REVERSE-TIME SCORES AND CONVEX-ORDER MONOTONICITY FOR ADDITIVE PROCESSES WITH LÉVY CLOCKS

Valentin Dendoncker, Michel Denuit, Christian Y. Robert

LIDAM Discussion Paper ISBA
2026 / 23

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication

Reverse-time scores and convex-order monotonicity for additive processes with Lévy clocks

Valentin Dendoncker
Institute of Statistics, Biostatistics and Actuarial Science – ISBA
Louvain Institute of Data Analysis and Modeling – LIDAM, UCLouvain
Quantitative and Artificial Intelligence Research Methods at Ichec – Quaresmi, ICHEC
valentin.dendoncker@uclouvain.be, valentin.dendoncker@ichec.be

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science – ISBA
Louvain Institute of Data Analysis and Modeling – LIDAM, UCLouvain
michel.denuit@uclouvain.be

Christian Y. Robert
Laboratoire de Sciences Actuarielle et Financière – LSAF
UCBL – Université Claude Bernard Lyon 1, ISFA
christian.robert@univ-lyon1.fr

June 8, 2026

Abstract

We study reverse-time conditional means, scores and drift quantities in additive noise models driven by deterministic Lévy clocks. Starting from an integrable signal Y and an independent centered additive process $(Z_t)_{t\in[0,T]}$, we consider the forward noising model $X_t = Y + Z_t$, $0 \le t \le T$, and its reverse-time version $\overline{X}_s = X_{T-s}$. The main examples are Gaussian noise with deterministic variance clock, compensated compound Poisson noise with deterministic intensity clock, and symmetric α-stable noise with stable clock, for $\alpha \in (1,2)$. We first derive Bartlett–Tweedie identities for additive noise models and characterize, within the infinitely divisible class, the cases in which the denoising map $x \mapsto \mathrm{E}[Y \mid Y + \varepsilon = x]$ is affine. This occurs precisely when the signal and the noise are compatible with the same convolution semigroup. We then prove that the reverse conditional mean and the denoising error normalized by the accumulated noise clock are martingales in reverse time. As a consequence, both families are increasing in the convex order along reverse time. The actual reverse drift, or reverse local mean drift in the jump and stable cases, is obtained by multiplying this normalized score by the instantaneous clock rate. It is therefore not a martingale in general, but it defines a peacock when the instantaneous rate is non-decreasing in reverse time. Explicit linear formulas are obtained for semigroup-compatible initial laws. The Cauchy case is treated separately, since the lack of integrability prevents an ordinary convex-order analysis of the reverse score drift.

Keywords: conditional expectation; convex order; peacocks; reverse-time dynamics; additive processes; Lévy clocks; infinitely divisible distributions; Tweedie formula.

MSC2020 subject classification: 60E15; 60E07; 60G44; 60G48; 60G51; 60G52; 60J76.

1 Introduction

In many applications, a probabilistic object or signal is progressively corrupted until it becomes unstructured, or pure noise with prescribed distribution. It is then reconstructed to generate a new object of the same type, statistically indistinguishable from the initial one. This relies on a forward-backward procedure where the forward process transforms the signal to a prescribed noise while the backward process reconstitutes the signal from noise. The backward step is achieved with the help of generative models that create the desired target from noise.

Diffusion models, also referred to as score-based generative models, have become the state-of-the-art approach for large-scale image and audio generation. These models use a forward noising process paired with a reverse denoising process: starting from random noise, diffusion models gradually refine it into structured data, generating a synthetic sample by the reverse process. Considering image data, the idea is to corrupt an image using a multi-step noise process to convert it into a sample from the target noise distribution. We refer the reader, e.g., to Chapter 20 in Bishop and Bishop [3] for more information about denoising diffusion probabilistic models.

A typical forward noising process consists in a Markov chain of length T. Starting from a real data sample, it produces a sequence of increasingly noisy samples to end up with pure noise (with prescribed distribution, often standard Gaussian). Initially described in discrete time, the forward-backward processes are now specified in continuous time, as the limits of the discrete models, and unified by stochastic differential equations (Song et al. [12]). This paper studies the backward construction, reversing the forward process in time, and associated score-matching to learn this process.

Score-based diffusion models have renewed interest in the probabilistic structure of time-reversed stochastic dynamics. Their basic mechanism is the forward–backward construction described previously: a forward process progressively destroys information by adding noise to an initial signal, before a backward process uses the score of the intermediate marginal laws to reconstruct the signal from noise. This point of view goes back to the time-reversal theory of diffusion processes developed by Anderson [1] and Haussmann and Pardoux [5]. It has become central in modern score-based generative modeling; see, for instance, the recent survey by Tang and Zhao [13].

A key step in this construction is the choice of the terminal noise distribution used to initialize the backward dynamics. In generative diffusion models, this noise distribution should be easy to sample and should not depend on the unknown target distribution. A standard way to obtain such a noise law is to decouple the initial state X_0 from the conditional distribution of X_T given X_0. For instance, in the variance-exploding Gaussian model, conditionally on $X_0 = x$, the terminal variable has the form $X_T = x + \varepsilon_T$, where ε_T is a centered Gaussian noise with large variance. Decoupling the location x from this conditional law leads to a tractable Gaussian reference distribution for the backward initialization. This decoupling principle is one of the reasons why time-dependent but state-independent noise coefficients are particularly convenient in score-based diffusion models: the conditional law of X_T given X_0 remains explicit and the noise component can be separated from the signal.

The present paper studies this mechanism from the viewpoint of conditional expectations, time reversal and stochastic orders. We consider an additive forward model $X_t = Y + Z_t$, $0 \leq t \leq T$, where $Y = X_0$ is an integrable signal and $(Z_t)_{t\in[0,T]}$ is an independent centered process with independent increments. This setting makes the decoupling explicit: conditionally on $Y = y$, the law of X_t is the law of $y + Z_t$, whereas the noise law itself is that of Z_t. The denoising map $m_t(x) = \mathrm{E}[Y \mid X_t = x]$ therefore measures how much information about the initial signal remains after adding the noise Z_t. In the Gaussian case, Tweedie's formula identifies $m_t(x) - x$ with the score of the marginal density of X_t, multiplied by the accumulated variance. Analogous Bartlett–Tweedie identities hold for several non-Gaussian noises, with local differential operators replaced by jump or fractional operators.

A second guiding principle is that the forward process should become increasingly variable as time increases. Since (Z_t) has centered independent increments, the forward dynamics preserve the mean, i.e. $\mathrm{E}[X_t] = \mathrm{E}[Y]$, while progressively adding independent noise. Thus, the forward noising mechanism should not only be viewed through the increase of a scalar variance. The natural probabilistic notion is the convex order: for $0 \le s \le t$, one expects X_s to be smaller than X_t in the convex order. This expresses the idea that the forward process becomes more dispersed, or less informative, as the noise level increases. The main question addressed in this paper is whether the reverse-time quantities that drive denoising inherit a corresponding convex-order structure.

We show that they do. In reverse time, with $\overline{X}_s = X_{T-s}$, the conditional mean $\overline{M}_s = \mathrm{E}[Y \mid \overline{X}_s]$ is a martingale. More importantly, after normalizing the denoising error by the correct deterministic noise clock, the corresponding score process is also a reverse-time martingale. The clock is the accumulated variance in the Gaussian case, the accumulated intensity for compensated compound Poisson noise, and the stable clock for symmetric α-stable noise with $\alpha \in (1, 2)$. Consequently, the reverse conditional mean and the normalized score are increasing in convex order along reverse time, or equivalently decreasing in convex order as the forward noise level increases.

This distinction between the normalized score and the actual reverse drift is important in the time-inhomogeneous case. When the noise coefficient is constant, the two differ only by a fixed multiplicative factor. When the Gaussian volatility, the Poisson intensity or the stable scale depends on time, the actual reverse drift is obtained by multiplying the normalized score by the instantaneous clock rate. It is therefore not a martingale in general. Nevertheless, if this instantaneous rate is non-decreasing in reverse time, then the actual reverse drift is also increasing in convex order. In this case, it defines a peacock and Kellerer's theorem yields a martingale with the same one-dimensional marginal laws (see e.g. [6]).

The paper is organized as follows. Section 2 derives Bartlett–Tweedie formulas for additive noise models and characterizes affine conditional means within the infinitely divisible class. Section 3 introduces the common additive framework, the centered independent-increments assumptions, the convex-order viewpoint, and the deterministic clocks used in the main examples. Section 4 isolates the general clock-based martingale principle: when the Bartlett–Tweedie correction is normalized by the accumulated noise clock, the resulting reverse-time score is a martingale and hence is increasing in the convex order along reverse time. It also explains how the actual reverse drift, or reverse local mean drift, is obtained by multiplying this normalized score by the instantaneous clock rate. Sections 5, 6 and 7 specialize this principle to Gaussian noise with deterministic variance clock, compensated compound Poisson noise with deterministic intensity clock, and symmetric α-stable noise with deterministic stable clock, respectively. In each case, we identify the corresponding reverse-time dynamics, the specific score operator, the actual reverse drift or local mean drift, and the affine formulas obtained for semigroup-compatible initial laws. Section 8 summarizes the main contributions and discusses possible extensions. The Cauchy case is treated separately in Appendix A, since the lack of integrability prevents an ordinary martingale or convex-order statement for the reverse score drift without truncation, localization or additional assumptions.

2 Conditional means in additive noise models

This section collects the conditional-mean identities that will be used throughout the paper. We consider the elementary additive observation model $X = Y + \varepsilon$, where Y is the signal and ε is an independent noise. The aim is to express the denoising map $x \mapsto \mathrm{E}[Y \mid X = x]$ in terms of the marginal density of X and of the characteristic exponent of the noise. This point of view is closely related to classical conditional limit results for expectations given sums; see, for instance, Zabell [15]. In the present paper, however, the focus is not asymptotic. The formulas below are used to identify the reverse-time score and drift quantities associated with additive processes.

We begin with the main identity of this section. It is a Fourier representation of the conditional mean in an additive noise model. The formula shows that the denoising correction $\mathrm{E}[Y \mid X = x] - x$ is entirely determined by the characteristic exponent of the noise and the marginal density of the noisy observation.

Proposition 2.1 (Bartlett representation for additive noise models)**.** *Let Y and ε be two real-valued independent random variables with $\mathrm{E}[|Y|] < \infty$, and define $X = Y + \varepsilon$. Assume that the characteristic function ϕ_X as well as its derivative ϕ'_X are integrable, and that the characteristic function ϕ_ε is such that $\phi_\varepsilon(t) = e^{\eta(t)}$, where η is locally absolutely continuous, with derivative η' understood almost everywhere (in particular, this holds true if ε is infinitely divisible with differentiable Lévy–Khintchine exponent). Assume moreover that $\eta'\phi_X$ is integrable.*

Then, for all x such that $f_X(x) > 0$,

$$\mathrm{E}[Y \mid X = x] = x - \frac{\mathcal{F}^{-1}\left[\frac{1}{i}\eta'(t)\phi_X(t)\right](x)}{f_X(x)}, \tag{2.1}$$

where $f_X(x) = \mathcal{F}^{-1}[\phi_X(t)]\,(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \phi_X(t)e^{-itx}\,dt$.

Proof. Since $X = Y + \varepsilon$, Y and ε being independent, we have

$$\mathrm{E}[Ye^{itX}] = \mathrm{E}[Ye^{itY}]\,\mathrm{E}[e^{it\varepsilon}] = \frac{1}{i}\phi'_Y(t)\phi_\varepsilon(t).$$

Using $\phi_X(t) = \phi_Y(t)\phi_\varepsilon(t)$ and $\phi'_X(t) = \phi'_Y(t)\phi_\varepsilon(t) + \eta'(t)\phi_X(t)$ for almost every $t \in \mathbb{R}$, we obtain

$$\mathrm{E}[Ye^{itX}] = \frac{1}{i}\phi'_Y(t)\phi_\varepsilon(t) = \frac{1}{i}\left(\phi'_X(t) - \eta'(t)\phi_X(t)\right),$$

which is thus integrable by assumptions. By Bartlett's formula (see, e.g., Zabell [14]), we get

$$\begin{aligned}
\mathrm{E}[Y \mid X = x] &= \frac{\int_{\mathbb{R}} \mathrm{E}[Ye^{itX}]e^{-itx}\,dt}{\int_{\mathbb{R}} \phi_X(t)e^{-itx}\,dt} \\
&= \frac{\mathcal{F}^{-1}\left[\frac{1}{i}(\phi'_X(t) - \eta'(t)\phi_X(t))\right](x)}{f_X(x)}.
\end{aligned}$$

Finally, the result follows from

$$\mathcal{F}^{-1}\left[\frac{1}{i}\phi'_X(t)\right](x) = xf_X(x).$$

This ends the proof. $\square$

Proposition 2.1 is the common source of all the formulas below. Its main message is that the operator appearing in the denoising correction is the adjoint, at the density level, of the operator encoded by the noise exponent. In the Gaussian case this operator is local and differential; in jump and stable cases it is non-local.

We begin with the Laplace distribution. This example is useful because it is still explicit, but already non-Gaussian. It shows that the correction term involves smoothing by the Laplace kernel rather than a purely local derivative of f_X.

Proposition 2.2 (Laplace noise)**.** *Under the assumptions of Proposition 2.1 with ε following the centered Laplace distribution with parameter $b > 0$, that is, $\phi_\varepsilon(t) = \frac{1}{1+b^2t^2}$, we have*

$$\mathrm{E}[Y \mid X = x] = x + \frac{2b^2(f_X * g_b)'(x)}{f_X(x)} \quad \textit{where } g_b(u) = \frac{1}{2b}e^{-|u|/b},$$

or, equivalently,

$$\mathrm{E}[Y \mid X = x] = x - \frac{\int_{\mathbb{R}} f_X(x-u)\,\mathrm{sign}(u) e^{-|u|/b}\,du}{f_X(x)}.$$

Proof. We have

$$\eta(t) = -\ln(1 + b^2 t^2) \text{ and } \eta'(t) = -\frac{2b^2 t}{1 + b^2 t^2} \Rightarrow \frac{1}{i}\eta'(t)\phi_X(t) = \frac{2ib^2 t}{1 + b^2 t^2}\phi_X(t).$$

Then,

$$\mathcal{F}^{-1}\big[(1 + b^2 t^2)^{-1}\big] = g_b \Rightarrow \mathcal{F}^{-1}\left[\frac{2ib^2 t}{1 + b^2 t^2}\phi_X(t)\right] = -2b^2 (f_X * g_b)'.$$

Substituting into the general formula (2.1) gives the announced result and ends the proof. □

The next example is central for the paper. Stable noises provide a continuum between the Gaussian case and the Cauchy case. Their denoising formula involves a fractional score operator. This is the non-local analogue of the usual score $\partial_x \ln f_X$.

Proposition 2.3 (Symmetric α-stable noise)**.** *Under the assumptions of Proposition 2.1 with ε following the symmetric α-stable distribution with $\alpha \in (1, 2]$ and scale parameter $c > 0$, that is, $\phi_\varepsilon(t) = e^{-c|t|^\alpha}$, we have*

$$\mathrm{E}[Y \mid X = x] = x - c\alpha \frac{(\mathcal{A}_{\alpha-1} f_X)(x)}{f_X(x)},$$

where the operator $\mathcal{A}_{\alpha-1}$ is defined by $\mathcal{F}[\mathcal{A}_{\alpha-1} f](t) = i\,\mathrm{sign}(t)|t|^{\alpha-1}\mathcal{F}[f](t)$.

Proof. We have

$$\eta(t) = -c|t|^\alpha \text{ and } \eta'(t) = -c\alpha|t|^{\alpha-1}\mathrm{sign}(t) \Rightarrow \frac{1}{i}\eta'(t)\phi_X(t) = ic\alpha|t|^{\alpha-1}\mathrm{sign}(t)\phi_X(t),$$

which gives the result by (2.1). □

Two boundary cases are worth emphasizing. When $\alpha = 2$, the fractional operator reduces to an ordinary derivative and one recovers the Gaussian Tweedie formula. When $\alpha = 1$, the operator becomes the Hilbert transform, which is the natural score-type object for Cauchy noise. This is formalized in the next result.

Corollary 2.4 (Gaussian boundary case and formal Cauchy limit)**.** *If $\alpha = 2$ then $\varepsilon \sim \mathrm{Normal}(0, \sigma^2)$, with $c = \sigma^2/2$, and*

$$\mathrm{E}[Y \mid X = x] = x + \sigma^2 (\log f_X)'(x).$$

Formally, when $\alpha \downarrow 1$, the fractional operator $\mathcal{A}_{\alpha-1}$ converges to the Hilbert-transform-type operator $\mathcal{A}_0$. This gives the Cauchy-type expression

$$\mathrm{E}[Y \mid X = x] = x - c\,\frac{\mathcal{A}_0 f_X(x)}{f_X(x)},$$

whenever the formula is meaningful. The Cauchy case is treated separately in Appendix A, because ordinary L^1 integrability fails for the noise and for the associated score in general.

We next consider finite-activity jump noise. This case is important because the reverse dynamics are not diffusive: the reverse drift must be interpreted as a local mean drift, or equivalently as the reverse generator applied to the identity function. The formula below is the jump analogue of Tweedie's identity.

Proposition 2.5 (Centered compound Poisson noise)**.** *Let $N \sim \text{Poisson}(\lambda)$. Let $J_1, J_2, \dots$ be independent random variables with common distribution F_J and mean m, such that $\mathrm{E}[|J_1|] < \infty$. Assume that $N, J_1, J_2, \dots$ are independent. Define*

$$\varepsilon = \sum_{k=1}^{N} J_k - \lambda m.$$

Under the assumptions of Proposition 2.1, we then have

$$\mathrm{E}[Y \mid X = x] = x - \lambda \frac{(f_X * \nu)(x)}{f_X(x)},$$

where $\nu(du) = uF_J(du) - m\delta_0(du)$.

Proof. We have

$$\phi_\varepsilon(t) = \exp\big(\lambda(\phi_J(t) - 1 - itm)\big) \Rightarrow \eta'(t) = \lambda(\phi'_J(t) - im).$$

Moreover,

$$\phi'_J(t) = i\mathrm{E}[J_1 e^{itJ_1}] \Rightarrow \frac{1}{i}\eta'(t) = \lambda\left(\mathrm{E}[J_1 e^{itJ_1}] - m\right) = \lambda \int_{\mathbb{R}} e^{itu}\nu(du).$$

Hence, substituting

$$\mathcal{F}^{-1}\left[\frac{1}{i}\eta'(t)\phi_X(t)\right](x) = \lambda(f_X * \nu)(x)$$

into the general formula (2.1) yields the result. □

The following special cases make the jump nature of the formula explicit. When the jump is deterministic, the correction is a finite difference of the density. When the jump is symmetric, the correction is a centered difference. This is summarized in the next result.

Corollary 2.6.

- *If $J_1 \equiv a$, then*

$$\mathrm{E}[Y \mid X = x] = x - \lambda a \frac{f_X(x-a) - f_X(x)}{f_X(x)}.$$

- *If $J_1 = \pm a$ symmetrically, then*

$$\mathrm{E}[Y \mid X = x] = x + \frac{\lambda a}{2}\frac{f_X(x+a) - f_X(x-a)}{f_X(x)}.$$

Examples in Corollary 2.6 can be summarized as follows: the conditional mean is always the observation x minus a noise-dependent score correction. The form of this correction is local for Gaussian noise, smoothed for Laplace noise, fractional for stable noise, and discrete for compound Poisson noise.

Remark 2.7 (Unified operator form)**.** In all cases, the conditional expectation can be written formally as

$$\mathrm{E}[Y \mid X = x] = x - \frac{\mathcal{L}_\varepsilon^* f_X(x)}{f_X(x)},$$

where the adjoint operator $\mathcal{L}_\varepsilon^*$ associated with the characteristic exponent of the noise is defined, at least formally, by the Fourier multiplier

$$\mathcal{F}[\mathcal{L}_\varepsilon^* f](t) = \frac{1}{i}\eta'(t)\mathcal{F}[f](t) \text{ with } \phi_\varepsilon(t) = e^{\eta(t)}.$$

Here, Gaussian noise yields a local differential operator, stable noise yields a fractional non-local operator, and compound Poisson noise yields a jump or finite-difference operator. Thus, the Gaussian case appears to be an exception: it is the only case in this list where the score correction is purely local. This distinction will be important when comparing diffusion reverse drifts with jump and stable reverse local mean drifts.

We now turn to a structural question. In the Gaussian case with a Gaussian prior, the conditional mean is linear in the observation. The same phenomenon appears for compound Poisson or stable variables when the signal and the noise belong to the same convolution semigroup. The next result shows that, within the infinitely divisible class, this is the only possible mechanism.

Proposition 2.8 (Characterization of proportional conditional means in the infinitely divisible class)**.** *Let Y and ε be two independent integrable infinitely divisible random variables, and set $X = Y + \varepsilon$. Assume that X is not deterministic. Let $\varphi_Y(t) = e^{\xi(t)}$ and $\varphi_\varepsilon(t) = e^{\eta(t)}$, where ξ and η are the corresponding Lévy exponents. Then, the following assertions are equivalent:*

(1) There exists a constant $a \in \mathbb{R}$ such that the identity $\mathrm{E}[Y \mid X] = aX$ a.s. holds.

(2) There exist a Lévy exponent ψ and constants $\lambda, \mu \geq 0$, not both zero, such that $\varphi_Y(t) = e^{\lambda\psi(t)}$ and $\varphi_\varepsilon(t) = e^{\mu\psi(t)}$, $t \in \mathbb{R}$. Equivalently, Y and ε have proportional Lévy triplets, that is, $(b_Y, \sigma_Y^2, \nu_Y) = \lambda(b, \sigma^2, \nu)$ and $(b_\varepsilon, \sigma_\varepsilon^2, \nu_\varepsilon) = \mu(b, \sigma^2, \nu)$, for some Lévy triplet (b, σ^2, ν).

In that case,

$$\mathrm{E}[Y \mid X] = \frac{\lambda}{\lambda + \mu} X \qquad a.s.$$

In particular, if both Y and ε are non-degenerate, then the proportionality constant belongs to $(0, 1)$.

Proof. We first prove (2) $\Rightarrow$ (1). To this end, assume that there exist a Lévy exponent ψ and constants $\lambda, \mu \geq 0$ such that $\varphi_Y(t) = e^{\lambda\psi(t)}$ and $\varphi_\varepsilon(t) = e^{\mu\psi(t)}$. Then,

$$\varphi_X(t) = \varphi_Y(t)\varphi_\varepsilon(t) = e^{(\lambda+\mu)\psi(t)}.$$

Since Y and ε are integrable, their characteristic functions are differentiable, and by independence,

$$\mathrm{E}[Ye^{itX}] = \mathrm{E}[Ye^{itY}]\mathrm{E}[e^{it\varepsilon}] = \frac{1}{i}\varphi_Y'(t)\varphi_\varepsilon(t).$$

Using the specific form of the exponents,

$$\varphi_Y'(t)\varphi_\varepsilon(t) = \lambda\psi'(t)e^{(\lambda+\mu)\psi(t)} = \frac{\lambda}{\lambda+\mu}\varphi_X'(t).$$

Hence,

$$\mathrm{E}[Ye^{itX}] = \frac{\lambda}{\lambda+\mu}\frac{1}{i}\varphi_X'(t) = \mathrm{E}\left[\frac{\lambda}{\lambda+\mu}Xe^{itX}\right] \Rightarrow \mathrm{E}\left[\left(Y - \frac{\lambda}{\lambda+\mu}X\right)e^{itX}\right] = 0, \ \ t \in \mathbb{R}.$$

Define the finite signed measure

$$\Gamma(B) := \mathrm{E}\left[\left(Y - \frac{\lambda}{\lambda+\mu}X\right)\mathrm{I}_{\{X \in B\}}\right].$$

Its Fourier transform is identically zero so that $\Gamma \equiv 0$. Therefore

$$\mathrm{E}\left[Y - \frac{\lambda}{\lambda+\mu}X \,\middle|\, X\right] = 0,$$

which proves that the announced identity is valid.

We now prove (1) $\Rightarrow$ (2). Assume that for some constant $a \in \mathbb{R}$, $\mathrm{E}[Y \mid X] = aX$ a.s. holds. Then, for every $t \in \mathbb{R}$,

$$\mathrm{E}[Ye^{itX}] = a\mathrm{E}[Xe^{itX}].$$

Using independence and

$$\mathrm{E}[Ye^{itX}] = \frac{1}{i}\varphi_Y'(t)\varphi_\varepsilon(t), \qquad \mathrm{E}[Xe^{itX}] = \frac{1}{i}\varphi_X'(t),$$

we get

$$\varphi_Y'(t)\varphi_\varepsilon(t) = a\varphi_X'(t) = a\left(\varphi_Y'(t)\varphi_\varepsilon(t) + \varphi_Y(t)\varphi_\varepsilon'(t)\right).$$

Since infinitely divisible characteristic functions do not vanish, we may divide by $\varphi_Y(t)\varphi_\varepsilon(t)$ and obtain

$$(1-a)\frac{\varphi_Y'(t)}{\varphi_Y(t)} = a\frac{\varphi_\varepsilon'(t)}{\varphi_\varepsilon(t)}.$$

In terms of the Lévy exponents ξ and η, this means $(1-a)\xi'(t) = a\eta'(t)$, $t \in \mathbb{R}$. If $a = 0$ then $\xi' \equiv 0$, hence $\xi \equiv 0$, so $Y \equiv 0$ a.s. If $a = 1$ then $\eta' \equiv 0$, hence $\eta \equiv 0$, so $\varepsilon \equiv 0$ a.s. In both cases, the representation is trivial. Assume now that $a \notin \{0, 1\}$. Then

$$\xi'(t) = \frac{a}{1-a}\eta'(t).$$

Since $\xi(0) = \eta(0) = 0$, integration yields

$$\xi(t) = \frac{a}{1-a}\eta(t), \qquad t \in \mathbb{R}.$$

With $\psi(t) := \xi(t) + \eta(t) = \ln \varphi_X(t)$, we have

$$\xi(t) = a\psi(t), \qquad \eta(t) = (1-a)\psi(t) \Rightarrow \varphi_Y(t) = e^{a\psi(t)}, \qquad \varphi_\varepsilon(t) = e^{(1-a)\psi(t)}.$$

Since ξ and η are Lévy exponents, this representation implies that $a\psi$ and $(1-a)\psi$ are Lévy exponents as well. Equivalently, the Lévy triplets of Y and ε are proportional to the Lévy triplet of X, that is, $(b_Y, \sigma_Y^2, \nu_Y) = a(b_X, \sigma_X^2, \nu_X)$ and $(b_\varepsilon, \sigma_\varepsilon^2, \nu_\varepsilon) = (1-a)(b_X, \sigma_X^2, \nu_X)$. Since X is not deterministic by assumption, the Lévy triplet of X has either $\sigma_X^2 > 0$ or $\nu_X \neq 0$. Since $a\psi$ and $(1-a)\psi$ are Lévy exponents, their Gaussian and jump parts must be non-negative. This forces $a \geq 0$ and $1 - a \geq 0$ to be valid. Hence, $a \in [0, 1]$. Setting $\lambda = a$ and $\mu = 1 - a$ gives the announced representation. This ends the proof. □

Corollary 2.9 (Affine conditional means)**.** *Let Y and ε be independent integrable infinitely divisible random variables and $X = Y + \varepsilon$. Assume that X is not deterministic. Then,*

$$\mathrm{E}[Y \mid X] = \alpha X + \beta$$

for some constants α, β if and only if, after centering, $Y - \mathrm{E}[Y]$ and $\varepsilon - \mathrm{E}[\varepsilon]$ have proportional Lévy exponents. In that case,

$$\mathrm{E}[Y \mid X] = \mathrm{E}[Y] + \frac{\lambda}{\lambda+\mu}\left(X - \mathrm{E}[X]\right).$$

Proof. Set

$$\widetilde{Y} := Y - \mathrm{E}[Y], \qquad \widetilde{\varepsilon} := \varepsilon - \mathrm{E}[\varepsilon], \qquad \widetilde{X} := X - \mathrm{E}[X].$$

Then,

$$\mathrm{E}[Y \mid X] = \alpha X + \beta \Leftrightarrow \mathrm{E}[\widetilde{Y} \mid \widetilde{X}] = \alpha \widetilde{X}.$$

Since translations preserve infinite divisibility and only modify the drift part of the Lévy exponent, Proposition 2.8 applies to $\widetilde{Y}$ and $\widetilde{\varepsilon}$. Hence the centered variables have proportional Lévy exponents if, and only if,

$$\mathrm{E}[\widetilde{Y} \mid \widetilde{X}] = \frac{\lambda}{\lambda+\mu}\widetilde{X}.$$

Returning to the original variables proves the result. □

Remark 2.10 (Connection with Bartlett representation and convolution semigroups)**.** Under the regularity assumptions of Proposition 2.1, Bartlett representation formula (2.1) shows that

$$\mathrm{E}[Y \mid X = x] = ax \Leftrightarrow \mathcal{F}^{-1}\left[\frac{1}{i}\eta'(t)\varphi_X(t)\right](x) = (1-a)xf_X(x).$$

Taking Fourier transforms, the right-hand side of the equivalence becomes $\eta'(t)\varphi_X(t) = (1-a)\varphi'_X(t)$. Since

$$\varphi'_X(t) = (\xi'(t) + \eta'(t))\varphi_X(t),$$

and since infinitely divisible characteristic functions do not vanish, we obtain $(1-a)\xi'(t) = a\eta'(t)$. This is precisely the differential relation between the Lévy exponents used in Proposition 2.8. Consequently, within the non-deterministic infinitely divisible class, $\mathrm{E}[Y \mid Y+\varepsilon]$ is proportional to $Y+\varepsilon$ if, and only if, Y and ε belong to the same convolution semigroup, up to deterministic scaling. This observation explains why the Gaussian–Gaussian, compound-Poisson–compound-Poisson and stable–stable examples treated later lead to linear conditional means. It also clarifies that such linearity is not a generic property of additive noise models, but a signature of semigroup compatibility between signal and noise.

3 Model setup and assumptions

We now introduce the additive forward model used throughout the paper. The setting is chosen to cover both diffusion and jump examples, including time-inhomogeneous models with deterministic volatility, intensity or stable scale. Thus, the driving noise is not necessarily a Lévy process with stationary increments. It is more generally an additive process: its increments are independent, but their distributions may depend on the time interval through a deterministic clock.

The central structural feature is the decomposition $X_t = Y + Z_t$, $0 \le t \le T$, where $Y = X_0$ is the initial signal and Z_t is an independent centered noise. Conditional on $Y = y$, the distribution of X_t is simply the distribution of $y + Z_t$. This explicit decoupling between the initial signal and the transition noise is the reason why the additive model is particularly transparent for denoising, score identities and time reversal.

Throughout the paper, we work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathrm{P})$ satisfying the usual conditions and we make the following assumptions.

3.1 Initial condition

The integrability of Y is a standing assumption because the main denoising quantity is the conditional mean $\mathrm{E}[Y \mid X_t]$, and its expectation must be finite for comparisons in convex order, since the convex order is defined for integrable random variables with equal means. This is precisely stated in the next assumption.

Assumption 3.1 (Initial distribution)**.** *Let Y be a $\mathcal{F}_0$-measurable real-valued random variable with distribution p_0, such that $\mathrm{E}[|Y|] < \infty$.*

3.2 Centered additive noise

The noising mechanism $(Z_t)_{t\in[0,T]}$ corrupting Y must fulfill the following conditions.

Assumption 3.2 (Centered independent-increments noise)**.** *Let $(Z_t)_{t\in[0,T]}$ be a $(\mathcal{F}_t)$-adapted real-valued process with independent increments such that:*

(1) $Z_0 = 0$;

(2) for every $0 \le s < t \le T$, the increment $Z_t - Z_s$ is independent of $\sigma(Y, Z_r : 0 \le r \le s)$;

(3) for every $t \in [0,T]$, Z_t is integrable (that is, $\mathrm{E}[|Z_t|] < \infty$) and centered (that is, $\mathrm{E}[Z_t] = 0$);

(4) $(Z_t)_{t\in[0,T]}$ is stochastically continuous, that is, for every $s \in [0,T]$, $Z_t \to Z_s$ in probability as $t \to s$.

The independence assumption means that the forward noising mechanism is external to the initial signal. The centering assumption ensures that the forward process preserves the mean: noise increases dispersion but does not change the first moment. This distinction is crucial for the use of convex order. Since $X_t - X_s = Z_t - Z_s$ for $0 \le s < t \le T$, the increments of (X_t) after time 0 inherit the independent-increment structure from (Z_t). More precisely, $X_t - X_s$ is independent of $\sigma(Y, Z_r : 0 \le r \le s)$, and hence independent of the past generated by the forward process.

3.3 Forward noising and convex-order increase

In differential form, we have $dX_t = dZ_t$ where the increment dZ_t is understood in the sense of processes with independent increments. In particular, stationarity of the increments is not required. The centering assumption gives $\mathrm{E}[X_t] = \mathrm{E}[Y] + \mathrm{E}[Z_t] = \mathrm{E}[Y]$, $0 \le t \le T$. Thus the forward process preserves the mean of the initial signal. At the same time, the process becomes more variable as independent centered noise is accumulated.

Let us recall the stochastic order used throughout the paper to compare dispersion (see, e.g., Shaked and Shanthikumar [9]).

Definition 3.3 (Convex order)**.** *Let U and V be integrable random variables. We say that U is smaller than V in the convex order, and write $U \preceq_{\mathrm{cx}} V$, if $\mathrm{E}[\Phi(U)] \le \mathrm{E}[\Phi(V)]$ holds for every convex function $\Phi : \mathbb{R} \to \mathbb{R}$ for which the expectations exist. Equivalently, $U \preceq_{\mathrm{cx}} V$ holds if, and only if, $\mathrm{E}[U] = \mathrm{E}[V]$ and $\mathrm{E}[(U-k)_+] \le \mathrm{E}[(V-k)_+]$ for all $k \in \mathbb{R}$.*

The convex order compares random variables with the same mean according to their dispersion. In the sequel, the forward process adds centered noise and hence preserves the mean. Convex order is therefore the natural stochastic order relation for formalizing the idea that the forward observation becomes more variable as the noise level increases.

Proposition 3.4 (Forward increase in convex order)**.** *Under Assumptions 3.1 and 3.2, the stochastic inequality $X_s \preceq_{\mathrm{cx}} X_t$ holds for every $0 \le s \le t \le T$.*

Proof. For $0 \le s \le t \le T$, we have $X_t = X_s + (Z_t - Z_s)$, where $Z_t - Z_s$ is independent of the past and satisfies $\mathrm{E}[Z_t - Z_s] = 0$. Hence, $\mathrm{E}[X_t \mid X_s] = X_s$. Equivalently, X_s is a conditional mean of X_t. By Jensen's inequality,

$$\Phi(X_s) = \Phi(\mathrm{E}[X_t \mid X_s]) \le \mathrm{E}[\Phi(X_t) \mid X_s].$$

Taking expectations gives $\mathrm{E}[\Phi(X_s)] \le \mathrm{E}[\Phi(X_t)]$. We can then conclude that $X_s \preceq_{\mathrm{cx}} X_t$ is valid, which ends the proof. □

Proposition 3.4 formalizes the forward deconstruction mechanism. As t increases, the observation X_t keeps the same mean but becomes more dispersed. The reverse-time results proved later in the paper show that the conditional mean and the normalized reverse score inherit a dual monotonicity property along reverse time.

3.4 Characteristic exponents and deterministic clocks

Under Assumption 3.2, (Z_t) is an integrable centered additive process. In particular, for every $t \in [0,T]$, the law of Z_t is infinitely divisible. We write its characteristic function as $\mathrm{E}[e^{iuZ_t}] = \exp\big(\Psi_t(u)\big)$, $u \in \mathbb{R}$. The family $(\Psi_t)_{t\in[0,T]}$ is additive in the sense that, for $0 \le s < t \le T$, the increment $Z_t - Z_s$ has characteristic function

$$\mathrm{E}[e^{iu(Z_t - Z_s)}] = \exp\big(\Psi_t(u) - \Psi_s(u)\big).$$

The centering condition is expressed at the level of the exponent by

$$\frac{1}{i}\partial_u \Psi_t(0) = 0, \qquad 0 \le t \le T.$$

When (Z_t) is a genuine Lévy process with stationary increments, one has $\Psi_t(u) = t\psi(u)$ for a single Lévy exponent ψ. In the time-inhomogeneous examples considered in this paper, the exponent instead takes the form $\Psi_t(u) = \kappa_t \psi(u)$ where κ_t is a deterministic increasing clock. These clocks are the correct normalizations in the Bartlett–Tweedie formulas and in the reverse-time martingale properties. The actual reverse drift or reverse local mean drift is obtained by multiplying the normalized score by the instantaneous clock rate.

3.5 Canonical examples

We now list the main examples used in the remainder of the paper.

3.5.1 Gaussian noise with deterministic volatility

Let $(W_t)_{0\le t\le T}$ be a standard Brownian motion and let $\sigma : [0,T] \to [0,\infty)$ be deterministic and square-integrable (that is, $\int_0^T \sigma^2(r)\,dr < \infty$). Define

$$Z_t = \int_0^t \sigma(r)\,dW_r \sim \mathrm{Normal}(0, A_t).$$

The relevant clock for Gaussian noise with deterministic volatility is given by

$$A_t = \int_0^t \sigma^2(r)\,dr. \tag{3.1}$$

Here, $\Psi_t(u) = -\frac{1}{2}A_t u^2$. The corresponding forward model is

$$X_t = Y + Z_t = Y + \int_0^t \sigma(r)\,dW_r, \tag{3.2}$$

where Y is independent of $(Z_t)_{0\le t\le T}$. This is the continuous-time analogue of a variance-exploding forward noising mechanism: the conditional law of X_t given $Y = y$ is $\mathrm{Normal}(y, A_t)$.

3.5.2 Compensated compound Poisson noise with deterministic intensity

Let $(N_t)_{0\le t\le T}$ be a non-homogeneous Poisson process with deterministic intensity $\lambda : [0,T] \to [0,\infty)$, where λ is measurable and integrable (that is, $\int_0^T \lambda(r)\,dr < \infty$). Denote the integrated intensity as

$$\Lambda_t = \int_0^t \lambda(r)\,dr. \tag{3.3}$$

Let $J_1, J_2, \ldots$ be independent integrable jumps with common distribution F_J, independent of N. Denote $m = \mathrm{E}[J_1]$. The compensated compound Poisson noise is then defined by

$$Z_t = \sum_{k=1}^{N_t} J_k - m\Lambda_t.$$

Then, (Z_t) has independent increments and is centered. Its characteristic exponent is

$$\Psi_t(u) = \Lambda_t \left(\varphi_J(u) - 1 - ium\right).$$

The corresponding forward model is

$$X_t = Y + Z_t = Y + \sum_{k=1}^{N_t} J_k - m\Lambda_t, \tag{3.4}$$

where Y is independent of $(Z_t)_{0\leq t\leq T}$. Here, the cumulative jump intensity Λ_t is the correct clock for the denoising formula, whereas the reverse local mean drift involves the instantaneous intensity $\lambda(t)$.

3.5.3 Symmetric α-stable noise with deterministic scale

Let $(L_t)_{0\leq t\leq T}$ be a symmetric α-stable Lévy process, with $\alpha \in (1,2)$, normalized by $\mathrm{E}[e^{iuL_t}] = \exp(-t|u|^\alpha)$. Let $\sigma_L : [0,T] \to \mathbb{R}$ be deterministic and assume $\int_0^T |\sigma_L(r)|^\alpha \, dr < \infty$. Define

$$Z_t = \int_0^t \sigma_L(r)\, dL_r.$$

Then, Z_t is symmetric α-stable with characteristic exponent $\Psi_t(u) = -\Gamma_t|u|^\alpha$. The relevant clock for symmetric α-stable noise with deterministic scale is given by

$$\Gamma_t = \int_0^t |\sigma_L(r)|^\alpha \, dr. \tag{3.5}$$

Since $\alpha > 1$, Z_t is integrable and centered. The corresponding forward model is

$$X_t = Y + Z_t = Y + \int_0^t \sigma_L(r)\, dL_r, \tag{3.6}$$

where Y is independent of $(Z_t)_{0\leq t\leq T}$. The clock Γ_t plays the same role as the variance clock in the Gaussian case, but the associated reverse score is fractional and non-local.

3.5.4 Cauchy noise with deterministic scale

Let $(L_t)_{0\leq t\leq T}$ be a standard symmetric Cauchy process and let $\sigma_C : [0,T] \to \mathbb{R}$ be deterministic with $\int_0^T |\sigma_C(r)|\, dr < \infty$. Define

$$Z_t = \int_0^t \sigma_C(r)\, dL_r. \tag{3.7}$$

Then, $\Psi_t(u) = -\Gamma_t|u|$ and the relevant clock is

$$\Gamma_t = \int_0^t |\sigma_C(r)|\, dr. \tag{3.8}$$

This case is structurally similar to the stable case with $\alpha = 1$, but it falls outside Assumption 3.2 because Z_t is not integrable. For this reason, the Cauchy case is treated separately in Appendix. The conditional mean $\mathrm{E}[Y \mid X_t]$ remains well defined when Y is integrable, but the reverse score drift is not integrable in general and cannot be compared directly in the usual convex order.

3.5.5 General centered additive process

More generally, $(Z_t)_{0\le t\le T}$ may be any integrable additive process with characteristic exponent

$$\Psi_t(u) = iub_t - \frac{1}{2}c_t u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathrm{I}_{\{|x|\le 1\}}\right)\nu_t(dx),$$

where (b_t, c_t, ν_t) is a time-dependent Lévy–Khintchine triplet. The centering condition is

$$b_t + \int_{|x|>1} x\,\nu_t(dx) = 0.$$

Examples in Sections 3.5.1, 3.5.2, 3.5.3, and 3.5.4 correspond to deterministic-clock subfamilies of this general additive framework.

4 A clock-based martingale principle for reverse-time scores

The three examples studied in the next sections share the same probabilistic structure. The forward noise is driven by a deterministic clock, the Bartlett–Tweedie identity expresses the denoising error through this accumulated clock, and the corresponding normalized score is a martingale in reverse time. The actual reverse drift, or reverse local mean drift in the jump and stable cases, is then obtained by multiplying this normalized score by the instantaneous clock rate. The purpose of this section is to isolate this common mechanism.

Let $(\mu_a)_{a\ge 0}$ be a centered convolution semigroup on $\mathbb{R}$, with $\mu_0 = \delta_0$ and $\mu_{a+b} = \mu_a * \mu_b$. Let $\kappa : [0,T] \to [0,\infty)$ be a deterministic nondecreasing clock such that $\kappa_0 = 0$ and $\kappa_t > 0$ for all $t > 0$. Consider a forward process $X_t = Y + Z_t$, $0 \le t \le T$, where the initial signal Y satisfies Assumption 3.1, and where the additive noise process $(Z_t)_{t\in[0,T]}$ satisfies Assumption 3.2. Assume moreover that the increments are of the form

$$Z_t - Z_s \sim \mu_{\kappa_t - \kappa_s}, \qquad 0 \le s \le t \le T.$$

We write p_t for the law of X_t and, whenever densities are used, we identify p_t with its density. Thus, formally, $p_t = p_0 * \mu_{\kappa_t}$. For $a \ge 0$, define the convolution operator

$$P_a q := q * \mu_a.$$

Hence, $p_t = P_{\kappa_t} p_0$ and, for $0 \le r \le t \le T$, $p_t = P_{\kappa_t - \kappa_r} p_r$.

We now impose the following regularity assumption.

Assumption 4.1 (Score representation regularity)**.** *In the setting above, assume that there exists a linear operator C, acting on a linear domain $\mathcal{D}(C)$ of densities, such that the following properties hold.*

1. *For every $t > 0$, the law of X_t admits a strictly positive density p_t, with $p_t \in \mathcal{D}(C)$, and Cp_t locally integrable.*

2. *For every $q \in \mathcal{D}(C)$ and every $a \ge 0$, one has $P_a q \in \mathcal{D}(C)$, the convolution $P_a(Cq)$ is well defined as a locally integrable function, and*

$$C(P_a q) = P_a(Cq).$$

3. *The clock-based Bartlett–Tweedie identity holds: for every $t > 0$,*

$$m_t(x) := \mathrm{E}[Y \mid X_t = x] = x - \kappa_t\,\frac{(Cp_t)(x)}{p_t(x)}.$$

The normalized score associated with the clock κ is defined by

$$S_t := \frac{M_t - X_t}{\kappa_t}, \qquad M_t := \mathrm{E}[Y \mid X_t], \qquad t > 0.$$

By Assumptions 3.1 and 3.2, Y and $X_t = Y + Z_t$ are integrable. Hence $M_t = \mathrm{E}[Y \mid X_t]$ is integrable, and since $\kappa_t > 0$ for $t > 0$, the random variable S_t is integrable. Moreover, by the preceding Bartlett–Tweedie identity,

$$S_t = -\frac{(Cp_t)(X_t)}{p_t(X_t)}.$$

This is the quantity which will be shown to be a reverse-time martingale.

Let

$$\mathcal{G}_t := \sigma(X_u : t \le u \le T), \qquad 0 \le t \le T,$$

be the decreasing filtration generated by the future of the forward process. In reverse time, we write

$$\overline{X}_s := X_{T-s}, \qquad \overline{\mathcal{F}}_s := \mathcal{G}_{T-s}, \qquad \overline{M}_s := M_{T-s}, \qquad \overline{S}_s := S_{T-s}, \qquad 0 \le s < T.$$

Proposition 4.2 (Clock-based reverse martingale principle)**.** *Under Assumption 4.1, for every* $0 < r < t \le T$*,*

$$\mathrm{E}[M_r \mid \mathcal{G}_t] = M_t, \qquad \mathrm{E}[S_r \mid \mathcal{G}_t] = S_t.$$

Equivalently, in reverse time, the processes $(\overline{M}_s)_{0 \le s < T}$ *and* $(\overline{S}_s)_{0 \le s < T}$ *are martingales with respect to* $(\overline{\mathcal{F}}_s)_{0 \le s < T}$*.*

Proof. Fix $0 < r < t \le T$. We first prove the reverse martingale property of the conditional mean. Recall that $M_u = \mathrm{E}[Y \mid X_u]$, $u \in [0, T]$. Since M_r is $\mathcal{H}_t$-measurable, Lemma B.1 gives $\mathrm{E}[M_r \mid \mathcal{G}_t] = \mathrm{E}[M_r \mid X_t]$. We now show that $\mathrm{E}[M_r \mid X_t] = M_t$. Indeed, since

$$X_t = X_r + (Z_t - Z_r),$$

and since $Z_t - Z_r$ is independent of $\sigma(Y, Z_u : 0 \le u \le r)$, the variables Y and X_t are conditionally independent given X_r. Indeed, conditionally on X_r, the variable X_t is obtained by adding the independent increment $Z_t - Z_r$, which is independent of $\sigma(Y, Z_u : 0 \le u \le r)$. Hence

$$\mathrm{E}[Y \mid X_r, X_t] = \mathrm{E}[Y \mid X_r] = M_r.$$

By the tower property,

$$\begin{aligned} M_t &= \mathrm{E}[Y \mid X_t] \\ &= \mathrm{E}\big[\mathrm{E}[Y \mid X_r, X_t] \big| X_t\big] \\ &= \mathrm{E}[M_r \mid X_t]. \end{aligned}$$

Therefore,

$$\mathrm{E}[M_r \mid \mathcal{G}_t] = \mathrm{E}[M_r \mid X_t] = M_t.$$

We now prove the reverse martingale property of the normalized score. Set

$$\Delta_{r,t} := \kappa_t - \kappa_r.$$

By Bayes' formula and the additive transition structure, the conditional law of X_r given $X_t = x$ is

$$\mathrm{P}[X_r \in dy \mid X_t = x] = \frac{p_r(y)\, \mu_{\Delta_{r,t}}(x - dy)}{p_t(x)}.$$

Using the representation

$$S_r = -\frac{(Cp_r)(X_r)}{p_r(X_r)},$$

we obtain, for every x such that $p_t(x) > 0$,

$$\begin{aligned} \mathrm{E}[S_r \mid X_t = x] &= -\frac{1}{p_t(x)} \int_{\mathbb{R}} (Cp_r)(y)\, \mu_{\Delta_{r,t}}(x - dy) \\ &= -\frac{P_{\Delta_{r,t}}(Cp_r)(x)}{p_t(x)}. \end{aligned}$$

Since C commutes with the convolution semigroup,

$$P_{\Delta_{r,t}}(Cp_r) = C(P_{\Delta_{r,t}} p_r) = Cp_t.$$

Consequently,

$$\mathrm{E}[S_r \mid X_t = x] = -\frac{(Cp_t)(x)}{p_t(x)}.$$

Evaluating at $x = X_t$, we get

$$\mathrm{E}[S_r \mid X_t] = -\frac{(Cp_t)(X_t)}{p_t(X_t)} = S_t.$$

Finally, since S_r is $\mathcal{H}_t$-measurable, Lemma B.1 again yields

$$\mathrm{E}[S_r \mid \mathcal{G}_t] = \mathrm{E}[S_r \mid X_t] = S_t.$$

Thus, for every $0 < r < t \le T$, $\mathrm{E}[M_r \mid \mathcal{G}_t] = M_t$ and $\mathrm{E}[S_r \mid \mathcal{G}_t] = S_t$. Replacing t by $T - s$ gives the reverse-time martingale formulation. □

The preceding result immediately implies the convex-order monotonicity of the reverse conditional mean and of the normalized score.

Corollary 4.3 (Convex-order monotonicity of normalized scores)**.** *Under Assumption 4.1, for every* $0 \le u \le s < T$,

$$\overline{M}_u \preceq_{\mathrm{cx}} \overline{M}_s, \qquad \overline{S}_u \preceq_{\mathrm{cx}} \overline{S}_s.$$

Thus both the reverse conditional mean and the reverse normalized score are increasing in convex order along reverse time.

Proof. The result follows directly from Proposition 4.2 and Jensen's inequality. Indeed, if (N_s) is an integrable martingale and $u \le s$, then

$$N_u = \mathrm{E}[N_s \mid \overline{\mathcal{F}}_u],$$

so that, for every convex function Φ for which the expectations exist,

$$\mathrm{E}[\Phi(N_u)] = \mathrm{E}\big[\Phi\big(\mathrm{E}[N_s \mid \overline{\mathcal{F}}_u]\big)\big] \le \mathrm{E}[\Phi(N_s)].$$

Applying this argument to $N = \overline{M}$ and $N = \overline{S}$ gives the announced convex-order inequalities. □

The previous result concerns the normalized score. The actual reverse drift, or reverse local mean drift, usually involves the instantaneous rate of the clock. We fix a non-negative measurable version a of κ_t'. When pointwise statements about the reverse drift are made, we assume that this version is finite at the considered times. In the examples below,

$$a_t = \sigma^2(t), \qquad a_t = \lambda(t), \qquad a_t = |\sigma_L(t)|^{\alpha},$$

respectively in the Gaussian, compensated compound Poisson and symmetric α-stable cases.

Suppose that the reverse drift, or reverse local mean drift, has the form

$$\overline{R}_s = a_{T-s}\,\overline{S}_s, \qquad 0 \le s < T.$$

Then $\overline{R}$ is generally not a martingale. Indeed, for $0 \le u \le s < T$,

$$\mathrm{E}[\overline{R}_s \mid \overline{\mathcal{F}}_u] = a_{T-s}\,\overline{S}_u \text{ whereas } \overline{R}_u = a_{T-u}\,\overline{S}_u.$$

These two quantities coincide only when the instantaneous rate is constant on the relevant interval, or in degenerate cases.

Nevertheless, the convex-order monotonicity can be recovered under a monotonicity condition on the instantaneous rate.

Proposition 4.4 (Clock-rate condition for drift peacocks)**.** *Assume that Assumption 4.1 holds. Assume moreover that κ is absolutely continuous, that*

$$\overline{R}_s = a_{T-s}\,\overline{S}_s, \qquad 0 \le s < T,$$

is integrable for every $s < T$, and that the map $s \longmapsto a_{T-s}$ is nondecreasing on $[0, T]$. Then

$$\overline{R}_u \preceq_{\mathrm{cx}} \overline{R}_s, \qquad 0 \le u \le s < T.$$

Thus $(\overline{R}_s)_{0\le s<T}$ is a peacock.

Proof. By Corollary 4.3, $\overline{S}_u \preceq_{\mathrm{cx}} \overline{S}_s$, $0 \le u \le s < T$. Moreover,

$$\mathrm{E}[\overline{S}_s] = \frac{\mathrm{E}[M_{T-s}] - \mathrm{E}[X_{T-s}]}{\kappa_{T-s}} = \frac{\mathrm{E}[Y] - \mathrm{E}[Y]}{\kappa_{T-s}} = 0.$$

Since $s \mapsto a_{T-s}$ is nondecreasing, we have

$$0 \le a_{T-u} \le a_{T-s}, \qquad 0 \le u \le s < T.$$

Multiplication by a nonnegative deterministic constant preserves the convex order, and increasing the deterministic scale of a centered integrable random variable increases it in the convex order. Hence

$$a_{T-u}\overline{S}_u \preceq_{\mathrm{cx}} a_{T-u}\overline{S}_s \preceq_{\mathrm{cx}} a_{T-s}\overline{S}_s.$$

That is, $\overline{R}_u \preceq_{\mathrm{cx}} \overline{R}_s$, which proves that $\overline{R}$ is increasing in convex order. □

Remark 4.5 (Normalized score versus actual reverse drift)**.** The distinction between $\overline{S}_s$ and $\overline{R}_s$ is essential in time-inhomogeneous models. The normalized score is governed by the accumulated clock κ_{T-s} and is a martingale in reverse time. The actual reverse drift is governed by the instantaneous rate a_{T-s}. Therefore it is, in general, only a deterministic time-dependent dilation of a martingale, not a martingale itself.

Remark 4.6 (Canonical specializations)**.** The three main models considered below fit the preceding framework as follows. The operator C is defined by the identity

$$m_t(x) - x = -\kappa_t\,\frac{(Cp_t)(x)}{p_t(x)}.$$

Table 4.1 lists the respective clock, operator, and instantaneous rate for the three main models considered in this paper.

Thus, the normalized score is always

$$S_t = -\frac{(Cp_t)(X_t)}{p_t(X_t)}.$$

Model	Clock κ_t	Operator C	Instantaneous rate $a_t = \kappa_t'$
Gaussian	$A_t = \int_0^t \sigma^2(r)\,dr$	$Cq = -\partial_x q$	$\sigma^2(t)$
Compensated compound Poisson	$\Lambda_t = \int_0^t \lambda(r)\,dr$	$Cq = q * \chi_J$	$\lambda(t)$
Symmetric α-stable	$\Gamma_t = \int_0^t \lvert\sigma_L(r)\rvert^\alpha\,dr$	$Cq = \alpha\mathcal{A}_{\alpha-1}q$	$\lvert\sigma_L(t)\rvert^\alpha$

Table 4.1: Clock, operator, and instantaneous rate for the three main models.

In the Gaussian case, $S_t = \partial_x \log p_t(X_t)$. In the compound Poisson case,

$$S_t = -\frac{(p_t * \chi_J)(X_t)}{p_t(X_t)}.$$

In the symmetric α-stable case,

$$S_t = -\alpha\,\frac{(\mathcal{A}_{\alpha-1}p_t)(X_t)}{p_t(X_t)}.$$

The following sections verify these identities and identify the corresponding reverse drift or reverse local mean drift in each model.

5 Gaussian forward dynamics with deterministic volatility and reverse-time representation

This section specializes the clock-based principle of Section 4 to the Gaussian additive model with deterministic time-dependent volatility. The relevant accumulated clock is the variance clock A_t in Table 4.1 whereas the actual reverse drift involves the instantaneous variance rate $\sigma^2(t)$. Thus the normalized reverse score is controlled by the accumulated variance, while the reverse drift is obtained by multiplying this score by the instantaneous rate.

5.1 Gaussian variance clock and reverse diffusion

Let $(W_t)_{0\le t\le T}$ be a standard Brownian motion independent of Y, and consider the forward process

$$X_t = Y + \int_0^t \sigma(r)\,dW_r, \qquad 0 \le t \le T,$$

where $\sigma : [0,T] \to [0,\infty)$ is deterministic and square-integrable. We assume that

$$0 < A_t < \infty, \qquad t > 0.$$

Equivalently, by deterministic time change, one may write

$$X_t = Y + B_{A_t},$$

where B is a standard Brownian motion independent of Y. Hence the Gaussian model is an additive noise model driven by the convolution semigroup

$$\mu_a = \mathrm{Normal}(0,a), \qquad a \ge 0,$$

and by the deterministic clock $\kappa_t = A_t$.

Let p_t denote the law of X_t. For $t > 0$, p_t admits a smooth strictly positive density, still denoted by p_t, and

$$p_t = p_0 * \varphi_{A_t}, \qquad \varphi_v(x) = \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{x^2}{2v}\right).$$

The forward generator is time-inhomogeneous and acts on smooth test functions f as

$$(\mathcal{L}_t f)(x) = \frac{\sigma^2(t)}{2} f''(x),$$

and its adjoint acts on densities as

$$(\mathcal{L}_t^* q)(x) = \frac{\sigma^2(t)}{2} \partial_{xx} q(x).$$

Consequently, the marginal densities solve the non-homogeneous heat equation

$$\partial_t p_t(x) = \frac{\sigma^2(t)}{2} \partial_{xx} p_t(x), \qquad p_{t=0} = p_0.$$

Let $\overline{X}_s := X_{T-s}$, $0 \le s \le T$, be the reverse-time process. Under conditions ensuring the validity of the classical time-reversal theorem for one-dimensional diffusions, for instance $p_t > 0$, $p_t \in C^{1,2}$ on $(0,T] \times \mathbb{R}$, and sufficient local integrability of $\partial_x \log p_t$, $(\overline{X}_s)_{0 \le s < T}$ is a time-inhomogeneous diffusion satisfying

$$d\overline{X}_s = b_s(\overline{X}_s)\, ds + \sigma(T-s)\, d\overline{W}_s,$$

where $(\overline{W}_s)$ is a Brownian motion in reverse time and

$$b_s(x) = \sigma^2(T-s) \partial_x \log p_{T-s}(x).$$

The corresponding backward generator is

$$(\overline{\mathcal{L}}_s f)(x) = b_s(x) f'(x) + \frac{\sigma^2(T-s)}{2} f''(x).$$

Thus, the Gaussian reverse drift is local: it is the score of the marginal density p_{T-s}, multiplied by the instantaneous variance rate $\sigma^2(T-s)$.

5.2 Tweedie formula, normalized score and reverse drift

We now identify the denoising correction and the Gaussian instance of the general clock-based score. The Gaussian convolution semigroup is generated by the heat semigroup

$$(P_a q)(x) = \int_{\mathbb{R}} q(y) \varphi_a(x-y)\, dy, \qquad a \ge 0.$$

In the notation of Section 4, the operator C is $Cq = -\partial_x q$. Indeed, C commutes with the heat semigroup, since

$$\partial_x P_a q = P_a(\partial_x q), \qquad a \ge 0.$$

Proposition 5.1 (Gaussian specialization of the clock-based score)**.** *In the Gaussian setting above, for every $t > 0$ and every $x \in \mathbb{R}$ such that $p_t(x) > 0$,*

$$m_t(x) := \mathrm{E}[Y \mid X_t = x] = x + A_t\, \partial_x \log p_t(x).$$

Consequently, with

$$M_t := \mathrm{E}[Y \mid X_t], \qquad S_t := \frac{M_t - X_t}{A_t},$$

one has

$$S_t = \partial_x \log p_t(X_t).$$

Moreover, the actual reverse drift satisfies

$$b_s(x) = \frac{\sigma^2(T-s)}{A_{T-s}} \left(m_{T-s}(x) - x \right), \qquad 0 \le s < T,$$

or, equivalently,

$$b_s(\overline{X}_s) = \sigma^2(T-s)\, \overline{S}_s, \qquad \overline{S}_s := S_{T-s}.$$

Proof. For fixed $t > 0$, the noise

$$Z_t = \int_0^t \sigma(r)\, dW_r$$

is Gaussian with variance A_t. The Gaussian Bartlett–Tweedie identity gives

$$\mathrm{E}[Y \mid X_t = x] = x + A_t\, \partial_x \log p_t(x).$$

Thus,

$$S_t = \frac{M_t - X_t}{A_t} = \partial_x \log p_t(X_t).$$

Taking $t = T - s$ in the preceding identity yields

$$\partial_x \log p_{T-s}(x) = \frac{m_{T-s}(x) - x}{A_{T-s}}.$$

Since the reverse drift is

$$b_s(x) = \sigma^2(T-s)\partial_x \log p_{T-s}(x),$$

the stated representation follows. $\square$

The Gaussian case therefore fits exactly the framework of Section 4 with

$$\kappa_t = A_t, \qquad a_t = \kappa_t' = \sigma^2(t), \qquad C = -\partial_x.$$

Consequently, the martingale and convex-order results follow directly from the general clock-based principle.

Corollary 5.2 (Reverse martingales and convex-order monotonicity). *Let*

$$\overline{M}_s := M_{T-s}, \qquad \overline{S}_s := S_{T-s}, \qquad \overline{B}_s := b_s(\overline{X}_s) = \sigma^2(T-s)\overline{S}_s.$$

Then, $(\overline{M}_s)_{0 \le s < T}$ and $(\overline{S}_s)_{0 \le s < T}$ are martingales with respect to the reverse filtration $\overline{\mathcal{F}}_s := \sigma(\overline{X}_u : 0 \le u \le s)$. In particular,

$$\overline{M}_u \preceq_{\mathrm{cx}} \overline{M}_s, \qquad \overline{S}_u \preceq_{\mathrm{cx}} \overline{S}_s, \qquad 0 \le u \le s < T.$$

If, in addition, the map $s \mapsto \sigma^2(T-s)$ is non-decreasing on $[0,T]$, then

$$\overline{B}_u \preceq_{\mathrm{cx}} \overline{B}_s, \qquad 0 \le u \le s < T.$$

Thus, the reverse drift $(\overline{B}_s)_{0 \le s < T}$ is a peacock whenever it is integrable.

Proof. The Gaussian heat semigroup and the operator $C = -\partial_x$ satisfy the commutation condition of Section 4. The Bartlett–Tweedie identity in Proposition 5.1 gives the required representation of the normalized score. The martingale and convex-order assertions therefore follow from Proposition 4.2 and Corollary 4.3. The peacock property of the reverse drift follows from Proposition 4.4 with $a_t = \sigma^2(t)$. $\square$

Remark 5.3 (Constant-volatility case)**.** If $\sigma(t) \equiv \sigma$, then $A_t = \sigma^2 t$. Hence

$$S_t = \frac{M_t - X_t}{\sigma^2 t}, \qquad B_t = \sigma^2 S_t = \frac{M_t - X_t}{t}.$$

Thus, the normalized score and the actual reverse drift differ only by the constant factor σ^2. In this homogeneous case, the reverse drift is itself a reverse-time martingale. In the time-inhomogeneous case, only the normalized score is automatically a martingale; the drift is a deterministic time-dependent dilation of it.

Remark 5.4 (Kellerer martingale associated with the Gaussian reverse drift)**.** When $(\overline{B}_s)_{0\leq s<T}$ is a peacock (see [6]), under the usual regularity assumptions required for the continuous-time version of Kellerer's theorem, this peacock admits a martingale realization, possibly defined on another filtered probability space, such that

$$\mathcal{M}_s \sim \overline{B}_s, \qquad 0 \leq s < T.$$

Equivalently, for every finite collection of times, the corresponding marginal laws are compatible with a martingale coupling. This martingale is a mimicking martingale: in general it has the same one-dimensional marginal distributions as the Gaussian reverse drift, but not the same finite-dimensional distributions. Indeed,

$$\mathrm{E}[\overline{B}_s \mid \overline{\mathcal{F}}_u] = \sigma^2(T-s)\overline{S}_u, \qquad \overline{B}_u = \sigma^2(T-u)\overline{S}_u,$$

so that $(\overline{B}_s)$ is itself a martingale only in the constant-volatility case, or in degenerate cases.

5.3 Linearity of the Gaussian reverse score and drift

We finally identify the Gaussian case in which the normalized score and the reverse drift are linear functions of the state. This is the Gaussian instance of the semigroup-compatibility principle: the signal and the noise must belong to the same Gaussian convolution semigroup, up to translation.

Proposition 5.5 (Linearity of the Gaussian reverse score and drift)**.** *Assume that $A_t > 0$ for every $t > 0$, and let*

$$s_t(x) := \partial_x \log p_t(x), \qquad t > 0.$$

The following assertions are equivalent:

(1) The initial distribution is Gaussian, possibly degenerate:

$$Y \sim \mathrm{Normal}(\mu, \tau^2), \qquad \tau \geq 0.$$

(2) For every $t > 0$, the Gaussian score s_t is affine in x.

(3) There exists $t_0 > 0$ such that the Gaussian score s_{t_0} is affine in x.

In that case, for every $t > 0$,

$$p_t = \mathrm{Normal}(\mu, \tau^2 + A_t) \ \textit{and} \ S_t = s_t(X_t) = -\frac{X_t - \mu}{\tau^2 + A_t}.$$

The actual reverse drift, indexed in forward time, is

$$B_t = \sigma^2(t) S_t = -\frac{\sigma^2(t)}{\tau^2 + A_t}(X_t - \mu).$$

Equivalently, in reverse time,

$$\overline{S}_s = -\frac{\overline{X}_s - \mu}{\tau^2 + A_{T-s}}, \qquad \overline{B}_s = -\frac{\sigma^2(T-s)}{\tau^2 + A_{T-s}}\left(\overline{X}_s - \mu\right).$$

Proof. If $Y \sim \mathrm{Normal}(\mu, \tau^2)$, then $X_t = Y + Z_t \sim \mathrm{Normal}(\mu, \tau^2 + A_t)$. Therefore,

$$\partial_x \log p_t(x) = -\frac{x-\mu}{\tau^2 + A_t},$$

which gives the displayed formulas for S_t, B_t, $\overline{S}_s$ and $\overline{B}_s$.

Conversely, assume that s_{t_0} is affine for some $t_0 > 0$. Then

$$\partial_x \log p_{t_0}(x) = ax + b$$

for some constants $a, b \in \mathbb{R}$. Hence,

$$p_{t_0}(x) = C \exp\left(\frac{a}{2}x^2 + bx\right).$$

Since p_{t_0} is a probability density on $\mathbb{R}$, necessarily $a < 0$, and X_{t_0} is Gaussian. But

$$X_{t_0} = Y + G_{A_{t_0}},$$

where $G_{A_{t_0}} \sim \mathrm{Normal}(0, A_{t_0})$ is independent of Y and $A_{t_0} > 0$. By Cramer's theorem, Y is Gaussian, possibly degenerate. This proves the equivalence. □

6 Compensated compound Poisson forward dynamics with deterministic intensity and reverse-time representation

This section specializes the clock-based principle of Section 4 to finite-activity jump noise. The forward perturbation is a compensated compound Poisson process with deterministic time-dependent intensity. The accumulated clock is the cumulative intensity Λ_t in Table 4.1 whereas the actual reverse local mean drift involves the instantaneous intensity $\lambda(t)$. Thus, as in the Gaussian case, the normalized score is governed by the accumulated clock, while the genuine reverse local mean drift is obtained by multiplying this score by the instantaneous clock rate.

6.1 Compound Poisson clock and time-reversed jump dynamics

Consider the setting of Section 3.5.2. Let $\lambda : [0, T] \to [0, \infty)$ be measurable and such that

$$0 < \Lambda_t < \infty, \qquad t > 0.$$

The forward noise is

$$Z_t = \sum_{k=1}^{N_t} J_k - m\Lambda_t, \qquad m := \mathrm{E}[J_1],$$

where (N_t) is a non-homogeneous Poisson process with integrated intensity Λ_t, and where the jumps $J_1, J_2, \ldots$ are independent with common distribution F_J. Then, Z has independent centered increments and, for $0 \le s < t \le T$,

$$Z_t - Z_s \overset{d}{=} \sum_{k=1}^{N_{\Lambda_t - \Lambda_s}} J_k - m(\Lambda_t - \Lambda_s).$$

Therefore, the model is an additive noise model driven by the centered compound Poisson convolution semigroup $(\mu_a)_{a \ge 0}$, where

$$\widehat{\mu}_a(u) = \exp\left\{a\left(\varphi_J(u) - 1 - ium\right)\right\}, \qquad u \in \mathbb{R}.$$

In the notation of Section 4, the deterministic clock is

$$\kappa_t = \Lambda_t, \qquad a_t = \kappa_t' = \lambda(t).$$

Let p_t denote the law of $X_t = Y + Z_t$. Throughout this section, when applying the clock-based principle of Section 4, we assume that p_0 admits a strictly positive density and that the regularity and integrability conditions of Assumption 4.1 are satisfied. Then, p_t admits a strictly positive density for every $t > 0$, and

$$p_t = p_0 * \mu_{\Lambda_t}.$$

The forward generator acts on smooth test functions f as

$$\begin{aligned}(\mathcal{L}_t f)(x) &= -\lambda(t)mf'(x) + \lambda(t)\int_{\mathbb{R}} \big(f(x+u) - f(x)\big)F_J(du) \\ &= \lambda(t)\int_{\mathbb{R}} \big(f(x+u) - f(x) - uf'(x)\big)F_J(du).\end{aligned}$$

Its adjoint acting on densities is

$$\begin{aligned}(\mathcal{L}_t^* q)(x) &= \lambda(t)m\partial_x q(x) + \lambda(t)\int_{\mathbb{R}} \big(q(x-u) - q(x)\big)F_J(du) \\ &= \lambda(t)\int_{\mathbb{R}} \big(q(x-u) - q(x) + u\partial_x q(x)\big)F_J(du).\end{aligned}$$

Thus, the marginal densities solve

$$\partial_t p_t(x) = \mathcal{L}_t^* p_t(x), \qquad p_{t=0} = p_0.$$

Let

$$\overline{X}_s := X_{T-s}, \qquad 0 \le s \le T,$$

be the reverse-time process. For $0 \le u \le s < T$, its transition kernel is

$$\overline{K}_{u,s}(x, dy) = \frac{p_{T-s}(y)}{p_{T-u}(x)}\,\mu_{\Lambda_{T-u}-\Lambda_{T-s}}(x - dy).$$

Equivalently, for every bounded measurable function f,

$$\mathrm{E}[f(\overline{X}_s) \mid \overline{X}_u = x] = \frac{1}{p_{T-u}(x)}\int_{\mathbb{R}} f(y)p_{T-s}(y)\mu_{\Lambda_{T-u}-\Lambda_{T-s}}(x - dy).$$

The reverse process is a time-inhomogeneous jump Markov process. Its backward generator is given by the reverse h-transform

$$(\overline{\mathcal{L}}_s f)(x) = \frac{1}{p_{T-s}(x)}\left(\mathcal{L}_{T-s}^*(fp_{T-s})(x) - f(x)\mathcal{L}_{T-s}^* p_{T-s}(x)\right).$$

Using the explicit expression of $\mathcal{L}_{T-s}^*$, this becomes

$$(\overline{\mathcal{L}}_s f)(x) = \lambda(T-s)mf'(x) + \lambda(T-s)\int_{\mathbb{R}} \big(f(x-u) - f(x)\big)\frac{p_{T-s}(x-u)}{p_{T-s}(x)}F_J(du).$$

Hence, in reverse time, jumps of size $-u$ occur with state-dependent kernel

$$\overline{\nu}_s(x, du) = \lambda(T-s)\frac{p_{T-s}(x-u)}{p_{T-s}(x)}F_J(du),$$

and there is a deterministic drift component $\lambda(T-s)m$ between jumps. The ratio $p_{T-s}(x-u)/p_{T-s}(x)$ is the jump analogue of the score factor in Gaussian time reversal.

6.2 Tweedie formula, normalized jump score and local mean drift

Define the signed measure

$$\chi_J(du) := uF_J(du) - m\delta_0(du).$$

It encodes the centered first-order jump operator associated with the compensated compound Poisson noise. In the notation of Section 4, the score operator is $Cq = q * \chi_J$. Since both C and the operators $P_a q = q * \mu_a$ are convolution operators, they commute:

$$C(P_a q) = P_a(Cq), \qquad a \geq 0.$$

Proposition 6.1 (Compound Poisson specialization of the clock-based score)**.** *In the setting of Section 3.5.2, for every $t > 0$ and every x such that $p_t(x) > 0$,*

$$m_t(x) := \mathrm{E}[Y \mid X_t = x] = x - \Lambda_t \frac{(p_t * \chi_J)(x)}{p_t(x)}.$$

Consequently, with

$$M_t := \mathrm{E}[Y \mid X_t], \qquad S_t := \frac{M_t - X_t}{\Lambda_t},$$

one has

$$S_t = -\frac{(p_t * \chi_J)(X_t)}{p_t(X_t)}.$$

Moreover, if

$$\beta_s(x) := (\overline{L}_s \mathrm{Id})(x), \qquad \mathrm{Id}(x) = x,$$

denotes the reverse local mean drift, then

$$\beta_s(x) = -\lambda(T-s)\frac{(p_{T-s} * \chi_J)(x)}{p_{T-s}(x)} = \frac{\lambda(T-s)}{\Lambda_{T-s}} \left(m_{T-s}(x) - x\right).$$

Equivalently,

$$\beta_s(\overline{X}_s) = \lambda(T-s)\overline{S}_s, \qquad \overline{S}_s := S_{T-s}.$$

Proof. For fixed $t > 0$, Z_t is a centered compound Poisson random variable with cumulative intensity parameter Λ_t, and characteristic function

$$\mathrm{E}[e^{iuZ_t}] = \exp\left\{\Lambda_t \left(\varphi_J(u) - 1 - ium\right)\right\}.$$

The compound Poisson Bartlett–Tweedie identity of Proposition 2.5 gives

$$\mathrm{E}[Y \mid X_t = x] = x - \Lambda_t \frac{(p_t * \chi_J)(x)}{p_t(x)}.$$

This proves the formula for m_t and hence the representation of S_t.

It remains to identify the reverse local mean drift. From the reverse generator,

$$\begin{aligned}(\overline{\mathcal{L}}_s \mathrm{Id})(x) &= \lambda(T-s)m + \lambda(T-s)\int_{\mathbb{R}} ((x-u) - x)\frac{p_{T-s}(x-u)}{p_{T-s}(x)} F_J(du) \\ &= \lambda(T-s)m - \lambda(T-s)\int_{\mathbb{R}} u\frac{p_{T-s}(x-u)}{p_{T-s}(x)} F_J(du).\end{aligned}$$

Since

$$(p_{T-s} * \chi_J)(x) = \int_{\mathbb{R}} u\, p_{T-s}(x-u) F_J(du) - m p_{T-s}(x),$$

we obtain

$$\beta_s(x) = -\lambda(T-s)\frac{(p_{T-s} * \chi_J)(x)}{p_{T-s}(x)}.$$

Combining this identity with the Tweedie formula at time $T-s$ yields

$$\beta_s(x) = \frac{\lambda(T-s)}{\Lambda_{T-s}} \left(m_{T-s}(x) - x\right).$$

Evaluating at $x = \overline{X}_s$ gives $\beta_s(\overline{X}_s) = \lambda(T-s)\overline{S}_s$ and ends the proof. $\square$

Thus the compensated compound Poisson model fits Section 4 with

$$\kappa_t = \Lambda_t, \qquad a_t = \lambda(t), \qquad Cq = q * \chi_J.$$

The martingale and convex-order properties are therefore direct consequences of the clock-based principle.

Corollary 6.2 (Reverse martingales and convex-order monotonicity). *Let*

$$\overline{M}_s := M_{T-s}, \qquad \overline{S}_s := S_{T-s}, \qquad \overline{R}_s := \beta_s(\overline{X}_s) = \lambda(T-s)\overline{S}_s.$$

Then, $(\overline{M}_s)_{0\le s<T}$ *and* $(\overline{S}_s)_{0\le s<T}$ *are martingales with respect to the reverse filtration*

$$\overline{\mathcal{F}}_s := \sigma(\overline{X}_u : 0 \le u \le s).$$

Consequently,

$$\overline{M}_u \preceq_{\mathrm{cx}} \overline{M}_s, \qquad \overline{S}_u \preceq_{\mathrm{cx}} \overline{S}_s, \qquad 0 \le u \le s < T.$$

If, in addition, the map $s \mapsto \lambda(T-s)$ *is non-decreasing on* $[0,T]$*, then*

$$\overline{R}_u \preceq_{\mathrm{cx}} \overline{R}_s, \qquad 0 \le u \le s < T.$$

Thus the reverse local mean drift $(\overline{R}_s)_{0\le s<T}$ *is a peacock whenever it is integrable.*

Proof. The convolution operator $Cq = q * \chi_J$ commutes with the compound Poisson convolution semigroup, and Proposition 6.1 gives the required clock-based Bartlett–Tweedie representation. The martingale and convex-order assertions follow from Proposition 4.2 and Corollary 4.3. The peacock property of the reverse local mean drift follows from Proposition 4.4 with $a_t = \lambda(t)$. $\square$

Remark 6.3 (Constant-intensity case). If $\lambda(t) \equiv \lambda$, then $\Lambda_t = \lambda t$. Hence

$$S_t = \frac{M_t - X_t}{\lambda t}, \qquad R_t = \lambda S_t = \frac{M_t - X_t}{t}.$$

Thus, the normalized compound Poisson score and the actual reverse local mean drift differ only by the constant factor λ. In this homogeneous case, the reverse local mean drift is itself a reverse-time martingale. In the time-inhomogeneous case, only the normalized score is automatically a martingale.

Remark 6.4 (Kellerer martingale associated with the reverse local mean drift). When $(\overline{R}_s)_{0\le s<T}$ is a peacock, Kellerer's theorem yields a martingale $(\mathcal{M}_s)_{0\le s<T}$, possibly defined on another filtered probability space, such that

$$\mathcal{M}_s \sim \overline{R}_s, \qquad 0 \le s < T.$$

This martingale is a mimicking martingale: in general it has the same one-dimensional marginal distributions as the reverse local mean drift, but not the same finite-dimensional distributions. Indeed,

$$\mathrm{E}[\overline{R}_s \mid \overline{\mathcal{F}}_u] = \lambda(T-s)\overline{S}_u, \qquad \overline{R}_u = \lambda(T-u)\overline{S}_u,$$

so that $(\overline{R}_s)$ is itself a martingale only in the constant-intensity case, or in degenerate cases.

6.3 Linearity of the compound Poisson reverse score and drift

We finally identify the semigroup-compatible case in which the conditional mean, the normalized score and the reverse local mean drift are affine functions of the state. This is the compound Poisson analogue of the Gaussian linear-score case: the initial law and the noise law must belong to the same centered compound Poisson convolution semigroup, up to translation.

Proposition 6.5 (Affine linearity under compound Poisson semigroup compatibility)**.** *In the setting of Section 3.5.2, assume that*

$$Y = y_0 + Y^{(0)}, \qquad y_0 \in \mathbb{R},$$

where $Y^{(0)}$ is a centered compound Poisson random variable with the same jump distribution F_J, namely

$$Y^{(0)} = \sum_{k=1}^{M} J'_k - \theta m,$$

where $M \sim \mathrm{Poisson}(\theta)$, $\theta > 0$, the random variables $J'_1, J'_2, \dots$ are independent with common distribution F_J, and M, (J'_k), (J_k), and (N_t) are mutually independent. Then, for every $t \in (0, T]$,

$$M_t := \mathrm{E}[Y \mid X_t] = y_0 + \frac{\theta}{\theta + \Lambda_t}(X_t - y_0).$$

Consequently,

$$S_t = \frac{M_t - X_t}{\Lambda_t} = -\frac{X_t - y_0}{\theta + \Lambda_t},$$

and the actual reverse local mean drift, indexed in forward time, is

$$R_t = \lambda(t) S_t = -\frac{\lambda(t)}{\theta + \Lambda_t}(X_t - y_0).$$

Equivalently, in reverse time,

$$\overline{M}_s = y_0 + \frac{\theta}{\theta + \Lambda_{T-s}}(\overline{X}_s - y_0),$$

$$\overline{S}_s = -\frac{\overline{X}_s - y_0}{\theta + \Lambda_{T-s}}, \qquad \overline{R}_s = -\frac{\lambda(T-s)}{\theta + \Lambda_{T-s}}(\overline{X}_s - y_0).$$

Thus, the normalized compound Poisson score and the reverse local mean drift are affine functions of the reverse state. If $y_0 = 0$, they are linear functions of $\overline{X}_s$.

Proof. The characteristic functions of $Y^{(0)}$ and Z_t are respectively

$$\varphi_{Y^{(0)}}(u) = \exp\left\{\theta\left(\varphi_J(u) - 1 - ium\right)\right\},$$

and

$$\varphi_{Z_t}(u) = \exp\left\{\Lambda_t\left(\varphi_J(u) - 1 - ium\right)\right\}.$$

Thus, $Y^{(0)}$ and Z_t belong to the same centered compound Poisson convolution semigroup, with parameters θ and Λ_t. Proposition 2.8 gives

$$\mathrm{E}[Y^{(0)} \mid Y^{(0)} + Z_t] = \frac{\theta}{\theta + \Lambda_t}(Y^{(0)} + Z_t).$$

Since $X_t - y_0 = Y^{(0)} + Z_t$,

$$\mathrm{E}[Y \mid X_t] = y_0 + \frac{\theta}{\theta + \Lambda_t}(X_t - y_0).$$

The formulas for S_t and R_t follow by subtracting X_t, dividing by Λ_t, and then multiplying by $\lambda(t)$. Replacing t by $T - s$ gives the reverse-time formulas. □

Remark 6.6 (Characterization within the infinitely divisible class)**.** Proposition 6.5 gives the semigroup-compatible case. Conversely, within the infinitely divisible class, affine linearity of the conditional mean forces this structure. Indeed, suppose that $Y - y_0$ is infinitely divisible and that $\mathrm{E}[Y \mid X_{t_0}]$ is an affine function of X_{t_0} for some $t_0 > 0$. After centering, one obtains

$$\mathrm{E}[Y - y_0 \mid X_{t_0} - y_0] = a(X_{t_0} - y_0)$$

for some $a \in [0, 1]$. Proposition 2.8 then implies that $Y - y_0$ and Z_{t_0} have proportional Lévy exponents. Since Z_{t_0} has exponent $\Lambda_{t_0}(\varphi_J(u) - 1 - ium)$, the centered variable $Y - y_0$ must have exponent $\theta(\varphi_J(u) - 1 - ium)$ for some $\theta \geq 0$. Thus, affine linearity is not generic; it reflects compatibility of the initial distribution with the same compound Poisson convolution semigroup as the noise.

7 Symmetric α-stable forward dynamics with deterministic scale and reverse-time representation

This section specializes the clock-based principle of Section 4 to symmetric α-stable noise, with $\alpha \in (1, 2)$. The restriction $\alpha > 1$ is essential: it ensures that the stable noise is integrable and centered, so that conditional means and convex-order comparisons are meaningful.

The accumulated stable clock is Γ_t in Table 4.1 whereas the instantaneous clock rate is $a_t := |\sigma_L(t)|^\alpha$. Thus the normalized fractional score is governed by the accumulated clock Γ_t, while the actual reverse local mean drift is obtained by multiplying this score by the instantaneous rate a_t.

7.1 Stable clock and reverse Lévy-type dynamics

Consider the setting of Section 3.5.3. Let $L = (L_t)_{t\geq 0}$ be a symmetric α-stable Lévy process normalized by

$$\mathrm{E}[e^{iuL_t}] = \exp(-t|u|^\alpha), \qquad u \in \mathbb{R}.$$

Let $\sigma_L : [0, T] \to \mathbb{R}$ be deterministic and measurable, and assume that

$$0 < \Gamma_t < \infty, \qquad t > 0.$$

The forward model is

$$X_t = Y + Z_t, \qquad Z_t = \int_0^t \sigma_L(r)\, dL_r, \qquad 0 \leq t \leq T,$$

where Y is independent of L. Since σ_L is deterministic, Z_t is symmetric α-stable with characteristic function

$$\mathrm{E}[e^{iuZ_t}] = \exp(-\Gamma_t|u|^\alpha).$$

Because $\alpha > 1$, Z_t is integrable and centered.

Let q_γ be the density of the symmetric α-stable law with characteristic function

$$\widehat{q}_\gamma(u) = \exp(-\gamma|u|^\alpha), \qquad \gamma \geq 0.$$

If p_t denotes the law of X_t, then

$$p_t = p_0 * q_{\Gamma_t}.$$

For $t > 0$, q_{Γ_t} is smooth and strictly positive on $\mathbb{R}$. Thus p_t is smooth and strictly positive whenever p_0 is a probability density.

The forward generator is time-inhomogeneous and acts on sufficiently smooth functions as

$$(\mathcal{L}_t f)(x) = -a_t(-\Delta)^{\alpha/2} f(x).$$

Equivalently, if

$$\nu_\alpha(dz) = \kappa_\alpha |z|^{-1-\alpha}\,dz$$

is the Lévy measure of the normalized symmetric α-stable process, then

$$(\mathcal{L}_t f)(x) = a_t \int_{\mathbb{R}} \big(f(x+z) - f(x) - zf'(x)\big)\,\nu_\alpha(dz).$$

The centered representation is well defined for $\alpha \in (1,2)$, since

$$\int_{\mathbb{R}} (|z|^2 \wedge |z|)\,\nu_\alpha(dz) < \infty.$$

Since the stable process is symmetric, the same operator acts on densities, and the marginal densities solve the fractional Fokker–Planck equation

$$\partial_t p_t(x) = -a_t(-\Delta)^{\alpha/2} p_t(x), \qquad p_{t=0} = p_0.$$

Let

$$\overline{X}_s := X_{T-s}, \qquad 0 \le s \le T,$$

be the reverse-time process. For $0 \le u \le s < T$, its transition kernel is

$$\overline{K}_{u,s}(x,dy) = q_{\Gamma_{T-u}-\Gamma_{T-s}}(x-y)\,\frac{p_{T-s}(y)}{p_{T-u}(x)}\,dy.$$

Equivalently, for every bounded measurable function f,

$$\mathrm{E}[f(\overline{X}_s) \mid \overline{X}_u = x] = \frac{1}{p_{T-u}(x)} \int_{\mathbb{R}} f(y) q_{\Gamma_{T-u}-\Gamma_{T-s}}(x-y) p_{T-s}(y)\,dy.$$

The reverse process is a time-inhomogeneous pure-jump Markov process with infinite activity. Its backward generator is the Doob h-transform

$$(\overline{\mathcal{L}}_s f)(x) = \frac{1}{p_{T-s}(x)}\,\big(\mathcal{L}_{T-s}(f p_{T-s})(x) - f(x)\mathcal{L}_{T-s} p_{T-s}(x)\big)\,.$$

Unlike the Gaussian reverse dynamics, this generator is non-local. The relevant first-order quantity is therefore the reverse local mean drift

$$\beta_s(x) := (\overline{\mathcal{L}}_s \mathrm{Id})(x), \qquad \mathrm{Id}(x) = x.$$

Proposition 7.1 (Reverse Lévy-type generator in centered form)**.** *Assume that p_{T-s} is strictly positive and C^1, and that for every $x \in \mathbb{R}$,*

$$\int_{\mathbb{R}} \big(|z|^2 \wedge |z|\big)\,\frac{p_{T-s}(x+z)}{p_{T-s}(x)}\nu_\alpha(dz) < \infty$$

and

$$\int_{\mathbb{R}} |z| \left| \frac{p_{T-s}(x+z)}{p_{T-s}(x)} - 1 \right| \nu_\alpha(dz) < \infty.$$

Then, the reverse generator admits the representation

$$(\overline{\mathcal{L}}_s f)(x) = \beta_s(x) f'(x) + \int_{\mathbb{R}} \big(f(x+z) - f(x) - zf'(x)\big)\,\overline{\nu}_s(x,dz),$$

where

$$\overline{\nu}_s(x,dz) = a_{T-s}\frac{p_{T-s}(x+z)}{p_{T-s}(x)}\nu_\alpha(dz),$$

and

$$\beta_s(x) = a_{T-s} \int_{\mathbb{R}} z \left(\frac{p_{T-s}(x+z)}{p_{T-s}(x)} - 1 \right) \nu_\alpha(dz).$$

Consequently, on a possibly enlarged filtered probability space, the reverse process admits the Lévy-type semimartingale representation

$$\overline{X}_s = \overline{X}_0 + \int_0^s \beta_r(\overline{X}_{r-})\, dr + \int_0^s \int_{\mathbb{R}} z\, \widetilde{\overline{N}}(dr, dz), \qquad 0 \le s < T,$$

where $\overline{N}(dr, dz)$ is the jump measure of $\overline{X}$, and

$$\widetilde{\overline{N}}(dr, dz) := \overline{N}(dr, dz) - \overline{\nu}_r(\overline{X}_{r-}, dz)\, dr.$$

The second integrability condition is the one ensuring that the compensated first moment of the tilted jump kernel is finite. Near the origin it is consistent with the C^1-regularity of p_{T-s}, since

$$\frac{p_{T-s}(x+z)}{p_{T-s}(x)} - 1 = O(z), \qquad z \to 0,$$

and $\int_{|z|\le 1} z^2 \nu_\alpha(dz) < \infty$. The condition at infinity is an integrability assumption on the tilted stable kernel.

Proof. Set $h_s(x) := p_{T-s}(x)$. By the reverse h-transform formula,

$$(\overline{\mathcal{L}}_s f)(x) = \frac{1}{h_s(x)} \left(\mathcal{L}_{T-s}(f h_s)(x) - f(x) \mathcal{L}_{T-s} h_s(x) \right).$$

Using the centered Lévy representation of $\mathcal{L}_{T-s}$, we obtain

$$(\overline{\mathcal{L}}_s f)(x) = a_{T-s} \int_{\mathbb{R}} \left((f(x+z) - f(x)) \frac{h_s(x+z)}{h_s(x)} - z f'(x) \right) \nu_\alpha(dz).$$

Adding and subtracting $z f'(x) \frac{h_s(x+z)}{h_s(x)}$ inside the integral gives

$$\begin{aligned}(\overline{\mathcal{L}}_s f)(x) &= a_{T-s} \int_{\mathbb{R}} \left(f(x+z) - f(x) - z f'(x) \right) \frac{h_s(x+z)}{h_s(x)} \nu_\alpha(dz) \\ &\quad + a_{T-s} f'(x) \int_{\mathbb{R}} z \left(\frac{h_s(x+z)}{h_s(x)} - 1 \right) \nu_\alpha(dz).\end{aligned}$$

Since $h_s = p_{T-s}$, this proves the announced expressions for $\overline{\nu}_s(x, dz)$ and $\beta_s(x)$. The semimartingale representation is the standard stochastic realization of the corresponding time-inhomogeneous Lévy-type generator. □

Remark 7.2 (Non-local nature of the reverse stable dynamics)**.** The reverse stable dynamics are driven by the state-dependent jump kernel

$$\overline{\nu}_s(x, dz) = a_{T-s} \frac{p_{T-s}(x+z)}{p_{T-s}(x)} \kappa_\alpha |z|^{-1-\alpha}\, dz.$$

The ratio $p_{T-s}(x+z)/p_{T-s}(x)$ tilts the forward stable jump measure according to the current marginal density. This is the stable analogue of the score factor in Gaussian time reversal, but it acts non-locally through the jump kernel.

7.2 Tweedie formula, normalized fractional score and local mean drift

Recall that the fractional score operator $\mathcal{A}_{\alpha-1}$ is defined by the Fourier multiplier

$$\mathcal{F}[\mathcal{A}_{\alpha-1}f](u) = i\,\mathrm{sign}(u)|u|^{\alpha-1}\widehat{f}(u), \qquad u \in \mathbb{R}.$$

The stable score is therefore non-local. In the notation of Section 4, the operator C is $Cq = \alpha\mathcal{A}_{\alpha-1}q$. Since both C and the stable convolution semigroup

$$P_\gamma q := q * q_\gamma, \qquad \gamma \geq 0,$$

are Fourier multiplier operators, they commute:

$$C(P_\gamma q) = P_\gamma(Cq), \qquad \gamma \geq 0.$$

Proposition 7.3 (Stable specialization of the clock-based score)**.** *In the setting of Section 3.5.3, for every $t > 0$ and every x such that $p_t(x) > 0$,*

$$m_t(x) := \mathrm{E}[Y \mid X_t = x] = x - \alpha\Gamma_t\frac{(\mathcal{A}_{\alpha-1}p_t)(x)}{p_t(x)}.$$

Consequently, with

$$M_t := \mathrm{E}[Y \mid X_t], \qquad S_t := \frac{M_t - X_t}{\Gamma_t},$$

one has

$$S_t = -\alpha\frac{(\mathcal{A}_{\alpha-1}p_t)(X_t)}{p_t(X_t)}.$$

Moreover, the reverse local mean drift satisfies

$$\beta_s(x) = -\alpha a_{T-s}\frac{(\mathcal{A}_{\alpha-1}p_{T-s})(x)}{p_{T-s}(x)} = \frac{a_{T-s}}{\Gamma_{T-s}}\left(m_{T-s}(x) - x\right), \qquad 0 \leq s < T.$$

Equivalently,

$$\beta_s(\overline{X}_s) = a_{T-s}\overline{S}_s, \qquad \overline{S}_s := S_{T-s}.$$

Proof. For fixed $t > 0$, the noise Z_t is symmetric α-stable with characteristic function

$$\mathrm{E}[e^{iuZ_t}] = \exp(-\Gamma_t|u|^\alpha).$$

The symmetric stable Bartlett–Tweedie identity of Proposition 2.3 gives

$$\mathrm{E}[Y \mid X_t = x] = x - \alpha\Gamma_t\frac{(\mathcal{A}_{\alpha-1}p_t)(x)}{p_t(x)}.$$

This proves the formula for m_t and hence the expression for S_t.

It remains to identify the local mean drift with the fractional score. Let $p = p_{T-s}$. Since

$$\mathcal{L}_{T-s} = -a_{T-s}(-\Delta)^{\alpha/2},$$

we have

$$\beta_s(x) = \frac{\mathcal{L}_{T-s}(xp)(x) - x\mathcal{L}_{T-s}p(x)}{p(x)}.$$

Using the Fourier convention

$$\widehat{xf}(u) = \frac{1}{i}\widehat{f}'(u),$$

we get

$$\mathcal{F}[\mathcal{L}_{T-s}(xp)](u) = -a_{T-s}|u|^\alpha\frac{1}{i}\widehat{p}'(u),$$

whereas

$$\mathcal{F}[x\mathcal{L}_{T-s}p](u) = \frac{1}{i}\partial_u\left(-a_{T-s}|u|^\alpha \widehat{p}(u)\right).$$

Subtracting yields

$$\mathcal{F}[\mathcal{L}_{T-s}(xp) - x\mathcal{L}_{T-s}p](u) = -\alpha a_{T-s} i\,\mathrm{sign}(u)|u|^{\alpha-1}\widehat{p}(u).$$

Thus,

$$\mathcal{L}_{T-s}(xp) - x\mathcal{L}_{T-s}p = -\alpha a_{T-s}\mathcal{A}_{\alpha-1}p,$$

and therefore

$$\beta_s(x) = -\alpha a_{T-s}\frac{(\mathcal{A}_{\alpha-1}p_{T-s})(x)}{p_{T-s}(x)}.$$

Combining this identity with the Tweedie formula at time $T-s$ gives

$$\beta_s(x) = \frac{a_{T-s}}{\Gamma_{T-s}}\left(m_{T-s}(x) - x\right).$$

Evaluating at $x = \overline{X}_s$ gives $\beta_s(\overline{X}_s) = a_{T-s}\overline{S}_s$ and ends the proof. $\square$

Thus the symmetric α-stable model fits Section 4 with

$$\kappa_t = \Gamma_t, \qquad a_t = \kappa_t' = |\sigma_L(t)|^\alpha, \qquad Cq = \alpha\mathcal{A}_{\alpha-1}q.$$

The martingale and convex-order properties follow directly from the clock-based principle.

Corollary 7.4 (Reverse martingales and convex-order monotonicity)**.** *Let*

$$\overline{M}_s := M_{T-s}, \qquad \overline{S}_s := S_{T-s}, \qquad \overline{R}_s := \beta_s(\overline{X}_s) = a_{T-s}\overline{S}_s.$$

Then, $(\overline{M}_s)_{0\le s<T}$ and $(\overline{S}_s)_{0\le s<T}$ are martingales with respect to the reverse filtration

$$\overline{\mathcal{F}}_s := \sigma(\overline{X}_u : 0 \le u \le s).$$

Consequently,

$$\overline{M}_u \preceq_{\mathrm{cx}} \overline{M}_s, \qquad \overline{S}_u \preceq_{\mathrm{cx}} \overline{S}_s, \qquad 0 \le u \le s < T.$$

If, in addition, the map $s \mapsto a_{T-s} = |\sigma_L(T-s)|^\alpha$ is non-decreasing on $[0,T]$, then

$$\overline{R}_u \preceq_{\mathrm{cx}} \overline{R}_s, \qquad 0 \le u \le s < T.$$

Thus the reverse local mean drift $(\overline{R}_s)_{0\le s<T}$ is a peacock whenever it is integrable.

Proof. The stable convolution semigroup and the operator $Cq = \alpha\mathcal{A}_{\alpha-1}q$ satisfy the commutation condition of Section 4. Proposition 7.3 gives the required clock-based Bartlett–Tweedie representation. The martingale and convex-order assertions follow from Proposition 4.2 and Corollary 4.3. The peacock property of the reverse local mean drift follows from Proposition 4.4 with $a_t = |\sigma_L(t)|^\alpha$. $\square$

Remark 7.5 (Constant-scale case)**.** If $a_t = |\sigma_L(t)|^\alpha \equiv a$, then $\Gamma_t = at$. Hence,

$$S_t = \frac{M_t - X_t}{at}, \qquad R_t = aS_t = \frac{M_t - X_t}{t}.$$

Thus, the normalized fractional score and the actual reverse local mean drift differ only by the constant factor a. In this homogeneous stable case, the reverse local mean drift is itself a reverse-time martingale. In the time-inhomogeneous case, only the normalized fractional score is automatically a martingale.

Remark 7.6 (Kellerer martingale associated with the reverse stable drift)**.** When $(\overline{R}_s)_{0\le s<T}$ is a peacock, Kellerer's theorem yields a martingale $(\mathcal{M}_s)_{0\le s<T}$, possibly defined on another filtered probability space, such that

$$\mathcal{M}_s \sim \overline{R}_s, \qquad 0 \le s < T.$$

This is a mimicking martingale: it has the same one-dimensional marginal distributions as the reverse local mean drift, but not its finite-dimensional distributions in general. Indeed,

$$\mathrm{E}[\overline{R}_s \mid \overline{\mathcal{F}}_u] = a_{T-s}\overline{S}_u, \qquad \overline{R}_u = a_{T-u}\overline{S}_u,$$

so that $(\overline{R}_s)$ is itself a martingale only in the constant-scale case, or in degenerate cases.

7.3 Linearity of the reverse fractional score and drift

We finally identify the semigroup-compatible case in which the conditional mean, the normalized fractional score and the reverse local mean drift are affine functions of the state. This is the stable analogue of the Gaussian and compound Poisson linear cases: the initial law and the noise law must belong to the same symmetric α-stable convolution semigroup, up to translation.

Proposition 7.7 (Linearity under stable semigroup compatibility)**.** *In the setting of Section 3.5.3, assume that*

$$Y = y_0 + Y^{(0)}, \qquad y_0 \in \mathbb{R},$$

where $Y^{(0)}$ is symmetric α-stable with characteristic function

$$\mathrm{E}[e^{iuY^{(0)}}] = \exp(-\theta|u|^\alpha), \qquad u \in \mathbb{R},$$

for some $\theta \ge 0$. Then, for every $t \in (0, T]$,

$$M_t := \mathrm{E}[Y \mid X_t] = y_0 + \frac{\theta}{\theta + \Gamma_t}(X_t - y_0).$$

Consequently,

$$S_t = \frac{M_t - X_t}{\Gamma_t} = -\frac{X_t - y_0}{\theta + \Gamma_t},$$

and the actual reverse local mean drift, indexed in forward time, is

$$R_t = a_t S_t = -\frac{a_t}{\theta + \Gamma_t}(X_t - y_0).$$

Equivalently, in reverse time,

$$\overline{M}_s = y_0 + \frac{\theta}{\theta + \Gamma_{T-s}}(\overline{X}_s - y_0),$$

$$\overline{S}_s = -\frac{\overline{X}_s - y_0}{\theta + \Gamma_{T-s}}, \qquad \overline{R}_s = -\frac{a_{T-s}}{\theta + \Gamma_{T-s}}(\overline{X}_s - y_0).$$

Thus, the normalized fractional score and the reverse local mean drift are affine functions of the reverse state. If $y_0 = 0$, they are linear functions of $\overline{X}_s$.

Proof. The characteristic functions of $Y^{(0)}$ and Z_t are respectively

$$\varphi_{Y^{(0)}}(u) = \exp(-\theta|u|^\alpha), \text{ and } \varphi_{Z_t}(u) = \exp(-\Gamma_t|u|^\alpha).$$

Thus $Y^{(0)}$ and Z_t belong to the same symmetric α-stable convolution semigroup, with parameters θ and Γ_t. Proposition 2.8 gives

$$\mathrm{E}[Y^{(0)} \mid Y^{(0)} + Z_t] = \frac{\theta}{\theta + \Gamma_t}(Y^{(0)} + Z_t).$$

Since $X_t - y_0 = Y^{(0)} + Z_t$, it follows that

$$\mathrm{E}[Y \mid X_t] = y_0 + \frac{\theta}{\theta + \Gamma_t}(X_t - y_0).$$

The formulas for S_t and R_t follow by subtracting X_t, dividing by Γ_t, and then multiplying by a_t. Replacing t by $T - s$ gives the reverse-time formulas. $\square$

Remark 7.8 (Characterization within the infinitely divisible class)**.** Proposition 7.7 gives the semigroup-compatible case. Conversely, within the infinitely divisible class, affine linearity of the conditional mean forces this structure. Indeed, suppose that $Y - y_0$ is infinitely divisible and that $\mathrm{E}[Y \mid X_{t_0}]$ is an affine function of X_{t_0} for some $t_0 > 0$. After centering, Proposition 2.8 implies that $Y - y_0$ and Z_{t_0} have proportional Lévy exponents. Since Z_{t_0} has exponent $-\Gamma_{t_0}|u|^\alpha$, the centered variable $Y - y_0$ must have exponent $-\theta|u|^\alpha$ for some $\theta \geq 0$. Thus, affine linearity of the stable reverse score and drift is not generic; it reflects compatibility with the symmetric α-stable convolution semigroup.

8 Conclusion

This paper develops a clock-based analysis of reverse-time conditional means, scores and drift quantities in additive noise models. The main contribution is to isolate a general martingale mechanism behind several apparently different reverse-time dynamics. When the forward noise is driven by a deterministic convolution clock κ_t, the denoising error $M_t - X_t$, where $M_t = \mathrm{E}[Y \mid X_t]$, must be normalized by the accumulated clock κ_t, not by calendar time nor by the instantaneous noise rate. Under the commutation condition between the score operator and the convolution semigroup, the normalized score

$$S_t = \frac{M_t - X_t}{\kappa_t}$$

is a martingale in reverse time. This gives a unified explanation for the convex-order monotonicity of reverse-time scores.

The Gaussian, compensated compound Poisson and symmetric α-stable models studied in this paper are then special cases of the same principle. In the Gaussian case, the clock is the accumulated variance $A_t = \int_0^t \sigma^2(r)\,dr$, and the normalized score is the usual logarithmic score $\partial_x \log p_t(X_t)$. In the compensated compound Poisson case, the clock is the cumulative intensity $\Lambda_t = \int_0^t \lambda(r)\,dr$, and the score is a non-local jump score involving the signed measure $uF_J(du) - m\delta_0(du)$. In the symmetric α-stable case, with $\alpha \in (1,2)$, the clock is $\Gamma_t = \int_0^t |\sigma_L(r)|^\alpha\,dr$, and the score is fractional, being defined through the multiplier $i\,\mathrm{sign}(u)|u|^{\alpha-1}$. These examples show that the Gaussian score is only the local member of a broader family of clock-normalized denoising scores.

The reverse-time martingale property has direct stochastic-order consequences. The reverse conditional mean and the reverse normalized score are increasing in convex order along reverse time. Equivalently, as the forward noise level increases, these quantities decrease in convex order. This provides a probabilistic ordering of the information retained by the denoising map and by the associated score as noise accumulates.

A second conclusion concerns the distinction between normalized scores and actual reverse drifts. In the time-inhomogeneous setting, the actual reverse drift in the Gaussian case, and the reverse local mean drift in the jump and stable cases, are obtained by multiplying the normalized score by the instantaneous clock rate. This rate is $\sigma^2(t)$, $\lambda(t)$, or $|\sigma_L(t)|^\alpha$ in the three main examples. Consequently, the actual reverse drift is not a martingale in general. It becomes increasing in convex order under a monotonicity condition on the instantaneous rate in reverse time. Under this condition, the reverse drift defines a peacock, and Kellerer's theorem yields a martingale realization with the same one-dimensional marginal distributions. This martingale is a mimicking martingale; it does not generally coincide with the genuine reverse drift process.

The paper also clarifies the role of infinitely divisible convolution semigroups in the linearity of denoising maps. Within the infinitely divisible class, proportionality of $\mathrm{E}[Y \mid Y+\varepsilon]$ to $Y+\varepsilon$ is equivalent to proportionality of the Lévy exponents of the signal and the noise. After allowing for translations, this explains the affine formulas obtained for Gaussian, compound Poisson and symmetric α-stable initial laws belonging to the same convolution semigroup as the corresponding noise. Thus, affine reverse scores and drifts are not generic; they are signatures of semigroup compatibility.

The Cauchy case illustrates the boundary of the theory. The Bartlett–Tweedie identity still has a natural Hilbert-transform form, but the Cauchy noise is not integrable. As a result, the conditional mean remains meaningful when Y is integrable, but the normalized Cauchy score and the associated reverse score drift are not integrable in general. They cannot therefore be treated as ordinary martingales, nor compared in the usual convex order, without truncation, localization or additional assumptions. This separates the Cauchy case from the stable regime $\alpha \in (1,2)$.

Several extensions remain open. A first direction is to formulate the clock-based martingale principle for broader additive processes, beyond the convolution-clock examples treated here, under minimal domain assumptions on the score operator and on the reverse generator. A second direction is to develop multidimensional analogues of the results. The identities derived in Dytso et al. [4] are relevant in that respect, and possible replacements for the univariate convex order include componentwise convex, upper-orthant convex, directionally convex, increasing convex or matrix convex orders. A third direction is to connect these results with reverse-time simulation and score-based generative methods beyond Gaussian diffusions. In non-Gaussian models, the genuine time-reversed dynamics and marginal-preserving reverse equations need not coincide, and understanding this distinction appears to be an important probabilistic problem.

Acknowledgements

Valentin Dendoncker and Michel Denuit gratefully acknowledge funding from the FWO and F.R.S.-FNRS under the Excellence of Science (EOS) programme, project ASTeRISK (40007517).

References

[1] Anderson, B.D.O. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications* **12**, 313–326.

[2] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, second edition. Cambridge University Press.

[3] Bishop, C.M., Bishop, H. (2024). *Deep Learning: Foundations and Concepts*. Springer.

[4] Dytso, A., Cardone, M., Zieder, I. (2023). Meta derivative identity for the conditional expectation. *IEEE Transactions on Information Theory* **69**, 4284-4302.

[5] Haussmann, U.G. and Pardoux, E. (1986). Time reversal of diffusions. *The Annals of Probability* **14**, 1188–1205.

[6] Hirsch, F., Profeta, C., Roynette, B. and Yor, M. (2011). *Peacocks and Associated Martingales, with Explicit Constructions*. Springer Science & Business Media.

[7] Ho, J., Jain, A. and Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851.

[8] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

[9] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer, New York.

[10] Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. and Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. *Proceedings of the 32nd International Conference on Machine Learning*, 2256–2265.

[11] Song, Y. and Ermon, S. (2019). Generative modeling by estimating gradients of the data distribution. *Advances in Neural Information Processing Systems* **32**, 11918–11930.

[12] Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S. and Poole, B. (2021). Score-based generative modeling through stochastic differential equations. *International Conference on Learning Representations*.

[13] Tang, W. and Zhao, H. (2025). Score-based diffusion models via stochastic differential equations. *Statistics Surveys* **19**, 28–64.

[14] Zabell, S.L. (1979). Continuous versions of regular conditional distributions. *Annals of Probability* **7**, 159-165.

[15] Zabell, S.L. (1993). A limit theorem for expectations conditional on a sum. *Journal of Theoretical Probability* **6**, 267–283.

Appendix

A Cauchy forward dynamics with deterministic scale and reverse-time representation

This section treats the borderline stable case $\alpha = 1$ described in Section 3.5.4, namely the case of a Cauchy driving noise with deterministic time-dependent scale. It is useful to present this case separately from the symmetric α-stable case with $\alpha \in (1, 2)$, because the Cauchy process has no first moment. This lack of integrability changes the nature of the reverse-time analysis.

The cumulative Cauchy clock is Γ_t defined in (3.8) whereas the instantaneous scale rate is $a_t := |\sigma_C(t)|$. As in the previous sections, the clock Γ_t is the correct normalization in the Bartlett–Tweedie formula. However, unlike the case $\alpha \in (1, 2)$, the normalized Cauchy score is not integrable in general. Consequently, the reverse Cauchy score drift cannot generally be treated as an integrable martingale or as a peacock in the usual convex order.

A.1 Model, forward generator and reverse Cauchy dynamics

Let $(L_t)_{t\geq 0}$ be a standard symmetric Cauchy Lévy process, normalized by $\mathrm{E}[e^{iuL_t}] = \exp(-t|u|)$, $u \in \mathbb{R}$. Let $\sigma_C : [0, T] \to \mathbb{R}$ be a deterministic measurable function such that $\Gamma_t < \infty$ and $\Gamma_t > 0$ for all $t > 0$. We consider the additive model $dX_t = \sigma_C(t)\, dL_t$, $t \in [0, T]$, with $X_0 = Y$, where Y is independent of (L_t). Thus, $X_t = Y + Z_t$ where Z_t is given in (3.7). Since L_t is Cauchy and σ_C is deterministic, the stochastic integral Z_t is again Cauchy, with characteristic function $\mathrm{E}[e^{iuZ_t}] = \exp(-\Gamma_t|u|)$. Contrary to the case $\alpha \in (1, 2)$, the noise Z_t is not integrable. Nevertheless, if Y is integrable then the conditional expectation $m_t(x) := \mathrm{E}[Y \mid X_t = x]$ is well defined.

Let p_t denote the law of X_t. Let q_γ be the Cauchy density with scale parameter $\gamma > 0$, that is,

$$q_\gamma(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + x^2} \text{ with } \widehat{q}_\gamma(u) = \exp(-\gamma|u|).$$

Then, $p_t = p_0 * q_{\Gamma_t}$. For every $t > 0$, q_{Γ_t} is smooth and strictly positive on $\mathbb{R}$. Hence, p_t is smooth and strictly positive whenever p_0 is a probability density.

The forward generator is time-inhomogeneous and acts on sufficiently regular test functions as

$$(\mathcal{L}_t f)(x) = -a_t(-\Delta)^{1/2} f(x).$$

Equivalently, in Lévy form,

$$(\mathcal{L}_t f)(x) = a_t \,\mathrm{p.v.} \int_{\mathbb{R}} (f(x + z) - f(x))\, \nu_C(dz),$$

where the notation p.v. stands for the Cauchy principal value and $\nu_C(dz) = \kappa_C |z|^{-2}\, dz$ for a suitable normalizing constant $\kappa_C > 0$. Since the Cauchy process is symmetric, the forward operator acting on densities coincides with L_t. Therefore the marginal densities solve

$$\partial_t p_t(x) = -a_t(-\Delta)^{1/2} p_t(x) \text{ subject to the boundary condition } p_{t=0} = p_0.$$

We introduce the reverse-time process $\overline{X}_s := X_{T-s}$, $s \in [0, T]$. For $0 \leq u \leq s < T$, the reverse transition kernel is

$$\overline{K}_{u,s}(x, dy) = q_{\Gamma_{T-u} - \Gamma_{T-s}}(x - y) \frac{p_{T-s}(y)}{p_{T-u}(x)}\, dy.$$

Equivalently, for every bounded measurable function f,

$$\mathrm{E}[f(\overline{X}_s) \mid \overline{X}_u = x] = \frac{1}{p_{T-u}(x)} \int_{\mathbb{R}} f(y)\, q_{\Gamma_{T-u} - \Gamma_{T-s}}(x - y) p_{T-s}(y)\, dy.$$

The reverse process is a time-inhomogeneous Markov jump process. Its backward generator is given by the Doob h-transform

$$(\overline{\mathcal{L}}_s f)(x) = \frac{1}{p_{T-s}(x)}\Big(\mathcal{L}_{T-s}(fp_{T-s})(x) - f(x)\mathcal{L}_{T-s}p_{T-s}(x)\Big).$$

Since Cauchy increments are not integrable, the identity function does not belong to the natural domain of the generator. Hence, the expression $(\overline{\mathcal{L}}_s\mathrm{Id})(x)$ does not have the same infinitesimal conditional-mean interpretation as in the stable case $\alpha \in (1, 2)$. The relevant object is instead a commutator drift associated with the fractional score.

Define the operator $\mathcal{A}_0$ by the Fourier multiplier

$$\mathcal{F}[\mathcal{A}_0 f](u) = i\,\mathrm{sign}(u)\widehat{f}(u), \qquad u \in \mathbb{R}.$$

Up to the sign convention, $\mathcal{A}_0$ is the Hilbert transform.

Proposition A.1 (Reverse Cauchy generator in commutator form)**.** *Let $h_s(x) := p_{T-s}(x)$. For sufficiently regular test functions f, the reverse generator satisfies*

$$(\overline{\mathcal{L}}_s f)(x) = \frac{1}{h_s(x)}\Big(-a_{T-s}(-\Delta)^{1/2}(fh_s)(x) + a_{T-s}f(x)(-\Delta)^{1/2}h_s(x)\Big).$$

Equivalently,

$$(\overline{\mathcal{L}}_s f)(x) = a_{T-s}\,\mathrm{p.v.}\int_{\mathbb{R}} (f(x+z) - f(x))\,\frac{h_s(x+z)}{h_s(x)}\nu_C(dz),$$

whenever the principal value is well defined.

The associated Cauchy commutator drift is

$$\beta_s(x) := -a_{T-s}\frac{(\mathcal{A}_0 h_s)(x)}{h_s(x)} = -a_{T-s}\frac{(\mathcal{A}_0 p_{T-s})(x)}{p_{T-s}(x)}.$$

Proof. The first identity follows directly from the reverse h-transform formula applied to $\mathcal{L}_{T-s} = -a_{T-s}(-\Delta)^{1/2}$. Using the principal-value Lévy representation of the Cauchy generator gives

$$\begin{aligned}(\overline{\mathcal{L}}_s f)(x) &= \frac{a_{T-s}}{h_s(x)}\mathrm{p.v.}\int_{\mathbb{R}} (f(x+z)h_s(x+z) - f(x)h_s(x))\,\nu_C(dz) \\ &\quad - \frac{a_{T-s}f(x)}{h_s(x)}\mathrm{p.v.}\int_{\mathbb{R}} (h_s(x+z) - h_s(x))\,\nu_C(dz) \\ &= a_{T-s}\,\mathrm{p.v.}\int_{\mathbb{R}} (f(x+z) - f(x))\,\frac{h_s(x+z)}{h_s(x)}\nu_C(dz).\end{aligned}$$

The commutator drift is obtained by the Fourier computation

$$\mathcal{L}_{T-s}(xh_s) - x\mathcal{L}_{T-s}h_s = -a_{T-s}\mathcal{A}_0 h_s.$$

Dividing by h_s gives the announced formula and ends the proof. □

Remark A.2 (Non-local nature of the reverse Cauchy dynamics)**.** The reverse Cauchy dynamics are genuinely non-local. The density ratio $\frac{p_{T-s}(x+z)}{p_{T-s}(x)}$ tilts the Cauchy jump kernel in the reverse generator. The quantity $\beta_s(x)$ in Proposition A.1 should therefore be understood as a fractional-score commutator drift, not as an ordinary local mean drift obtained by applying the generator to an integrable identity function.

A.2 Conditional mean and reverse martingale properties

We now derive the Cauchy analogue of Tweedie's formula. The formula is meaningful as long as Y is integrable. It expresses the conditional mean in terms of the Hilbert-transform score of the marginal density p_t.

Proposition A.3 (Cauchy Tweedie formula with deterministic Cauchy clock)**.** *Assume that Y is integrable. In the setting of Section 3.5.4, for every $t > 0$ and every x such that $p_t(x) > 0$,*

$$m_t(x) := \mathrm{E}[Y \mid X_t = x] = x - \Gamma_t \frac{(\mathcal{A}_0 p_t)(x)}{p_t(x)}.$$

Consequently, the Cauchy reverse score drift is

$$\beta_s(x) = -a_{T-s} \frac{(\mathcal{A}_0 p_{T-s})(x)}{p_{T-s}(x)} = \frac{a_{T-s}}{\Gamma_{T-s}} \left(m_{T-s}(x) - x \right), \qquad 0 \leq s < T.$$

In particular, $\beta_s(\overline{X}_s) = \frac{a_{T-s}}{\Gamma_{T-s}} \left(\mathrm{E}[Y \mid \overline{X}_s] - \overline{X}_s \right)$.

Proof. For fixed $t > 0$, the noise Z_t is Cauchy with characteristic function

$$\mathrm{E}[e^{iuZ_t}] = \exp(-\Gamma_t |u|).$$

The Cauchy Bartlett–Tweedie formula gives

$$\mathrm{E}[Y \mid X_t = x] = x - \Gamma_t \frac{(\mathcal{A}_0 p_t)(x)}{p_t(x)},$$

which proves the first identity.

Using the same identity at time $T - s$, we get

$$m_{T-s}(x) - x = -\Gamma_{T-s} \frac{(\mathcal{A}_0 p_{T-s})(x)}{p_{T-s}(x)}.$$

Multiplying by a_{T-s}/Γ_{T-s} yields

$$\frac{a_{T-s}}{\Gamma_{T-s}} \left(m_{T-s}(x) - x \right) = -a_{T-s} \frac{(\mathcal{A}_0 p_{T-s})(x)}{p_{T-s}(x)} = \beta_s(x)$$

and ends the proof. □

Define $M_t := \mathrm{E}[Y \mid X_t]$, $t \in [0, T]$. Since Y is integrable, the process M_t is integrable. By contrast, the normalized Cauchy score

$$S_t := \frac{M_t - X_t}{\Gamma_t} = -\frac{(\mathcal{A}_0 p_t)(X_t)}{p_t(X_t)}$$

is not integrable in general, because X_t contains a non-integrable Cauchy component. The reverse score drift is formally

$$R_t := a_t S_t = \frac{a_t}{\Gamma_t} (M_t - X_t),$$

and in reverse time we write $\overline{M}_s := M_{T-s}$, $\overline{S}_s := S_{T-s}$, and $\overline{R}_s := R_{T-s} = \beta_s(\overline{X}_s)$.

The next result records the reverse martingale property that survives without any additional integrability assumption on the Cauchy score: the reverse conditional mean is a martingale. The proof is similar to the one of Proposition 4.2.

Proposition A.4 (Reverse martingale property of the conditional mean)**.** *In the setting of Section 3.5.4, let $\mathcal{G}_t := \sigma(X_u : u \in [t, T])$, $t \in [0, T]$, be the decreasing filtration generated by the future of the forward process. Then, $\mathrm{E}[M_s \mid \mathcal{G}_t] = M_t$ for $0 < s < t \leq T$. Equivalently, in reverse time, $(\overline{M}_s)_{0 \leq s < T}$ is a martingale with respect to $\overline{\mathcal{F}}_s := \mathcal{G}_{T-s} = \sigma(\overline{X}_u : 0 \leq u \leq s)$.*

Remark A.5 (The reason why the Cauchy score is not a martingale in the usual sense)**.** The formal normalized Cauchy score is

$$S_t = \frac{M_t - X_t}{\Gamma_t}.$$

Since M_t is integrable and $X_t = Y + Z_t$ contains a Cauchy component, X_t is not integrable for every $t > 0$. Hence, S_t is not integrable in general. For instance, if $Y \equiv 0$, then $X_t = Z_t$ is Cauchy and

$$S_t = -\frac{X_t}{\Gamma_t}, \qquad R_t = -\frac{a_t}{\Gamma_t} X_t,$$

which are not integrable. Thus, $(\overline{S}_s)$ and $(\overline{R}_s)$ cannot be treated as ordinary martingales without truncation, localization, or another generalized notion of integrability.

A.3 Convex-order comparisons and peacock processes

The ordinary convex order is defined for integrable random variables. Therefore, in the Cauchy setting, the conditional mean $\overline{M}_s$ admits a genuine convex-order comparison, but the Cauchy score drift $\overline{R}_s$ does not in general.

Corollary A.6 (Convex-order monotonicity of the reverse conditional mean)**.** *In the setting of Section 3.5.4, the stochastic inequality $\overline{M}_u \preceq_{\mathrm{cx}} \overline{M}_s$ holds for all $0 \le u \le s < T$. Hence, $(\overline{M}_s)_{0 \le s < T}$ is a peacock.*

Proof. The process $(\overline{M}_s)_{0 \le s < T}$ is a martingale by Proposition A.4. The result follows from conditional Jensen's inequality. Since a martingale is increasing in convex order, the family $(\overline{M}_s)$ is a peacock. □

Remark A.7 (No ordinary peacock statement for the Cauchy score drift)**.** The reverse Cauchy score drift

$$\overline{R}_s = -a_{T-s} \frac{(\mathcal{A}_0 p_{T-s})(\overline{X}_s)}{p_{T-s}(\overline{X}_s)}$$

does not generally define a peacock in the ordinary sense, because the usual convex order requires integrability and equality of means. In the Cauchy case, integrability fails for the score and for the score drift in general. A peacock-type statement for $(\overline{R}_s)$ would therefore require a modified framework, for example a truncated convex order, a localized martingale problem, or additional assumptions replacing ordinary L^1-integrability.

Remark A.8 (Contrast with $\alpha \in (1, 2)$)**.** For symmetric α-stable noise with $\alpha \in (1, 2)$, the stable noise is integrable, and the normalized fractional score is an L^1-martingale in reverse time. This makes it possible to compare both the score and the reverse local mean drift in convex order. The Cauchy case is the borderline case where the Bartlett–Tweedie identity still exists, but the integrability required for the convex-order theory of the score drift is lost.

A.4 Linearity of the Cauchy score and drift

In the Gaussian, compound Poisson and α-stable cases with $\alpha \in (1, 2)$, linearity of the reverse score and drift is obtained when the initial distribution belongs to the same convolution semigroup as the noise. The Cauchy case is more delicate. The semigroup-compatible initial distribution is itself Cauchy, but then Y is not integrable and the ordinary conditional mean $\mathrm{E}[Y \mid X_t]$ is not defined.

The next result makes this distinction explicit.

Proposition A.9 (No non-trivial affine conditional mean in the integrable Cauchy setting)**.** *Assume that Y is integrable. In the setting of Section 3.5.4, if*

$$\mathrm{E}[Y \mid X_t] = \alpha_t X_t + \beta_t$$

for some constants $\alpha_t, \beta_t \in \mathbb{R}$, $t > 0$, then necessarily $\alpha_t = 0$. Thus, no non-trivial affine shrinkage of X_t can represent the ordinary conditional mean in the Cauchy setting.

Proof. Since Y is integrable, the conditional mean $M_t = \mathrm{E}[Y \mid X_t]$ is integrable. If $M_t = \alpha_t X_t + \beta_t$ and $\alpha_t \neq 0$, then X_t would be integrable. This is impossible because $X_t = Y + Z_t$, where Y is integrable and Z_t is Cauchy with scale $\Gamma_t > 0$. Hence, Z_t is not integrable, and therefore $\alpha_t = 0$. □

The formal linear Cauchy formulas can nevertheless be recovered at the level of characteristic exponents when the initial distribution is Cauchy. These formulas should not be interpreted as ordinary L^1-conditional expectation identities.

Proposition A.10 (Formal Cauchy semigroup-compatible linearity)**.** *Assume formally that $Y = y_0 + Y^{(0)}$, where $Y^{(0)}$ is symmetric Cauchy with scale parameter $\theta > 0$, namely*

$$\mathrm{E}[e^{iuY^{(0)}}] = \exp(-\theta|u|), \qquad u \in \mathbb{R}.$$

Then, $X_t - y_0 = Y^{(0)} + Z_t$ is Cauchy with scale parameter $\theta + \Gamma_t$. At the level of Cauchy characteristic exponents, the formal analogue of the linear conditional mean is

$$M_t^{\text{form}} = y_0 + \frac{\theta}{\theta + \Gamma_t}(X_t - y_0).$$

The corresponding formal normalized score is

$$S_t^{\text{form}} = -\frac{X_t - y_0}{\theta + \Gamma_t},$$

and the formal reverse score drift is

$$R_t^{\text{form}} = -\frac{a_t}{\theta + \Gamma_t}(X_t - y_0).$$

Equivalently, in reverse time,

$$\overline{M}_s^{\text{form}} = y_0 + \frac{\theta}{\theta + \Gamma_{T-s}}(\overline{X}_s - y_0), \quad \overline{S}_s^{\text{form}} = -\frac{\overline{X}_s - y_0}{\theta + \Gamma_{T-s}}, \text{ and } \overline{R}_s^{\text{form}} = -\frac{a_{T-s}}{\theta + \Gamma_{T-s}}(\overline{X}_s - y_0).$$

Proof. The characteristic function of $Y^{(0)}$ is $\varphi_{Y^{(0)}}(u) = \exp(-\theta|u|)$, whereas the characteristic function of Z_t is $\varphi_{Z_t}(u) = \exp(-\Gamma_t|u|)$. Thus, $Y^{(0)}$ and Z_t belong to the same Cauchy convolution semigroup, with parameters θ and Γ_t. Formally, the part of the total Cauchy exponent associated with $Y^{(0)}$ is $\frac{\theta}{\theta+\Gamma_t}$. This yields the displayed linear shrinkage formula. The formulas for the formal score and formal score drift follow by subtracting X_t, dividing by Γ_t, and multiplying by a_t. □

Remark A.11 (Interpretation of the formal linear formulas)**.** Proposition A.10 lies outside the ordinary conditional-mean framework used in the rest of the paper. If the initial distribution is Cauchy, then Y is not integrable and $\mathrm{E}[Y \mid X_t]$ is not defined as an L^1-conditional expectation. The displayed formulas are therefore formal identities at the level of characteristic exponents, or equivalently the limiting $\alpha \downarrow 1$ form of the stable semigroup-compatible formulas.

B An additional lemma

Lemma B.1 (Future increments and conditional reduction)**.** *Assume that Assumption 3.2(2) holds, and suppose that*

$$X_s = Y + Z_s, \qquad 0 \le s \le T.$$

Fix $t \in [0,T]$ and set

$$\mathcal{H}_t := \sigma(Y, Z_v : 0 \le v \le t), \qquad \mathcal{I}_t := \sigma(Z_u - Z_t : t \le u \le T).$$

Then $\mathcal{I}_t$ is independent of $\mathcal{H}_t$. Moreover,

$$\mathcal{G}_t := \sigma(X_u : t \le u \le T) = \sigma(X_t) \vee \mathcal{I}_t.$$

Consequently, for every integrable $\mathcal{H}_t$-measurable random variable U,

$$\mathrm{E}[U \mid \mathcal{G}_t] = \mathrm{E}[U \mid X_t].$$

In particular, if $0 < r < t$ and if $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable with $g(X_r)$ integrable, then

$$\mathrm{E}[g(X_r) \mid \mathcal{G}_t] = \mathrm{E}[g(X_r) \mid X_t].$$

Proof. For $u \ge t$, we have

$$X_u = X_t + (Z_u - Z_t).$$

Hence

$$\mathcal{G}_t = \sigma(X_t) \vee \sigma(Z_u - Z_t : t \le u \le T) = \sigma(X_t) \vee \mathcal{I}_t.$$

By Assumption 3.2(2), applied successively to the increments over disjoint intervals after time t, every finite-dimensional future-increment vector

$$(Z_{u_1} - Z_t, \dots, Z_{u_n} - Z_t), \qquad t \le u_1 \le \cdots \le u_n \le T,$$

is independent of $\mathcal{H}_t$. Hence, by a monotone-class argument, the sigma-field

$$\mathcal{I}_t = \sigma(Z_u - Z_t : t \le u \le T)$$

is independent of $\mathcal{H}_t$.

Let U be integrable and $\mathcal{H}_t$-measurable. Let H be a bounded $\sigma(X_t)$-measurable random variable and let K be a bounded $\mathcal{I}_t$-measurable random variable. Since U, H, and $\mathrm{E}[U \mid X_t]$ are $\mathcal{H}_t$-measurable, whereas K is independent of $\mathcal{H}_t$, we get

$$\begin{aligned} \mathrm{E}[UHK] &= \mathrm{E}[K]\,\mathrm{E}[UH] \\ &= \mathrm{E}[K]\,\mathrm{E}\big[\mathrm{E}[U \mid X_t]H\big] \\ &= \mathrm{E}\big[\mathrm{E}[U \mid X_t]HK\big]. \end{aligned}$$

By a monotone-class argument, the same identity holds for every bounded $\mathcal{G}_t$-measurable random variable in place of HK. Therefore

$$\mathrm{E}[U \mid \mathcal{G}_t] = \mathrm{E}[U \mid X_t].$$

Finally, if $0 < r < t$, then $X_r = Y + Z_r$ is $\mathcal{H}_t$-measurable, and so is $g(X_r)$. This gives the last assertion. $\square$