Stochastic Knothe–Rosenblatt:

Light-speed Calibration of Stochastic Local Volatility Models

M. Beiglböck, M. Hasenbichler, G. Pammer

Abstract

European option smiles determine the risk-neutral marginal laws of an asset, but not their intertemporal coupling, which is decisive for many applications. The Bass martingale construction selects, among all calibrated martingales, the one closest to Bachelier dynamics; it permits fast calibration at discrete maturities and recovers the Dupire local-volatility (LV) model as the maturity grid is refined.

This article develops a modular calibration overlay for existing stochastic and path-dependent volatility models. We recursively construct a martingale that matches all prescribed marginals exactly while remaining as close as possible, in an adapted Knothe–Rosenblatt sense, to the reference dynamics. As with the Bass LV model, each calibration step is amenable to an efficient Martingale Sinkhorn algorithm. We develop the theoretical foundations, numerical implementation and consider convergence to the SLV model. We also benchmark the method for Heston and Bergomi finite-factor path-dependent volatility dynamics and develop the multi-asset extension.

Keywords: Bass martingale; martingale Sinkhorn algorithm; calibration; stochastic local volatility; Dupire local volatility; martingale optimal transport; weak optimal transport; adapted Wasserstein distance; Knothe–Rosenblatt coupling.
MSC2020 subject classifications: Primary 60G42, 49Q22, 91G20; secondary 60G44, 91G60.

1 Introduction

1.1 Outline

Optimal transport provides variational tools for model calibration when the risk-neutral marginal laws $\mu_0, \dots, \mu_n$ of the asset price are available at finitely many maturities $0 = T_0 < T_1 < \ldots < T_n$ [10, 12, 24, 25, 28, 37, 44, 48, 50]. A prominent example is the Bass LV model [19]. It solves the Bachelier *minimal distortion calibration* problem

$$\inf_{\mathrm{d}X_t = \sigma_t^X \mathrm{d}B_t, X_{t_i} \sim \mu_i} \mathbb{E}\Big[\int_0^{T_n} |\sigma_t^X - \sigma^{\mathrm{Bach}}|^2 \,\mathrm{d}t \Big], \tag{Bass-MDC}$$

and is closest to the Bachelier model, subject to the marginal constraints. Crucially, (Bass-MDC) can be *separated* into n independent martingale transport problems, which allows for efficient numerics based on the *Martingale Sinkhorn algorithm* [19]. For many applications, however, one would like to calibrate relative to richer reference models.

Challenge (MDC). Given a reference model R, possibly driven by stochastic factors, construct a model X
1. that exactly reproduces the prescribed marginals $\mu_0, \dots, \mu_n$, and
2. matches the dynamics of the reference model R as closely as possible.

To obtain a numerically feasible approach to this challenge, a tempting simplification is to mimic R on each interval $[T_i, T_{i+1}]$ and then concatenate the corresponding solutions. However, this effectively enforces stepwise Markovianity. In the regular diffusion setting, any continuous strong Markov martingale limit with the prescribed marginals is the Dupire LV model [21], just as for the much simpler Bass model; cf. [40, 41]. To preserve the reference dynamics while retaining tractable computation, we pursue the following alternative.

Stochastic Knothe–Rosenblatt approach to MDC. We propose a *Stochastic Knothe–Rosenblatt* (SKR) construction, which builds the calibrated model forward in time. Its modular design allows quant desks to retain their existing stochastic volatility models as references and add SKR as an exact calibration overlay. At each maturity, we retain the joint law of the calibrated model and the reference for the next calibration step. This preserves the dependence structure while reducing the optimization to single-period problems. These are solved using a fast and robust Martingale Sinkhorn algorithm that is simple to implement [31, 32].

We develop the theoretical foundations and benchmark the method for Heston, Bergomi, and finite-factor path-dependent volatility dynamics. Under the regularity assumptions stated in Appendix A, the SKR model approaches the SLV model associated with the same reference model R as the maturity grid becomes finer. In particular, the SKR construction provides an efficient and stable numerical scheme for SLV models, in the same way that the Bass construction does for the Dupire LV model.

Funding. MB gratefully acknowledges financial support from FWF through P34743, P35197, and FW506064 and from OeNB through AB1898311. All authors are grateful to Valentin Tissot-Daguette for many insightful discussions and recommendations.
In the appendix we report a proof of convergence of SKR to SLV which was established with very substantial support by ChatGPT 6.0.

1.2 Existence and uniqueness of SKR

Let E be a Polish space and let $R = (Y, L)$ be a Markov process on $\mathbb{R} \times \mathsf{E}$. Throughout we assume that for every starting point $r \in \mathbb{R} \times \mathsf{E}$, the first coordinate Y is a square integrable martingale, with $\mathrm{Law}(Y_t \mid R_s = r)$ continuous for $s < t$. Let $\delta_{x_0} = \mu_0 \leq_c \ldots \leq_c \mu_n$ and assume that all marginals have finite second moment.

Suppose that calibration has reached T_i. The state passed to the next step is $\kappa_i := \mathrm{Law}(X_{T_i}, L_{T_i})$; its first marginal is the prescribed law μ_i. Starting from κ_i, we calibrate to the next marginal μ_{i+1} by choosing the dynamics of X as close as possible to the dynamics of Y. We formulate this as minimizing the squared difference of the volatilities of X and Y:

$$V_i(\kappa_i, \mu_{i+1}) := \min_{X \in \mathcal{A}(\kappa_i, \mu_{i+1})} \mathbb{E}[[X - Y]_{T_{i+1}} - [X - Y]_{T_i}], \tag{Bridge}$$

where $\mathcal{A}(\kappa_i, \mu_{i+1})$ consists of square-integrable $(\mathcal{F}_t)$-martingales X satisfying $X_{T_i} = Y_{T_i}$ and $X_{T_{i+1}} \sim \mu_{i+1}$. We work with the usual augmentation $(\mathcal{F}_t)$ of the natural filtration generated by $R = (Y, L)$, started with law κ_i at T_i, and assume that R is Markov with respect to this filtration.

Thus, for each T_i the reference process is started in κ_i. The bridge starts from the same random state as the reference process and minimizes the quadratic variation accumulated by $X - Y$ over $[T_i, T_{i+1}]$.[1]

The joint evolution of the optimizer X and the reference model then determines $\kappa_{i+1} = \mathrm{Law}(X_{T_{i+1}}, L_{T_{i+1}})$, which serves as the initial condition for the next step. By constructing κ_i sequentially from the preceding state, our approach follows the *Knothe–Rosenblatt* paradigm from classical transport: building complex couplings through tractable conditional transports [14, 17].

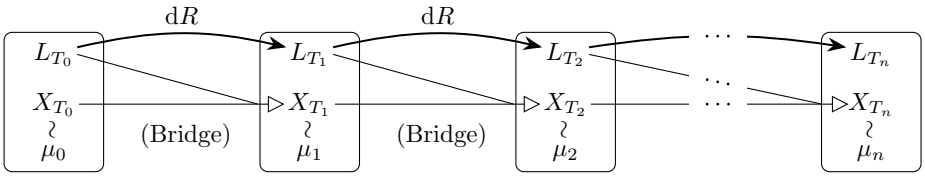


Figure 1: Minimal Distortion Calibration via Stochastic Knothe–Rosenblatt. At each maturity, the propagated state has law $\kappa_i = \mathrm{Law}(X_{T_i}, L_{T_i})$, whose first marginal is μ_i.

To solve (Bridge), a key observation is that this *dynamic* problem admits an equivalent *static* formulation as a *weak martingale optimal transport* (WMOT) problem [13, 18]. Unlike in classical martingale transport, where the cost is assigned to a pair of points (x, y), WMOT costs may depend on x and the terminal law of a bridge started at x.

Under the above assumptions and the regularity conditions specified in Section 2 for each calibration interval, we obtain

Theorem 1.1 (Existence and uniqueness of the SKR construction)**.** *The one-period optimizers exist and are unique in law. The stochastic Knothe–Rosenblatt model exists and is unique in law.*

1.3 Bass-structure

Fix one calibration interval and rescale it to $[0, 1]$. Let μ, ν be probabilities on $\mathbb{R}$ with finite second moments, $\mu \leq_c \nu$, where (μ, ν) is irreducible, i.e. the call potentials c_μ, c_ν satisfy that $I := \{c_\mu < c_\nu\}$ is an interval and $\mu(I) = 1$. Let κ be a probability on $\mathbb{R} \times \mathsf{E}$ with first marginal μ and for $r \in \mathbb{R} \times \mathsf{E}$, let

$$\gamma_r := \mathrm{Law}(Y_1 \mid R_0 = r)$$

be the conditional terminal law of the reference martingale. We assume that γ_r is continuous for every r and that the map $r \mapsto \gamma_r$ is $\mathcal{W}_2$ continuous.

Starting the reference process (Y, L) in κ, the unit interval bridge problem is to find the martingale X solving

$$\inf_{X: X_0 = Y_0, X_1 \sim \nu} \mathbb{E}[X - Y]_1. \tag{1}$$

We will give a structural characterization in the spirit of Bass martingales. To this end we introduce some notation.

Given a function $g : \mathbb{R} \to [-\infty, \infty]$ and a probability γ on $\mathbb{R}$ we denote the 'reflected' convolution by $g \star \gamma(a) := \int g(a + y)\, d\gamma(y)$. If g is increasing and γ is atom-free, then $g \star \gamma$ is increasing and continuous on the interval where it is finite. We use $(g \star \gamma)^{-1}$ to denote an inverse of $g \star \gamma$.

Definition 1.2 (R-Bass martingale)**.** A martingale X is called an R-Bass martingale if for some increasing $g : \mathbb{R} \to [-\infty, \infty]$ and $A := a(R_0) := (g \star \gamma_{R_0})^{-1}(X_0)$

$$X_t = \mathbb{E}\big[g(A + Y_1) \,\big|\, \sigma(A, R_s : 0 \leq s \leq t)\big] = \mathbb{E}[g(A + Y_1) \mid A, R_t], \qquad 0 \leq t \leq 1. \tag{2}$$

[1] When X and Y can be represented as Itô processes $\mathrm{d}X_t = \sigma_t^X\, \mathrm{d}W_t$, $\mathrm{d}Y_t = \sigma_t^Y\, \mathrm{d}W_t$, the optimization objective can also be written in the form $\mathbb{E}\Big[[X - Y]_{T_{i+1}} - [X - Y]_{T_i}\Big] = \mathbb{E}\Big[\int_{T_i}^{T_{i+1}} |\sigma_t^X - \sigma_t^Y|^2\, dt\Big]$, as in (Bass-MDC). However, the formulation in (Bridge) is technically more convenient, in particular in view of applications to non-continuous models.

The second equality follows from the Markov property of the reference process and the fact that $A = a(R_0)$ is measurable at the initial time.

We briefly comment on Definition 1.2. First, note that the martingale condition $X_0 = \mathbb{E}[X_1|X_0, L_0] = \mathbb{E}[X_1|R_0]$ amounts to

$$X_0 = \mathbb{E}[g(a(R_0) + Y_1)|R_0] \quad \Longleftrightarrow \quad x_0 = (g \star \gamma_{r_0})(a(r_0)),\ \kappa\text{-a.s.}$$

This explains the choice $a(r_0) = (g \star \gamma_{r_0})^{-1}(x_0)$. In the definition of R-Bass martingales we tacitly assume that $(g \star \gamma_{r_0})^{-1}(x_0)$ exists κ-a.s. While the inverse used here is not necessarily unique, it is straightforward to see that $g(A + Y_1)$ does not depend on the specific choice of the inverse.

We observe also that shifting the function g to the right by a constant c amounts to adding c to the respective A. This degree of freedom will be used to normalize the shifts.

Theorem 1.3 (Bass characterization)**.** *Under the standing assumptions, an admissible martingale X from κ to ν solves the unit interval bridge problem* (1) *if and only if it is an R-Bass martingale.*

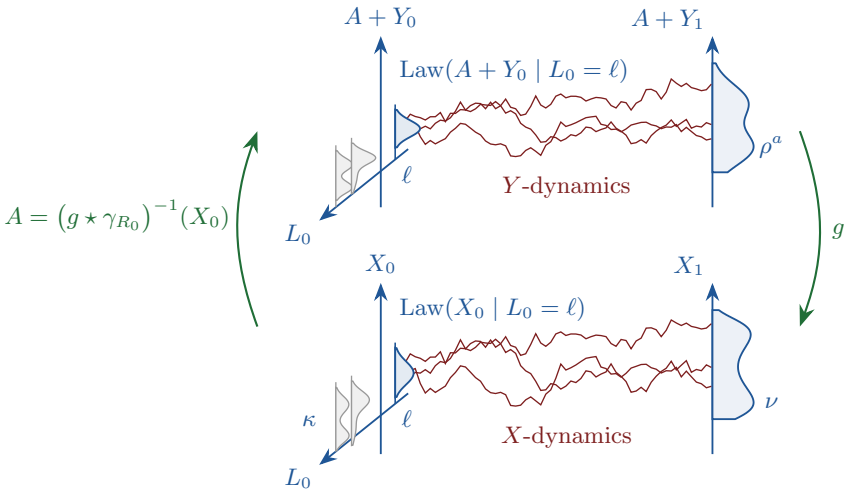


Figure 2: An R-Bass martingale. The reference model is started in κ and shifted by $A = a(R_0)$; the increasing map g carries the shifted terminal law ρ^a onto ν, and $X_t = \mathbb{E}[g(A + Y_1) \mid A, R_t]$ is the resulting martingale. The shift is pinned down by the martingale condition $X_0 = (g \star \gamma_{R_0})(A)$. Each fibre over the latent axis carries the conditional law of the initial value given L_0; the paths are started on the front fibre $L_0 = \ell$.

In the case where R is Brownian motion, the construction underlying our one-period steps goes back to Bass [9], who transported a Gaussian law onto a prescribed one by a monotone map in order to solve the Skorokhod embedding problem. While Bass was only concerned with the deterministic starting case, Bass martingales between general marginals were introduced by Backhoff-Veraguas–Beiglböck–Huesmann–Källblad [6] and Guo–Loeper–Wang [25]. Extensions beyond the Brownian case are considered in [2, 3, 15, 19, 32, 49]. The R-Bass martingales of Definition 1.2 continue this line of thought: we replace the Brownian increment by the conditional increment of a general Markovian reference model, and it is precisely this flexibility which allows the calibrated model to inherit the dynamics of R.

1.4 The Martingale Sinkhorn algorithm

The structural characterization of optimal bridges in Theorem 1.3 forms the basis for a particularly simple, yet powerful numerical algorithm, which we introduce below. For this it is useful to define for a shift $a : \mathbb{R} \times \mathsf{E} \to \mathbb{R}$, the shifted reference terminal law

$$\rho^a := \int (a(r) + \cdot\,)_{\#}\gamma_r\, \kappa(\mathrm{d}r) = \mathrm{Law}(a(R_0) + Y_1). \tag{3}$$

The initial and terminal constraints of an R-Bass martingale can then be expressed concisely in terms of a and g.

Terminal marginal constraint. $g_{\#}(\rho^a) = \nu$.

Martingale constraint. $x = (g \star \gamma_r)(a(r))$ for κ-a.e. $r = (x, \ell)$.

This observation suggests a natural fixed-point algorithm. Starting from a shift $a_0 : \mathbb{R} \times \mathsf{E} \to \mathbb{R}$, perform the following two updates.

Terminal-marginal update. Let g_k be the increasing transport from ρ^{a_k} to ν, i.e. $g_k := F_\nu^{-1} \circ F_{\rho^{a_k}}$,

Martingale update. For κ-a.e. $r = (x, \ell)$, set $a_{k+1}(x, \ell) = (g_k \star \gamma_{(x,\ell)})^{-1}(x)$.

Thus the first update enforces the terminal marginal and the second enforces the martingale constraint.

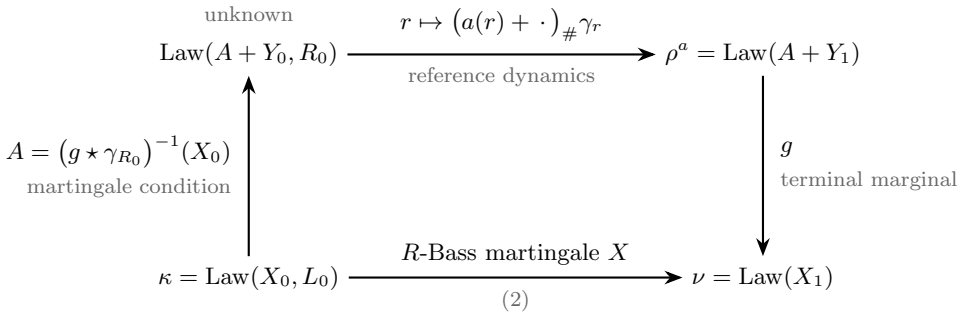


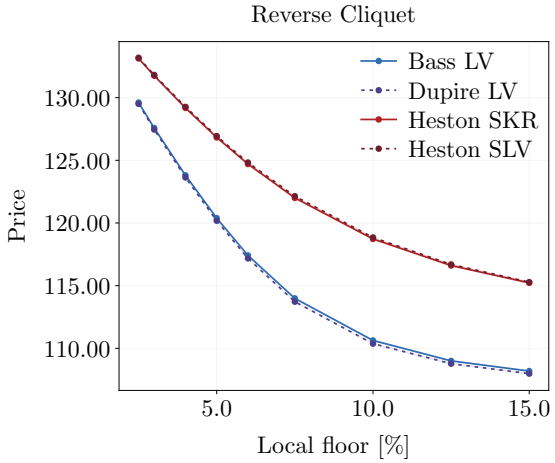
Figure 3: The Bass diagram behind Definition 1.2, in the notation of Figure 2. The prescribed data are the marginals κ and ν; the unknown is the shift a, equivalently the initial law of the *shifted* reference process. The martingale condition determines A from the initial state, running the reference dynamics from $A+Y_0$ produces ρ^a, and the increasing map g carries ρ^a onto ν. Composing the three outer arrows gives the R-Bass martingale along the bottom.

In Theorem 4.1 below we prove convergence of the Martingale Sinkhorn algorithm to the R-Bass martingale in the case where (μ,ν) is strictly irreducible and ν is compactly supported, under the additional reference assumptions and normalizations stated there. In numerical experiments, we observe linear convergence for generic reference processes.

For the classical Brownian Bass construction the above fixed point scheme was introduced by Conze–Henry-Labordère [19] who demonstrate its striking numerical properties for the approximation of the (Bass) LV model. In the Brownian case, linear convergence was proved by Acciaio–Marini–Pammer [4], and a measure-preserving, multidimensional variant is given in [31]. In the case where the reference dynamics are Brownian, the iterative scheme above reduces precisely to the Conze–Henry-Labordère [19] algorithm.

1.5 An instantaneous calibration principle — from SKR to SLV

Since local volatility is well known to generate unrealistic forward smiles and spot-volatility dynamics, it is common practice to superimpose a leverage function or leverage adjustment on a stochastic volatility reference, which leads to stochastic local volatility (SLV) models. Their calibration amounts to a McKean–Vlasov equation whose well-posedness and numerics are somewhat delicate, see [1, 20, 27, 29, 35, 36]. The present SKR construction addresses the same calibration task, but replaces the McKean–Vlasov equation by a sequence of 1-d WMOT problems.



	Valuation		Time [s]		
Model	Price	MC SE	Calibration	Pricing	**Total**
Bass LV	113.987	0.038	0.079	0.018	**0.097**
Dupire LV	113.879	0.038	0.005	4.290	**4.295**
Heston SKR	122.000	0.039	0.197	0.076	**0.273**
Heston SLV	122.139	0.039	18.802	6.976	**25.778**

Figure 4: Comparison of Bass LV with Dupire LV and Heston SKR with Heston SLV for the reverse cliquet of Section 6.2. All models use the SPX-based ESSVI targets of Section 6.1, with identical reference parameters for the two Heston models. The plot varies the loss cap L; the table reports prices, Monte Carlo standard errors, and timings at $L=7.5\%$ for 10^5 pricing paths.[2]

There is a simple probabilistic heuristic which suggests convergence of SKR to SLV. We start by interpreting SLV as the *instantaneously least-distorting Markovian calibration*. Consider a diffusion reference $\mathrm{d}Y_t=\sigma_R(t,Y_t,L_t)\mathrm{d}W_t$ with an autonomous factor L, and suppose that a smooth family of target marginals (μ_t) is prescribed. Write σ_{Dup} for the corresponding Dupire diffusion coefficient, i.e., the volatility depending only on time and state for which the local-volatility martingale $\mathrm{d}Z_t=\sigma_{\mathrm{Dup}}(t,Z_t)\mathrm{d}B_t$ has $\mathrm{Law}(Z_t)=\mu_t$.

At a fixed time, retain the joint law of (X_t,L_t) and ask how to modify the reference volatility as little as possible while matching the prescribed marginal evolution. For calibrated dynamics $\mathrm{d}X_t=a(t,X_t,L_t)\mathrm{d}W_t$, the latter requires

$$\mathbb{E}[a(t,X_t,L_t)^2\mid X_t=x]=\sigma_{\mathrm{Dup}}(t,x)^2. \tag{4}$$

[2]Dupire local volatility is computed analytically, without numerical calibration, from five ESSVI smiles at 0.25, 0.5, 1, 2, and 3 years (SPX, 23 June 2020 [22]). Heston SLV is calibrated to this local-volatility target using 250 000 particles. The ESSVI surface calibration from raw quotes is excluded from the timings.

We use QuantLib's C++ implementations of Dupire LV and Heston SLV with 64 time steps per year, and our C++ implementations of Bass LV and Heston SKR. Computations were performed on an Intel Core i5-12500 processor with 32 GiB RAM under Debian 13.

Among coefficients satisfying (4), the conditional mean-square distortion $\mathbb{E}[|a(t,x,L_t) - \sigma_R(t,x,L_t)|^2 \mid X_t = x]$ is minimized by

$$a(t,x,l) = \lambda(t,x)\sigma_R(t,x,l), \qquad \lambda(t,x) = \frac{\sigma_{\mathrm{Dup}}(t,x)}{\sqrt{\mathbb{E}[\sigma_R(t,X_t,L_t)^2 \mid X_t = x]}}; \tag{5}$$

this is a consequence of Cauchy–Schwarz. Thus the least-distorting instantaneous adjustment rescales the reference volatility by a factor λ depending only on time and price. This is precisely the SLV structure $\mathrm{d}X_t = \lambda(t,X_t)\sigma_R(t,X_t,L_t)\mathrm{d}W_t$.

SKR implements the same principle over finite calibration intervals. Each bridge in (Bridge) compares the calibrated and reference martingale increments, starting from the same price and factor state. Over a short interval, their mean-square difference is governed, to leading order, by the difference of their instantaneous volatilities. Matching the next marginal imposes the variance constraint (4) to first order. The optimal bridge should therefore select the leverage adjustment (5). The leverage in (5) depends on the distribution of the factor L conditional on the current calibrated price. By propagating the joint law of (X_{T_i}, L_{T_i}), SKR carries this information forward to the next calibration step.

This suggests viewing SKR as a variational time discretization of SLV, with exact calibration at every grid point. We develop this heuristic in Section 5 and establish convergence under regularity assumptions in Appendix A.

1.6 Multi-asset calibration with SKR

The SKR construction extends to several assets while retaining the scalar structure of the calibration problems. We consider a Markov reference model $R = (Y, L)$ on $\mathbb{R}^d \times \mathsf{E}$, whose price coordinates $Y = (Y^1, \dots, Y^d)$ form a square integrable martingale. The calibration data consist of the single-asset marginal laws $X^j_{T_i} \sim \mu^j_i$, where $\delta_{x^j_0} = \mu^j_0 \leq_c \cdots \leq_c \mu^j_n$ have finite second moments. These constraints leave the inter-asset dependence undetermined; the reference model provides the dynamics against which the joint calibration is compared.

As in the single-asset construction, we propagate the full joint state law $\kappa_i := \mathrm{Law}(X^1_{T_i}, \dots, X^d_{T_i}, L_{T_i})$. To describe one calibration step, rescale the interval to $[0,1]$, write $\kappa = \kappa_i$ and $\nu^j = \mu^j_{i+1}$, and start the reference process with $R_0 \sim \kappa$. The bridge problem becomes

$$\inf_{X:\, X_0 = Y_0, X^j_1 \sim \nu^j} \mathbb{E}\Big[\sum_{j=1}^{d} [X^j - Y^j]_1\Big], \tag{6}$$

with the same martingale and filtration requirements as in the single-asset case.

The reason the preceding theory applies is that, for fixed κ, both the distortion objective and the terminal constraints separate across assets. For asset j, we use the single-asset construction with the remaining reference price coordinates included in the latent state. In particular, its conditional reference laws are taken given the full state $r = (x^1, \dots, x^d, \ell)$. The resulting scalar optimizers can all be realized on the same reference process, so they simultaneously attain the minimum of the sum in (6).

More explicitly, put $\gamma^j_r := \mathrm{Law}(Y^j_1 \mid R_0 = r)$. Under the assumptions of the single-asset Bass characterization, imposed coordinatewise, the unique optimizer has the representation

$$X^j_t = \mathbb{E}\Big[g^j\big(a^j(R_0) + Y^j_1\big) \,\Big|\, \mathcal{F}_t\Big], \qquad a^j(r) \;\; = (g^j \star \gamma^j_r)^{-1}(x^j), \qquad j = 1, \dots, d, \tag{7}$$

where $(\mathcal{F}_t)$ is the reference filtration and the increasing maps $g^j : \mathbb{R} \to [-\infty, \infty]$ enforce the terminal laws ν^j. The shifts a^j enforce the martingale constraints and may depend on all components of the initial state.

For fixed incoming state law, the Martingale Sinkhorn iteration therefore applies separately to each coordinate, using the kernels γ^j_r. The coordinate updates can be performed in parallel, and no d-dimensional transport problem needs to be solved. The precise formulation and the existence and uniqueness result are given in Section 3.

1.7 Closely related literature

SLV models combine the exact fit of local volatility with the richer dynamics of stochastic volatility models. Early constructions include those of Jex–Henderson–Wang [33], Lipton [38], and Ren–Madan–Qian [46]. The leverage relation connects Dupire's local volatility formula [21] with the mimicking theorem of Gyöngy [30]. Guyon–Henry-Labordère [27, 29] developed a particle method for the resulting conditional McKean–Vlasov problem. Its well-posedness is delicate; rigorous results in restricted settings include [1, 35, 36]. A variational approach to calibration was developed by Guo–Loeper–Wang, first for local volatility [25] and subsequently for local-stochastic volatility [26]; see also Guo–Loeper [24] for path-dependent calibration.

The Bass construction goes back to Bass [9], who used a monotone transform of Brownian motion to solve the Skorokhod embedding problem. The Bass LV model for general initial laws was introduced independently in [6] and in the work of Loeper and Guo–Loeper–Wang [25, 39], through probabilistic and PDE approaches to the martingale Benamou–Brenier problem. Conze and Henry-Labordère [19] proposed the fixed-point scheme which makes its calibration particularly efficient. Acciaio–Marini–Pammer [4] proved existence and uniqueness of the fixed point, up to translation, and geometric convergence under suitable assumptions. The Martingale Sinkhorn formulation and

its extension to higher dimensions are developed by Hasenbichler–Joseph–Loeper–Obłój–Pammer [31]. The relation between increasingly fine Bass interpolations and the Dupire model is studied in [4, 12].

Closer to the present work are extensions of the Brownian reference. Conze–Henry-Labordère [19] already discuss non-Brownian kernels for a single Bass step, in a setting different from our bridge problem. Geometric Bass martingales and their change-of-numeraire representation are studied in [7, 15], while Tschiderer [49] introduced q-Bass martingales, replacing the Gaussian reference law by a more general measure. Acciaio–Antonio [2, 3] established existence first in the semidiscrete setting and then for general one-dimensional irreducible marginals with absolutely continuous reference law, together with stability and convergence results under additional assumptions.

2 The bridge problem as a weak martingale transport problem

As in the introduction, we rescale the calibration interval to $[0,1]$. The reference process $R = (Y, L)$ has initial law $\kappa \in \mathcal{P}_2(\mathbb{R} \times \mathsf{E})$, whose first marginal we denote by $\mu := \mathrm{pr}^X_{\#}\kappa$. We work with the usual augmentation $(\mathcal{F}_t)$ of the natural filtration generated by R, and assume that R is Markov with respect to this filtration. For a prescribed $\nu \in \mathcal{P}_2(\mathbb{R})$, the class $\mathcal{A}(\kappa, \nu)$ consists of square-integrable $(\mathcal{F}_t)$-martingales X satisfying $X_0 = Y_0$ and $X_1 \sim \nu$.

For $r = (x, \ell) \in \mathbb{R} \times \mathsf{E}$, write $\gamma_r := \mathrm{Law}(Y_1 \mid R_0 = r)$ for the conditional terminal law of the reference process. We work under the following assumptions on the conditional reference laws:

(A1) $(\rho, r) \mapsto \mathcal{W}_2(\rho, \gamma_r)$ is lower semicontinuous on $\mathcal{P}_2(\mathbb{R}) \times (\mathbb{R} \times \mathsf{E})$.

(A2) $\iint z^2 \, \gamma_r(\mathrm{d}z)\, \kappa(\mathrm{d}r) < \infty$, and $\bar{\gamma}_r = x$ for κ-a.e. $r = (x, \ell)$.

(A3) γ_r is atomless for κ-a.e. r.

Since $X_0 = Y_0$, integration by parts shows that the bridge problem is equivalent to

$$P(\kappa, \nu) := \sup_{X \in \mathcal{A}(\kappa,\nu)} \mathbb{E}[X_1 Y_1]. \tag{8}$$

To obtain a static formulation, consider an admissible bridge X and set $\pi := \mathrm{Law}(R_0, X_1)$, with disintegration $\pi(\mathrm{d}r, \mathrm{d}y) = \kappa(\mathrm{d}r)\pi_r(\mathrm{d}y)$. The martingale property implies

$$\int y \, \pi_r(\mathrm{d}y) = x \quad \text{for } \kappa\text{-a.e. } r = (x, \ell). \tag{9}$$

Thus π belongs to $\mathcal{M}(\kappa, \nu)$, the set of couplings in $\mathrm{Cpl}(\kappa, \nu)$ satisfying the martingale property (9).

Conditionally on $R_0 = r$, the pair (X_1, Y_1) is a coupling of π_r and γ_r leading us to consider the maximal covariance

$$\mathrm{MCov}(\rho, \gamma) := \sup_{q \in \mathrm{Cpl}(\rho,\gamma)} \int yz \, q(\mathrm{d}y, \mathrm{d}z).$$

Adapting the dynamic-to-static reduction used for the martingale Benamou–Brenier problem in [6], we obtain the following static formulation of the bridge problem.

Theorem 2.1 (Existence and uniqueness for the primal problem)**.** *Suppose that* (A1), (A2) *and* (A3) *hold, and let* $\mu \leq_{\mathrm{cvx}} \nu$. *Then*

$$P(\kappa, \nu) = \max_{\pi \in \mathcal{M}(\kappa,\nu)} \int \mathrm{MCov}(\pi_r, \gamma_r)\, \kappa(\mathrm{d}r). \tag{wBridge}$$

Moreover, both the dynamic and static problems admit optimizers. The optimizer of (wBridge) *is unique, and the optimizer of* (8) *is unique in law.*

The maximization problem on the right-hand side of (wBridge) fits into the weak optimal transport framework of [23]. More specifically, it falls within the extended weak martingale optimal transport framework of [34, Section 2.1], with ℓ as the additional information parameter.

Proof. Conditioning on R_0 and using the definition of MCov yields

$$P(\kappa, \nu) \leq \sup_{\pi \in \mathcal{M}(\kappa,\nu)} \int \mathrm{MCov}(\pi_r, \gamma_r)\, \kappa(\mathrm{d}r). \tag{10}$$

The static objective equals $C - \frac{1}{2}\int \mathcal{W}_2^2(\pi_r, \gamma_r)\, \kappa(\mathrm{d}r)$, where $C < \infty$ is independent of π by (A2). Since $(r, \rho) \mapsto \mathcal{W}_2^2(\rho, \gamma_r)$ is nonnegative, jointly lower semicontinuous by (A1), and convex in ρ, its integral is $\mathcal{W}_2$-lower semicontinuous by [5, Proposition 2.8]. Thus the objective is upper semicontinuous on the nonempty $\mathcal{W}_2$-compact set $\mathcal{M}(\kappa, \nu)$ and admits a maximizer $\hat{\pi}$.

By (A3), the conditional monotone couplings can be realized on the reference probability space by setting

$$X_1 := F^{-1}_{\hat{\pi}_{R_0}}\big(F_{\gamma_{R_0}}(Y_1)\big), \qquad X_t := \mathbb{E}[X_1 \mid \mathcal{F}_t],$$

where F_ρ and F_ρ^{-1} denote the distribution and quantile functions of ρ. The barycenter constraint gives $X_0 = Y_0$, while $X_1 \sim \nu$, so $X \in \mathcal{A}(\kappa, \nu)$. By construction,

$$\mathbb{E}[X_1 Y_1] = \int \mathrm{MCov}(\hat{\pi}_r, \gamma_r)\, \kappa(\mathrm{d}r).$$

Hence equality holds in (10), and both problems are attained.

By (A3), strict concavity of $\rho \mapsto \mathrm{MCov}(\rho, \gamma_r)$ gives uniqueness of $\hat{\pi}$. Any dynamic optimizer X must satisfy $\mathrm{Law}(R_0, X_1) = \hat{\pi}$, and its conditional terminal coupling must attain $\mathrm{MCov}(\hat{\pi}_r, \gamma_r)$. This is the unique monotone coupling, so X_1 is uniquely determined, up to null sets, as a function of (R_0, Y_1). Since $X_t = \mathbb{E}[X_1 \mid \mathcal{F}_t]$, the optimizer is unique, and in particular unique in law. □

In particular, applying Theorem 2.1 to every bridge between maturities establishes Theorem 1.1, the existence and uniqueness of the SKR model.

We next characterize optimal bridges as R-Bass martingales through the dual problem, following the approach of [8] for the martingale Benamou–Brenier problem. Under the assumptions below, this gives a concrete representation of the optimal bridge in terms of a single convex dual potential ψ, with generating function $g = \partial_x \psi^*$. Here, ψ^* denotes the convex conjugate of ψ. This representation provides the basis for the martingale Sinkhorn algorithm and its convergence analysis.

Here and below, the inverse in Definition 1.2 denotes a measurable choice of a finite shift $a(r)$ satisfying $(g \star \gamma_r)(a(r)) = x$ for κ-a.e. $r = (x, \ell)$. We require $g(a(R_0) + Y_1) \in L^2$; in particular, the convolution at $a(r)$ is absolutely convergent for κ-a.e. r.

To derive the dual problem, first let $\psi \in L^1(\nu)$ be proper, convex and lower semicontinuous. For $r = (x, \ell)$, set

$$(\psi^* \star \gamma_r)(a) := \int \psi^*(a + z)\, \gamma_r(\mathrm{d}z), \qquad \psi^C(r) := (\psi^* \star \gamma_r)^*(x).$$

Fenchel's inequality, together with the barycenter constraint $\int y\, \pi_r(\mathrm{d}y) = x$, yields

$$\mathrm{MCov}(\pi_r, \gamma_r) \le \int \psi\, \mathrm{d}\pi_r - \psi^C(r).$$

Consequently,

$$\sup_{\pi \in \mathcal{M}(\kappa,\nu)} \int \mathrm{MCov}(\pi_r, \gamma_r)\, \kappa(\mathrm{d}r) \le \int \psi\, \mathrm{d}\nu - \int \psi^C\, \mathrm{d}\kappa.$$

Taking the infimum over such ψ gives the associated dual problem. However, a limiting potential obtained from a minimizing sequence need not belong to $L^1(\nu)$, so the difference $\int \psi\, \mathrm{d}\nu - \int \psi^C\, \mathrm{d}\kappa$ may no longer be well defined. We therefore use the generalized-integral construction of Beiglböck and Juillet [10, Appendix A.3], further developed by Beiglböck, Nutz and Touzi [11, Section 4]: instead of evaluating the two integrals separately, we integrate the conditional difference.

Write $I_\nu := \operatorname{int}\operatorname{co}(\operatorname{supp}\nu)$. A *dual potential* is a proper, convex and lower semicontinuous function ψ that is finite on I_ν and equal to $+\infty$ outside $\bar{I}_\nu$. We refer to this domain restriction as the *support convention*. Adapting the relaxed dual functional in [8, Proposition 4.2] to the present setting, we define, for an irreducible pair (μ, ν), a dual potential ψ and $\pi \in \mathcal{M}(\kappa, \nu)$,

$$\mathcal{E}(\psi) := \int \Big(\int \psi\, \mathrm{d}\pi_r - \psi^C(r) \Big)\, \kappa(\mathrm{d}r). \tag{11}$$

Lemma 2.4 shows that $\mathcal{E}(\psi)$ is well-defined and independent of the choice of $\pi \in \mathcal{M}(\kappa, \nu)$. For $\psi \in L^1(\nu)$, it reduces to $\int \psi\, \mathrm{d}\nu - \int \psi^C\, \mathrm{d}\kappa$.

Theorem 2.2 (Duality and characterization of R-Bass martingales)**.** *Suppose that* (A1)*,* (A2) *and* (A3) *hold, and let* (μ, ν) *be irreducible. Then*

$$P(\kappa, \nu) = \min_{\psi \text{ dual potential}} \mathcal{E}(\psi). \tag{dBridge}$$

A martingale $X \in \mathcal{A}(\kappa, \nu)$ *solves* (8) *if and only if it is an R-Bass martingale.*

Moreover, suppose that $\operatorname{supp}\gamma_r$ *is either* $\mathbb{R}$ *or* $[0, \infty)$ *for* κ*-a.e.* r*. Then the dual minimizer is unique up to addition of an affine function on* I_ν*.*

In particular, under these assumptions, to every (κ, ν) there is at most one R-Bass martingale $X \in \mathcal{A}(\kappa, \nu)$. To show Theorem 2.2, we begin by identifying the value of the static problem with the dual problem over integrable convex potentials.

Lemma 2.3 (No duality gap)**.** *Suppose that* (A1) *and* (A2) *hold and that* $\mu \le_{\mathrm{cvx}} \nu$*. Then*

$$\sup_{\pi \in \mathcal{M}(\kappa,\nu)} \int \mathrm{MCov}(\pi_r, \gamma_r)\, \kappa(\mathrm{d}r) = \inf_{\substack{v \in L^1(\nu) \\ v \text{ cvx, lsc}}} \Big\{ \int v\, \mathrm{d}\nu - \int v^C\, \mathrm{d}\kappa \Big\}. \tag{12}$$

Proof. For a Gaussian reference law, this dual representation was established in [8, Sections 3–4]. We adapt that argument to the conditional reference laws γ_r.

The cost $C(r,\rho) := \frac{1}{2}\mathcal{W}_2^2(\rho,\gamma_r)$ when $\bar{\rho} = x$, and $+\infty$ otherwise, is nonnegative, jointly lower semicontinuous and convex in ρ. For such costs, [5, Theorem 1.3] identifies the primal value with a dual over continuous test functions. After cancelling the second-moment terms, this gives

$$\sup_{\pi\in\mathcal{M}(\kappa,\nu)} \int \mathrm{MCov}(\pi_r,\gamma_r)\,\kappa(\mathrm{d}r) = \inf_v \Big\{ \int v\,\mathrm{d}\nu - \int v_r(x)\,\kappa(\mathrm{d}r) \Big\}, \quad v_r(z) := \inf_{\substack{\rho\in\mathcal{P}_2(\mathbb{R}) \\ \bar{\rho}=z}} \Big\{ \int v\,\mathrm{d}\rho - \mathrm{MCov}(\rho,\gamma_r) \Big\},$$

where v is continuous and satisfies $c + \frac{1}{2}|\cdot|^2 \le v \le a + b|\cdot|^2$ for some $a, b, c \in \mathbb{R}$.

The argument of Pramenković [45, Theorem 3.2] shows that, for weak costs decreasing in convex order, replacing a dual potential by its convex envelope cannot worsen the dual objective. Since $\mathrm{MCov}(\cdot,\gamma_r)$ is increasing in convex order, this argument allows us to restrict the infimum to convex v.

It remains to identify v_r. The calculation in the proof of [16, Proposition 5.6] removes the barycenter constraint by conjugation and yields, under our quadratic bounds,

$$v_r^*(a) = \int v^*(a+y)\,\gamma_r(\mathrm{d}y) = (v^* \star \gamma_r)(a).$$

The quadratic bounds make v_r finite, while mixing competitors gives convexity. Thus Fenchel–Moreau yields $v_r = (v^* \star \gamma_r)^*$ and hence $v_r(x) = v^C(r)$.

Finally, Fenchel's inequality gives weak duality for every convex lower semicontinuous $v \in L^1(\nu)$, so enlarging the class of potentials leaves the value unchanged. In the irreducible case used below, setting $v = +\infty$ outside $\bar{I}_\nu$ preserves $\int v\,\mathrm{d}\nu$ and can only increase v^C, so the support convention also leaves the infimum unchanged. □

The preceding lemma identifies the static value with the dual problem over ν-integrable convex potentials. To pass to the relaxed dual problem, we first verify that the functional $\mathcal{E}$ is well defined and does not depend on the coupling π used in its definition.

Lemma 2.4. *Suppose that (μ,ν) is irreducible and that* (A2) *holds. Let $v : \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ be proper, convex and lower semicontinuous with $I_\nu = \operatorname{int}\operatorname{dom}(v)$. Then $\mathcal{E}(v) \in [0,\infty]$ is well defined and is independent of the choice of $\pi \in \mathcal{M}(\kappa,\nu)$. Moreover, $\mathcal{E}$ is invariant under addition of affine functions.*

Proof. Let $\pi \in \mathcal{M}(\kappa,\nu)$. For κ-a.e. $r = (x,\ell)$, we have $(v^* \star \gamma_r)^*(x) \le v(x) - x^2$. By Jensen's inequality,

$$\int v(y)\,\pi_r(\mathrm{d}y) - (v^* \star \gamma_r)^*(x) \ge x^2,$$

and so $\mathcal{E}(v) \in [0,\infty]$. Moreover, define $\chi(\mathrm{d}x,\mathrm{d}y) := \int \pi_{x,\ell}(\mathrm{d}y)\,\kappa(\mathrm{d}x,\mathrm{d}\ell)$, which is a martingale coupling of (μ,ν). Then

$$\begin{aligned}\mathcal{E}(v) &= \int \Big(\int v(y)\,\pi_{x,\ell}(\mathrm{d}y) - v(x) \Big) \kappa(\mathrm{d}x,\mathrm{d}\ell) + \int \big(v(x) - v^C(x,\ell) \big)\,\kappa(\mathrm{d}x,\mathrm{d}\ell) \\ &= \int \Big(\int v(y)\,\chi_x(\mathrm{d}y) - v(x) \Big) \mu(\mathrm{d}x) + \int \big(v(x) - v^C(x,\ell) \big)\,\kappa(\mathrm{d}x,\mathrm{d}\ell).\end{aligned}$$

[10, Lemma A.15] asserts that the first integral depends only on v and the marginals (μ,ν), but not on the particular coupling χ. Since the second integral is independent of π as well, $\mathcal{E}(v)$ is independent of the choice $\pi \in \mathcal{M}(\kappa,\nu)$.

Finally, let $c(x) := kx + d$ for some $k, d \in \mathbb{R}$. Then $(v+c)^*(y) = v^*(y-k) - d$, and hence

$$((v+c)^* \star \gamma_r)(y) = (v^* \star \gamma_r)(y-k) - d.$$

Taking conjugates gives

$$((v+c)^* \star \gamma_r)^*(x) = (v^* \star \gamma_r)^*(x) + kx + d = v^C(r) + c(x).$$

Since $\bar{\pi}_r = x$, we therefore obtain

$$\int (v+c)\,\mathrm{d}\pi_r - (v+c)^C(r) = \int v\,\mathrm{d}\pi_r - v^C(r),$$

for κ-a.e. $r = (x,\ell)$. Integrating with respect to κ yields $\mathcal{E}(v+c) = \mathcal{E}(v)$. □

To obtain a dual optimizer, we will pass to a locally uniform limit of normalized convex potentials on I_ν. The next lemma shows that $\mathcal{E}$ is lower semicontinuous with respect to such sequences.

Lemma 2.5. *Suppose that (μ,ν) is irreducible and that* (A2) *holds. Let $\hat{\psi}$ and $(\hat{\psi}_n)_{n\in\mathbb{N}}$ be dual potentials with $\hat{\psi}, \hat{\psi}_n \ge 0$ and $\hat{\psi}(\bar{\mu}) = \hat{\psi}_n(\bar{\mu}) = 0$ for every n. Suppose that $\hat{\psi}_n \to \hat{\psi}$ locally uniformly on I_ν. Then*

$$\mathcal{E}(\hat{\psi}) \le \liminf_{n\to\infty} \mathcal{E}(\hat{\psi}_n).$$

Proof. Fix $r = (x, \ell)$ with $x \in I_\nu$, $\gamma_r \in \mathcal{P}_2(\mathbb{R})$ and $\bar{\gamma}_r = x$, as holds for κ-a.e. r. We first show that

$$\limsup_{n\to\infty} (\hat{\psi}_n^* \star \gamma_r)^*(x) \le (\hat{\psi}^* \star \gamma_r)^*(x).$$

Pass to a subsequence along which the limsup is attained as a limit. If this limit is $-\infty$, there is nothing to show. Otherwise, $(\hat{\psi}_n^* \star \gamma_r)^*(x)$ is bounded below for all sufficiently large n.

Choose $\delta > 0$ such that $[x-\delta, x+\delta] \subset I_\nu$. Since $(\hat{\psi}_n^* \star \gamma_r)^*(z) \le \hat{\psi}_n(z) - xz$ for every $z \in I_\nu$, $(\hat{\psi}_n^* \star \gamma_r)^*$ is uniformly bounded above on $[x-\delta, x+\delta]$. Let a_n such that $(\hat{\psi}_n^* \star \gamma_r)^*(x) \le a_n x - (\hat{\psi}_n^* \star \gamma_r)(a_n) + \frac{1}{n}$ for every $n \in \mathbb{N}$. Since $(\hat{\psi}_n^* \star \gamma_r)^*(z) \ge a_n z - (\hat{\psi}_n^* \star \gamma_r)(a_n)$ for every $z \in I_\nu$, the sequence $(a_n)_{n\in\mathbb{N}}$ is bounded. Passing to a further subsequence, let $a_n \to a$. Then, for every z and every $y \in I_\nu$,

$$\liminf_{n\to\infty} \hat{\psi}_n^*(a_n + z) \ge \liminf_{n\to\infty} \big((a_n + z)y - \hat{\psi}_n(y)\big) = (a+z)y - \hat{\psi}(y),$$

and so $\liminf_{n\to\infty} \hat{\psi}_n^*(a_n + z) \ge \hat{\psi}^*(a+z)$. Moreover, $\hat{\psi}_n^*(y) \ge y\bar{\mu}$, and therefore Fatou's lemma gives

$$\int \hat{\psi}^*(a+y)\,\gamma_r(\mathrm{d}y) \le \liminf_{n\to\infty} \int \hat{\psi}_n^*(a_n+y)\,\gamma_r(\mathrm{d}y).$$

Consequently,

$$\limsup_{n\to\infty} (\hat{\psi}_n^* \star \gamma_r)^*(x) \le ax - \int \hat{\psi}^*(a+y)\,\gamma_r(\mathrm{d}y) \le (\hat{\psi}^* \star \gamma_r)^*(x).$$

Let $\pi \in \mathcal{M}(\kappa, \nu)$. By convexity and lower semicontinuity, $\hat{\psi}(y) \le \liminf_n \hat{\psi}_n(y)$ also at the endpoints of I_ν. Applying Fatou's lemma two more times then yields

$$\mathcal{E}(\hat{\psi}) = \int \Big(\int \hat{\psi}\,\mathrm{d}\pi_{x,\ell} - (\hat{\psi}^* \star \gamma_r)^*(x) \Big)\,\kappa(\mathrm{d}x, \mathrm{d}\ell) \le \liminf_{n\to\infty} \mathcal{E}(\hat{\psi}_n). \qquad \square$$

The following elementary lemma will be used for showing uniqueness of the dual optimizer.

Lemma 2.6. *Let $\gamma \in \mathcal{P}(\mathbb{R})$ be atomless and suppose that $\operatorname{supp}\gamma$ is an interval. Let $\psi, v : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous. Suppose that, for some $a, b \in \mathbb{R}$ and $\rho \in \mathcal{P}(\mathbb{R})$,*

$$\big(\partial_x \psi^*(a + \cdot)\big)_{\#}\gamma = \big(\partial_x v^*(b + \cdot)\big)_{\#}\gamma = \rho.$$

Then $\psi - v$ is affine on $I_\rho := \operatorname{int}\operatorname{co}(\operatorname{supp}\rho)$.

Proof. Set $\psi_a := \psi - a\,\mathrm{id}$ and $v_b := v - b\,\mathrm{id}$, so that $\psi_a^* = \psi^*(a+\cdot)$ and $v_b^* = v^*(b+\cdot)$. Since γ is atomless, the monotone transport from γ to ρ is γ-a.e. unique [42]. Hence $\partial_x \psi_a^* = \partial_x v_b^*$, γ-a.e. As $J := \operatorname{supp}\gamma$ is an interval, every set of full γ-measure is dense in J. Monotonicity of the derivatives of convex functions therefore yields $\psi_a^* = v_b^* + c$ on $\operatorname{int} J$ for some $c \in \mathbb{R}$.

Now fix $y \in I_\rho$. Then $\rho((-\infty, y)) > 0$ and $\rho((y, \infty)) > 0$. Since $\rho = (\partial_x \psi_a^*)_{\#}\gamma$, monotonicity of $\partial\psi_a^*$ yields some $z \in \operatorname{int} J$ such that $y \in \partial\psi_a^*(z)$. As $\psi_a^* = v_b^* + c$ on $\operatorname{int} J$, also $y \in \partial v_b^*(z)$. Fenchel equality therefore gives

$$\psi_a(y) + \psi_a^*(z) = yz = v_b(y) + v_b^*(z),$$

and hence $\psi_a(y) - v_b(y) = -c$. Thus $\psi(y) - v(y) = (a-b)y - c$ for every $y \in I_\rho$. $\square$

The shift in the R-Bass representation will be obtained from the subgradient of $(\psi^* \star \gamma_r)^*$. Whenever $\operatorname{supp}\gamma_r$ is either $\mathbb{R}$ or $[0, \infty)$ for κ-a.e. r, this subgradient is in fact a singleton on the relevant interval, as shown next.

Lemma 2.7. *Let $\gamma \in \mathcal{P}_1(\mathbb{R})$ have support either $\mathbb{R}$ or $[0, \infty)$, and let $\psi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be proper, convex and lower semicontinuous. Suppose that $(\psi^* \star \gamma)^*$ is proper. Then $(\psi^* \star \gamma)^*(z) \le \psi(z) - z\bar{\gamma}$ for every $z \in \mathbb{R}$ and, in particular, $\operatorname{dom}\psi \subset \operatorname{dom}(\psi^* \star \gamma)^*$. Moreover, $(\psi^* \star \gamma)^*$ is differentiable on $\operatorname{int}\operatorname{dom}\psi$.*

Proof. Since $(\psi^* \star \gamma)^*$ is proper, so is $\psi^* \star \gamma$. For every $a \in \mathbb{R}$, Jensen's inequality gives $(\psi^* \star \gamma)(a) \ge \psi^*(a + \bar{\gamma})$. Taking convex conjugates yields $(\psi^* \star \gamma)^*(z) \le \psi(z) - z\bar{\gamma}$ for every $z \in \mathbb{R}$. In particular, $\operatorname{dom}\psi \subset \operatorname{dom}(\psi^* \star \gamma)^*$.

Suppose that $(\psi^* \star \gamma)^*$ is not differentiable at some $x \in \operatorname{int}\operatorname{dom}\psi$, and choose $a_1 < a_2$ in $\partial(\psi^* \star \gamma)^*(x)$. By Fenchel–Moreau,

$$(\psi^* \star \gamma)(a_1) - a_1 x = -(\psi^* \star \gamma)^*(x) = (\psi^* \star \gamma)(a_2) - a_2 x.$$

Since $a \mapsto (\psi^* \star \gamma)(a) - ax$ is convex and attains its minimum at both a_1 and a_2, it is constant on $[a_1, a_2]$. Thus there exists $b \in \mathbb{R}$ such that

$$(\psi^* \star \gamma)(a) = xa + b, \qquad a \in [a_1, a_2]. \tag{13}$$

In particular,

$$\int \psi^*\Big(\frac{a_1 + a_2}{2} + z\Big)\,\gamma(\mathrm{d}z) = \int \Big(\frac{1}{2}\psi^*(a_1 + z) + \frac{1}{2}\psi^*(a_2 + z)\Big)\,\gamma(\mathrm{d}z).$$

By convexity, equality holds pointwise for γ-a.e. z, and hence ψ^* is affine on $[a_1+z, a_2+z]$ for every such z.

Set $J := a_1 + \operatorname{supp}\gamma$. The set of such z is dense in $\operatorname{supp}\gamma$, so the interiors of these intervals cover $\operatorname{int} J$. The affine representations agree on overlaps, and therefore ψ^* is affine on $\operatorname{int} J$ and, by lower semicontinuity, on J. Write $\psi^*(u) = su + d$ for $u \in J$. Since J is unbounded above, Fenchel–Moreau gives $\psi(y) \geq \sup_{u\in J}\{u(y-s)-d\} = +\infty$ for every $y > s$. Hence $\sup\operatorname{dom}\psi \leq s$. On the other hand, for every $u \in \operatorname{int} J$, we have $s \in \partial\psi^*(u)$ and hence $u \in \partial\psi(s)$, so $s \in \operatorname{dom}\psi$. It follows that $s = \sup\operatorname{dom}\psi$.

Since $x \in \operatorname{int}\operatorname{dom}\psi$, we have $x < s$. Moreover, $a + \operatorname{supp}\gamma \subset J$ for every $a \in [a_1, a_2]$, so

$$(\psi^* \star \gamma)(a) = \int \psi^*(a+z)\,\gamma(\mathrm{d}z) = sa + s\bar{\gamma} + d, \qquad a \in [a_1, a_2].$$

Comparing with (13) gives $s = x$, a contradiction. □

We now turn to the proof of Theorem 2.2. We first prove attainment of the relaxed dual problem and show that a dual optimizer ψ generates, through $g = \partial_x\psi^*$, an R-Bass martingale attaining the static value. The same structure also yields the converse: if an admissible R-Bass martingale is generated by an increasing function g, then the corresponding convex potential ψ with $g = \partial_x\psi^*$ is a dual optimizer.

Proof of Theorem 2.2. By Theorem 2.1, there exists $\hat{\pi} \in \mathcal{M}(\kappa,\nu)$ attaining the static problem. Together with Lemma 2.3, this gives

$$P(\kappa,\nu) = \int \mathrm{MCov}(\hat{\pi}_r, \gamma_r)\,\kappa(\mathrm{d}r) = \inf_{\substack{v\in L^1(\nu)\\ v \text{ cvx, lsc}}} \mathcal{E}(v). \tag{14}$$

We first prove dual attainment in (dBridge). Let $(\psi_n)_{n\in\mathbb{N}}$ be a minimizing sequence in (14), chosen according to the support convention. Since $\mathcal{E}$ is invariant under addition of affine functions by Lemma 2.4, we may normalize it as follows: Choose $s_n \in \partial\psi_n(\bar{\mu})$ and set $c_n(x) := \psi_n^*(s_n) - s_n x$ as well as $\hat{\psi}_n := \psi_n + c_n$. Then $\hat{\psi}_n \geq 0$, $\hat{\psi}_n(\bar{\mu}) = 0$, and $\mathcal{E}(\hat{\psi}_n) = \mathcal{E}(\psi_n)$.

To extract a converging subsequence, we use that (μ,ν) is irreducible. By [31, Proposition 3.3], there exists a martingale coupling π' of (μ,ν) whose conditional laws charge every open set charged by ν. We regard it as an element of $\mathcal{M}(\kappa,\nu)$ by setting $\pi'_{x,\ell} := \pi'_x$. Since $(\hat{\psi}_n^* \star \gamma_{x,\ell})^*(x) \leq \hat{\psi}_n(x) - x^2$, Lemma 2.4 gives

$$0 \leq \sup_n \int \Big(\int \hat{\psi}_n\,\mathrm{d}\pi'_x - \hat{\psi}_n(x) \Big)\,\mu(\mathrm{d}x) \leq \sup_n \mathcal{E}(\hat{\psi}_n) - \int x^2\,\mu(\mathrm{d}x) < \infty. \tag{15}$$

The tightness result [31, Lemma 3.9] states that normalized proper, convex and lower semicontinuous functions satisfying this bound are locally bounded on I_ν. After passing to a subsequence, we therefore obtain a dual potential ψ such that $\hat{\psi}_n \to \psi$ locally uniformly on I_ν. By Lemma 2.5,

$$\mathcal{E}(\psi) \leq \liminf_{n\to\infty} \mathcal{E}(\hat{\psi}_n) = \int \mathrm{MCov}(\hat{\pi}_r, \gamma_r)\,\kappa(\mathrm{d}r).$$

Conversely, Fenchel's inequality gives $\mathrm{MCov}(\hat{\pi}_r,\gamma_r) \leq \int \psi\,\mathrm{d}\hat{\pi}_r - \psi^C(r)$, and hence the reverse inequality after integration. Thus ψ attains (dBridge), and moreover

$$\mathrm{MCov}(\hat{\pi}_r, \gamma_r) = \int \psi\,\mathrm{d}\hat{\pi}_r - (\psi^* \star \gamma_r)^*(x) \tag{16}$$

for κ-a.e. $r = (x,\ell)$.

We next use (16) to construct the optimal R-Bass martingale. Since $x \in I_\nu \subset \operatorname{int}\operatorname{dom}\psi$, the subdifferential $\partial(\psi^* \star \gamma_r)^*(x)$ is nonempty. Choose measurably $a(r) \in \partial(\psi^* \star \gamma_r)^*(x)$. Fenchel equality gives

$$\mathrm{MCov}(\hat{\pi}_r,\gamma_r) = \int \big(\psi(y) - a(r)y\big)\,\hat{\pi}_r(\mathrm{d}y) + \int \psi^*(a(r)+z)\,\gamma_r(\mathrm{d}z).$$

Hence $\psi - a(r)\,\mathrm{id}$ is a dual optimizer of $\mathrm{MCov}(\hat{\pi}_r,\gamma_r)$. Let g be the left derivative of ψ^* on $J := \operatorname{int}\operatorname{dom}\psi^*$, extended monotonically to $\mathbb{R}$ by the values $-\infty$ and $+\infty$ to the left and right of J, respectively. The preceding equality implies that $\psi^*(a(r)+z)$ is finite for γ_r-a.e. z. By (A3), $a(r)+z$ therefore belongs to J and avoids the countable set of nondifferentiability points of ψ^* for γ_r-a.e. z. Equality in the preceding covariance duality thus yields, for κ-a.e. r,

$$\big(g(a(r)+\cdot)\big)_{\#}\gamma_r = \hat{\pi}_r. \tag{17}$$

Set $A := a(R_0)$ and $X_t := \mathbb{E}[g(A+Y_1) \mid A, R_t]$. By (17), $\mathrm{Law}(R_0, X_1) = \hat{\pi}$ and $g(A+Y_1) \in L^2$. Since $\int y\,\hat{\pi}_r(\mathrm{d}y) = x$, we have $X_0 = Y_0$ a.s., and $X_1 \sim \nu$. Thus $X \in \mathcal{A}(\kappa,\nu)$ and is an R-Bass martingale. Moreover, $\mathrm{Law}(X_1, Y_1 \mid R_0 = r)$ is the monotone coupling of $\hat{\pi}_r$ and γ_r, and hence

$$\mathbb{E}[X_1 Y_1] = \int \mathrm{MCov}(\hat{\pi}_r,\gamma_r)\,\kappa(\mathrm{d}r) = P(\kappa,\nu).$$

In particular, X attains (8).

Conversely, let $X \in \mathcal{A}(\kappa,\nu)$ be an R-Bass martingale, generated by g and $A = a(R_0)$, and set $\pi := \mathrm{Law}(R_0, X_1)$. Let h be the lower semicontinuous convex extension of a primitive of g on $J := \mathrm{int}\{g \in \mathbb{R}\}$, with $h = +\infty$ outside $\bar{J}$. By (A3) and finiteness of $g(A+Y_1)$, J is nonempty and $a(r)+z \in J$ for γ_r-a.e. z, for κ-a.e. r. Writing $y = g(a(r)+z)$, we have $y \in \partial h(a(r)+z)$ and hence

$$h^*(y) + h(a(r)+z) = (a(r)+z)y.$$

In particular, h^* is finite ν-a.e., and hence on I_ν. Define ψ to agree with h^* on $\bar{I}_\nu$ and to equal $+\infty$ elsewhere. Then ψ is a dual potential. Since $y \in \bar{I}_\nu$ almost surely, the preceding equality implies $\psi^*(a(r)+z) = h(a(r)+z)$ and

$$\psi(y) + \psi^*(a(r)+z) = (a(r)+z)y, \qquad y \in \partial\psi^*(a(r)+z). \tag{18}$$

Fix $r = (x,\ell)$ outside a κ-null set. The product $(a(r)+z)y$ is γ_r-integrable by the conditional second-moment bounds. Since ψ and ψ^* admit affine lower bounds, (18) shows that both terms on its left-hand side are integrable. Integrating the subgradient inequality and using $\int y\,\pi_r(\mathrm{d}y) = x$ gives

$$x \in \partial(\psi^* \star \gamma_r)(a(r)), \qquad \psi^C(r) = a(r)x - \int \psi^*(a(r)+z)\,\gamma_r(\mathrm{d}z).$$

Moreover, the increasing map $z \mapsto g(a(r)+z)$ induces the monotone coupling of γ_r and π_r. Integrating (18) therefore yields

$$\mathrm{MCov}(\pi_r,\gamma_r) = \int \psi\,\mathrm{d}\pi_r - \psi^C(r).$$

By Lemma 2.4, we may evaluate $\mathcal{E}(\psi)$ using this particular π. Hence

$$\int \mathrm{MCov}(\pi_r,\gamma_r)\,\kappa(\mathrm{d}r) = \mathcal{E}(\psi).$$

Since the static and dual values coincide, ψ is a dual minimizer and π is the static maximizer. Finally, conditionally on $R_0 = r$, the pair (X_1, Y_1) is precisely the monotone coupling above, and therefore

$$\mathbb{E}[X_1Y_1] = \int \mathrm{MCov}(\pi_r,\gamma_r)\,\kappa(\mathrm{d}r) = P(\kappa,\nu).$$

Thus X attains (8).

It remains to prove uniqueness of the dual minimizer under the additional assumption that $\mathrm{supp}\,\gamma_r$ is either $\mathbb{R}$ or $[0,\infty)$. Let v be another minimizer of (dBridge). By Lemma 2.7, $b(r) := \partial_x(v^* \star \gamma_r)^*(x)$ is well defined for κ-a.e. $r = (x,\ell)$. As above,

$$\big(\partial_x v^*(b(r)+\cdot)\big)_{\#}\gamma_r = \hat{\pi}_r = \big(\partial_x\psi^*(a(r)+\cdot)\big)_{\#}\gamma_r.$$

By Lemma 2.6, $\psi - v$ is affine on $I_r := \mathrm{int}\,\mathrm{co}(\mathrm{supp}\,\hat{\pi}_r)$. Fix $y \in I_\nu$. Irreducibility gives

$$0 < \int |z-y|\,\nu(\mathrm{d}z) - \int |z-y|\,\mu(\mathrm{d}z) = \int\Big(\int |z-y|\,\hat{\pi}_{x,\ell}(\mathrm{d}z) - |x-y|\Big)\kappa(\mathrm{d}x,\mathrm{d}\ell).$$

The integrand is nonnegative by Jensen's inequality, so it is strictly positive on a set of positive κ-measure. For every such r, $\hat{\pi}_r(-\infty,y) > 0$ and $\hat{\pi}_r(y,\infty) > 0$ as otherwise $z \mapsto |z-y|$ is affine on its support and Jensen's inequality is an equality. Therefore $y \in I_r$ for some r, and thus every $y \in I_\nu$ has a neighbourhood on which $\psi - v$ is affine. Since I_ν is an interval, $\psi - v$ is affine on I_ν. By convexity and lower semicontinuity, this identifies the two dual optimizers on $\mathrm{co}\,\mathrm{supp}(\nu)$ up to addition of an affine function. □

3 Multi-asset calibration with SKR

Let $R = (Y, L)$ be a reference process on $\mathbb{R}^d \times \mathsf{E}$, Markov with respect to the reference filtration $(\mathcal{F}_t)$, whose price coordinates $Y = (Y^1,\dots,Y^d)$ form a square-integrable martingale. We fix one calibration interval, rescale it to $[0,1]$, and start R with law $\kappa \in \mathcal{P}_2(\mathbb{R}^d \times \mathsf{E})$. As in the single-asset case, $(\mathcal{F}_t)$ is the usual augmentation of the natural filtration generated by R. Write $\mu^j := \mathrm{pr}^{X^j}_{\#}\,\kappa$ and let $\nu^1,\dots,\nu^d \in \mathcal{P}_2(\mathbb{R})$ be the prescribed terminal marginals.

The class $\mathcal{A}(\kappa,\nu^1,\dots,\nu^d)$ consists of square-integrable $\mathbb{R}^d$-valued martingales X on the same filtered probability space as R, satisfying $X_0 = Y_0$ and $X_1^j \sim \nu^j$ for every $j = 1,\dots,d$. We emphasize that only the 1-dimensional terminal marginals are prescribed, but not the joint law of X_1. As before, minimizing $\mathbb{E}[\sum_{j=1}^d [X^j - Y^j]_1]$ is equivalent to

$$P(\kappa,\nu^1,\dots,\nu^d) := \sup_{X \in \mathcal{A}(\kappa,\nu^1,\dots,\nu^d)} \mathbb{E}[X_1 \cdot Y_1]. \tag{multi-Bridge}$$

For $r = (x^1,\dots,x^d,\ell) \in \mathbb{R}^d \times \mathsf{E}$, write

$$\gamma_r := \mathrm{Law}(Y_1 \mid R_0 = r), \qquad \gamma_r^j := \mathrm{Law}(Y_1^j \mid R_0 = r).$$

Thus the reference law and its one-dimensional projections are conditioned on the full initial state.

For an admissible bridge X, the coupling $\pi^j := \mathrm{Law}(R_0, X_1^j)$ has disintegration $\pi^j(\mathrm{d}r, \mathrm{d}y) = \kappa(\mathrm{d}r)\pi_r^j(\mathrm{d}y)$ with

$$\int y\, \pi_r^j(\mathrm{d}y) = x^j \quad \text{for } \kappa\text{-a.e. } r.$$

We denote by $\mathcal{M}^j(\kappa, \nu^j)$ the set of couplings in $\mathrm{Cpl}(\kappa, \nu^j)$ satisfying this martingale property. For fixed κ, both the objective and the terminal constraints separate across assets. We therefore apply the single-asset construction to each coordinate, including the remaining reference price coordinates in the latent state, to obtain the following result.

Theorem 3.1 (Existence and uniqueness for the primal problem)**.** *Suppose that* (A1)*,* (A2) *and* (A3) *hold coordinatewise, and that $\mu^j \leq_{\mathrm{cvx}} \nu^j$ for every $j = 1, \dots, d$. Then*

$$P(\kappa, \nu^1, \dots, \nu^d) = \sum_{j=1}^{d} \max_{\pi^j \in \mathcal{M}^j(\kappa, \nu^j)} \int \mathrm{MCov}(\pi_r^j, \gamma_r^j)\, \kappa(\mathrm{d}r). \tag{multi-wBridge}$$

Every single-asset problem admits a unique optimizer $\hat{\pi}^j$. The dynamic problem admits an optimizer, which is unique in law.

Proof. Apply Theorem 2.1 to each single-asset problem, and denote its unique static optimizer by $\hat{\pi}^j$. Each coordinate of an admissible multi-asset bridge is admissible for the corresponding single-asset problem, so

$$P(\kappa, \nu^1, \dots, \nu^d) \leq \sum_{j=1}^{d} \int \mathrm{MCov}(\hat{\pi}_r^j, \gamma_r^j)\, \kappa(\mathrm{d}r). \tag{19}$$

On the common filtered probability space, define

$$T_r^j := F_{\hat{\pi}_r^j}^{-1} \circ F_{\gamma_r^j}, \qquad X_t^j := \mathbb{E}[T_{R_0}^j(Y_1^j) \mid \mathcal{F}_t], \qquad j = 1, \dots, d.$$

By (A3), $\mathrm{Law}(X_1^j \mid R_0 = r) = \hat{\pi}_r^j$. The barycenter constraints give $X_0 = Y_0$, and each coordinate has terminal law ν^j and attains its single-asset problem. Thus $X = (X^1, \dots, X^d)$ is admissible and attains the bound in (19).

Finally, let $\tilde{X}$ be another optimizer. It must attain every single-asset optimum, so $\mathrm{Law}(R_0, \tilde{X}_1^j) = \hat{\pi}^j$ for every j. Its conditional terminal coupling with Y_1^j must also be optimal. By (A3), this is the unique monotone coupling, and hence

$$\tilde{X}_1^j = T_{R_0}^j(Y_1^j) = X_1^j \quad \text{a.s.}, \qquad j = 1, \dots, d.$$

Since X and $\tilde{X}$ are martingales with respect to the same filtration, their càdlàg versions are indistinguishable. Hence, the optimizer is unique in law. □

We next characterize optimal bridges as multi-R-Bass martingales through the coordinatewise dual problems. Write $I_{\nu^j} := \mathrm{int}\,\mathrm{co}(\mathrm{supp}\,\nu^j)$. A dual potential for ν^j is a proper, convex and lower semicontinuous function v that is finite on I_{ν^j} and equal to $+\infty$ outside $\bar{I}_{\nu^j}$. For such a potential, set

$$(v^* \star \gamma_r^j)(a) := \int v^*(a + z)\, \gamma_r^j(\mathrm{d}z), \qquad v^{C,j}(r) := (v^* \star \gamma_r^j)^*(x^j).$$

For an irreducible pair (μ^j, ν^j), the corresponding relaxed objective is

$$\mathcal{E}^j(v) := \int \Big(\int v\, \mathrm{d}\pi_r^j - v^{C,j}(r) \Big)\, \kappa(\mathrm{d}r), \qquad \pi^j \in \mathcal{M}^j(\kappa, \nu^j).$$

By Lemma 2.4, $\mathcal{E}^j(v)$ is well-defined, independent of π^j and invariant under addition of affine functions.

Definition 3.2 (Multi-R-Bass martingale)**.** An admissible martingale $X \in \mathcal{A}(\kappa, \nu^1, \dots, \nu^d)$ is called a *multi-R-Bass martingale*, or a *separable R-Bass martingale*, if there exist increasing functions $g^1, \dots, g^d : \mathbb{R} \to [-\infty, \infty]$ and measurable shifts $a^j : \mathbb{R}^d \times \mathsf{E} \to \mathbb{R}$ such that

$$(g^j \star \gamma_r^j)(a^j(r)) = x^j \quad \text{for } \kappa\text{-a.e. } r,$$

and, writing $A^j := a^j(R_0)$ and $A := (A^1, \dots, A^d)$, we have

$$X_t^j = \mathbb{E}[g^j(A^j + Y_1^j) \mid \mathcal{F}_t] = \mathbb{E}[g^j(A^j + Y_1^j) \mid A, R_t], \qquad 0 \leq t \leq 1, \tag{20}$$

for every $j = 1, \dots, d$.

The shifts enforce the martingale property. As in the single-asset case, we write $a^j(r) = (g^j \star \gamma_r^j)^{-1}(x^j)$, where the inverse denotes a measurable choice of the inverse. The L^2 condition ensures that the convolution at this shift is absolutely convergent for κ-a.e. r. The second equality in (20) follows from the Markov property of the reference process. Although each generating function g^j is one-dimensional, its shift a^j may depend on all components of the initial state.

Theorem 3.3 (Duality and characterization of multi-R-Bass martingales)**.** *Suppose that* (A1), (A2) *and* (A3) *hold coordinatewise, and that* (μ^j, ν^j) *is irreducible for every* $j = 1, \dots, d$*. Then*

$$P(\kappa, \nu^1, \dots, \nu^d) = \sum_{j=1}^{d} \min_{v^j \text{ dual potential}} \mathcal{E}^j(v^j). \tag{multi-dBridge}$$

An admissible martingale solves (multi-Bridge) *if and only if it is a multi-R-Bass martingale.*

Moreover, suppose that $\operatorname{supp} \gamma_r^j$ *is either* $\mathbb{R}$ *or* $[0, \infty)$ *for* κ*-a.e.* r *and every* j*. Then each dual minimizer is unique up to addition of an affine function on* I_{ν^j}*.*

Proof. Apply Theorem 2.2 to each single-asset problem, and denote its static and dual optimizers by $\hat{\pi}^j$ and ψ^j. Together with Theorem 3.1, this gives

$$P(\kappa, \nu^1, \dots, \nu^d) = \sum_{j=1}^{d} \int \mathrm{MCov}(\hat{\pi}_r^j, \gamma_r^j)\, \kappa(\mathrm{d}r) = \sum_{j=1}^{d} \mathcal{E}^j(\psi^j).$$

This proves (multi-dBridge).

Let a^j be the shifts obtained in the single-asset construction and set $A := (a^1(R_0), \dots, a^d(R_0))$. On the common filtered probability space, define

$$X_t^j := \mathbb{E}[\partial_x(\psi^j)^*(A^j + Y_1^j) \mid \mathcal{F}_t], \qquad j = 1, \dots, d.$$

Each coordinate has terminal law ν^j and attains its single-asset problem. Thus $X = (X^1, \dots, X^d)$ is admissible and attains the bound in (19). The Markov property gives the representation (20), with $g^j = \partial_x(\psi^j)^*$.

An admissible bridge attains this bound precisely when each coordinate attains its single-asset optimum. The single-asset Bass characterization therefore gives the claimed equivalence.

Under the additional support assumption, uniqueness of each dual minimizer follows from Theorem 2.2. □

We observe that the optimizer has joint conditional terminal law

$$\mathrm{Law}(X_1 \mid R_0 = r) = (G_r)_{\#}\gamma_r, \qquad G_r(z^1, \dots, z^d) := \big(g^1(a^1(r) + z^1), \dots, g^d(a^d(r) + z^d)\big).$$

Thus the single-asset reduction does not impose independence and the assets are coupled through the reference process.

Coordinatewise martingale Sinkhorn iteration. For fixed κ, the martingale Sinkhorn iteration applies separately to each coordinate. For a shift $a^j : \mathbb{R}^d \times \mathsf{E} \to \mathbb{R}$, define

$$\rho^{a^j} := \int (a^j(r) + \cdot)_{\#}\gamma_r^j\, \kappa(\mathrm{d}r) = \mathrm{Law}(a^j(R_0) + Y_1^j). \tag{21}$$

Starting from shifts $a_0^1, \dots, a_0^d \in L^2(\kappa)$, the coordinate updates, whenever well defined, are

$$g_k^j := F_{\nu^j}^{-1} \circ F_{\rho^{a_k^j}},$$
$$a_{k+1}^j(r) := (g_k^j \star \gamma_r^j)^{-1}(x^j), \qquad j = 1, \dots, d.$$

The first update chooses the increasing transport to ν^j; the second enforces the martingale constraint. The coordinate updates can be performed in parallel, using the kernels γ_r^j conditioned on the full initial state, and no d-dimensional transport problem needs to be solved. The additional assumptions and normalizations for convergence are given in Section 4 and apply asset-wise.

Calibration across maturities. For maturities $0 = T_0 < T_1 < \cdots < T_n$, prescribe marginal laws in $\mathcal{P}_2(\mathbb{R})$ satisfying

$$\delta_{x_0^j} = \mu_0^j \leq_{\mathrm{cvx}} \cdots \leq_{\mathrm{cvx}} \mu_n^j, \qquad j = 1, \dots, d.$$

After calibration up to T_i, restart the reference dynamics from the full joint state law

$$\kappa_i := \mathrm{Law}(X_{T_i}, L_{T_i}).$$

On $[T_i, T_{i+1}]$, solve the multi-bridge problem with starting law κ_i and terminal marginals $\mu_{i+1}^1, \dots, \mu_{i+1}^d$. The joint evolution of the calibrated bridge and the reference latent process determines $\kappa_{i+1} = \mathrm{Law}(X_{T_{i+1}}, L_{T_{i+1}})$, which supplies the initial state law for the next step.

4 Convergence of the Martingale Sinkhorn algorithm

We next study convergence of the Martingale Sinkhorn algorithm. Throughout this section, we assume that $\mu \leq_{\mathrm{cvx}} \nu$ and that (μ, ν) is irreducible. In addition to (A1), (A2) and (A3), we impose the following assumptions:

(A4) For κ-a.e. r, $\operatorname{supp} \gamma_r$ is either $\mathbb{R}$ or $[0, \infty)$.

(A5) $\operatorname{supp} \nu$ is compact, and $\operatorname{supp} \mu \subset I_\nu = \operatorname{int} \operatorname{co} \operatorname{supp}(\nu)$.

Assumption (A5) gives the uniform bounds on the shifts needed in the convergence proof. For related estimates without compact support in the Brownian case, see [31]. For an irreducible pair (μ, ν), the additional condition $\operatorname{supp} \mu \subset I_\nu$, equivalently $\operatorname{co} \operatorname{supp} \mu \subset I_\nu$, is often referred to as *strict irreducibility*.

The iteration is as follows.

Algorithm 1 Martingale Sinkhorn algorithm

1: Choose an initial shift $a_0 \in L^2(\kappa)$ with $\mathbb{E}[a_0(R_0)] = 0$.
2: **for** $i = 1, 2, \ldots$ **do**
3: **Terminal-marginal update:**
Set $\rho_i := \operatorname{Law}\big(a_{i-1}(R_0) + Y_1\big)$, and choose a dual potential ψ_i, satisfying the support convention above, such that $(\partial_x \psi_i^*)_{\#} \rho_i = \nu$.
Normalize by $\psi_i \leftarrow \psi_i - \int (\psi_i^* \star \gamma_r)^*(\bar{\mu})\, \kappa(\mathrm{d}r)$.
4: **Martingale update:**
For $r = (x, \ell)$, set $a_i(r) := \partial_x (\psi_i^* \star \gamma_r)^*(x)$.
Normalize by $a_i(r) \leftarrow a_i(r) - \mathbb{E}[a_i(R_0)]$.
5: **end for**

The terminal-marginal update enforces the prescribed terminal law. In one dimension it is simply the monotone rearrangement from ρ_i to ν:

$$\partial_x \psi_i^* = F_\nu^{-1} \circ F_{\rho_i}.$$

Here we use $F_\nu^{-1}(0) := \min \operatorname{supp} \nu$ and $F_\nu^{-1}(1) := \max \operatorname{supp} \nu$. In particular, if $\operatorname{supp} \rho_i$ is bounded below, the rearrangement is extended constantly by $\min \operatorname{supp} \nu$ below that support. Thus this step is explicit once ρ_i is known. The potential ψ_i itself is determined only up to an additive constant. Since the implementation uses only $\partial_x \psi_i^*$, this constant has no effect on the iteration; the normalization in Algorithm 1 merely fixes a convenient representative. The martingale update then restores the martingale constraint. Write

$$b_i(r) := \partial_x (\psi_i^* \star \gamma_r)^*(x), \qquad c_i := \int b_i \, \mathrm{d}\kappa, \qquad a_i = b_i - c_i.$$

Before centering the shift, Fenchel–Moreau gives, for κ-a.e. $r = (x, \ell)$,

$$x = \int \partial_x \psi_i^*(b_i(r) + z)\, \gamma_r(\mathrm{d}z),$$

so the conditional law of $\partial_x \psi_i^*(b_i(r) + Y_1)$ given $R_0 = r$ has barycenter x.

The conditional laws depend on the shift and the potential through the map

$$z \longmapsto \partial_x \psi_i^*(b_i(r) + z).$$

Replacing b_i by $a_i = b_i - c_i$ subtracts c_i from the argument of this map. We compensate by replacing ψ_i with

$$\tilde{\psi}_i(y) := \psi_i(y) - c_i(y - \bar{\mu}),$$

which gives $\partial_x \tilde{\psi}_i^*(u) = \partial_x \psi_i^*(u + c_i)$. The two changes therefore cancel, leaving the conditional laws unchanged. The dual value is also unchanged because the correction to ψ_i is affine; see Lemma 2.4.

We are allowed to center the shifts because the dual objective is invariant under addition of affine functions. Indeed, if $c(y) = sy + d$, then $(\psi + c)^*(z) = \psi^*(z - s) - d$; see Lemma 2.4 and the proof thereof. Thus the potential paired with the centered shift is $\tilde{\psi}_i(y) := \psi_i(y) - c_i(y - \bar{\mu})$, since

$$\partial_x \tilde{\psi}_i^*(a_i(r) + z) = \partial_x \psi_i^*(b_i(r) + z).$$

The pairs $(\tilde{\psi}_i, a_i)$ and (ψ_i, b_i) therefore represent the same conditional laws, and affine invariance gives $\mathcal{E}(\tilde{\psi}_i) = \mathcal{E}(\psi_i)$. The algorithm retains ψ_i as the terminal-marginal potential and uses the centered shift a_i in the next iteration.

The main result of this section is the convergence of these iterates to the solution of the dual problem (dBridge).

Theorem 4.1 (Convergence of the Martingale Sinkhorn algorithm)**.** *Let (μ,ν) be irreducible. Suppose further that* (A1)*,* (A2)*,* (A3)*,* (A4) *and* (A5) *hold. Then $\mathcal{E}(\psi_i)\downarrow P(\kappa,\nu)$, and the potentials $(\psi_i)_i$ generated by Algorithm 1 converge locally uniformly on I_ν to a dual minimizer of* (dBridge)*.*

By Theorem 2.2, the dual minimizer is unique up to addition of an affine function. The normalizations in Algorithm 1 select a unique limiting representative, as shown in the proof below. We denote this normalized dual minimizer by ψ. Setting

$$g := \partial_x \psi^*, \qquad a(r) := \partial_x(\psi^* \star \gamma_r)^*(x), \quad r=(x,\ell),$$

this yields the R-Bass martingale $X_t = \mathbb{E}\big[g(a(R_0)+Y_1) \mid a(R_0), R_t\big]$, and hence the unique optimal bridge in law; see Theorem 2.1.

The main mechanism behind the convergence proof is the following observation: each Martingale Sinkhorn iteration strictly decreases $\mathcal{E}$ unless the current potential already attains (dBridge). Before proving this, we record the regularity properties of the iterates that will be used throughout the argument.

Lemma 4.2 (Regularity of the iterates)**.** *Suppose that the assumptions of Theorem 4.1 hold. Let*

$$L := \max\{|y| : y \in \operatorname{supp}\nu\}, \qquad m(r) := \int |y|\,\gamma_r(\mathrm{d}y), \qquad \bar m := \int m(r)\,\kappa(\mathrm{d}r).$$

Then ψ_i^* is L-Lipschitz for every i, and for $u \in I_\nu$ and κ-a.e. r,

$$\big|(\psi_i^* \star \gamma_r)^*(u) - \psi_i(u)\big| \le L m(r). \tag{22}$$

Moreover, there exist a compact set $K \subset I_\nu$ and $\delta > 0$, independent of i, such that for κ-a.e. r,

$$|a_i(r)| \le \frac{1}{\delta}\Big(4 \sup_{u\in K}|\psi_i(u)| + 2L(m(r)+\bar m)\Big). \tag{23}$$

Consequently, $a_i \in L^2(\kappa)$, $\rho_{i+1} \in \mathcal{P}_2(\mathbb{R})$ is atomless, and $\psi_i \in L^1(\nu)$, $\psi_i^C \in L^1(\kappa)$, and $\psi_i^*, \psi_{i+1}^* \in L^1(\rho_{i+1})$. In particular, $\mathcal{E}(\psi_i) < \infty$.

Proof. By the support convention, $\operatorname{dom}\psi_i \subset \operatorname{co}\operatorname{supp}\nu \subset [-L,L]$, so the conjugate ψ_i^* is L-Lipschitz. Hence, for every $a \in \mathbb{R}$ and κ-a.e. r,

$$\big|(\psi_i^* \star \gamma_r)(a) - \psi_i^*(a)\big| \le L m(r).$$

Taking conjugates gives (22). Since $\operatorname{supp}\mu$ is a compact subset of I_ν, choose $\delta > 0$ such that $K := \operatorname{supp}\mu + [-\delta,\delta] \subset I_\nu$. Set $M_i := \sup_{u\in K}|\psi_i(u)|$. By (22),

$$\big|(\psi_i^* \star \gamma_r)^*(u)\big| \le M_i + L m(r), \qquad u \in K.$$

In particular, taking $u = x$ shows that $\psi_i^C \in L^1(\kappa)$.

The shift $\tilde a_i(r) := \partial_x(\psi_i^* \star \gamma_r)^*(x)$ before normalization is well defined by Lemma 2.7. The subgradient inequality at $x \pm \delta$ gives

$$\delta|\tilde a_i(r)| \le 2M_i + 2Lm(r).$$

Since $a_i(r) = \tilde a_i(r) - \int \tilde a_i(e)\,\kappa(\mathrm{d}e)$ we obtain

$$\delta|a_i(r)| \le 4M_i + 2L\big(m(r)+\bar m\big),$$

which is (23). By Jensen's inequality and (A2), $m \in L^2(\kappa)$, and hence $a_i \in L^2(\kappa)$. It follows that $\rho_{i+1} = \operatorname{Law}(a_i(R_0)+Y_1) \in \mathcal{P}_2(\mathbb{R})$. Moreover, $\rho_{i+1}(\{z\}) = \int \gamma_r(\{z - a_i(r)\})\,\kappa(\mathrm{d}r) = 0$, so ρ_{i+1} is atomless.

Since ψ_j^* is L-Lipschitz,

$$|\psi_j^*(z)| \le |\psi_j^*(0)| + L|z|,$$

and therefore $\psi_i^*, \psi_{i+1}^* \in L^1(\rho_{i+1})$. By construction, $(\partial_x\psi_{i+1}^*)_\#\rho_{i+1} = \nu$, and Fenchel–Moreau yields

$$\int |\psi_{i+1}|\,\mathrm{d}\nu = \int \big|z\partial_x\psi_{i+1}^*(z) - \psi_{i+1}^*(z)\big|\,\rho_{i+1}(\mathrm{d}z) \le L\int |z|\,\rho_{i+1}(\mathrm{d}z) + \int |\psi_{i+1}^*|\,\mathrm{d}\rho_{i+1} < \infty.$$

Since $a_0 \in L^2(\kappa)$, the same argument applies to $\rho_1 = \operatorname{Law}(a_0(R_0)+Y_1)$ and gives $\psi_1 \in L^1(\nu)$. Thus $\psi_i \in L^1(\nu)$ for every i. Together with $\psi_i^C \in L^1(\kappa)$, this also shows that $\mathcal{E}(\psi_i) < \infty$. □

Lemma 4.3 (Strict descent)**.** *Under the assumptions of Theorem 4.1, $\mathcal{E}(\psi_{i+1}) \le \mathcal{E}(\psi_i)$ for every $i \in \mathbb{N}$. Moreover, equality holds if and only if ψ_i attains* (dBridge)*.*

Proof. We follow the argument of [31, Lemma 3.5]. Fix $i \in \mathbb{N}$. By Lemma 4.2, all integrals below are finite. Set

$$c_i := \int \partial_x(\psi_i^* \star \gamma_r)^*(x)\, \kappa(\mathrm{d}r), \qquad \tilde{\psi}_i(y) := \psi_i(y) - c_i(y - \bar{\mu}).$$

By construction of the normalized shift, $a_i(r) = \partial_x(\tilde{\psi}_i^* \star \gamma_r)^*(x)$ for κ-a.e. $r = (x, \ell)$, while affine invariance yields $\mathcal{E}(\tilde{\psi}_i) = \mathcal{E}(\psi_i)$. Hence Fenchel–Moreau and the definition of ρ_{i+1} give

$$\mathcal{E}(\psi_i) = \underbrace{\Big(\int \tilde{\psi}_i \,\mathrm{d}\nu + \int \tilde{\psi}_i^* \,\mathrm{d}\rho_{i+1}\Big)}_{\text{(I)}} - \underbrace{\int x a_i(r)\, \kappa(\mathrm{d}r)}_{\text{(II)}}.$$

Since $(\partial_x \psi_{i+1}^*)_{\#}\rho_{i+1} = \nu$,

$$\text{(I)} \geq \mathrm{MCov}(\nu, \rho_{i+1}) = \int \psi_{i+1} \,\mathrm{d}\nu + \int \psi_{i+1}^* \,\mathrm{d}\rho_{i+1}. \tag{24}$$

Moreover, Fenchel's inequality gives

$$\text{(II)} \leq \int \Big((\psi_{i+1}^* \star \gamma_r)(a_i(r)) + (\psi_{i+1}^* \star \gamma_r)^*(x)\Big)\, \kappa(\mathrm{d}r). \tag{25}$$

Because $\int \psi_{i+1}^* \,\mathrm{d}\rho_{i+1} = \int (\psi_{i+1}^* \star \gamma_r)(a_i(r))\, \kappa(\mathrm{d}r)$, combining (24) and (25) yields

$$\mathcal{E}(\psi_i) \geq \mathcal{E}(\psi_{i+1}).$$

Suppose that equality holds. Then equality holds in (24), so $\tilde{\psi}_i$ also attains the dual problem for $\mathrm{MCov}(\nu, \rho_{i+1})$. Since ρ_{i+1} is atomless, $(\partial_x \tilde{\psi}_i^*)_{\#}\rho_{i+1} = \nu$. On the other hand, by the definition of a_i and Lemma 4.2,

$$x = \partial_x(\tilde{\psi}_i^* \star \gamma_r)(a_i(r)) = \int \partial_x \tilde{\psi}_i^*(a_i(r) + z)\, \gamma_r(\mathrm{d}z)$$

for κ-a.e. $r = (x, \ell)$. Thus $\partial_x \tilde{\psi}_i^*$ and a_i generate an admissible R-Bass martingale, and Theorem 2.2 shows that $\tilde{\psi}_i$ attains (dBridge). Since $\mathcal{E}(\tilde{\psi}_i) = \mathcal{E}(\psi_i)$, ψ_i attains (dBridge) too. Conversely, if ψ_i attains (dBridge), then

$$P(\kappa, \nu) = \mathcal{E}(\psi_i) \geq \mathcal{E}(\psi_{i+1}) \geq P(\kappa, \nu),$$

and hence equality holds. □

We now prove convergence. The bound on the dual values $(\mathcal{E}(\psi_i))_i$ gives local boundedness of the normalized potentials $(\psi_i)_i$, while the difference between successive dual values controls their Fenchel gaps. This will identify every accumulation point of $(\psi_i)_i$ as a dual optimizer.

Proof of Theorem 4.1. The proof has three steps. We first obtain relative compactness of the iterates. We then use the vanishing descent gaps to identify every accumulation point as the appropriately normalized dual minimizer. Finally, we pass to the limit in the dual objective. By Lemma 4.3, $(\mathcal{E}(\psi_i))_{i \in \mathbb{N}}$ is decreasing and bounded below by $P(\kappa, \nu)$. Hence

$$\mathcal{E}(\psi_i) - \mathcal{E}(\psi_{i+1}) \longrightarrow 0. \tag{26}$$

We first prove relative compactness. Choose $p_i \in \partial\psi_i(\bar{\mu})$ and set $\varphi_i(y) := \psi_i(y) - \psi_i(\bar{\mu}) - p_i(y - \bar{\mu})$. Then $\varphi_i \geq 0$ and $\varphi_i(\bar{\mu}) = 0$. By [31, Proposition 3.3], there exists a martingale coupling π' of (μ, ν) whose conditional laws π'_x charge every open set charged by ν. Since affine functions cancel from the corresponding Jensen gap,

$$0 \leq \int \Big(\int \varphi_i \,\mathrm{d}\pi'_x - \varphi_i(x)\Big)\, \mu(\mathrm{d}x) = \int \psi_i \,\mathrm{d}\nu - \int \psi_i \,\mathrm{d}\mu \leq \mathcal{E}(\psi_i) - \int x^2 \,\mu(\mathrm{d}x).$$

The right-hand side is uniformly bounded by Lemma 4.3. The compactness result [31, Lemma 3.9] states that normalized convex functions satisfying such a uniform Jensen-gap bound are locally bounded on I_ν. Hence $(\varphi_i)_i$ is locally bounded there.

Retain the notation $L, m, \bar{m}$ from Lemma 4.2. By (22) and the additive normalization in Algorithm 1, $|\psi_i(\bar{\mu})| \leq L\bar{m}$. Moreover, applying (23) to the affine representatives φ_i, which induce the same normalized shifts a_i, gives

$$|a_i(r)| \leq C(1 + m(r)) \tag{27}$$

for some $C > 0$, all i, and κ-a.e. r. Thus the second moments of $(\rho_i)_i$ are uniformly bounded.

It remains to control p_i. Choose $y_- < \bar{\mu} < y_+$ with $\nu((-\infty, y_-]) > 0$ and $\nu([y_+, \infty)) > 0$. Since $(\partial_x \psi_i^*)_{\#}\rho_i = \nu$ and $\bar{\mu} \in \partial\psi_i^*(p_i)$, monotonicity gives

$$\rho_i((-\infty, p_i]) \geq \nu((-\infty, y_-]), \qquad \rho_i([p_i, \infty)) \geq \nu([y_+, \infty)).$$

The uniform second-moment bound on ρ_i therefore implies $\sup_i |p_i| < \infty$. Hence $(\psi_i)_i$ is locally bounded on I_ν, and convexity gives local equicontinuity. It is therefore relatively compact for locally uniform convergence.

The same estimates also give a uniform linear-growth bound

$$|\psi_i^*(z)| \le C + L|z|, \qquad i \in \mathbb{N},\ z \in \mathbb{R}. \tag{28}$$

Let $\psi_{i_k} \to \psi$ locally uniformly on I_ν. We extend ψ lower semicontinuously to $\bar{I}_\nu$ and set it equal to $+\infty$ elsewhere. By [47, Theorem 7.17], this implies epi-convergence on $\mathbb{R}$; convex conjugation preserves epi-convergence by [47, Theorem 11.34]. Since ψ^* is finite on $\mathbb{R}$, another application of [47, Theorem 7.17] yields $\psi_{i_k}^* \to \psi^*$ locally uniformly on $\mathbb{R}$.

As in the proof of Lemma 4.3, write

$$b_i(r) := \partial_x(\psi_i^* \star \gamma_r)^*(x), \qquad c_i := \int b_i(r)\,\kappa(\mathrm{d}r), \qquad a_i = b_i - c_i.$$

The subgradient estimate in the proof of Lemma 4.2, together with the local bounds above, gives $|b_i(r)| \le C(1+m(r))$. By (28), dominated convergence gives local uniform convergence $\psi_{i_k}^* \star \gamma_r \to \psi^* \star \gamma_r$ for κ-a.e. r. Conjugation therefore gives locally uniform convergence of $(\psi_{i_k}^* \star \gamma_r)^*$ to $(\psi^* \star \gamma_r)^*$ on I_ν. Hence Lemma 2.7 and [47, Exercise 12.40(b)] give

$$b_{i_k}(r) \longrightarrow b(r) := \partial_x(\psi^* \star \gamma_r)^*(x).$$

Dominated convergence therefore yields $c_{i_k} \to c := \int b\,\mathrm{d}\kappa$ and $a_{i_k} \to a := b - c$ in $L^2(\kappa)$. Consequently, with $\rho := \mathrm{Law}(a(R_0) + Y_1)$, $\mathcal{W}_2(\rho_{i_k+1}, \rho) \to 0$. The law ρ is atomless. Set $\tilde{\psi}_i(y) := \psi_i(y) - c_i(y - \bar{\mu})$ and $\tilde{\psi}(y) := \psi(y) - c(y - \bar{\mu})$. Then $\tilde{\psi}_{i_k} \to \tilde{\psi}$ locally uniformly and $a(r) = \partial_x(\tilde{\psi}^* \star \gamma_r)^*(x)$.

We now identify the limit using the two nonnegative gaps in the proof of Lemma 4.3. Denote by G_i the gap in (24) and by H_i the Fenchel gap in (25). That proof gives

$$\mathcal{E}(\psi_i) - \mathcal{E}(\psi_{i+1}) = G_i + H_i, \qquad G_i, H_i \ge 0.$$

Hence (26) implies $G_i, H_i \to 0$.

Let $\chi_k \in \mathrm{Cpl}(\nu, \rho_{i_k+1})$ attain $\mathrm{MCov}(\nu, \rho_{i_k+1})$. Passing to a further subsequence, assume $\chi_k \to \chi \in \mathrm{Cpl}(\nu, \rho)$ weakly. Since

$$G_{i_k} = \int \big(\tilde{\psi}_{i_k}(x) + \tilde{\psi}_{i_k}^*(y) - xy\big)\,\chi_k(\mathrm{d}x, \mathrm{d}y),$$

epi-convergence and Fatou's lemma, applied to the nonnegative Fenchel gaps on a Skorokhod representation of the weakly convergent couplings, imply $\tilde{\psi}(x) + \tilde{\psi}^*(y) = xy$ for χ-a.e. (x, y). Since ρ is atomless, $(\partial_x \tilde{\psi}^*)_{\#}\rho = \nu$. On the other hand, the definition of a and the regularity established in Lemma 4.2 give, for κ-a.e. $r = (x, \ell)$,

$$x = \partial_x(\tilde{\psi}^* \star \gamma_r)(a(r)) = \int \partial_x \tilde{\psi}^*(a(r) + z)\,\gamma_r(\mathrm{d}z).$$

Thus $\partial_x \tilde{\psi}^*$ and a generate an admissible R-Bass martingale. By Theorem 2.2, $\tilde{\psi}$, and hence ψ, attains (dBridge).

It remains only to identify the affine representative. We claim that $c_i \to 0$. Indeed, from any subsequence we may pass further so that $\psi_i \to u$ and $\psi_{i+1} \to v$ locally uniformly on I_ν. Along this subsequence, the preceding argument gives $a_i \to a$ κ-a.e. and in $L^2(\kappa)$, with $\int a\,\mathrm{d}\kappa = 0$. The same argument applies to ψ_{i+1} and gives locally uniform convergence of the convolutions and their conjugates. Since $H_i \to 0$, Fatou's lemma applied to its nonnegative integrand yields

$$0 \le \int \Big((v^* \star \gamma_r)(a(r)) + (v^* \star \gamma_r)^*(x) - x a(r)\Big)\,\kappa(\mathrm{d}r) \le \liminf_i H_i = 0.$$

Hence Fenchel equality and Lemma 2.7 give $a(r) = \partial_x(v^* \star \gamma_r)^*(x)$ for κ-a.e. $r = (x, \ell)$. The convergence of the raw shifts in $L^2(\kappa)$ therefore implies $c_{i+1} \to \int a\,\mathrm{d}\kappa = 0$ along this subsequence. Since the original subsequence was arbitrary, $c_{i+1} \to 0$, and thus $c_i \to 0$.

Thus every accumulation point ψ satisfies

$$\int (\psi^* \star \gamma_r)^*(\bar{\mu})\,\kappa(\mathrm{d}r) = 0, \qquad \int \partial_x(\psi^* \star \gamma_r)^*(x)\,\kappa(\mathrm{d}r) = 0.$$

The first identity follows from the normalization in Algorithm 1 and dominated convergence, while the second is $c = 0$. By Theorem 2.2, dual minimizers differ only by an affine function, and these two conditions fix one affine normalization. Hence all accumulation points coincide, so $(\psi_i)_i$ converges locally uniformly on I_ν.

Finally, let ψ denote this limit. Since $c_i \to 0$, the preceding argument gives $a_i \to a$ in $L^2(\kappa)$ and $\rho_{i+1} \to \rho$ in $\mathcal{W}_2$, where $a(r) = \partial_x(\psi^* \star \gamma_r)^*(x)$ and $\rho = \mathrm{Law}(a(R_0) + Y_1)$. Moreover, $\tilde{\psi}_i^* \to \psi^*$ locally uniformly on $\mathbb{R}$, with a uniform linear-growth bound inherited from (28). Consequently,

$$\int \tilde{\psi}_i^*\,\mathrm{d}\rho_{i+1} \longrightarrow \int \psi^*\,\mathrm{d}\rho.$$

Since $(\partial_x \psi^*)_{\#}\rho = \nu$, Fenchel–Moreau and the same growth bound give $\psi \in L^1(\nu)$ and

$$\int \psi \,\mathrm{d}\nu + \int \psi^* \,\mathrm{d}\rho = \mathrm{MCov}(\nu, \rho).$$

As $\int \tilde{\psi}_i \,\mathrm{d}\nu = \int \psi_i \,\mathrm{d}\nu$, the definition of G_i gives

$$\int \psi_i \,\mathrm{d}\nu = \mathrm{MCov}(\nu, \rho_{i+1}) - \int \tilde{\psi}_i^* \,\mathrm{d}\rho_{i+1} + G_i \longrightarrow \int \psi \,\mathrm{d}\nu,$$

using continuity of maximal covariance in $\mathcal{W}_2$ and $G_i \to 0$. Moreover, (22), local uniform convergence on the compact set $\operatorname{supp}\mu \subset I_\nu$, and dominated convergence give

$$\int (\psi_i^* \star \gamma_r)^*(x)\, \kappa(\mathrm{d}r) \longrightarrow \int (\psi^* \star \gamma_r)^*(x)\, \kappa(\mathrm{d}r).$$

Therefore $\mathcal{E}(\psi_i) \downarrow \mathcal{E}(\psi) = P(\kappa, \nu)$. □

5 Comparison with SLV models

We explain formally how the Bass structure of a short SKR bridge leads to the instantaneous calibration principle described in the introduction. Let $(\mu_t)_{0\le t\le T}$ be a smooth family of target marginals, and write σ_{Dup} for the corresponding Dupire diffusion coefficient. Consider a diffusion reference

$$\mathrm{d}Y_t = \sigma_R(t, Y_t, L_t)\,\mathrm{d}W_t,$$

with an autonomous factor L. For $h = T/N$, let X^h be the SKR model calibrated at the maturities $t_j = jh$. The factor and Brownian drivers are retained throughout the construction.

A short Bass bridge. Fix a calibration interval $[t, t+h]$ and write $x = X_t^h$, $l = L_t$, and $\mathbb{E}_t = \mathbb{E}[\cdot \mid \mathcal{F}_t]$. On this interval the reference price is restarted at x:

$$Y_t^{t,h} = x, \qquad \mathrm{d}Y_s^{t,h} = \sigma_R(s, Y_s^{t,h}, L_s)\,\mathrm{d}W_s.$$

Its increment $M_{t,h} = Y_{t+h}^{t,h} - x$ has conditional mean zero and, to first order, is given by $\sigma_R(t, x, l)(W_{t+h} - W_t)$.

Let $G_{t,h}$ be the increasing Bass map for this bridge. For the following informal calculation, assume that the map is smooth and invertible, with the spatial bounds needed for the short-time expansions below. Write $\xi_{t,h} = G_{t,h}^{-1}$ and $k_{t,h} = \xi'_{t,h}$. The terminal value of the bridge can then be written as

$$X_{t+h}^h = G_{t,h}\big(\xi_{t,h}(x) + b_{t,h}(x, l) + M_{t,h}\big), \qquad \mathbb{E}_t[X_{t+h}^h] = x. \tag{29}$$

Here the shift $b_{t,h}$ is chosen to preserve the conditional mean.

The reference fluctuations are of order $\sqrt{h}$. Applying the curved map $G_{t,h}$ changes their conditional mean by order h, which is compensated by the centering shift. The leading random fluctuation is therefore determined by the slope of $G_{t,h}$. Since $G'_{t,h}(\xi_{t,h}(x)) = 1/k_{t,h}(x)$, linearization of (29) suggests

$$X_{t+h}^h - X_t^h \simeq \frac{\sigma_R(t, X_t^h, L_t)}{k_{t,h}(X_t^h)}(W_{t+h} - W_t). \tag{30}$$

Thus the reciprocal inverse Bass slope plays the role of a leverage function.

Identification of the leverage. The bridge reproduces the next prescribed marginal exactly. To first order, this imposes the conditional variance constraint (4). Hence, if the joint laws and inverse slopes admit a sufficiently regular limit, (30) gives

$$\frac{\mathbb{E}[\sigma_R(t, x, L_t)^2 \mid X_t = x]}{k(t, x)^2} = \sigma_{\mathrm{Dup}}(t, x)^2.$$

Consequently, the candidate limiting dynamics are

$$\mathrm{d}X_t = \lambda(t, X_t)\sigma_R(t, X_t, L_t)\,\mathrm{d}W_t, \qquad \lambda(t, x) = \frac{1}{k(t, x)} = \frac{\sigma_{\mathrm{Dup}}(t, x)}{\sqrt{\mathbb{E}[\sigma_R(t, x, L_t)^2 \mid X_t = x]}}. \tag{31}$$

This is the SLV prescription (5). The reference volatility is evaluated at the calibrated price because each bridge starts there. Propagating the joint law of price and factor supplies the conditional distribution needed to update the leverage.

Accumulation of the errors. The martingale condition also explains why a local approximation can be useful over many steps. Once a smooth SLV target X is specified, put $f(t,x,l) = \lambda(t,x)\sigma_R(t,x,l)$ and consider the endpoint errors

$$\epsilon_j = X^h_{t_{j+1}} - X^h_{t_j} - f(t_j, X^h_{t_j}, L_{t_j})(W_{t_{j+1}} - W_{t_j}).$$

Each error has conditional mean zero given $\mathcal{F}_{t_j}$. Errors from different intervals are therefore orthogonal:

$$\mathbb{E}\left|\sum_{j=0}^{N-1} \epsilon_j\right|^2 = \sum_{j=0}^{N-1} \mathbb{E}|\epsilon_j|^2. \tag{32}$$

Moreover, within each interval the Bass interpolation is the conditional expectation of its terminal value. Endpoint estimates consequently control the intervening paths as well. Together with stability of the target dynamics, these observations reduce convergence to quantitative estimates for the centered bridge endpoints; see Proposition A.2.

Appendix A carries out this program for a specified SLV target with one autonomous real factor and constant correlation in $(-1,1)$. It compares the exact bridges with smooth centered maps whose inverse slopes are $1/\lambda(t_j,\cdot)$, and controls the difference through quantitative calibration. Under the coefficient and regular-start assumptions stated in Subsection A.1, Theorem A.1 yields

$$\mathbb{E}\sup_{0\le t\le T}|X^h_t - X_t|^2 \le C_T h.$$

This gives strong order one half on the common Brownian drivers. The general factor result requires the stated regular initial joint density. When the reference price coefficient is independent of the factor, Corollary A.4 gives the same rate for any square-integrable initial pair, including a deterministic initial price.

6 Numerical experiments

We compare Bass LV with the Heston, 3/2, and one-factor Bergomi SKR models, all calibrated to the same European option prices. We report the vanilla fits and the prices of three-year reverse cliquets and memory autocallables with quarterly observations.

6.1 Vanilla calibration

We fit European call prices corresponding to the SPX smiles of 23 June 2020 reported by Farkas, Ferrari and Ulrych [22], at maturities $(T_1,\dots,T_5) = (0.25, 0.5, 1, 2, 3)$ years. We work with forward-normalized prices $X_t = S_t/F_t$.

Reference calibration. We first fit the Heston, 3/2, and one-factor Bergomi reference models to these vanilla prices, obtaining the parameters in Table 1. For Bergomi, the price volatility is $\sqrt{V_t}$, where

$$V_t = \xi_0(t)\exp\Big(\eta Z_t - \frac{\eta^2}{4\kappa}\big(1 - e^{-2\kappa t}\big)\Big), \tag{33}$$

with $\mathrm{d}Z_t = -\kappa Z_t\,\mathrm{d}t + \mathrm{d}B_t$ and $Z_0 = 0$. The correlation between the price and volatility drivers is ρ. The initial forward variance ξ_0 is constant on each interval listed in Table 1. Its levels are obtained from successive differences of the total variance implied by log-contract prices, divided by the corresponding interval lengths. This curve is fixed when fitting the remaining Bergomi parameters.

Heston		3/2		Bergomi-1F	
κ	1.2484	κ	0.3600	κ	2.4971
θ	0.0988	θ	5.0050	η	3.6191
η	1.1175	η	7.6870	ρ	−0.8796
ρ	−0.8038	ρ	−1.0000		
v_0	0.1020	v_0	0.0877		

t [years]	$[0, 0.25)$	$[0.25, 0.5)$	$[0.5, 1)$	$[1, 2)$	$[2, 3]$
$\xi_0(t)$	0.10209961	0.11963007	0.09081965	0.08462666	0.08383247

Table 1: Fitted reference parameters and the piecewise-constant initial forward-variance curve for Bergomi. Forward-variance levels are annualized variances.

Marginal calibration. With the reference parameters fixed, we apply Algorithm 1 successively between calibration maturities, alternating the terminal-marginal and martingale updates. At each maturity, the joint law $\kappa_i = \mathrm{Law}(X_{T_i}, L_{T_i})$ is passed to the next calibration interval. Values at intermediate observation dates are obtained from the conditional-expectation representation (2). Only the five stated maturities impose marginal constraints.

Figure 5 reports the repriced calls in implied-volatility terms. All four models closely reproduce the target smiles over the displayed strike ranges.

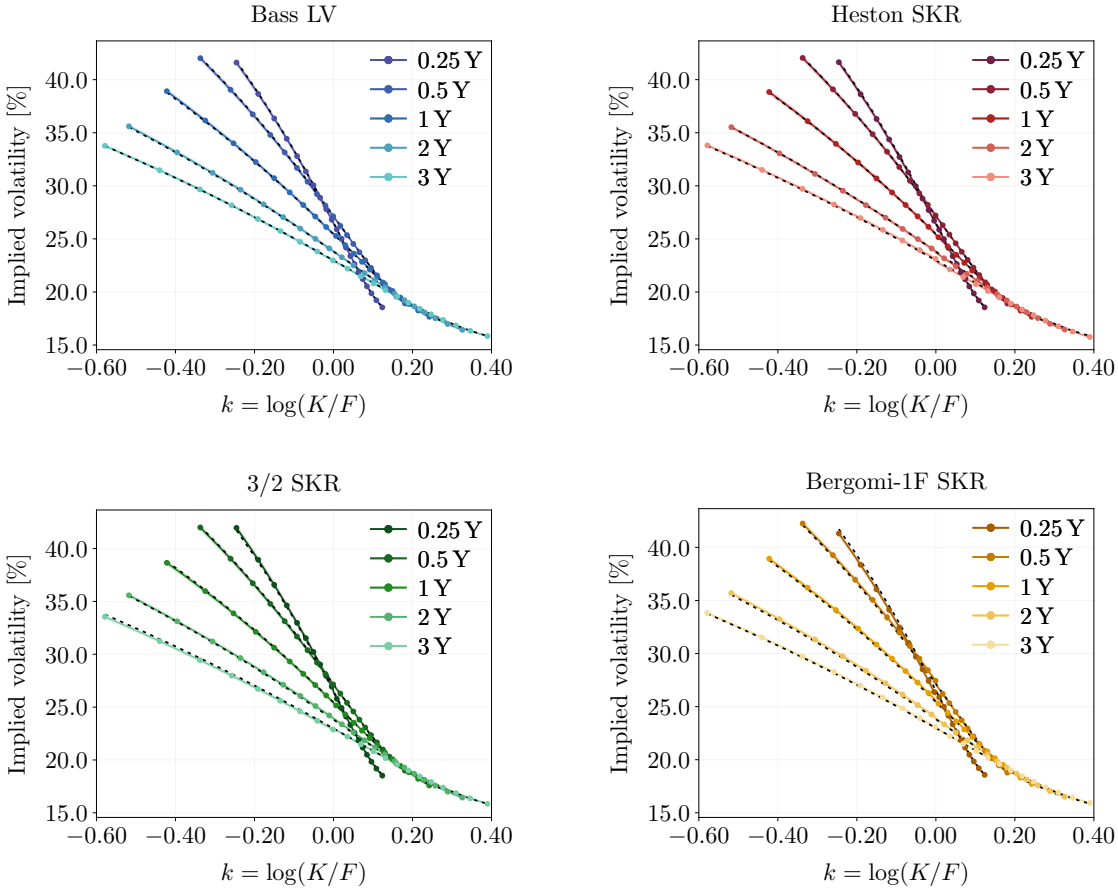


Figure 5: Fit of Bass LV and the three SKR models to the SPX implied-volatility smiles. Black dashed curves show the ESSVI targets; coloured curves with markers show implied volatilities recovered from Monte Carlo option prices under the calibrated models. Within each panel, darker to lighter shades indicate maturities 0.25, 0.5, 1, 2, and 3 years. Shading denotes pointwise 95% Monte Carlo confidence bands.

For both products below, we take $S_0 = 1$, a constant interest rate $r = 2\%$, and no dividends, with discount factor $D(t) = e^{-rt}$. All product prices use 10^5 Monte Carlo paths per model. Table 2 reports reference path simulation and fixed-point iteration separately, together with the pricing times for both products.

	Calibration [s]		Pricing [s]	
Model	Reference simulation	Fixed-point iteration	Cliquet	Autocall
Bass LV	0.0011	0.0801	0.0220	0.0219
Heston SKR	0.0334	0.1626	0.0757	0.0756
3/2 SKR	0.1421	0.1952	0.0753	0.0753
Bergomi-1F SKR	0.0313	0.9204	1.0074	1.0073

Table 2: Calibration and pricing times using 10^5 paths.[3]Pricing reuses the reference paths generated during calibration. Each product's pricing time includes evaluating the model price process at the observation dates and averaging the discounted payoffs.

6.2 Reverse cliquets

We consider a three-year reverse cliquet with notional $N = 100$ and quarterly observation dates $t_j = j/4$, $j = 0, \dots, 12$. Writing $R_j = S_{t_j}/S_{t_{j-1}} - 1$, its terminal payoff is

$$\Pi_T^{\text{cliq}} = N\Big[1 + \Big(C_0 - \sum_{j=1}^{12} \min\{(-R_j)^+, L\}\Big)^+\Big], \qquad T = t_{12}, \tag{34}$$

where C_0 is the aggregate coupon budget and $L > 0$ is the local loss cap, corresponding to a local return floor of $-L$. The price is $D(T)\mathbb{E}[\Pi_T^{\text{cliq}}]$. Table 3 reports prices and Monte Carlo standard errors for $C_0 = 50\%$ and $L = 7.5\%$.

[3]Computations were performed on an Intel Core i5-12500 processor with 32 GiB RAM under Debian 13.

Model	Price	MC SE
Bass LV	113.987	0.038
Heston SKR	122.000	0.039
3/2 SKR	116.991	0.035
Bergomi-1F SKR	120.144	0.036

Table 3: Three-year quarterly reverse cliquet with $N = 100$, $C_0 = 50\%$, and $L = 7.5\%$. MC SE denotes the Monte Carlo standard error.

Figure 6 varies L at fixed $C_0 = 50\%$. Each quarterly deduction is a put spread on R_j with strikes 0 and $-L$, so the comparison tests the downside part of the forward smile. Raising L increases deductions in quarters whose loss exceeds the cap and hence lowers the price. Once $12L > C_0$, the aggregate coupon floor also makes the price sensitive to how losses accumulate across quarters. At $L = 7.5\%$, the SKR models generate fewer negative quarters and lower expected clipped losses than Bass LV. The reduction in expected clipped losses explains their higher prices; the smaller aggregate-floor contribution partly offsets this effect.

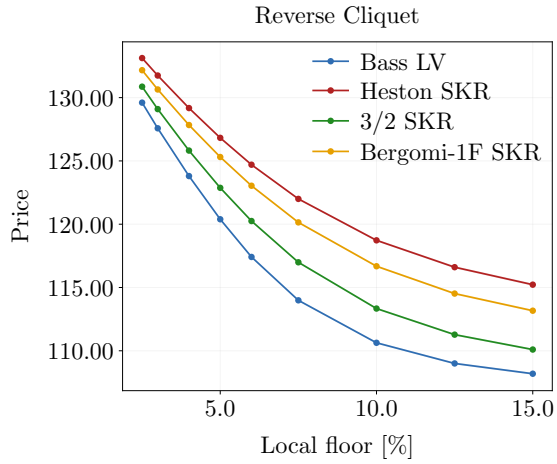


Figure 6: Local-floor sensitivity of the three-year quarterly reverse cliquet under Bass LV and the three SKR models, with $N = 100$ and $C_0 = 50\%$. The horizontal axis shows the loss cap L, corresponding to the return floor $-L$.

6.3 Memory autocallables

We consider memory autocallables with the same notional, maturity, and quarterly observations, and a quarterly coupon rate $c = 5\%$. While the note remains outstanding, the coupon Nc and all previously unpaid coupons are paid at an observation if $S_{t_j} \geq bS_0$, where b is the coupon barrier. The note redeems at par at the first quarterly observation before maturity for which $S_{t_j} \geq S_0$, including any coupon due on that date. The protection barrier $b_{\mathrm{P}} = 60\%$ is tested only at maturity; any coupons still unpaid then expire.

Let τ be the redemption date, equal to T if the note is not called, and let C_j, $j = 1, \ldots, 12$, be the coupon payment at t_j, including arrears. Writing $s_T = S_T/S_0$, the discounted payoff is

$$\Pi^{\mathrm{ac}} = \sum_{t_j \leq \tau} D(t_j) C_j + ND(\tau) - ND(T)(1 - s_T)\mathbf{1}_{\{\tau = T,\ s_T < b_{\mathrm{P}}\}}. \tag{35}$$

The final term is the capital loss for notes surviving to maturity below the protection barrier. The price is $\mathbb{E}[\Pi^{\mathrm{ac}}]$. Table 4 reports results for $b = 60\%$.

Model	Price	MC SE
Bass LV	104.899	0.069
Heston SKR	106.834	0.070
3/2 SKR	106.444	0.066
Bergomi-1F SKR	106.800	0.067

Table 4: Three-year quarterly memory autocallable with $N = 100$, coupon 5%, coupon barrier 60%, autocall barrier 100%, and terminal protection barrier 60%. Autocalling is possible from the first quarterly observation. MC SE denotes the Monte Carlo standard error.

Figure 7 varies b while keeping the autocall and protection barriers fixed. This isolates coupon-payment and recovery risk: raising b delays coupon payments or leaves more coupons unpaid at maturity. The sensitivity depends on reaching bS_0 while the note remains outstanding, particularly after missed coupons; varying b changes neither the redemption date nor the principal payoff on a given path.

Across models, price levels additionally depend on call times and terminal downside on uncalled paths. At $b = 60\%$, the higher SKR prices reflect larger expected discounted coupon receipts and smaller expected discounted capital losses, which outweigh the effect of later principal repayment.

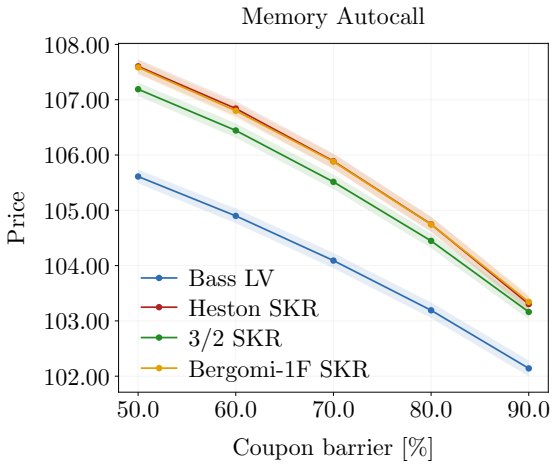


Figure 7: Coupon-barrier sensitivity of the three-year quarterly memory autocallable under Bass LV and the three SKR models, with $N = 100$ and quarterly coupon $c = 5\%$. The autocall and terminal protection barriers are 100% and 60% of S_0, respectively. Shading denotes pointwise 95% Monte Carlo confidence bands.

A Convergence to the stochastic local volatility model

This appendix makes the single-asset comparison in Section 5 quantitative. Existence and uniqueness of each bridge, together with its R-Bass representation, are supplied by Theorems 2.1 and 2.2; what remains is to control the effect of concatenating the bridges. Centered endpoint errors give a direct martingale criterion for path convergence. The quantitative argument controls those errors through the calibrated inverse slopes for a smooth, uniformly nondegenerate model with a regular initial law.

A.1 The model and the quantitative conclusion

Fix a finite horizon T and a deterministic constant $\rho \in (-1, 1)$. On a probability space carrying independent Brownian motions $W, W^{\perp}$, the reference model has dynamics

$$\mathrm{d}Y_t = \alpha(t, L_t)\vartheta(t, Y_t)\,\mathrm{d}W_t, \tag{36}$$

$$\mathrm{d}L_t = b(t, L_t)\,\mathrm{d}t + \beta(t, L_t)\big(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t^{\perp}\big). \tag{37}$$

Here the factor is real-valued. The target model and its price marginals are specified by

$$\mathrm{d}X_t = f(t, X_t, L_t)\,\mathrm{d}W_t, \qquad \sigma_R(t, x, l) = \alpha(t, l)\vartheta(t, x), \qquad f(t, x, l) = \lambda(t, x)\sigma_R(t, x, l), \qquad \mu_t = \mathrm{Law}(X_t). \tag{38}$$

The factor evolves autonomously and is shared by the reference and target. The initial pair (X_0, L_0) is square-integrable and independent of the Brownian increments. We write $\bar{\kappa}_t = \mathrm{Law}(X_t, L_t)$ for the target joint law, and use the completed filtration generated by the initial pair and the two Brownian motions.

The following coefficient assumptions are imposed for the SLV convergence statements: α, b, β are smooth on $[0, T] \times \mathbb{R}$, ϑ, λ are smooth on $[0, T] \times \mathbb{R}$, the functions and all their derivatives are bounded, and

$$0 < c \le \alpha, \beta, \vartheta, \lambda \le C < \infty.$$

Constants below may depend on these bounds, on T, and on the fixed correlation, but are independent of the number of resets. A fixed-order estimate uses only finitely many derivative bounds. The independent WMOT theorem in Subsection A.3 uses its own stated reference and base-density assumptions; it does not use the target leverage or the SLV marginal curve.

For $h = T/N$ and $t_j = jh$, let X^h be the SKR construction of the main text for the marginals $(\mu_{t_j})_{j=0}^N$. It propagates the incoming joint state law $\kappa_j = \mathrm{Law}(X^h_{t_j}, L_{t_j})$ and starts each reference interval at $Y_{t_j} = X^h_{t_j}$, retaining the factor L. Its initial pair agrees with that of the target, and all reference intervals use the same Brownian motions as the target. The bounded Lipschitz coefficients give continuous transition kernels in W_2 and finite second moments. Uniform nondegeneracy gives atomless reference price transitions with support $\mathbb{R}$. Moreover, successive target marginals are irreducible: conditional strict Jensen for $|X_{t_{j+1}} - k|$, using the full support of the target transitions, gives strict convex order at every $k \in \mathbb{R}$. These properties verify (A1), (A2) and (A3) at every reset, so Theorem 2.1 supplies the unique bridge. Irreducibility additionally gives its R-Bass representation by Theorem 2.2. This realizes the bridge on the common drivers with the conditional-expectation interpolation (2). No new existence argument is needed.

The general quantitative conclusion uses one additional *regular-start assumption*. Define

$$Z_w = \int_{\mathbb{R}} e^{-\sqrt{1+x^2}}\,\mathrm{d}x, \qquad w(x) = Z_w^{-1} e^{-\sqrt{1+x^2}}, \qquad \omega(l) = Z_w^{-1} e^{-\sqrt{1+l^2}}, \qquad \mathfrak{w}(x, l) = w(x)\omega(l).$$

The initial joint law is assumed to have a density $p_0 = \mathfrak{w}\mathsf{r}_0$, where r_0 is smooth, all its derivatives are bounded, and $c \le \mathsf{r}_0 \le C$; the density has total mass one. This is an additional assumption for the general quantitative theorem. The endpoint criterion and the autonomous-price result below also apply to the deterministic initial price of the main text. Constants in estimates using this assumption may also depend on these initial-density bounds and on finitely many derivatives of r_0.

Under this additional assumption, write $p(t, x, l)$ for the target joint density and

$$m_t(x) = \int_{\mathbb{R}} p(t, x, l)\,\mathrm{d}l, \qquad n_t(l) = \int_{\mathbb{R}} p(t, x, l)\,\mathrm{d}x.$$

In particular, $\mu_t(\mathrm{d}x) = m_t(x)\,\mathrm{d}x$. The weighted Feynman–Kac formula in Subsection A.12 establishes

$$p(t, x, l) = \mathfrak{w}(x, l)\mathsf{r}(t, x, l), \qquad 0 < c_T \le \mathsf{r} \le C_T, \tag{39}$$

with bounded space-time derivatives of every order.

Theorem A.1 (Strong convergence of SKR)**.** *Under the regular-start assumption, for all sufficiently small h,*

$$\mathbb{E} \sup_{0 \le t \le T} |X^h_t - X_t|^2 \le Ch. \tag{40}$$

Consequently, (X^h, L) converges to (X, L) in probability in the uniform path topology, and hence in distribution.

The rate is strong order one half. In particular, the theorem concerns bounded uniformly positive coefficients and a regular initial joint law; it does not supply the general stochastic-factor rate from a singular initial law. The role of that extra assumption is confined to the quantitative calibration argument. We first give the probabilistic reduction for an arbitrary square-integrable initial pair.

Structure of the proof and the techniques used. The argument has three parts. Centered endpoint comparisons first reduce path convergence to a bound on the calibrated inverse slopes. A standalone WMOT theorem then controls the joint-density error in the explicit weighted norm

$$\mathcal{N}_b(z)^2 = \sup_n w(n)^{-1} \int \frac{\chi_n(x)^2 z(x,l)^2}{\sigma_R(x,l)^2 p(x,l)}\, dx\, dl.$$

Here p is the known comparison density and σ_R^2 is the reference price variance at the beginning of the step. Finally, an explicit smooth comparison step applies that theorem along the SLV flow. The central cancellation is simple: inverse-variance weighting makes the highest-order scalar calibration term independent of the factor; the common terminal price marginal then makes that term vanish. A full two-dimensional heat interpolation implements this cancellation without replacing the transport of a rough factor density by a frozen factor approximation.

Probability and WMOT: from endpoints to paths.

- **Subsection A.1: model and scope.** The assumptions specify the regular initial law and the separated reference and shared autonomous factor used in the SLV application. The conclusion is a strong, uniform-path estimate on every fixed finite horizon. The later standalone WMOT theorem permits a more general reference diffusion; this does not enlarge the scope of the stated SLV convergence theorem.

- **Subsection A.2: a centered endpoint criterion.** A known smooth comparison slope produces a centered endpoint with an $O(h^2)$ squared Euler error. Martingale orthogonality and Doob's inequality accumulate the endpoint errors; Lipschitz-SDE estimates then give the maximal path bound. Comparing two centered inverse maps shows that a uniform slope error ε costs only $Ch\varepsilon^2$ per step. These are probabilistic estimates and an elementary monotone-map calculation.

- **The autonomous-price corollary in the same subsection.** A target continuation is then an admissible WMOT competitor. Optimality and a known uniformly convex potential bound its endpoint error by Ch^2. This short WMOT argument also covers deterministic initial price, without the regular-density assumption.

The standalone WMOT theorem: a density energy and its quantitative proof.

- **Subsection A.3: the law-level statement.** A reference experiment on $[0,1]$ represents physical duration δ: it is a known smooth nondegenerate diffusion with martingale price coordinate. A known smooth joint law and positive inverse slope generate its terminal price law. Nearby incoming laws have the same price marginal. The theorem controls their WMOT endpoints in the explicit variance-weighted density distance above, with amplification at most $1 + C\delta$ on a neighborhood of radius $\delta^{3/4}$, which contains the $O(\delta)$ errors needed later. The metric uses only the known density and reference variance; it requires no regularity assumption on an unknown optimizer.

- **Subsection A.4: endpoint calculus.** Rescaling the exact endpoint and calendar formulas gives C^2 bounds in elementary L^2 and L^∞ spaces. Scalar Gaussian averaging supplies a factor $h^{-1/4}$; a neighborhood of radius $h^{1/4}$ pays for it. The fixed price marginal cancels the sole unsmoothed term in the calendar equation. This is analytic work, using Gaussian bounds and product estimates without density derivatives.

- **Subsection A.5: the exact density derivative.** Tilting the endpoint experiment writes the incoming-density derivative as a conditional expectation plus a small mass-preserving error. The mean constraint expresses the slope derivative as a scalar calibration term. The associated known-kernel price commutator is exactly symmetric. These identities follow from the conditional endpoint formulas in physical coordinates, retaining the frozen initial shift label.

- **Subsection A.6: full heat jumps and centered kernels.** An auxiliary reversible diffusion retains the full known price–factor covariance. Its jump form is comparable to a two-dimensional Gaussian jump form, and a weighted Poincaré inequality gives a gap of order h. A Gaussian-midpoint estimate converts exact column cancellation and localized kernel bounds into an inverse-noise estimate. The diffusion and resampling interpretations are probabilistic; the Gaussian bounds, weighted inequalities and kernel estimates are quantitative analysis.

- **Subsection A.7: the scalar inverse.** A crossing-flux identity turns price calibration into an equation for the scalar slope. An occupation series for a Gaussian-mixture walk with positive killing probability constructs its inverse. The representation is probabilistic; the Gaussian parametrix and weighted Schur bounds that make it uniform are analytic.

- **Subsection A.8: exact calibration.** The scalar inverse and the Banach implicit-function theorem give the local calibrated branch on the $h^{1/4}$ neighborhood. This analytic construction permits signed densities and uses only the known smooth base and the explicit endpoint formulas.

- **Subsection A.9: the weighted heat cancellation.** A forward–backward heat interpolation connects a density perturbation to a price-calibrated endpoint. Differentiating its weighted squared norm twice gives a nonnegative principal part and a known first-order commutator. At the terminal time, the price-marginal constraint cancels the leading scalar calibration term. Integration by parts then yields an energy bound with output jump dissipation. Full Gaussian kernel comparison, cutoff identities and a global scalar correction transfer this bound to the exact

density derivative at cost $O(h)$. This is an analytic energy argument for a known probabilistic experiment; it neither differentiates an input density nor freezes its factor coordinate.

- **Subsection A.10: the nonlinear step.** Changing the slope leaves the reference terminal factor unchanged. This gives exact column cancellation in the first and second density derivatives. The centered-kernel estimate and integration of the C^2 second derivative along the signed branch give a quadratic remainder. Pairing it with the retained output dissipation and using Young's inequality produces a fourth-order energy error, hence a cubic norm error. A uniformly convex WMOT identity finally identifies positive calibrated endpoints as the unique optimizers. The only differentiability used here comes from the explicit scalar and endpoint formulas.

Application to the specified smooth SLV target.

- **Subsection A.11: smooth consistency.** A strong first-order endpoint expansion identifies the target generator. A first directional calendar expansion identifies the leading response to a smooth slope change. Gaussian derivative transfer establishes these estimates at known smooth inputs, in the finite orders needed by the scalar preconditioner and density norm. No regularity of an unknown calibrated endpoint is used.

- **Subsection A.12: bounds for the target.** Conjugating the target forward equation by the fixed density weight gives a Feynman–Kac formula with bounded coefficients and potential. Positivity gives the upper and lower density bounds; ordinary forward SDE variations give the required derivatives. This is a probabilistic representation after an elementary differential-operator calculation.

- **Subsection A.13: comparison and iteration.** An explicit pointwise slope correction gives a known smooth comparison step with joint and calibration defects $O(h^2)$. The existing local branch corrects its target at a cost $O(h^2)$, even at the actual rough incoming density. The triangle inequality combines these defects with stability. Adjacent norms compare by the pointwise ratio of the known weights $p_t\sigma_R(t)^2$, independently of the comparison slope. The standalone theorem therefore yields

$$e_{j+1} \le (1 + Ch)e_j + Ch^2.$$

  An $O(h)$ bootstrap closes for every fixed finite horizon, well inside the calibration neighborhood. The slope bound and the initial endpoint criterion then give the strong path rate.

Probability and WMOT supply the endpoint comparison, the conditional identities, the scalar occupation representation and the path closure. The quantitative proof also uses Gaussian kernel estimates, a weighted Poincaré inequality, a heat energy calculation, localization and the Banach implicit-function theorem. The regular-start and coefficient assumptions above remain part of the convergence statement.

A.2 From centered endpoint errors to path convergence

Write $x_j = X^h_{t_j}$, $\mathbb{E}_j = \mathbb{E}[\cdot \mid \mathcal{F}_{t_j}]$, $\Delta W_j = W_{t_{j+1}} - W_{t_j}$ and $f_j = f(t_j, x_j, L_{t_j})$. On the jth reference interval let $M_j = Y_{t_{j+1}} - x_j$ and $U_j = X^h_{t_{j+1}}$. Introduce the known comparison inverse slope and its centered endpoint

$$k^0_j(x) = 1/\lambda(t_j, x), \qquad V_j = G_j\big(\xi_j(x_j) + b_j + M_j\big), \qquad \mathbb{E}_j V_j = x_j,$$

where $\xi_j' = k^0_j$ and $G_j = \xi_j^{-1}$. This endpoint need not have the prescribed terminal marginal.

Only the known comparison map is differentiated in the following elementary estimate:

$$\mathbb{E}_j|V_j - x_j - f_j\Delta W_j|^2 \le Ch^2, \qquad \mathbb{E}_j[V_j - x_j - f_j\Delta W_j] = 0. \tag{41}$$

Indeed reference Itô isometry gives $\mathbb{E}_j|M_j - \sigma_R(t_j, x_j, L_{t_j})\Delta W_j|^2 \le Ch^2$ and $\mathbb{E}_j|M_j|^4 \le Ch^2$. Since $\mathbb{E}_j M_j = 0$, Taylor expansion of the smooth G_j first gives $|b_j| \le Ch$ and then the displayed squared error. Its conditional mean is exactly zero by construction.

Proposition A.2 (An endpoint criterion)**.** *For any square-integrable initial pair and its genuine Bass interpolation,*

$$\mathbb{E}\sup_{t\le T}|X^h_t - X_t|^2 \le C_T\left(h + \sum_{j<N}\mathbb{E}|U_j - V_j|^2\right). \tag{42}$$

Proof. The residuals $\epsilon_j = U_j - x_j - f_j\Delta W_j$ have $\mathbb{E}_j\epsilon_j = 0$ and $\mathbb{E}\epsilon_j^2 \le C\mathbb{E}|U_j - V_j|^2 + Ch^2$. The within-bridge process is $X^h_t = \mathbb{E}[U_j \mid \mathcal{F}_t]$. Thus its difference from the corresponding frozen-coefficient stochastic integral is the martingale

$$R_t = \sum_{i<j}\epsilon_i + \mathbb{E}[\epsilon_j \mid \mathcal{F}_t], \qquad t_j \le t \le t_{j+1}.$$

Orthogonality and Doob's inequality bound $\mathbb{E}\sup|R|^2$ by $4\sum_j \mathbb{E}\epsilon_j^2$. Subtract the target equation from

$$X^h_t = X_0 + \int_0^t f(t_{\lfloor s/h\rfloor}, X^h_{t_{\lfloor s/h\rfloor}}, L_{t_{\lfloor s/h\rfloor}})\,dW_s + R_t.$$

The target coefficient is globally Lipschitz, and the target price and common autonomous factor have squared increments Ch. Burkholder–Davis–Gundy and Gronwall prove (42). This uses the frozen-label conditional expectation, not a Markov assumption on current physical price and factor. Only unconditional bounds on the squared residuals are needed. □

Lemma A.3 (Stability of a centered Bass endpoint)**.** *Condition on an incoming state with price x. Let M have mean zero and second moment at most Ch. For $c \le k_i \le C$, $i = 0, 1$, put*

$$R_i(u) = \int_0^u k_i(x+s)\,ds, \qquad J_i = R_i^{-1}, \qquad U_i = x + J_i(b_i + M), \quad \mathbb{E}U_i = x.$$

Then

$$\mathbb{E}|U_1 - U_0|^2 \le Ch\|k_1 - k_0\|_\infty^2. \tag{43}$$

Proof. For $\varepsilon = \|k_1 - k_0\|_\infty$, the uniform inverse bounds give $|J_1(s) - J_0(s)| \le C\varepsilon|s|$. Each mean map $b \mapsto \mathbb{E}J_i(b+M)$ has all difference quotients bounded above and below by positive constants. Consequently $|b_i| \le C\mathbb{E}|M| \le C\sqrt{h}$. At b_0, the two means differ by at most $C\varepsilon\sqrt{h}$, so $|b_1 - b_0| \le C\varepsilon\sqrt{h}$. Hence $|U_1 - U_0| \le C\varepsilon(|M| + \sqrt{h})$, which proves (43). No spatial derivative of either slope is used. □

Corollary A.4 (Autonomous price volatility)**.** *If α is independent of the factor, the strong bound* (40) *holds for any square-integrable initial pair, including a deterministic initial price.*

Proof. Start the target price equation from x_j with the same Brownian increments, and let Z_j be its terminal value. Its price dynamics are autonomous, so its terminal law is $\mu_{t_{j+1}}$ and $\mathbb{E}_j Z_j = x_j$. It is therefore an admissible martingale endpoint for the static WMOT comparison.

Choose $\psi_j' = \xi_j$ and write $B_j(z,v) = \psi_j(z) - \psi_j(v) - \psi_j'(v)(z-v)$. The identity $\xi_j(V_j) = Y_{t_{j+1}} + \zeta_j(x_j, L_{t_j})$ and the conditional means show that, among admissible endpoints Z, $\mathbb{E}B_j(Z, V_j)$ is a fixed constant minus $\mathbb{E}[ZY_{t_{j+1}}]$. WMOT optimality therefore gives

$$\mathbb{E}|U_j - V_j|^2 \le C\mathbb{E}B_j(U_j, V_j) \le C\mathbb{E}B_j(Z_j, V_j) \le C\mathbb{E}|Z_j - V_j|^2 \le Ch^2.$$

Here the curvature bounds for the known ψ_j give the two quadratic comparisons; all terms are integrable by their quadratic growth. The final estimate follows from the centered Euler approximations for the smooth target and for V_j. Summation and Proposition A.2 prove the assertion. No density assumption is used in this corollary. □

The factor-dependent price case will use the centered-slope estimate after quantitative calibration. We now prove a standalone theorem for the short WMOT update that supplies that calibration bound.

A.3 A one-period WMOT calibration and stability theorem

We isolate the one-step result from the convergence problem. Its data are a reference experiment, a smooth incoming density and a smooth positive inverse Bass slope. The terminal price law is initially the one generated by this known base. The smooth target application uses an explicit comparison slope and a small prescribed-target correction in Subsection A.13. Neither a target SLV model nor a subdivision enters the statement.

The reference experiment on $[0,1]$**.** Fix the probability weights w, ω defined above and let $0 < \delta \le \delta_0$. Conditional on $(Y_0, L_0) = (x, l)$, the reference experiment is a two-dimensional diffusion on $[0,1]$ with generator

$$\delta\mathcal{L}_{\delta u}, \qquad \mathcal{L}_s = \tfrac{1}{2}A_s^R : D^2 + b_s^R\partial_l, \qquad 0 \le u \le 1. \tag{44}$$

The covariance matrix $A_s^R(x,l)$ and drift $b_s^R(x,l)$ are smooth with bounded derivatives in physical time and space, b^R is bounded, and $cI \le A^R \le CI$. The first coordinate has zero drift and is therefore a martingale. Write $\sigma_R^2 = (A^R)_{11}$. Thus δ is the physical duration represented by the unit interval. Time-translated data are permitted; arbitrary fast time dependence on u is not included. Neither a separated price volatility nor an autonomous second coordinate is required here.

The known base is a probability density $p = \mathfrak{w}r$ and a slope k, where $c \le r, k \le C$ and all derivatives of r, k are bounded. A *uniform class* means that these bounds and the coefficient bounds are fixed. The data may depend on δ without any regularity in δ, provided the bounds remain uniform. Only finitely many of these smooth bounds enter the theorem.

Set

$$\xi_k(x) = \int_0^x k(v)\,dv, \qquad G_k = \xi_k^{-1}.$$

There is a unique centering shift $b_{\delta,k}(x,l)$ such that

$$U^b = G_k\big(\xi_k(x) + b_{\delta,k}(x,l) + Y_1 - x\big), \qquad \mathbb{E}[U^b \mid x, l] = x. \tag{45}$$

In the notation of Section 2, the generator is $g = G_k = \partial_x \psi_k^*$, where $\psi_k' = \xi_k$ and hence $\psi_k'' = k$. The bounds on k make G_k globally finite and Lipschitz. The corresponding shift is $a_k(x,l) = \xi_k(x) + b_{\delta,k}(x,l) - x$, so that $U^b = G_k(a_k(x,l) + Y_1)$; this is the shift denoted by ζ_k below. With incoming density p, denote the density of (U^b, L_1) by $q^b = \mathcal{O}_\delta(p,k)$ and its price marginal by ν^b. The full joint reference kernel $(x,l) \mapsto \mathrm{Law}(Y_1, L_1 \mid x, l)$ is part of the data: it determines the joint price–factor output of the WMOT optimizer.

For a nonnegative incoming density $\widetilde{p}$ and terminal law $\nu \in \mathcal{P}_2(\mathbb{R})$, the WMOT problem can equivalently be written as

$$\sup\big\{\mathbb{E}[UY_1] :\ \mathrm{Law}(U) = \nu, \quad \mathbb{E}[U \mid Y_0, L_0] = Y_0\big\},$$

where the reference experiment has initial density $\widetilde{p}$ and the supremum is over couplings of U with that experiment. The martingale constraint and the prescribed reference law are fixed in this optimization.

Joint laws and the distance used below. Write $\gamma_b = p(x,l)\,dx\,dl$, $\mu = p^X$, and

$$\mathcal{P}_\mu = \{\gamma \in \mathcal{P}_2(\mathbb{R}^2) : \gamma^X = \mu\}.$$

Here γ denotes an incoming joint state law, corresponding to κ in Section 2. The WMOT update sends the incoming joint law to the joint law of the optimal price endpoint and the terminal reference factor:

$$\Phi_\delta : \mathcal{P}_\mu \longrightarrow \mathcal{P}_{\nu^b}.$$

Only its restriction to the neighborhood below is asserted. The compared incoming laws have the same price marginal; the factor law and the dependence between price and factor may change. The output factor law is the terminal second-coordinate law of the reference experiment with that full incoming joint law. It is independent of the choice of the endpoint slope. The base endpoint U^b is itself optimal for its generated target, as the convex-potential identity below verifies.

Fix $\chi \in C_c^\infty(\mathbb{R})$ with $0 \le \chi \le 1$, $\chi = 1$ on $[-1,1]$ and support in $[-2,2]$, and put $\chi_n(x) = \chi(x-n)$ for $n \in \mathbb{Z}$. For joint laws absolutely continuous with respect to γ_b, write $f_\gamma = d\gamma/d\gamma_b$ and define

$$d_b(\gamma, \bar\gamma)^2 = \sup_n \frac{1}{w(n)}\,\mathbb{E}_{\gamma_b}\left[\frac{\chi_n^2}{\sigma_R(0)^2}(f_\gamma - f_{\bar\gamma})^2\right]. \tag{46}$$

All functions in the expectation are evaluated at the incoming pair. This distance measures the squared relative-density error, locally in price, weighted by the reciprocal reference price variance. It is defined for arbitrary joint price marginals and is used at both endpoints of the update. Its formula uses only the known input law and reference variance; the optimizing slope does not enter.

Theorem A.5 (Stability of a short WMOT update)**.** *Fix a uniform class of known reference experiments and smooth bases as above. There are constants $C, \delta_* > 0$ such that, for $0 < \delta \le \delta_*$ and every incoming law $\gamma \in \mathcal{P}_\mu$, $\gamma \ll \gamma_b$, with $d_b(\gamma, \gamma_b) \le \delta^{3/4}$, the WMOT update exists uniquely and*

$$d_b(\Phi_\delta\gamma, \Phi_\delta\gamma_b) \le (1 + C\delta) d_b(\gamma, \gamma_b). \tag{47}$$

Thus a short regular WMOT update has amplification at most $1 + C\delta$ throughout a specified shrinking neighborhood. This neighborhood contains all errors of order δ. The estimate compares each perturbed input with the known base; it does not assert pairwise contraction on a fixed neighborhood.

The density norm behind the distance. For the proof, write the distance as $d_b(\gamma, \bar\gamma) = \mathcal{N}_b(p_\gamma - p_{\bar\gamma})$, where

$$\mathcal{N}_b(z)^2 = \sup_n w(n)^{-1} \int_{\mathbb{R}^2} \frac{\chi_n(x)^2 z(x,l)^2}{\sigma_R(0,x,l)^2 p(x,l)}\,dx\,dl. \tag{48}$$

For signed density differences, put

$$\|z\|_E = \sup_n \left\| \chi_n(x) \frac{z(x,l)}{w(x)\omega(l)} \right\|_{L^2(dx\,\omega(l)dl)}, \qquad \mathcal{Z} = \{z \in E : z^X = 0\},$$

Then $\mathcal{N}_b$ is a norm on E, uniformly equivalent to $\|\cdot\|_E$. No density derivative is required. The fixed-price constraint is used for calibration; the norm itself is defined on all E.

For a family of known bases b_a, parameterized by an interval, a uniform bound on $|\partial_a \log(p_a \sigma_{R,a}(0)^2)|$ gives

$$e^{-C|a-a'|}\mathcal{N}_{b_a}(z) \le \mathcal{N}_{b_{a'}}(z) \le e^{C|a-a'|}\mathcal{N}_{b_a}(z), \qquad z \in E. \tag{49}$$

The constant may additionally depend on these family bounds.

What the proof constructs. The proof first constructs the calibrated density map $\mathcal{S}_\delta$ on the larger signed neighborhood $\|z\|_E \le r\delta^{1/4}$. Its scalar inverse slope is measured in

$$\|v\|_{K_\delta} = \|v\|_{L^2_{\mathrm{uloc}}} + \delta^{1/4}\|v\|_\infty.$$

An auxiliary reversible diffusion keeps the full joint covariance. Its weighted heat interpolation has a lower bound on its second energy derivative, and terminal price calibration gives dissipation. The exact density channel differs by centered short-jump errors. Conservation of the terminal factor law places the quadratic remainder in the dual of the same jump energy. Pairing it with the linear output and applying Young's inequality gives

$$\mathcal{N}_b(\mathcal{S}_\delta(p+z) - \mathcal{S}_\delta p) \le (1 + C\delta)\mathcal{N}_b(z) + C\delta^{-1/2}\mathcal{N}_b(z)^3.$$

For $\mathcal{N}_b(z) \le \delta^{3/4}$, the last term is at most $C\delta\mathcal{N}_b(z)$. The physical distance is the explicit one in (48). Scalar branch differentiability is proved as part of the argument.

Subsections A.4–A.8 construct the scalar branch. The auxiliary heat geometry and linear dissipation are proved in Subsections A.5–A.9. The centered-kernel C^2 argument in Subsection A.10 completes the theorem. The subsequent consistency and SLV application are separate results.

A.4 Calculus for one endpoint experiment

We use physical time, writing $h = \delta$ on an interval $[t, t+h]$. The case $t = 0$ is the experiment on $[0, 1]$ in Theorem A.5; another t translates the coefficient data. Write $q^{x,l}_{t,h}(m, l')$ for the density of $(Y_{t+h} - x, L_{t+h})$ conditional on $(Y_t, L_t) = (x, l)$ and $r^{x,l}_{t,h}(m) = \int q^{x,l}_{t,h}(m, l')\, dl'$. The initial-price dependence is retained throughout. We use the reference generator $\mathcal{L}_t$ in (44). In the SLV application it specializes to

$$A^R = \begin{pmatrix} \sigma_R^2 & \rho\sigma_R\beta \\ \rho\sigma_R\beta & \beta^2 \end{pmatrix}, \qquad b^R = b, \qquad \sigma_R = \alpha\vartheta. \tag{50}$$

All one-period estimates below use only the general reference assumptions of the standalone theorem.

For the positive inverse slope k, put $R_k(x, u) = \int_0^u k(x + v)\, dv$. The centering shift is characterized by

$$\int u\, r^{x,l}_{t,h}(R_k(x, u) - b_k(x, l))k(x + u)\, du = 0. \tag{51}$$

If $J_{k,x}$ is the inverse of $u \mapsto R_k(x, u)$ and M has density $r^{x,l}_{t,h}$, this says $U = x + J_{k,x}(b_k + M)$ and $\mathbb{E}[U \mid x, l] = x$. Moreover,

$$\partial_b \mathbb{E}[J_{k,x}(b + M)] = \int r^{x,l}_{t,h}(R_k(x, u) - b)\, du \in [C^{-1}, c^{-1}]. \tag{52}$$

Thus the shift is unique. For a known smooth slope, the martingale identity $\mathbb{E}[M \mid x, l] = 0$ and Taylor's formula give

$$\|b_k\|_\infty \le Ch. \tag{53}$$

The exact joint endpoint density is

$$\mathcal{O}_{t,h}(p, k)(z, l') = k(z) \int p(x, l) q^{x,l}_{t,h} \left(\int_x^z k(v)\, dv - b_k(x, l), l' \right) dx\, dl. \tag{54}$$

It is linear in p, preserves mass and global price mean, and has exactly the reference experiment's terminal factor marginal, independently of k. Indeed, changing variables in the terminal price gives

$$\int \mathcal{O}_{t,h}(P, k)(u, l')\, du = \int P(x, l) q^L_{t,h}(x, l; l')\, dx\, dl, \qquad q^L_{t,h}(x, l; l') = \int q^{x,l}_{t,h}(m, l')\, dm.$$

The factor follows the reference path, even if its coefficients depend on the reference price. The endpoint is $G_k(Y_{t+h} + \zeta_k(x, l))$, where $\zeta_k = \xi_k(x) + b_k - x$. This initial label stays frozen within the step; the physical interpolation need not be Markov in price and factor alone.

Write $h = \delta$, $\varepsilon = \sqrt{h}$ and $\tau = h^{1/4}$. For the relative density and the scalar slope use

$$\|P\|_E = \sup_n \left\| \chi_n(x) \frac{P(x, l)}{w(x)\omega(l)} \right\|_{L^2(dx\, \omega(l) dl)}, \qquad \|v\|_{K_h} = \|v\|_{L^2_{\mathrm{uloc}}} + \tau\|v\|_\infty.$$

Thus K_h is L^∞ equipped with this h-dependent norm. No derivative of an input density or a slope is included. Slopes are measurable representatives bounded above and below; their primitives are absolutely continuous, strictly increasing bi-Lipschitz maps. Fix a known smooth base (p, k_0) with the usual positive weighted bounds. Throughout, density perturbations belong to the single space $\mathcal{Z} = \{z \in E : z^X = 0\}$. No factor-marginal constraint is required.

Define the normalized centered calendar by

$$J_h(P,k)(z) = h^{-1}\int P(x,l)\mathbb{E}\big[(U-z)^+ - (x-z)^+ - \mathbf{1}_{\{x>z\}}(U-x)\mid x,l\big]\,dx\,dl.$$

The subtracted linear term has conditional mean zero. Set $J_b = J_h(p,k_0)$ and $m_b = \mathcal{O}_{t,h}(p,k_0)^X/k_0$; keep this smooth base weight fixed when varying the density and slope. Choose once for the uniform class $a_* = (1+2\overline{v}^{\,2})^{-1}$, where $\overline{v}$ bounds σ_R/k on the permitted positive slope range, and put

$$\mathcal{F}_b(P,k) = (a_*I - h\partial_{xx})\frac{J_h(P,k)-J_b}{m_b},$$

where J_b is the known base calendar. The derivatives in this preconditioner are distributional; the identity used below identifies the result as a function in K_h.

Lemma A.6 (Scaled finite-step calculus)**.** *There is a fixed small $r>0$ such that on*

$$P\in p+\mathcal{Z},\qquad \|P-p\|_E + \|k-k_0\|_{K_h} < r\tau,$$

the endpoint map into E and the scalar map into K_h are C^2, and

$$\|D^j\mathcal{O}_h\| + \|D^j\mathcal{F}_b\| \le C\tau^{1-j},\qquad j=1,2.$$

The constants are independent of small h. Density arguments may be signed. The fixed incoming price marginal in the scalar assertion is essential.

Proof. Choose r small enough that $\|k-k_0\|_\infty \le \tau^{-1}\|k-k_0\|_{K_h} < r$ preserves uniform positive slope bounds. Only these bounds, not slope derivatives, will enter the kernel estimates in this proof.

The conditional mean equation. Use the scaled reference kernels

$$Q_h(x,l;r,\eta) = hq^{x,l}_{t,h}(\varepsilon r, l+\varepsilon\eta),\qquad \bar{Q}_h(x,l;r) = \varepsilon r^{x,l}_{t,h}(\varepsilon r),$$

which have uniform polynomial Gaussian majorants for the finitely many fast derivatives required below. Define

$$A_k(x,r) = \int_0^r k(x+\varepsilon s)\,ds,\qquad \beta_k = b_k/\varepsilon.$$

The exact mean equation is

$$\int r\,\bar{Q}_h(x,l;A_k(x,r)-\beta_k(x,l))k(x+\varepsilon r)\,dr = 0.$$

Its derivative in β is

$$\int \bar{Q}_h(x,l;A_k(x,r)-\beta_k(x,l))\,dr \in [C^{-1},c^{-1}].$$

The pointwise mean root satisfies $|\beta_k|\le C$: the scaled reference increment has bounded first moment and A_k^{-1} has uniform slope bounds. Implicit differentiation is therefore legitimate in L^∞ and yields the required parameter derivatives.

More precisely, for a fixed enlarged Gaussian averaging operator $\mathcal{G}_hf(x) = \int G(s)|f(x+\varepsilon s)|\,ds$, the explicit implicit formulas give, for $1\le j\le 2$,

$$\sup_l |D_k^j\beta_k[v_1,\dots,v_j](x,l)| \le C\sum_{i=1}^{j}\left(\prod_{a\ne i}\|v_a\|_\infty\right)\mathcal{G}_h|v_i|(x). \tag{55}$$

Indeed, every slope direction is a fixed translation or an integral of fixed translations. Put all directions except one in L^∞; Gaussian Fubini bounds the remaining direction by the displayed average. Products of lower mean-shift derivatives have the same form, and division by the positive mean derivative preserves it. This is a zeroth-order averaged-direction estimate. No spatial derivative of β_k, k or a density is taken.

Three elementary kernel estimates. The endpoint formula after $x = u-\varepsilon r$ and $l = l'-\varepsilon\eta$ contains only fixed translations of P and of the slope directions. Its kernel is bounded by a fixed Gaussian in (r,η) because $|A_k(x,r)-\beta_k| \ge c|r| - C$. Weight ratios are absorbed by Gaussian moments. The same description holds after parameter differentiation, with polynomial displacement factors and the averages in (55).

For a signed density Z, a kernel of this type has operator norm $C : E\to E$ and price-supremum norm $C\tau^{-1} : E\to L^\infty_x(L^2_l)$ for the relative output. To verify the latter, first use the weighted factor Schur estimate at fixed incoming and outgoing prices. The remaining price kernel is a one-dimensional Gaussian at scale ε; its L^2 norm is $C\varepsilon^{-1/2} = C\tau^{-1}$.

Cauchy–Schwarz in price proves the claim. Marginal integration in the factor has the same bound. No factor-uniform supremum of Z is used.

Consequently a term with one arbitrary density factor Z and j scalar directions obeys

$$\|\text{term}\|_E + \tau\|\text{price marginal}/w\|_\infty \le C\|Z\|_E \prod_i \|v_i\|_\infty \le C\tau^{-j}\|Z\|_E \prod_i \|v_i\|_{K_h}. \tag{56}$$

For the fixed known density p, use $|p| \le Cw\omega$ and put one of the scalar directions in L^2_{uloc}. Formula (55), Gaussian translation estimates and the bounded relative density give instead

$$\|\text{term}\|_E + \tau\|\text{price marginal}/w\|_\infty \le C\tau^{1-j} \prod_i \|v_i\|_{K_h}, \qquad j \ge 1. \tag{57}$$

For the supremum part of (57), put all directions in L^∞ and use its prefactor τ. For the E part, the one L^2 direction may occur at the terminal price, at an integrated price, or in a mean-shift average; each case is bounded by the same Gaussian translation estimate.

These estimates also cover the calendar kernel J_h. Its centered payoff has the exact scaled form

$$B(v,r) = (r-v)^+ - (-v)^+ - r\mathbf{1}_{\{v<0\}}, \quad 0 \le B \le |r|, \quad B = 0 \text{ unless } |v| \le |r|.$$

The payoff and incoming-price Jacobian cancel the normalization by h. Parameter derivatives act on the scaled density, not on B; only Gaussian moments are added. The same estimates cover $h\partial_x J_h$. One can differentiate the payoff once: the conditional distribution function minus the incoming indicator is Gaussian-localized. After one or more slope derivatives, the ordinary chain rule gives fast reference-density derivatives times phase increments of size ε; their powers cancel. The extra factor h ensures that this kernel has no larger operator bound than the preceding ones. This argument differentiates the strike payoff, not a rough slope.

The exact preconditioner and the rough-density supremum. Put $s_b = (\log m_b)'$. The distributional identity $hJ_h'' = \mathcal{O}_h^X - P^X$ gives exactly

$$(a_* I - h\partial_{xx})\frac{J_h}{m_b} = (a_* + h(s_b' - s_b^2))\frac{J_h}{m_b} + 2hs_b\frac{\partial_x J_h}{m_b} - \frac{\mathcal{O}_h^X - P^X}{m_b}.$$

All parameter derivatives of P^X vanish on the fixed-price affine space. The other terms are covered by (56)–(57), including their scaled output supremum. This cancellation is why the assertion does not hold on an unrestricted E density domain.

Finally decompose $P = p + (P-p)$. A pure jth slope derivative at p has the bound (57). The term with $P-p$ has the bound (56), and its size $r\tau$ improves τ^{-j} to $r\tau^{1-j}$. A full jth parameter derivative containing a density direction has only $j-1$ slope directions, because the maps are linear in the density, and (56) gives the same τ^{1-j} bound. This proves the estimates for $1 \le j \le 2$.

The formulas are Fréchet differentiable: their kernels are smooth functions of bounded slope and mean-shift parameters. One further fixed fast derivative supplies a common polynomial Gaussian majorant for differences of the differentiated kernels. Since convergence in K_h implies uniform convergence at fixed h, the same Schur bounds give operator-norm continuity of the second derivatives. No Gaussian derivative estimate for an arbitrary rough-slope endpoint kernel is assumed; only fast derivatives of the fixed smooth reference kernel were used. □

A.5 Density variation in the tilted endpoint experiment

Fix a known smooth positive step with endpoint $U = G_k(Y_{t+h} + \zeta(x,l))$ and conditional mean $\mathbb{E}[U \mid x,l] = x$. Put

$$a(u) = 1/k(u), \qquad \bar{a}(x,l) = \mathbb{E}[a(U) \mid x,l].$$

Tilt the joint initial/terminal law by $a(U)$. Its raw marginal densities are

$$r_- = p\bar{a}, \qquad r_+ = p^+/k, \qquad p^+ = \mathcal{O}_{t,h}(p,k). \tag{58}$$

Let $C : L^2(r_+) \to L^2(r_-)$ be conditional expectation given the initial pair under the normalized tilted law. Its adjoint C^* is the reverse conditional expectation. Put $Q_+ = I - C^*C$.

For an inverse-potential variation q with slope variation $v = q'$, differentiating the conditional mean gives

$$\delta U = a(U)(Cq - q(U)), \qquad \delta_k p^+ = \partial_u(r_+ Q_+ q).$$

Define

$$D = r_+^{-1}\partial_u(r_+ \cdot), \qquad K_x = DQ_+ - Q_+\partial_u.$$

Thus the full physical density derivative is exactly

$$D\mathcal{S}_h[p]z = \mathcal{O}_{t,h}(z,k) + r_+(Q_+ v + K_x q). \tag{59}$$

Here the first term is linear in z. The calibrated output has zero price marginal. The initial shift label stays frozen throughout.

The commutator K_x is self-adjoint in $L^2(r_+)$ and kills constants, because $D^* = -\partial_u$. If $\mathcal{J}_h^G$ is the symmetric physical conductance of C^*C, its physical kernel is

$$(\partial_x r_+)(y)\delta_y(dy') - (\partial_x + \partial_{x'})\mathcal{J}_h^G(y, y')\, dy'.$$

The diagonal gives exact zero row mass. The known slow Gaussian bounds below control its off-diagonal part and hence its symmetric form by a product of two Gaussian jump seminorms.

Comparison of the known endpoint densities. We record the smooth-base estimates independently of the later second-order consistency lemma:

$$p^+ - p = O(h)\mathfrak{w}, \qquad \bar{a} - 1/k = O(h), \qquad r_+ - r_- = O(h)\mathfrak{w}. \tag{60}$$

They hold with each required finite number of relative spatial derivatives. The centered mean equation and the reference Itô formula give bounded $\partial_h b_{t,h}$ at a known smooth slope. Horizon differentiation here holds the coefficient data, p and k fixed; it imposes no regularity on their permitted dependence on δ. Apply the terminal forward equation once to the exact endpoint kernel. Terminal derivatives transfer by integration by parts to the smooth input and to simultaneous translation derivatives of the reference kernel; the price-phase coefficient $k + b_x$ is bounded away from zero. The Gaussian bounds give $\partial_h \mathcal{O}_{t,h}(p, k) = O(\mathfrak{w})$ with the required fixed derivatives. Integrating from $h = 0$ proves the first bound. Taylor's formula for $a(U)$, using the exact conditional mean and $\mathbb{E}[|U - x|^2 \mid x, l] = O(h)$, proves the second. The same differentiated mean and kernel formulas give its fixed derivatives. The definitions of $r_\pm$ prove the third.

A.6 An auxiliary diffusion and its jump energy

For the known smooth base put

$$r_\circ = p/k, \qquad \Gamma = \mathsf{D}_k A_t^R \mathsf{D}_k, \qquad \mathsf{D}_k = \operatorname{diag}(k^{-1}, 1).$$

All coefficients in this subsection are frozen at the physical start of the step. The auxiliary diffusion has generator $-H$, where

$$H = -\frac{1}{2r_\circ} \operatorname{div}(r_\circ \Gamma \nabla), \qquad P_s = e^{-sH}, \qquad Q_0 = I - P_{2h}. \tag{61}$$

Its drift is $\operatorname{div}(r_\circ\Gamma)/(2r_\circ)$. It is smooth and bounded, and Γ is uniformly positive. The semigroup is reversible in $L^2(r_\circ)$; the finite invariant measure $r_\circ$ need not have mass one. Write $\|s\|_{Q_0}^2 = \langle s, Q_0 s\rangle_{r_\circ}$. For a centered function F, its inverse-noise norm is

$$\|F\|_{Q_0^{-1}} = \sup_{\substack{\int s r_\circ = 0 \\ \|s\|_{Q_0} \le 1}} |\langle F, s\rangle_{r_\circ}|.$$

Let G_s be the centered Gaussian density with covariance sI on $\mathbb{R}^2$. The known reference Gaussian estimates proved in Subsection A.7 apply also to this auxiliary diffusion. Its jump form satisfies

$$\|s\|_{Q_0}^2 \asymp \iint r_\circ(z) G_h(z' - z) |s(z') - s(z)|^2 \, dz \, dz'. \tag{62}$$

Any fixed positive Gaussian covariance can be used on the right. Indeed, split a larger displacement into a fixed number of equal pieces and telescope. After translating the integration point, the bounded logarithmic score of $r_\circ$ costs at most an exponential of the displacement, absorbed by the Gaussian. Two-sided heat-kernel bounds then compare the result with the actual jump form.

The product exponential weight has a Poincaré inequality: the one-dimensional weight is comparable to $e^{-|x|}$, for which the weighted Hardy inequality on each half-line gives $\int |f - f(0)|^2 e^{-|x|} \le 4 \int |f'|^2 e^{-|x|}$; tensorization gives the product inequality. Bounded comparison transfers it to $r_\circ$, and uniform ellipticity gives a positive spectral gap for H. The spectral theorem therefore gives

$$\|s\|_{Q_0}^2 \ge ch \|s\|_{L^2(r_\circ)}^2, \qquad \int s r_\circ = 0. \tag{63}$$

In particular multiplication by a known bounded smooth function a satisfies

$$\|as\|_{Q_0} \le C\big(\|s\|_{Q_0} + \sqrt{h}\|s\|_{L^2(r_\circ)}\big). \tag{64}$$

This is just the product rule for a Gaussian jump and its second moment. A price-only function is still a function on both coordinates; the full factor and cross covariance remain in P_s and Q_0.

The tilted operator C from the preceding subsection also has a centered leading Gaussian of covariance $h\Gamma$, and C^*C has leading covariance $2h\Gamma$. To see this directly, its Lebesgue kernel is

$$C((x,l),(u,l')) = \frac{q^R_{t,h}((x,l),(\xi(u)-\zeta(x,l),l'))}{\bar{a}(x,l)}.$$

The terminal tilt cancels the transport Jacobian. Expand the known smooth phase at x and use (60). The remainder is $O(\sqrt{h})$ times a broader Gaussian. The adjoint has the same leading term because the density ratio and change of freezing point cost $O(|z-z'|+h)$. Composition gives the asserted Gibbs kernel. The required simultaneous spatial derivatives have the same bounds. In particular its symmetric physical conductance has bounded slow price derivatives. These are estimates only at the known base.

Centered short-jump kernels. Let $X=(x,l)$ and $Y=(y,\ell)$. Suppose a signed physical density kernel $D(X,Y)$, allowing kernel measures and diagonal terms, has zero column mass and the exponential Schur bounds

$$\int e^{a|X-Y|^2/h}|D(X,Y)|\,dY \le V r_\circ(X), \qquad \int e^{a|X-Y|^2/h}|D(X,Y)|\,dX \le V r_\circ(Y) \tag{65}$$

for a fixed $a>0$. Then

$$\left\| r_\circ^{-1}\int D(\cdot,Y)f(Y)\,dY \right\|_{Q_0^{-1}} \le CV\|f\|_{L^2(r_\circ)}. \tag{66}$$

Indeed, the Schur bounds imply the jump estimate

$$\iint |D(X,Y)|\,|s(X)-s(Y)|^2\,dX\,dY \le CV\|s\|^2_{Q_0}.$$

To prove it, insert a Gaussian midpoint Z with density $G_{bh}(Z-X)$, where b is sufficiently large. The first leg uses the row bound and native Gaussian jump comparison. For the second,

$$G_{bh}(Z-X) \le CG_{2bh}(Z-Y)e^{|X-Y|^2/(2bh)},$$

so the column bound applies when $1/(2b)\le a$. Column cancellation then writes the pairing with a test as

$$\iint D(X,Y)[s(X)-s(Y)]f(Y)\,dX\,dY.$$

Cauchy–Schwarz proves (66).

A.7 A localized scalar calibration inverse

Let K_h be the reversible scalar Gibbs operator in the derivative of the price calibration equation, with invariant density $m(x)\asymp w(x)$. Its symmetric conductance is $\mathcal{J}_h(x,y)=m(x)\mathsf{K}_h(x,\mathrm{d}y)/\mathrm{d}y$. For $(I-\mathsf{K}_h)\varphi=f$, integration and symmetry give the exact flux identity

$$-\int_{-\infty}^{x} m(y)f(y)\,\mathrm{d}y = \int_{\mathbb{R}} \mathcal{T}_h(x,z)\varphi'(z)\,\mathrm{d}z, \tag{67}$$

where

$$\mathcal{T}_h(x,z) = \int_{u\le\min(x,z),\ v\ge\max(x,z)} \mathcal{J}_h(u,v)\,\mathrm{d}u\,\mathrm{d}v.$$

The derivative in x has a jump $-m(x)$ at $z=x$. Thus the second derivative contains exactly $-m(x)\delta_{x=z}$. Define

$$B_h v(x) = \frac{1}{hm(x)}\int \mathcal{T}_h(x,z)v(z)\,\mathrm{d}z, \quad L_h = a_*I - h\partial_{xx}, \quad D_h = L_hB_h.$$

The diagonal jump gives $D_h = I+\mathcal{R}_h$ with a localized integral remainder. This exact singular coefficient is essential.

This is the derivative needed for calibration. Indeed $m=r_+^X$ and $I-\mathsf{K}_h=\Pi_+Q_+\Pi_+$ on price-only functions, where Π_+ denotes conditional expectation given price under r_+. For a slope direction $v=q'$, the density identity gives $\delta\nu=\partial_x[m(I-\mathsf{K}_h)q]$. Integrating with the call boundary conditions and using (67) yields $\delta J_h=-mB_hv$. Consequently

$$D_k\mathcal{F}_b = -D_h.$$

This derivation precedes, and does not assume, a solution of the scalar calibration equation.

A.7.1 Exact conductance and scalar density remainder

Let $\xi' = k$, let $\theta(x,l) = \xi(x) + b_k(x,l)$, and abbreviate the incoming density by p. The terminal-tilted initial/price-end measure has density

$$p(x,l) r_{t,h}^{x,l}(\xi(u) - \theta(x,l))\,\mathrm{d}x\,\mathrm{d}l\,\mathrm{d}u.$$

Its initial marginal is $p\bar{a}$ and its scalar terminal marginal is $m = p^X_{\mathrm{out}}/k$. Thus the exact Gibbs conductance is

$$\mathcal{J}_h(u,v) = \int_{\mathbb{R}^2} \frac{p(x,l)}{\bar{a}(x,l)} r_{t,h}^{x,l}(\xi(u) - \theta(x,l)) r_{t,h}^{x,l}(\xi(v) - \theta(x,l))\,\mathrm{d}x\,\mathrm{d}l. \tag{68}$$

This formula involves no terminal factor and no change of variables through the mean shift. In particular its denominator has not been canceled by a translation-invariance argument.

Let γ_v denote the centered scalar Gaussian of variance v. Uniformly in the initial pair,

$$r_{t,h}^{x,l}(m) = h^{-1/2}\left[\gamma_{\sigma_R(t,x,l)^2}(m/\sqrt{h}) + \sqrt{h}\, e_{t,h}^{x,l}(m/\sqrt{h})\right]. \tag{69}$$

For $j = 0, 1$ and every fixed number of slow initial-pair derivatives,

$$|\partial_z^j \partial_{x,l}^\alpha e_{t,h}^{x,l}(z)| \le C_\alpha (1 + |z|)^{M_\alpha} e^{-c_\alpha z^2}. \tag{70}$$

Here an initial-price derivative fixes the increment argument.

We give the kernel argument used both here and in Subsection A.6. Standard two-sided density and first-gradient bounds for this uniformly elliptic class are also given by [43, Theorem 1.2]; the simultaneous-translation estimates needed here are explained explicitly below. Freeze the full two-dimensional covariance at the terminal point in a Gaussian parametrix Z_s. Its residual, including time variation and drift, satisfies

$$|(\mathcal{L} - \mathcal{L}_{\mathrm{frozen}}) Z_s| \le C s^{-1/2} G_{cs}(z' - z).$$

A coefficient difference is $O(|z - z'| + s)$, a second Gaussian derivative costs s^{-1}, and its spatial factor supplies $\sqrt{s}$. Gaussian convolution and the beta integral bound the convergent series $Z * \sum_{n\ge1} H^{*n}$ by $C\sqrt{h}\, G_{ch}$. For a first fast derivative place an initial-coordinate derivative on the outer Gaussian. The resulting time convolution has the integrable singularities $s^{-1/2}(h-s)^{-1/2}$ and is bounded by $C G_{ch}$.

Simultaneous translation of initial and terminal coordinates differentiates smooth coefficients but not the fast coordinate difference. Every differentiated coefficient difference retains first-order vanishing, so the same estimates hold after any fixed number of simultaneous derivatives. A terminal fast derivative is a simultaneous derivative minus an initial one; price translation invariance is neither available nor used. For the scalar marginal write the terminal factor as $l + v$ before integration. Initial-factor differentiation is then simultaneous translation in that coordinate too. Replacing terminal by initial freezing costs the same remainder. Gaussian integration in $v \in \mathbb{R}$ proves (69)–(70). Higher fixed fast derivative estimates follow by splitting the time interval and differentiating separate Gaussian factors, or recursively from the semigroup identity. All constants are global on both axes.

Insert this expansion in (68). On its Gaussian scale,

$$\xi(u) - \theta(x,l) = k(u)(u - x) + O((u-x)^2 + h), \qquad \bar{a}(x,l) = 1/k(x) + O(h).$$

Freezing all slow coefficients at u and convolving the two leading Gaussians in the incoming price gives

$$\frac{\mathcal{J}_h(u, u + \sqrt{h}\, z)}{m(u)} = h^{-1/2}[j_u(z) + \sqrt{h}\, E_h(u,z)]. \tag{71}$$

Here j_u is the mixture of centered Gaussians with variances $2\sigma_R(t,u,l)^2/k(u)^2$, and mixing law $p(u,l)\,\mathrm{d}l/p^X(u)$. The remainder, its first fast derivative, and all required simultaneous slow derivatives have the bounds (70). The factor mixing law is comparable to ω. Normalization gives $m = p^X/k + O(h)w$ in relative smooth norms. Gaussian upper and lower bounds give $m \asymp w$, and symmetry supplies the estimates with endpoints exchanged.

Only one fast derivative is needed. Off the diagonal,

$$\partial_x^2 \mathcal{T}_h(x,y) = \begin{cases} \displaystyle\int_{v \ge y} \partial_1 \mathcal{J}_h(x,v)\,\mathrm{d}v, & x < y, \\ \displaystyle -\int_{u \le y} \partial_2 \mathcal{J}_h(u,x)\,\mathrm{d}u, & x > y. \end{cases} \tag{72}$$

The diagonal derivative jump is exactly $-m(x)$. If $s_m = m'/m$, the full kernel is

$$D_h = a_* \frac{\mathcal{T}_h}{hm} - \frac{\partial_x^2 \mathcal{T}_h}{m} + 2 s_m \frac{\partial_x \mathcal{T}_h}{m} - (s_m^2 - s_m') \frac{\mathcal{T}_h}{m}. \tag{73}$$

Its singular part is exactly the identity. Integrating the Gaussian estimates in (72) proves that the exact regular kernel differs from its frozen regular kernel by E_h^D with

$$\sup_x \int e^{\eta|x-y|/\sqrt{h}} |E_h^D(x,y)| \, \mathrm{d}y + \sup_y \int e^{\eta|x-y|/\sqrt{h}} |E_h^D(x,y)| \, \mathrm{d}x \le C\sqrt{h} \tag{74}$$

for some $\eta > 0$, and the same bound after every required simultaneous derivative $\partial_x + \partial_y$. In (73), the last two terms have norms $O(\sqrt{h})$ and $O(h)$ respectively. In the first two, coefficient displacement under the crossing integrals has Gaussian size $\sqrt{h}$ and gives the remaining error. This proves the differentiated comparison directly from a first, not a second, fast derivative of the heat-kernel remainder.

Lemma A.7 (Scalar localized pressure inverse)**.** *At a known smooth base, D_h^{-1} is a bounded multiplication operator plus an integral kernel with uniformly bounded exponentially weighted Schur norm on scale $\sqrt{h}$. In particular, it acts uniformly on L^∞ and L^2_{uloc} for sufficiently small h.*

Proof. We construct the inverse using a Gaussian-mixture walk with a uniform probability of killing. The small exact-kernel remainder will be handled in the same real-variable kernel norm.

Use the centered Gaussian mixture j_x from (71). Its component variances are $2v^2$, with $0 < \underline{v} \le v \le \overline{v} < \infty$. Its crossing profile is

$$t_x(s) = \int_{|s|}^\infty (u - |s|) j_x(u) \, \mathrm{d}u.$$

For a centered Gaussian density γ_{2v^2} and $s \ge 0$, Gaussian integration by parts gives

$$\int_s^\infty (u-s)\gamma_{2v^2}(u) \, \mathrm{d}u = 2v^2 \gamma_{2v^2}(s) - s \int_s^\infty \gamma_{2v^2}(u) \, \mathrm{d}u.$$

Taking mixtures and using Fubini consequently gives

$$0 \le t_x(s) \le 2\overline{v}^{\,2} j_x(s), \qquad \underline{v}^{\,2} \le \int_{\mathbb{R}} t_x(s) \, \mathrm{d}s = \frac{1}{2} \int_{\mathbb{R}} u^2 j_x(u) \, \mathrm{d}u \le \overline{v}^{\,2}. \tag{75}$$

Thus the nonnegative kernel

$$\mathcal{K}_h(x,y) = h^{-1/2} \left[j_x\left(\frac{y-x}{\sqrt{h}}\right) - a_* t_x\left(\frac{y-x}{\sqrt{h}}\right) \right]$$

has row mass at most $1 - \epsilon$, where $\epsilon = a_* \underline{v}^{\,2} > 0$. It is the transition kernel of a walk that is killed with probability at least ϵ at every step.

For an integral kernel K write

$$\|K\|_{\eta,h} = \max\left\{ \sup_x \int e^{\eta|x-y|/\sqrt{h}} |K(x,y)| \, \mathrm{d}y, \sup_y \int e^{\eta|x-y|/\sqrt{h}} |K(x,y)| \, \mathrm{d}x \right\}.$$

These norms are submultiplicative under kernel composition. The Gaussian bounds give an $O(\eta)$ change from the unweighted row mass. For a column integral, replacing the profile frozen at $x = y - \sqrt{h}\, s$ by the one frozen at y costs $O(\sqrt{h})$, by its bounded slow derivatives and Gaussian tails. Hence

$$\|\mathcal{K}_h\|_{\eta,h} \le 1 - \epsilon + C\eta + C\sqrt{h}. \tag{76}$$

The frozen crossing operator has kernel $h^{-1/2} t_x((y-x)/\sqrt{h})$. The diagonal jump and (74) therefore give directly

$$D_h = I - \mathcal{K}_h + \mathcal{E}_h, \qquad \|\mathcal{E}_h\|_{\eta,h} \le C\sqrt{h}. \tag{77}$$

Choose a fixed $\eta > 0$ small enough and then h small enough. Equations (76) and (77) give $\|\mathcal{K}_h - \mathcal{E}_h\|_{\eta,h} \le 1 - \epsilon/2$. Thus

$$D_h^{-1} = \sum_{n=0}^\infty (\mathcal{K}_h - \mathcal{E}_h)^n \tag{78}$$

is a two-sided inverse. Without the remainder, this is the expected occupation operator of the killed walk. The exact inverse is dominated in absolute value by the occupation series with nonnegative kernel $\mathcal{K}_h + |\mathcal{E}_h|$, which is still substochastic for the same choice of parameters. The walk has a uniformly bounded expected lifetime, and the exponentially weighted norm controls its displacement before killing. Thus the integral part of the inverse has the required uniform localization. Exponential density weights are absorbed by these kernels and their bounded logarithmic scores.

Preconditioning adds no solutions: the homogeneous equation $(a_* I - h\partial_{xx})u = 0$ has only combinations of $e^{\sqrt{a_*/h}x}$ and $e^{-\sqrt{a_*/h}x}$. Their growing tails are excluded by the stated spaces for small h. □

A.8 A local calibrated branch on a shrinking neighborhood

Lemma A.8 (Scaled local scalar calibration)**.** *At a known smooth positive base with its own terminal price law, there are fixed $r, C > 0$ such that, for*

$$P \in p + \mathcal{Z}, \qquad \|P - p\|_E + \|R\|_{K_h} < r h^{1/4},$$

the equation $\mathcal{F}_b(P, k) + R = 0$ has a unique local positive slope $k = \kappa_h(P, R)$ near the base slope. The slope and endpoint maps are C^2 with

$$\|D\kappa_h\| + \|D\mathcal{S}_h\| \le C, \qquad \|D^2\kappa_h\| + \|D^2\mathcal{S}_h\| \le C h^{-1/4}.$$

The displacement of the slope in K_h is at most $C(\|P - p\|_E + \|R\|_{K_h})$. Signed densities are allowed.

Proof. At the known base $D_k\mathcal{F}_b = -D_h$. The already constructed inverse is bounded in L^2_{uloc} and L^∞ by its weighted Schur kernel, hence uniformly in K_h. The preceding second derivative bound times the radius $rh^{1/4}$ is Cr; choose r so that the slope derivative remains invertible and the slope stays positive. The quantitative implicit-function theorem gives the branch and the uniform first derivative. Its second implicit derivative identity gives $Ch^{-1/4}$. Composition with the endpoint map gives the same estimates. The residual is an additive Banach parameter, so no regularity of it beyond K_h is assumed. □

A.9 Stability in an explicit weighted density norm

The auxiliary reversible diffusion retains the full joint covariance; its difference from the exact density channel is a centered short-jump error. Its heat interpolation is an auxiliary analytic construction, not a Markov representation of the physical Bass bridge.

Lemma A.9 (A weighted heat interpolation)**.** *Let $r_\circ \asymp w\omega$ be a known smooth density with bounded relative derivatives, let Γ be a known smooth covariance matrix with $cI \le \Gamma \le CI$ and bounded derivatives of the required orders, and put*

$$H = -\frac{1}{2r_\circ}\operatorname{div}(r_\circ \Gamma \nabla), \qquad P_s = e^{-sH}, \qquad Q_h = I - P_{2h}.$$

Thus P_s is the reversible Markov semigroup in $L^2(r_\circ)$. Let Π denote conditional expectation given price. Suppose W is positive, bounded above and below with the required bounded derivatives, and

$$W(x, l)\Gamma_{11}(x, l) = \kappa(x)$$

is price-only. For $f, v \in L^2(r_\circ)$ with v price-only, set

$$g = P_h f + Q_h v, \qquad \Pi g = 0.$$

There are uniform $c, C, h_0 > 0$ such that, for $h \le h_0$,

$$\int W g^2 r_\circ + c\|g\|^2_{Q_h} \le \int W f^2 r_\circ + Ch\big(\|f\|^2_{L^2(r_\circ)} + \|v\|^2_{L^2(r_\circ)}\big). \tag{79}$$

The estimate retains the full two-dimensional heat transport.

Proof. All unlabelled L^2 norms use $r_\circ$. Write

$$a_s = P_s(f - P_h v), \qquad b_s = P_{h-s}v, \qquad u_s = a_s + b_s, \qquad 0 \le s \le h.$$

Then $u_0 = f$, $u_h = g$, $u_s' = -Ha_s + Hb_s$, and $u_s'' = H^2 u_s$. Put $E(s) = \langle W u_s, u_s\rangle_{r_\circ}$ and $K = [H, W]$. Self-adjointness of H gives the exact second-variation identity

$$E''(s) = 4\|Ha_s\|^2_{L^2(Wr_\circ)} + 4\|Hb_s\|^2_{L^2(Wr_\circ)} + 2\langle K u_s, H u_s\rangle_{r_\circ}. \tag{80}$$

Indeed, $E'' = 2\|u_s'\|^2_{L^2(Wr_\circ)} + 2\langle W u_s, H^2 u_s\rangle_{r_\circ}$. Move one H past W in the second term and use $Hu_s = Ha_s + Hb_s$; the two squared norms cancel their cross terms.

The commutator is the known first-order operator

$$Ku = -\Gamma\nabla W \cdot \nabla u - \frac{\operatorname{div}(r_\circ \Gamma \nabla W)}{2r_\circ} u.$$

Its coefficients are bounded. Moreover

$$\|\nabla u\|_2^2 \le C\langle u, Hu\rangle_{r_\circ} \le C\|u\|_2\|Hu\|_2.$$

Use this inequality and Young's inequality to absorb the commutator into the first two terms of (80). Since $Hu_s = Ha_s + Hb_s$, this proves

$$E''(s) \ge -CE(s). \tag{81}$$

Only the L^2 size of the scalar pressure enters the resulting bound. The ordinary $L^2(r_\circ)$ contraction of P_s also gives

$$\sup_{0\le s\le h} E(s) \le C(\|f\|_2^2 + \|v\|_2^2).$$

At the terminal time,

$$E'(h) = -2\langle Wg, Hg\rangle_{r_\circ} + 4\langle Wg, Hv\rangle_{r_\circ}.$$

The first term is at most $-c\|\nabla g\|_2^2 + C\|g\|_2^2$, by one integration by parts. For a price-only v, put

$$d = \frac{\partial_x(r_\circ\Gamma_{11}) + \partial_l(r_\circ\Gamma_{21})}{r_\circ}.$$

Then

$$WHv = -\tfrac{1}{2}\kappa(x)v_{xx} - \tfrac{1}{2}Wd\,v_x.$$

The first term pairs to zero with g because $\Pi g = 0$. Integrating the second once in price moves its derivative onto g and known coefficients. Cauchy–Schwarz and Young therefore give

$$E'(h) \le -c\|\nabla g\|_2^2 + C\big(E(h) + \|v\|_2^2\big). \tag{82}$$

Integrating (81) backward from h gives

$$E(h) \le E(0) + hE'(h) + C\int_0^h sE(s)\,ds.$$

Use the interior bound and (82), and absorb $ChE(h)$ for small h. This yields

$$E(h) + ch\|\nabla g\|_2^2 \le E(0) + Ch(\|f\|_2^2 + \|v\|_2^2).$$

Finally,

$$\|g\|_{Q_h}^2 = 2\int_0^h \langle P_s g, HP_s g\rangle_{r_\circ}\,ds \le 2h\langle g, Hg\rangle_{r_\circ} \le Ch\|\nabla g\|_2^2,$$

which proves (79).

These computations extend to the stated L^2 inputs. For fixed $h > 0$, the calibration identity reads

$$v = \Pi P_{2h}v - \Pi P_h f.$$

The known heat-kernel derivative bounds make its right-hand side smooth in price, with each required weighted derivative finite at that fixed h. Thus v and g have the regularity used at s=h; no uniform higher derivative estimate is asserted. On $0 < s < h$ both heat factors are smooth. Apply the integrated inequality from $s = \epsilon$ to h and let $\epsilon \downarrow 0$, using strong semigroup continuity and the uniform interior bound. □

For the known base use the auxiliary density $r_\circ = p/k$ and covariance Γ from (61), and set

$$W = \frac{1}{k\sigma_R^2}.$$

Then $W\Gamma_{11} = k^{-3}$ is price-only and

$$\int W(z/r_\circ)^2 r_\circ = \int \frac{z^2}{\sigma_R^2 p}.$$

This is the explicit physical weight in the standalone theorem.

Comparison of the exact kernels. Retain $r_+ = p^+/k$, the known tilted conditional operator C and $Q_+ = I - C^*C$ solely to write the exact slope derivative. Define

$$\mathcal{A}f = r_\circ^{-1}\mathcal{O}_h(r_\circ f, k), \qquad L = r_+/r_\circ, \qquad \mathcal{B} = LQ_+ - Q_0, \qquad \mathcal{J} = LK_x.$$

The exact derivative, with $f = z/r_\circ$, $v = D\kappa_h[p]z$ and $q' = v$, is

$$g := \frac{D\mathcal{S}_h[p]z}{r_\circ} = P_h f + Q_0 v + (\mathcal{A} - P_h)f + \mathcal{B}v + \mathcal{J}q. \tag{83}$$

Here $\Pi f = \Pi g = 0$, where Π is conditional averaging on price under $r_\circ$, and the scalar branch bound gives $\|v\|_{L^2_{\mathrm{uloc}}} \le C\|z\|_E$.

Both $\mathcal{A}$ and P_h preserve the integral against $r_\circ$. Their kernels have the same leading Gaussian of covariance $h\Gamma$. Thus $\mathcal{A} - P_h$ has exactly zero column mass and an $O(\sqrt{h})$ Gaussian majorant. This comparison keeps the full factor and cross covariance; it does not replace the action on a rough density by a frozen-factor convolution.

Likewise C^*C and P_{2h} have the same leading Gaussian of covariance $2h\Gamma$, and $L = 1 + O(h)$ with known smooth bounds. Consequently $\mathcal{B}$ has an $O(\sqrt{h})$ Gaussian regular kernel and the diagonal term $(L-1)I$, of size $O(h)$. Its column mass is exactly zero, because

$$\int r_\circ L Q_+ v = \int r_+ Q_+ v = 0 = \int r_\circ Q_0 v$$

for every test function v. The diagonal term must be retained in this cancellation.

Finally $\mathcal{J}$ is self-adjoint in $L^2(r_\circ)$ and kills constants:

$$\langle s, \mathcal{J} q\rangle_{r_\circ} = \langle s, K_x q\rangle_{r_+}.$$

Its symmetric off-diagonal conductance is bounded by a known full Gaussian conductance. Therefore its form is bounded by the product of the two Gaussian jump seminorms. Since q is price-only,

$$q(x+d) - q(x) = d\int_0^1 v(x+sd)\,ds,$$

and its jump energy on any translated price window is at most $Chw(n)\|v\|^2_{L^2_{\mathrm{uloc}}}$. Only these increments occur; no choice or estimate of a local pressure constant is required.

The complete localized error. Put $e = \|z\|_E$, $\chi = \chi_n$ and $w_n = w(n)$. From (83),

$$\chi g = P_h(\chi f) + Q_0(\chi v) + E_n, \tag{84}$$

where

$$E_n = [\chi, P_h]f + [\chi, Q_0]v + \chi(\mathcal{A} - P_h)f + \chi\mathcal{B}v + \chi\mathcal{J}q.$$

The complete error is centered against $r_\circ$: both χf and χg are price-centered, and P_h preserves $r_\circ$ while Q_0 has zero integral. Moreover,

$$\|E_n\|_{Q_0^{-1}} \le C\sqrt{h}\sqrt{w_n}\,e. \tag{85}$$

Here and below the inverse norm is on centered functions.

We give the localization details. The two Markov commutators have kernels $O(\sqrt{h})$ times a Gaussian. Their row and column masses are $O(h)$, since $P_h\chi - \chi = O(h)$, and similarly at time $2h$. The bounds are supported near the price window, with Gaussian tails. If D is the physical density kernel (including its output weight $r_\circ$) of either centered operator $\mathcal{A} - P_h$ or $\mathcal{B}$, then the column mass created by multiplication by χ is

$$\int D(X,Y)[\chi(X) - \chi(Y)]\,dX.$$

Its size is $O(h)$: D already costs $\sqrt{h}$, and the cutoff displacement supplies another $\sqrt{h}$. A diagonal term contributes zero to this expression. Subtract this column mass on the diagonal; the remaining centered kernel has $O(\sqrt{h})$ exponential Schur norm. The midpoint bound controls it, while the $O(h)$ multiplication term is bounded using (63). Weighted Schur estimates and the global uniformly local input bounds supply exactly $\sqrt{w_n}e$. Far windows are summable by Gaussian tails and the bounded logarithmic weight derivatives.

For the last term, pair $\chi\mathcal{J}q$ with a centered test s and use the symmetric form on χs and q. Only pairs touching the price window remain. The local q-jump bound above supplies $C\sqrt{h}\sqrt{w_n}e$. The product rule bounds the jump of χs by $C(\|s\|_{Q_0} + \sqrt{h}\|s\|)$, and the gap absorbs the second term. This proves (85) without a factor derivative of an input density or a localized projection assertion.

A scalar correction for the auxiliary heat experiment. Solve, with zero $r_\circ$-mean, the global scalar equation

$$\Pi Q_0 \Pi w_n^\sharp = \Pi E_n. \tag{86}$$

The gap gives existence by the elementary coercive Hilbert-space argument. Testing with $w_n^\sharp$ yields

$$\|w_n^\sharp\|_{Q_0} \le \|E_n\|_{Q_0^{-1}}, \qquad \|w_n^\sharp\|_{L^2(r_\circ)} \le C\sqrt{w_n}e.$$

Set

$$v_n = \chi_n v + w_n^\sharp, \qquad u_n = P_h(\chi_n f) + Q_0 v_n, \qquad R_n = E_n - Q_0 w_n^\sharp.$$

Then $\Pi u_n = \Pi R_n = 0$, $\chi_n g = u_n + R_n$, and $\|v_n\|_{L^2(r_\circ)} \le C\sqrt{w_n}e$. Orthogonality gives

$$\|R_n\|_{Q_0^{-1}} \le \|E_n\|_{Q_0^{-1}}. \tag{87}$$

This is a single global projection applied separately to each complete centered window error. No sharp pressure bound or localized projection theorem is needed.

The explicit density energy. Apply Lemma A.9 to the input $\chi_n f$ and the scalar v_n. With

$$d_n = \|u_n\|_{Q_0}^2,$$

it gives

$$\|u_n\|_{L^2(Wr_\circ)}^2 + cd_n \le \|\chi_n f\|_{L^2(Wr_\circ)}^2 + Chw_n e^2. \tag{88}$$

Multiplication by the known smooth W obeys

$$\|Wu_n\|_{Q_0} \le C\big(\|u_n\|_{Q_0} + \sqrt{h}\|u_n\|_{L^2(r_\circ)}\big).$$

Pair R_n with Wu_n, use (87), and absorb a fraction of d_n by Young's inequality. Since $Q_0 \le I$, the squared L^2 norm of R_n is at most its squared inverse norm. Therefore

$$\|\chi_n g\|_{L^2(Wr_\circ)}^2 + cd_n \le \|\chi_n f\|_{L^2(Wr_\circ)}^2 + Chw_n e^2, \qquad \|\chi_n g\|_{Q_0} \le C\big(d_n^{1/2} + \sqrt{h}\sqrt{w_n}e\big). \tag{89}$$

Because $W/r_\circ = 1/(\sigma_R^2 p)$, division by w_n and the supremum give the linear $1 + Ch$ estimate in (48).

Thus

$$\mathcal{N}_b(D\mathcal{S}_h[p]z) \le (1 + Ch)\mathcal{N}_b(z). \tag{90}$$

The retained output dissipation in (89) will absorb the nonlinear remainder. Along a known family, metric comparison follows directly from the pointwise ratio of the two weights $(p\sigma_R^2)^{-1}$. No regularity of a slope family is needed for that comparison.

A.10 The nonlinear remainder from centered density kernels

All noise norms in this subsection use the original known smooth base's auxiliary density $r_\circ = p/k$ and heat noise $Q = Q_0$. Write $\tau = h^{1/4}$. Evaluated slopes need only be measurable and bounded above and below by fixed positive constants. The input coefficient of a kernel will either be the known density $p \le Cw\omega$, or an arbitrary signed density in E.

For scalar price inputs use a kernel $D_l(x, y)$ with

$$\int D_l(x,y)\,dx = 0, \qquad \begin{aligned} \int e^{a|x-y|^2/h}|D_l(x,y)|\,dy &\le Vw(x)\omega(l), \\ \int e^{a|x-y|^2/h}|D_l(x,y)|\,dx &\le Vw(y)\omega(l). \end{aligned} \tag{91}$$

Embed it as a joint kernel supported on equal factor coordinates and apply the preceding argument to $f(y, \ell) = v(y)$. Thus its inverse-noise bound is $CV\|v\|_{L^2(w)}$. The column bound and cancellation in (91) hold separately for every output factor; averaging them over the factor would not suffice.

These estimates localize to any fixed translated price cutoff χ_n. For an input cut off on a fixed larger interval, test the uncut centered operator against $\chi_n s$. For centered s,

$$\|\chi_n s\|_Q \le C(\|s\|_Q + \sqrt{h}\|s\|_{L^2(r_\circ)}) \le C\|s\|_Q$$

by the native gap. The near-input norm is at most $C\sqrt{w(n)}$ times its uniformly local norm. Other input windows have an extra factor $e^{-c\,\mathrm{dist}^2/h}$ in Schur norm. An ordinary L^2 bound followed by the centered gap estimates these terms; the Gaussian tail absorbs its $h^{-1/2}$ factor and all price-weight ratios. Consequently each local centered-test norm is bounded by $CV\sqrt{w(n)}$ times the corresponding global uniformly local input norm. No separate localized raw source is declared centered. This argument uses the single cutoffs defining the density energy, without derivative components or an additional window sum.

Exact phase derivatives and their localization. Write the conditional endpoint density as

$$K_k(x, l; u, l') = k(u)q_{t,h}^{x,l}\left(\int_x^u k(y)\,dy - b_k(x, l), l'\right).$$

For slope directions v_1, v_2 put $B_i = \int_x^u v_i(y)\,dy - D_k b_k[v_i]$ and $b_{12} = D_k^2 b_k[v_1, v_2]$. Fast derivatives of the fixed reference kernel give the exact formulas

$$\begin{aligned} D_k K_k[v_1] &= v_1(u)q + k(u)q_m B_1, \\ D_k^2 K_k[v_1, v_2] &= v_1(u)q_m B_2 + v_2(u)q_m B_1 + k(u)q_{mm}B_1B_2 - k(u)q_m b_{12}. \end{aligned} \tag{92}$$

No spatial derivative of a slope occurs.

The mean-root derivatives have actual Gaussian kernels in the slope directions. To see this, in rescaled variables put

$$A_i(r) = \int_0^r v_i(x + \sqrt{h}s)\,ds, \qquad \mathfrak{q}(r) = \bar{Q}_h(x, l; A_k(r) - \beta_k), \qquad \beta_k = b_k/\sqrt{h}.$$

Integration by parts in the first variation of the mean equation cancels the terminal slope term. Its first two derivatives are exactly

$$\beta_i = \frac{\int \mathfrak{q} A_i}{\int \mathfrak{q}}, \qquad \beta_{12} = \frac{\int \mathfrak{q}_m (A_1 - \beta_1)(A_2 - \beta_2)}{\int \mathfrak{q}}. \tag{93}$$

The denominator is uniformly positive. Gaussian-tail Fubini in the first formula, and the bound $|A_1 - \beta_1| \le C\|v_1\|_\infty(1 + |r|)$ in the second, show

$$|D_k b_k[v]| \le C\sqrt{h}\, \mathcal{G}_h|v|, \qquad |D_k^2 b_k[v_1, v_2]| \le C\sqrt{h}\|v_1\|_\infty \mathcal{G}_h|v_2|, \tag{94}$$

where $\mathcal{G}_h$ is a fixed Gaussian price average about the initial price. These inequalities follow from the explicit signed kernels, not from an inference that an operator bound implies localization.

When all scalar directions are bounded, $|B_i| \le C\|v_i\|_\infty(|u - x| + \sqrt{h})$. The fast Gaussian bounds in (92) therefore give joint Gaussian majorants with constants $C\|v_1\|_\infty$ and $C\|v_1\|_\infty\|v_2\|_\infty$, respectively. Every column has zero mass: changing a slope preserves the conditional terminal factor law, separately for each initial state. After weighting a signed input Z by $Z/r_\circ$, the joint kernel estimate thus gives, in each normalized local centered-test norm,

$$\begin{aligned} \|D_k \mathcal{O}_h[Z, k]v\|_{Q^{-1}} &\le C\|v\|_\infty \|Z\|_E, \\ \|D_k^2 \mathcal{O}_h[Z, k](v_1, v_2)\|_{Q^{-1}} &\le C\|v_1\|_\infty \|v_2\|_\infty \|Z\|_E. \end{aligned} \tag{95}$$

Here and below a density source is divided by the fixed $r_\circ$ in an inverse-noise norm.

For the known input $p \le Cw\omega$, integrate (92) against $p(x, l)$ and keep one direction in local L^2. Its possible placements are the terminal value $v(u)$, a segment average $v(x + \theta(u - x))$, or the Gaussian mean-root average in (94). The first gives a diagonal atom. In the second, change variables $y = (1 - \theta)x + \theta u$ after the Gaussian input integration; the resulting price kernel has variance $C(1 - \theta)^2 h$ about u. The third convolves two fixed-scale Gaussian price kernels. Polynomial displacement factors are absorbed into broader Gaussians before this change of variables. All positive mixing measures have uniformly bounded mass, including the endpoint $\theta = 1$. Gaussian integration of the known factor weight leaves $C\omega(l')$ uniformly on the whole factor axis.

These are exactly the row and column exponential Schur bounds in (91): Gaussian mixtures of scales at most Ch and diagonal atoms have these bounds uniformly. For the second derivative fix v_1 in L^∞; its kernel constant is multiplied by at most $C\|v_1\|_\infty$. Factor conservation for every scalar direction gives zero price mass of every scalar column separately at each output factor, including its diagonal part. Consequently

$$\begin{aligned} \|D_k \mathcal{O}_h[p, k]v\|_{Q^{-1}} &\le C\|v\|_{L^2_{\mathrm{uloc}}}, \\ \|D_k^2 \mathcal{O}_h[p, k](v_1, v_2)\|_{Q^{-1}} &\le C\|v_1\|_\infty \|v_2\|_{L^2_{\mathrm{uloc}}}. \end{aligned} \tag{96}$$

The evaluated k may be rough; only its uniform positive bounds have been used. All target norms remain those of the original known base.

Lemma A.10 (Second-order density remainder in inverse noise)**.** *Let $\mathcal{S}_h$ be the calibrated branch from the C^2 scalar implicit theorem. If $z \in \mathcal{Z}$ and $e = \|z\|_E \le c\tau$, then*

$$\mathcal{S}_h(p + z) - \mathcal{S}_h(p) = D\mathcal{S}_h[p]z + R_2, \qquad \left\| \frac{\chi_n R_2}{r_\circ} \right\|_{Q^{-1}} \le C\sqrt{w(n)}\, \tau^{-1} e^2. \tag{97}$$

The complete localized source is centered. The constants and the noise form belong to the original known base.

Proof. Put $P_s = p + sz$, $k_s = \kappa_h(P_s, 0)$, $v_s = \partial_s k_s$ and $a_s = \partial_s^2 k_s$. The uniform first and second implicit derivative bounds imply

$$\|v_s\|_{L^2_{\mathrm{uloc}}} \le Ce, \quad \|v_s\|_\infty \le C\tau^{-1}e, \quad \|a_s\|_{L^2_{\mathrm{uloc}}} \le C\tau^{-1}e^2, \quad \|a_s\|_\infty \le C\tau^{-2}e^2.$$

The exact second derivative is

$$\partial_s^2 \mathcal{S}_h(P_s) = 2D_{pk}\mathcal{O}_h[z, v_s] + D_{kk}\mathcal{O}_h[P_s](v_s, v_s) + D_k \mathcal{O}_h[P_s]a_s.$$

Decompose $P_s = p + sz$ in the last two terms. The joint and scalar kernel estimates bound the resulting five terms by

$$C\sqrt{w(n)}\{\tau^{-1}e^2 + \tau^{-2}e^3\} \le C\sqrt{w(n)}\tau^{-1}e^2.$$

This split uses the known bounded p only in the scalar estimate; the other input sz is allowed to be signed and need not be pointwise close to p. Integrating the second derivative with weight $1 - s$ proves the C^2 remainder identity and its inverse-noise estimate. Exact calibration makes each full second derivative have zero price marginal, so multiplying it by χ_n preserves its total mass zero. This justifies the actual inverse-noise norm of the sum; no nonlinear pressure projection is needed. □

Absorption by the output dissipation. Normalize d_n from (89) by $w(n)$. The localized linear estimate retains this nonnegative output dissipation, and multiplication by W gives output Q norm at most $C(d_n^{1/2} + \sqrt{h}e)$. Pair (97) with the linear output in the explicit physical energy. Young's inequality gives

$$C\tau^{-1}e^2(d_n^{1/2} + \sqrt{h}e) \le \epsilon d_n + Che^2 + C\tau^{-2}e^4.$$

The squared remainder is at most $C\tau^{-2}e^4$, because W is bounded and $Q \le I$. Absorb the chosen fraction of the same-window output dissipation, divide by the price weight and take the supremum. Norm equivalence yields

$$\mathcal{N}_b(\mathcal{S}_h(p+z) - \mathcal{S}_h(p))^2 \le (1+Ch)\mathcal{N}_b(z)^2 + Ch^{-1/2}\mathcal{N}_b(z)^4.$$

Taking square roots gives

$$\mathcal{N}_b(\mathcal{S}_h(p+z) - \mathcal{S}_h(p)) \le (1+Ch)\mathcal{N}_b(z) + Ch^{-1/2}\mathcal{N}_b(z)^3. \tag{98}$$

Only the output jump, C^2 endpoint calculus and column conservation are needed; no derivative of an incoming density is taken.

Completion of the proof of Theorem A.5. Use $h = \delta$ and $t = 0$. Lemmas A.6 and A.8 construct the signed calibrated branch on a ball of radius $r\delta^{1/4}$ in $p + \mathcal{Z}$. Exact price calibration puts all fixed-target endpoint differences in $\mathcal{Z}$. Estimate (89) and Lemma A.10 give (98).

Use $\mathcal{N}_b$ at both endpoints. For small δ, norm equivalence puts the ball $\mathcal{N}_b(z) \le \delta^{3/4}$ inside the signed branch's $r\delta^{1/4}$ neighborhood. On this ball,

$$\delta^{-1/2}\mathcal{N}_b(z)^3 \le \delta\mathcal{N}_b(z).$$

Thus (98) proves (47); pointwise comparison of the explicit weights proves (49).

It remains to identify the calibrated endpoint for a nonnegative incoming density. The kernel is positive, its conditional price mean is the incoming price, and its terminal price law is the prescribed one. Choose $\psi' = \xi_\kappa$ for its calibrated slope κ. The identity $\xi_\kappa(U) = Y_1 + \zeta_\kappa(x,l)$ gives, for any competing martingale endpoint Z with the same terminal price law, coupled with the reference endpoint conditional on (x,l),

$$\mathbb{E}[UY_1] - \mathbb{E}[ZY_1] = \mathbb{E}\big[\psi(Z) - \psi(U) - \psi'(U)(Z-U)\big] \ge \frac{c}{2}\mathbb{E}|Z-U|^2.$$

Indeed, the common terminal marginal cancels the potential terms, and the conditional martingale means cancel the frozen shift. All terms are integrable because the slope is bounded and the laws have second moments. Maximizing the reference covariance is the WMOT problem of the main text. The displayed identity proves optimality and uniqueness, without an additional dual-attainment assertion. The same argument applies whenever the local scalar construction succeeds.

□

A.11 Smooth consistency at a known input

The consistency statement concerns only known smooth input data. Write $h = \delta$ and allow a translated physical start t.

Lemma A.11 (First-order smooth consistency)**.** *For known smooth p, k with the stated positive bounds,*

$$\mathcal{O}_{t,h}(p,k) = p + h\mathcal{G}_t(k)^* p + O(h^2), \tag{99}$$

where

$$\mathcal{G}_t(k) = \tfrac{1}{2}(\mathsf{D}_k A_t^R \mathsf{D}_k) : D^2 + b_t^R \partial_l, \qquad \mathsf{D}_k = \mathrm{diag}(k^{-1}, 1). \tag{100}$$

The endpoint remainder is $O(h^2)$ in E and in every fixed finite relative smooth norm needed below. The normalized calendar satisfies

$$J_h(p,k)(x) = \frac{1}{2k(x)^2}\int \sigma_R(t,x,l)^2 p(x,l)\, dl + O(h), \tag{101}$$

with the remainder divided by w in each required fixed smooth norm. For each required output order s, a sufficiently large fixed r also gives the directional expansion

$$D_k J_h(p,k)[v] = -\frac{1}{k^3}\left(\int \sigma_R(t,\cdot,l)^2 p(\cdot,l)\, dl\right) v + O(h) w \|v\|_{C_b^r}, \tag{102}$$

where the relative remainder is measured in C_b^s. The constants are uniform on known bounded positive smooth sets. Only finitely many known spatial and physical-time derivatives are used.

Proof. The ordinary reference semigroup and the scalar mean equation give a smooth centering shift at $h = 0$, with bounded first two horizon derivatives in every required finite spatial norm, and

$$b_{t,h}(x,l) = \frac{h}{2}\sigma_R(t,x,l)^2\frac{k'(x)}{k(x)^2} + O(h^2).$$

To obtain a strong remainder, write the reference transition density as $q_h^R((x,l),(y,l'))$ and substitute $y = x+\xi_k(z)-\xi_k(x)-b_{t,h}(x,l)$. For the substituted kernel $\widetilde{q}$, put $\mathcal{D}_x = \partial_{\text{initial }x} + \partial_y$ and $\mathcal{D}_l = \partial_{\text{initial }l} + \partial_{l'}$. The exact identities

$$\partial_x\widetilde{q} = \mathcal{D}_x q - (k+b_x)\partial_y q, \qquad \partial_l\widetilde{q} = \mathcal{D}_l q - \partial_{l'} q - b_l\partial_y q$$

transfer terminal derivatives to the known incoming density or to simultaneous kernel derivatives. Here $k + b_x$ stays positive, and the simultaneous derivatives obey the previously proved Gaussian bounds without a fast short-time loss. Apply the reference forward equation for time derivatives of the kernel and repeat these two identities. Gaussian Schur bounds and the exponential weights then bound the first two h derivatives of the endpoint density, with any fixed required output derivatives, uniformly down to zero. The first h derivative therefore has a strong right limit at zero. Itô's formula identifies that limit as $\mathcal{G}_t(k)^*p$, and Taylor's formula gives (99).

Apply the call primitive to the price marginal of (99). Mass and first-moment cancellation permit the two-tail weighted bound. Since $hJ_h'' = \mathcal{O}_{t,h}^X - p^X$, this gives (101), including its fixed scalar derivative bounds. Differentiate this calculation once in a smooth slope direction. The derivative-transfer identities still apply; their coefficients and slope derivatives have bounded finite smooth norms. Thus the same two horizon-derivative bounds hold after this parameter differentiation. Differentiating the leading calendar gives (102), including its $O(h)$ remainder. No expansion at a rough input density is used. □

A.12 The smooth target density by a weighted Feynman–Kac formula

Write the target generator in the form

$$L_t\varphi = B_t\cdot\nabla\varphi + \tfrac{1}{2}A_t : D^2\varphi, \qquad A_t = \Sigma_t\Sigma_t^\top,$$

where, for the present model,

$$B = \begin{pmatrix}0\\ b\end{pmatrix}, \qquad \Sigma = \begin{pmatrix} f & 0 \\ \rho\beta & \sqrt{1-\rho^2}\,\beta\end{pmatrix}.$$

Let $W(x,l) = w(x)\omega(l) = \mathfrak{w}(x,l)$ and write $p_0 = W\mathsf{r}_0$, as in the regular-start assumption. All positive-order derivatives of $\log W$ are bounded. Expanding the forward operator gives exactly

$$W^{-1}L_t^*(Wr) = \widetilde{L}_t r + c_t r, \qquad \widetilde{L}_t = \tfrac{1}{2}A_t : D^2 + \widetilde{B}_t\cdot\nabla, \tag{103}$$

where

$$\widetilde{B}_t = \operatorname{div} A_t + A_t\nabla\log W - B_t, \qquad c_t = W^{-1}L_t^*W. \tag{104}$$

Here $(\operatorname{div} A)_j = \sum_i \partial_i A_{ij}$. The functions $\widetilde{B}, c, \Sigma$ and all their required derivatives are bounded.

Lemma A.12 (Propagation of the relative density bounds)**.** *The target density is $p_t = W\mathsf{r}_t$, where*

$$0 < c_T \le \mathsf{r}_t(z) \le C_T, \qquad \sup_{0\le t\le T, z}|\partial_t^j\partial_z^\alpha \mathsf{r}_t(z)| < \infty$$

for every fixed j and spatial multi-index α.

Proof. For fixed t and starting point z, run the ordinary auxiliary SDE

$$dZ_s = \widetilde{B}_{t-s}(Z_s)\,ds + \Sigma_{t-s}(Z_s)\,d\widetilde{W}_s, \qquad Z_0 = z, \quad 0\le s\le t.$$

The coefficient time is reversed; the spatial process is an ordinary forward SDE. Define

$$\mathsf{r}_t(z) = \mathbb{E}_z\left[\mathsf{r}_0(Z_t)\exp\left(\int_0^t c_{t-s}(Z_s)\,ds\right)\right]. \tag{105}$$

Since c is bounded and r_0 is bounded above and below by positive constants, the same is true of r_t, uniformly on $[0,T]$.

Only ordinary SDE variations are needed for its derivatives. The first derivative of Z_s with respect to z satisfies the linear equation obtained by differentiating the bounded smooth coefficients. Every higher derivative satisfies the same linear equation with a forcing made of lower derivatives. Burkholder–Davis–Gundy and Gronwall therefore give every required moment of these derivatives, uniformly over $t \le T$ and z. Differentiating (105), including its time-integrated bounded potential, yields the claimed uniform spatial bounds.

For clarity, the reversed coefficient time in (105) produces the forward equation

$$\partial_t\mathsf{r}_t = (\widetilde{L}_t + c_t)\mathsf{r}_t, \qquad \mathsf{r}_{t=0} = \mathsf{r}_0.$$

One can verify it directly. The Markov property splits the first short interval of the auxiliary experiment for r_{t+h} and leaves r_t as its terminal function. Itô's formula on this first interval gives

$$\mathsf{r}_{t+h} - \mathsf{r}_t = h(\widetilde{L}_t + c_t)\mathsf{r}_t + o(h).$$

SDE stability supplies continuity in t, also after each fixed spatial differentiation. The established bounds make the remainder and right derivative locally uniform. The right derivative is continuous, hence is the time derivative. Repeated use of this equation gives all fixed time-derivative bounds, including right derivatives at zero.

Equation (103) now shows that $p_t = W\mathsf{r}_t$ solves the target forward equation with initial density p_0. It is the density of the target SDE: test this forward equation against the backward expectations of a smooth compactly supported terminal function. Their backward Itô equation makes the integral constant in time, identifying it with the target expectation. Integration by parts is justified by the exponential tails of W and the established derivative bounds. This also proves mass one without a separate assumption on the auxiliary process. □

Uniform ellipticity is not needed for this density estimate. The positive smooth initial ratio is propagated by an ordinary diffusion with a bounded multiplicative weight.

A.13 Composition along the smooth SLV target

Lemma A.13 (An explicit smooth comparison slope)**.** *Under the coefficient and regular-start assumptions of the convergence theorem, put $p_t = p(t,\cdot)$, $k_0(t,x) = 1/\lambda(t,x)$, and*

$$M_t(x) = \int \sigma_R(t,x,l)^2 p_t(x,l)\,dl, \quad J^\star_{t,h} = \frac{\mathrm{Call}(\mu_{t+h}) - \mathrm{Call}(\mu_t)}{h},$$
$$E^0_{t,h} = J_h(p_t, k_0(t)) - J^\star_{t,h}, \qquad D^0_t = \frac{M_t}{k_0(t,\cdot)^3}.$$

Define the known slope pointwise by

$$k^{\mathrm{app}}_{t,h} = k_0(t) + E^0_{t,h}/D^0_t. \tag{106}$$

It is positive for small h. For each fixed required finite r,

$$\|k^{\mathrm{app}}_{t,h} - k_0(t)\|_{C^r_b} \le C_{r,T}h, \qquad 0 \le t \le T-h. \tag{107}$$

Write

$$q^{\mathrm{app}}_{t,h} = \mathcal{O}_{t,h}(p_t, k^{\mathrm{app}}_{t,h}), \qquad m_{\mathrm{app}} = (q^{\mathrm{app}}_{t,h})^X / k^{\mathrm{app}}_{t,h}.$$

Then

$$\|q^{\mathrm{app}}_{t,h} - p_{t+h}\|_E \le C_T h^2, \qquad \left\| L_h \frac{J_h(p_t, k^{\mathrm{app}}_{t,h}) - J^\star_{t,h}}{m_{\mathrm{app}}} \right\|_{K_h} \le C_T h^2. \tag{108}$$

The second expression is a prescribed target residual at the known base $(p_t, k^{\mathrm{app}}_{t,h})$.

Proof. The target density and coefficient bounds give $D^0_t \asymp w$; D^0_t/w and its reciprocal have bounded derivatives. Known smooth calendar consistency and the target forward equation give $E^0_{t,h} = O(h)w$ in every required finite relative spatial norm. The leading terms agree because $k_0 = 1/\lambda$. Thus (107) follows by division, and positivity follows from the positive lower bound for k_0.

Set $d = E^0_{t,h}/D^0_t = O(h)$ in the required smooth norms. The fundamental theorem of calculus and (102) give

$$\begin{aligned} J_h(p_t, k_0 + d) - J^\star_{t,h} &= E^0_{t,h} + \int_0^1 D_k J_h(p_t, k_0 + sd)[d]\,ds \\ &= E^0_{t,h} - D^0_t d + O(h)\, w\, \|d\|_{C^r_b} + O(1)\, w\, \|d\|^2_{C^r_b} = O(h^2)w. \end{aligned}$$

Here the last quadratic term comes simply from $M_t/(k_0+sd)^3 - D^0_t = O(h)w$ in the same smooth norms; no second derivative of the calendar map is needed. Take the output order at least two higher than needed for the preconditioner. The known-base endpoint estimates give $m_{\mathrm{app}} \asymp w$ with the same required smooth relative bounds. Division and $L_h = a_* I - h\partial_{xx}$ therefore give an $O(h^2)$ residual in both L^2_{uloc} and L^∞, hence in K_h.

The first-order strong endpoint expansion applied directly at the known smooth slope $k^{\mathrm{app}}_{t,h}$ gives

$$q^{\mathrm{app}}_{t,h} = p_t + h\mathcal{G}_t(k^{\mathrm{app}}_{t,h})^* p_t + O(h^2).$$

Since $k^{\mathrm{app}}_{t,h} - k_0 = O(h)$ in the required smooth norms, $\mathcal{G}_t(k^{\mathrm{app}}_{t,h})^* p_t - \mathcal{G}_t(k_0)^* p_t = O(h)$ in E. The target forward equation has generator $\mathcal{G}_t(k_0)$, so this proves the first bound in (108). No implicit regularity theorem for an optimizing slope is used. □

Corollary A.14 (Quantitative calibration)**.** *Under the assumptions of Theorem A.1, let $h = T/N$, $t_j = jh$, and let p_j be the joint density at the jth SKR endpoint. For sufficiently small h,*

$$\max_{0\le j\le N} \|p_j - p(t_j)\|_E \le C_T h, \qquad \max_{0\le j<N} \|k_j - 1/\lambda(t_j,\cdot)\|_{K_h} \le C_T h. \tag{109}$$

Here k_j is the positive inverse slope of the exact Bass update. In particular its L^2_{uloc} error is $O(h)$ and its uniform error is $O(h^{3/4})$.

Proof. At step j use the known smooth base $b_j = (p(t_j), k^{\mathrm{app}}_{t_j,h})$ from Lemma A.13. Write $q^{\mathrm{app}}_j = \mathcal{S}_0(p(t_j))$ for its own endpoint and set

$$R_j = L_h \frac{J_h(p(t_j), k^{\mathrm{app}}_{t_j,h}) - J^\star_{t_j,h}}{m_{\mathrm{app},j}}.$$

The predictor gives $\|R_j\|_{K_h} \le Ch^2$ and $\|q^{\mathrm{app}}_j - p(t_{j+1})\|_E \le Ch^2$. The joint local branch of Lemma A.8 gives $\mathcal{S}_{R_j}$, uniformly for incoming $P = p(t_j) + z$ in its shrinking neighborhood. Its bounded first residual derivative implies

$$\|\mathcal{S}_{R_j}(P) - \mathcal{S}_0(P)\|_E \le Ch^2. \tag{110}$$

The equation $\mathcal{F}_{b_j}(P,k) + R_j = 0$ says exactly $L_h((J_h(P,k) - J^\star_{t_j,h})/m_{\mathrm{app},j}) = 0$. The exponential calendar tails exclude its growing homogeneous modes, so this branch has price marginal $\mu_{t_{j+1}}$. For nonnegative P the positive endpoint kernel and the WMOT identity in the standalone theorem identify it as the unique optimizer.

The explicit norm is defined on all density differences. Its triangle inequality, the standalone theorem for the known base's own target, and the two $O(h^2)$ defects give

$$\begin{aligned}\mathcal{N}_{b_j}(\mathcal{S}_{R_j}(P) - p(t_{j+1})) &\le \mathcal{N}_{b_j}(\mathcal{S}_{R_j}(P) - \mathcal{S}_0(P)) \\ &\quad + \mathcal{N}_{b_j}(\mathcal{S}_0(P) - \mathcal{S}_0(p(t_j))) + \mathcal{N}_{b_j}(q^{\mathrm{app}}_j - p(t_{j+1})) \\ &\le (1 + Ch)\mathcal{N}_{b_j}(P - p(t_j)) + Ch^2.\end{aligned} \tag{111}$$

Starting from $p_0 = p(0)$, iterate these positive calibrated maps. Put $e_j = \mathcal{N}_{b_j}(p_j - p(t_j))$ for $j < N$, and use the last step's physical norm at $j = N$. Thus no reference interval beyond T is required. All these differences have both marginals zero. The bounded time derivative of $\log(p_t \sigma_R(t)^2)$ gives metric comparison at cost $1 + Ch$ between consecutive steps. This compares the same physical density difference and uses no time regularity of the predictor slope. As long as $e_j \le h^{3/4}$, Theorem A.5 and (111) give

$$e_{j+1} \le (1 + Ch)e_j + Ch^2, \qquad e_0 = 0. \tag{112}$$

Discrete Gronwall gives $e_j \le A_T h$. For sufficiently small h, $A_T h < \frac{1}{2}h^{3/4}$, so induction at the first possible exit closes on every fixed finite horizon.

The local optimizer identity identifies every constructed iterate with the actual SKR update. The scalar branch bound gives

$$\|k_j - k_0(t_j)\|_{K_h} \le C\big(\|p_j - p(t_j)\|_E + \|R_j\|_{K_h}\big) + \|k^{\mathrm{app}}_{t_j,h} - k_0(t_j)\|_{K_h} \le C_T h.$$

Norm equivalence proves (109). □

Proof of Theorem A.1. Quantitative calibration gives $\|k_j - 1/\lambda(t_j,\cdot)\|_\infty \le C_T h^{3/4}$. The reference conditional second moment is Ch, so Lemma A.3 yields $\mathbb{E}|U_j - V_j|^2 \le C_T h^{5/2}$. Proposition A.2 now gives $\mathbb{E}\sup_{t\le T}|X^h_t - X_t|^2 \le C_T h$. □

The convergence statement retains the regular-start and coefficient assumptions of Theorem A.1. The autonomous-price corollary permits any square-integrable initial pair, including a deterministic initial price. A general Dirac-start convergence result and existence of an SLV limit for arbitrary prescribed marginal curves are not asserted. Composition uses the standalone WMOT stability theorem, an explicit smooth comparison slope and first-order consistency; no regularity theorem for an exactly calibrated comparison optimizer is required.

References

[1] Frédéric Abergel and Rémi Tachet. A nonlinear partial integro-differential equation from mathematical finance. *Discrete and Continuous Dynamical Systems. Series A*, 27(3):907–917, 2010. doi: 10.3934/dcds.2010.27.907.

[2] Beatrice Acciaio and Antonio Marini. Existence of q-bass martingales in the semidiscrete setting, 2026. URL `https://arxiv.org/abs/2607.15872`.

[3] Beatrice Acciaio and Antonio Marini. Fixed points for the q-bass martingale: Existence, stability, and convergence, 2026. URL `https://arxiv.org/abs/2609.07351`.

[4] Beatrice Acciaio, Antonio Marini, and Gudmund Pammer. Calibration of the Bass local volatility model. *SIAM Journal on Financial Mathematics*, 16(3):803–833, 2025.

[5] Julio Backhoff-Veraguas, Mathias Beiglböck, and Gudmund Pammer. Existence, duality, and cyclical monotonicity for weak transport costs. *Calculus of Variations and Partial Differential Equations*, 58(6):203, 2019.

[6] Julio Backhoff-Veraguas, Mathias Beiglböck, Martin Huesmann, and Sigrid Källblad. Martingale Benamou-Brenier: A probabilistic perspective. *Ann. Probab.*, 48(5):2258–2289, 2020.

[7] Julio Backhoff-Veraguas, Gregoire Loeper, and Jan Obłój. Geometric martingale Benamou–Brenier transport and geometric Bass martingales. *Proceedings of the American Mathematical Society*, 2025.

[8] Julio Backhoff-Veraguas, Mathias Beiglböck, Walter Schachermayer, and Bertram Tschiderer. Existence of Bass martingales and the martingale Benamou–Brenier problem in $\mathbb{R}^d$. *The Annals of Probability*, 54(5):2216–2263, 2026. doi: 10.1214/25-AOP1800.

[9] Richard F. Bass. Skorokhod imbedding via stochastic integrals. In *Séminaire de Probabilités XVII 1981/82*, volume 986 of *Lecture Notes in Mathematics*, pages 221–224. Springer, Berlin, 1983. doi: 10.1007/BFb0068318.

[10] Mathias Beiglböck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *The Annals of Probability*, 44(1):42–106, 2016.

[11] Mathias Beiglböck, Marcel Nutz, and Nizar Touzi. Complete duality for martingale optimal transport on the line. *The Annals of Probability*, 45(5):3038–3074, 2017. doi: 10.1214/16-AOP1131.

[12] Mathias Beiglböck, Gudmund Pammer, and Walter Schachermayer. From Bachelier to Dupire via optimal transport. *Finance and Stochastics*, 26(1):59–84, 2022.

[13] Mathias Beiglböck, Benjamin Jourdain, William Margheriti, and Gudmund Pammer. Stability of the weak martingale optimal transport problem. *The Annals of Applied Probability*, 33(6B):5382–5412, 2023.

[14] Mathias Beiglböck, Gudmund Pammer, and Alexander Posch. The Knothe-Rosenblatt distance and its induced topology. *arXiv preprint arXiv:2312.16515*, 2023.

[15] Mathias Beiglböck, Gudmund Pammer, and Lorenz Riess. Change of numeraire for weak martingale transport. *arXiv preprint arXiv:2406.07523*, 2024.

[16] Mathias Beiglböck, Gudmund Pammer, Lorenz Riess, and Stefan Schrott. The fundamental theorem of weak optimal transport. *arXiv preprint arXiv:2501.16316*, 2025.

[17] Guillaume Carlier, Alfred Galichon, and Filippo Santambrogio. From Knothe's transport to Brenier's map and a continuation method for optimal transport. *SIAM Journal on Mathematical Analysis*, 41(6):2554–2576, 2010.

[18] Guillaume Carlier, Hugo Malamut, and Maxime Sylvestre. Weak optimal transport with moment constraints: constraint qualification, dual attainment and entropic regularization. *arXiv preprint arXiv:2511.16211*, 2025.

[19] Antoine Conze and Pierre Henry-Labordere. Bass construction with multi-marginals: Lightspeed computation in a new local volatility model. *SSRN Electronic Journal*, 2021.

[20] Andrei Cozma, Matthieu Mariapragassam, and Christoph Reisinger. Calibration of a hybrid local-stochastic volatility stochastic rates model with a control variate particle method. *SIAM Journal on Financial Mathematics*, 10(1):181–213, 2019.

[21] Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[22] Walter Farkas, Francesco Ferrari, and Urban Ulrych. Pricing autocallables under local-stochastic volatility. *Frontiers of Mathematical Finance*, 1(4):575–610, 2022.

[23] Nathael Gozlan, Cyril Roberto, Paul-Marie Samson, and Prasad Tetali. Kantorovich duality for general transport costs and applications. *J. Funct. Anal.*, 273(11):3327–3405, 2017.

[24] Ivan Guo and Grégoire Loeper. Path dependent optimal transport and model calibration on exotic derivatives. *The Annals of Applied Probability*, 31(3):1232–1263, 2021.

[25] Ivan Guo, Grégoire Loeper, and Shiyi Wang. Local volatility calibration by optimal transport. In *2017 MATRIX Annals*, pages 51–64. Springer, 2019.

[26] Ivan Guo, Grégoire Loeper, and Shiyi Wang. Calibration of local-stochastic volatility models by optimal transport. *Mathematical Finance*, 32(1):46–77, 2022.

[27] J. Guyon and P. Henry-Labordère. Nonlinear option pricing. *CRC Press*, 2013.

[28] Julien Guyon. Dispersion-constrained martingale Schrödinger problems and the exact joint S&P 500/VIX smile calibration puzzle. *Finance and Stochastics*, 28(1):27–79, 2024.

[29] Julien Guyon and Pierre Henry-Labordere. The smile calibration problem solved. *Available at SSRN 1885032*, 2011.

[30] István Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability Theory and Related Fields*, 71:501–516, 1986. doi: 10.1007/BF00699039.

[31] Manuel Hasenbichler, Benjamin Joseph, Gregoire Loeper, Jan Obloj, and Gudmund Pammer. The martingale Sinkhorn algorithm. *arXiv preprint arXiv:2310.13797v5*, 2026.

[32] Manuel Hasenbichler, Gudmund Pammer, and Stefan Thonhauser. A weak transport approach to the Schrödinger–Bass bridge. *arXiv preprint arXiv:2604.02312*, 2026.

[33] Mark Jex, Robert Henderson, and David Wang. Pricing exotics under the smile. *Risk*, pages 72–75, November 1999.

[34] Benjamin Jourdain and Gudmund Pammer. An extension of martingale transport and stability in robust finance. *Electronic Journal of Probability*, 29(57):1–30, 2024. doi: 10.1214/24-EJP1114.

[35] Benjamin Jourdain and Alexandre Zhou. Existence of a calibrated regime switching local volatility model. *Mathematical Finance*, 30(2):501–546, 2020.

[36] Daniel Lacker, Mykhaylo Shkolnikov, and Jiacheng Zhang. Inverting the markovian projection, with an application to local stochastic volatility models. *The Annals of Probability*, 48(5), 2020.

[37] Bryan Liang, Marcel Nutz, Shunan Sheng, and Valentin Tissot-Daguette. Bid–ask martingale optimal transport. *arXiv preprint arXiv:2603.24605*, 2026.

[38] Alexander Lipton. The vol smile problem. *Risk*, 15(2):61–65, 2002.

[39] Gregoire Loeper. Option pricing with linear market impact and nonlinear Black-Scholes equations. *Ann. Appl. Probab.*, 28(5):2664–2726, 2018. ISSN 1050-5164,2168-8737. doi: 10.1214/17-AAP1367. URL `https://doi.org/10.1214/17-AAP1367`.

[40] George Lowther. Fitting martingales to given marginals. *arXiv preprint arXiv:0808.2319*, 2008.

[41] George Lowther. Limits of one-dimensional diffusions. *The Annals of Probability*, 37(1):78–106, 2009.

[42] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Mathematical Journal*, 80(2):309–323, 1995.

[43] Stéphane Menozzi, Antonello Pesce, and Xicheng Zhang. Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift. *Journal of Differential Equations*, 272:330–369, 2021. doi: 10.1016/j.jde.2020.09.004.

[44] Marcel Nutz, Johannes Wiesel, and Long Zhao. Martingale Schrödinger bridges and optimal semistatic portfolios. *Finance and Stochastics*, 27(1):233–254, 2023.

[45] Filip Pramenković. Weak optimal transport: When is the dual potential convex? *arXiv preprint arXiv:2507.07200*, 2025. URL `https://arxiv.org/abs/2507.07200`.

[46] Yong Ren, Dilip Madan, and M. Qian. Calibrating and pricing with embedded local volatility models. *Risk*, pages 138–143, September 2007.

[47] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1998.

[48] Xiaolu Tan and Nizar Touzi. Optimal transportation under controlled stochastic dynamics. *The Annals of Probability*, 41(5):3201–3240, 2013.

[49] Bertram Tschiderer. q-bass martingales. *arXiv preprint arXiv:2402.05669*, 2024.

[50] ShengQuan Zhou. Markov-functional models with local drift. *arXiv preprint arXiv:2411.15053*, 2024.