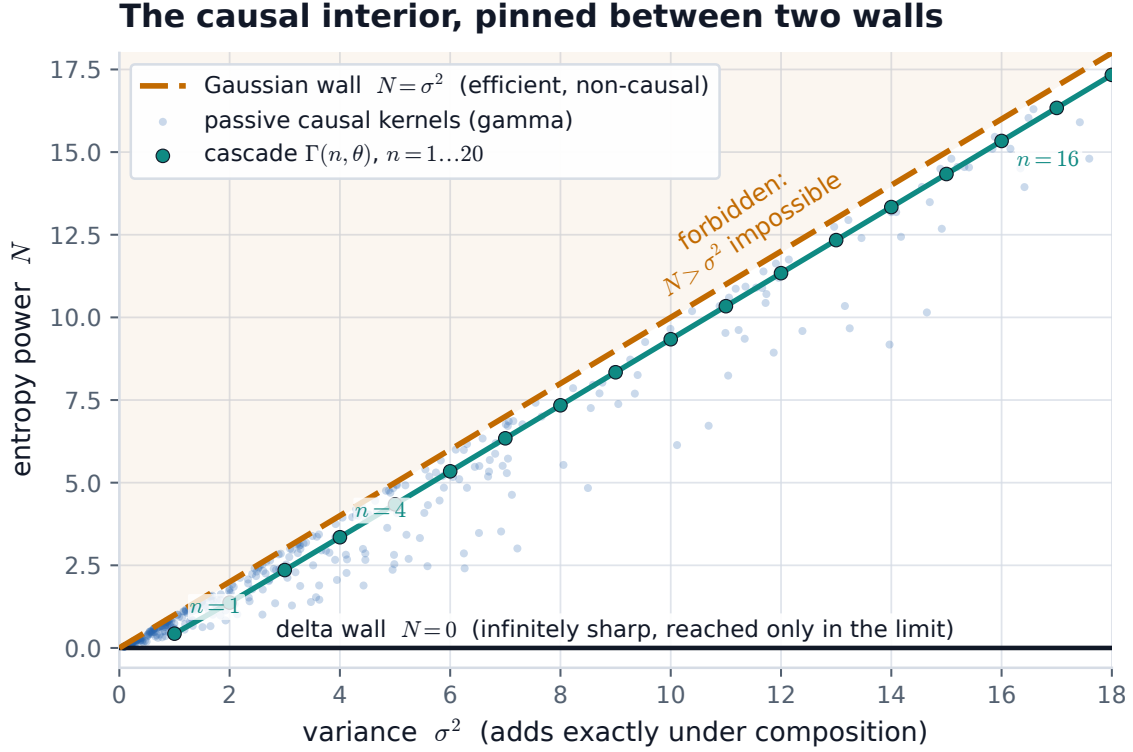


Figure 4: The causal interior between two walls. Entropy power N vs variance σ^2 . The Gaussian locus $N = \sigma^2$ (dashed) and delta locus $N = 0$ bound the admissible region; causal gamma kernels fill the strict interior. The cascade $\Gamma(n, \theta)$ climbs toward the Gaussian wall, reaching $N/\sigma^2 \approx 0.96$ at $n = 16$ without equality.

the continuous one—diffusion is the generator of entropy increase—are the same irreversibility read at two resolutions.

6 Finite-speed corollary and interpretation

For space-time kernels, let the causal cone condition be

$$G(\xi, \tau) = 0 \quad \text{unless} \quad \tau \geq 0, \quad \|\xi\| \leq c\tau. \quad (15)$$

Corollary 1 (Cone closure). *If G_1 and G_2 satisfy (15), then every convex mixture $(1 - \lambda)G_1 + \lambda G_2$ also satisfies (15). Moreover, their convolution satisfies (15).*

Proof. The mixture statement is immediate from support inclusion. For convolution, any nonzero contribution factors through increments (ξ_1, τ_1) and (ξ_2, τ_2) with $\xi = \xi_1 + \xi_2$ and $\tau = \tau_1 + \tau_2$. Since $\|\xi_i\| \leq c\tau_i$,

$$\|\xi\| \leq \|\xi_1\| + \|\xi_2\| \leq c(\tau_1 + \tau_2) = c\tau.$$

Thus the composed kernel remains inside the cone. \square

7 Conclusion

The resulting picture is a small geometry of mediation kernels. The instantaneous local kernel $\delta(\xi)\delta(\tau)$ is a singular idealization. Ballistic kernels concentrate response on causal arrival surfaces. Diffusive kernels distribute response across many small randomized delays. Lossy kernels reduce total gain; active kernels require an external energy source. But passive causal kernels have stable algebraic structure: they may be mixed, composed, and decomposed into channels and stages without violating causality.

A jump in acceleration is therefore the boundary case obtained when the mediating kernel collapses to a delta. Real finite-response dynamics correspond to interior points of a constrained kernel space. The proposition gives one precise sense in which moving into that interior is irreversible under passive composition: additional hidden mediation stages add spread rather than remove it. The entropy-power refinement makes the irreversibility strict and gives it a second wall to face: the response can approach neither the infinitely sharp delta nor the maximally efficient Gaussian, and diffusion is exactly the flow that drives entropy upward between them. Sharpness is not destroyed; it is revealed as an ideal boundary of causal response.

References

- [1] D. Applebaum, *Levy Processes and Stochastic Calculus*, 2nd ed., Cambridge University Press, 2009.
- [2] A. R. Barron, “Entropy and the central limit theorem,” *Annals of Probability* **14**(1), 336–342, 1986.
- [3] N. M. Blachman, “The convolution inequality for entropy powers,” *IEEE Transactions on Information Theory* **11**(2), 267–271, 1965.
- [4] A. Dembo, T. M. Cover, and J. A. Thomas, “Information theoretic inequalities,” *IEEE Transactions on Information Theory* **37**(6), 1501–1518, 1991.
- [5] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal* **27**, 379–423 and 623–656, 1948.
- [6] A. J. Stam, “Some inequalities satisfied by the quantities of information of Fisher and Shannon,” *Information and Control* **2**(2), 101–112, 1959.