The Quadratic Gaussian Model

the smile a short rate can make

Martin Burke (with Claude)

Draft

Abstract

The one-factor Hull–White model is a beautiful machine, and it has two holes in it. Because its short rate is Gaussian, that rate can wander below zero with real probability; and because it is Gaussian, the implied-volatility smile it makes is *flat*—a fixed shape, with no parameter left to bend it across strike. Both limitations have the same root: a Gaussian rate. This note develops a model that keeps almost everything good about Hull–White while fixing both, through a single change of variable: let the short rate be a *quadratic* function of a Gaussian factor, $r_t = x_t^2 + \beta x_t + \phi(t)$. Squaring a Gaussian gives a rate that cannot fall below a hard, tunable floor, and whose own volatility grows with its level—which is exactly a volatility *skew*. Remarkably, the tractability survives: bond prices are no longer exponential-affine but exponential-*quadratic*, with coefficients that solve a Riccati equation, and vanilla options reduce to one-dimensional Gaussian integrals because the driving factor stays Gaussian under the pricing measure. We recap only what we reuse from the companion Hull–White note, then build the quadratic Gaussian model from its two-layer spine—a drift that fits today's curve exactly, and a few volatility parameters for the dynamics—show the floor and the skew it produces, price caps and swaptions, calibrate it, and work through the products a single Gaussian factor simply cannot price: out-of-the-money caps and floors, the swaption skew, convexity-sensitive CMS, and floored or capital-guaranteed notes.

Contents

1 Where Hull–White leaves off

We pick up exactly where the Hull–White note ended. There, the short rate was a Gaussian process,

$$\mathrm{d}r_t = \big(\theta(t) - a\,r_t\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \tag{1}$$

or, in the cleaner state-variable form, $r_t = x_t + \alpha(t)$ with x_t a mean-reverting Gaussian and $\alpha(t)$ a deterministic shift chosen to reprice today's curve. That model earned its place: it fits the curve exactly, prices caps and swaptions in closed form, and runs on a lattice for early-exercise products. Its whole character—good and bad—follows from one fact: *the short rate is normally distributed*. Two consequences of that fact are the reason this note exists.

Gap 1: rates can go negative. A normal random variable has support on the whole real line. However well it is centred, a Gaussian short rate places positive probability on every negative value (Figure 1, left). For decades this was treated as a blemish to ignore—rates were comfortably positive, the left tail was thin, and the analytic convenience was worth it. In a low-rate world it stops being ignorable: when the curve sits near zero, the Gaussian tail below zero carries material mass, and any product whose value turns on rates *not* falling through a floor—a guaranteed-minimum coupon, a capital-protected note—is mispriced by a model that lets them.

Gap 2: the smile is flat, and fixed. Everything in Hull–White is priced at, or near, the money. But options trade across strikes, and their implied volatilities are not flat: the market shows a *smile*, usually a pronounced *skew*. A Gaussian short rate produces forward rates that are also (nearly) Gaussian, and a normal model has an almost flat smile when implied vols are quoted in normal (Bachelier) terms (Figure 1, right). That flatness is not a knob you can turn—it is forced by Gaussianity. Hull–White's two volatility parameters a, σ are already spent on the *level and term structure* of at-the-money volatility; nothing is left to reshape the smile across strike. A market with a real skew cannot be matched.

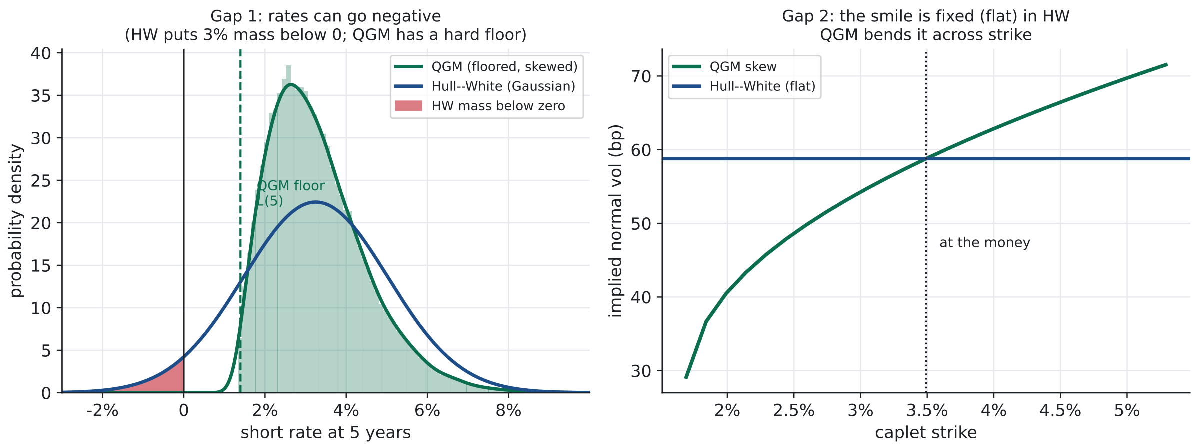

Figure 1: The two things a one-factor Gaussian model cannot do. *Left:* the distribution of the short rate at five years. Hull–White (blue) is a symmetric Gaussian with genuine mass below zero (red); the quadratic Gaussian model (green) is skewed and sits entirely above a hard floor $L(5)$. *Right:* the implied normal volatility of a caplet across strike. Hull–White is flat (blue)—a fixed shape set by Gaussianity; the quadratic Gaussian model (green) makes a genuine skew. Both panels are computed from the models of this note.

Both gaps share a cause, so they can share a cure. The cause is that the rate *is* the Gaussian. The cure is to keep a Gaussian *factor*—and so keep the tractability—but let the rate be a *nonlinear* function of it. The simplest nonlinear function that bounds the rate below and tilts its volatility is the *square*. That is the whole idea of the quadratic Gaussian model, and the rest of this note is its consequences.

2 The idea: square a Gaussian

Keep the familiar engine. Let x_t be a scalar, mean-reverting Gaussian factor under the risk-neutral measure—an Ornstein–Uhlenbeck process, exactly the object that drove Hull–White,

$$\mathrm{d}x_t = -\kappa\, x_t\, \mathrm{d}t + \sigma\, \mathrm{d}W_t, \qquad x_0 = 0. \tag{2}$$

The two parameters (κ, σ) play precisely the role Hull–White's (a, σ) did: κ is the mean-reversion speed, σ the volatility of the factor. What changes is the link from the factor to the rate. Hull–White set $r_t = x_t + \alpha(t)$, a *linear* link. We make it *quadratic*:

$$r_t \;=\; x_t^2 + \beta\, x_t + \phi(t). \tag{3}$$

Here $\phi(t)$ is a deterministic function—the analogue of Hull–White's $\alpha(t)$—chosen later to fit today's curve exactly, and β is a constant we will read shortly as the knob that sets the smile. Completing the square makes the geometry plain:

$$r_t \;=\; \left(x_t + \tfrac{\beta}{2}\right)^2 + L(t), \qquad L(t) := \phi(t) - \tfrac{\beta^2}{4}. \tag{4}$$

The short rate is a *squared Gaussian sitting on a floor*: a parabola in the factor x, with its vertex—its minimum—at the level $L(t)$. Because a square is never negative, the rate can never fall below $L(t)$. Choose ϕ so that $L(t) \geq 0$ and the model has *guaranteed non-negative rates*—something no Gaussian short rate can offer. Figure 2 shows the mechanism in one picture: a symmetric Gaussian density for x, pushed through the parabola, comes out as a floored, right-skewed density for r.

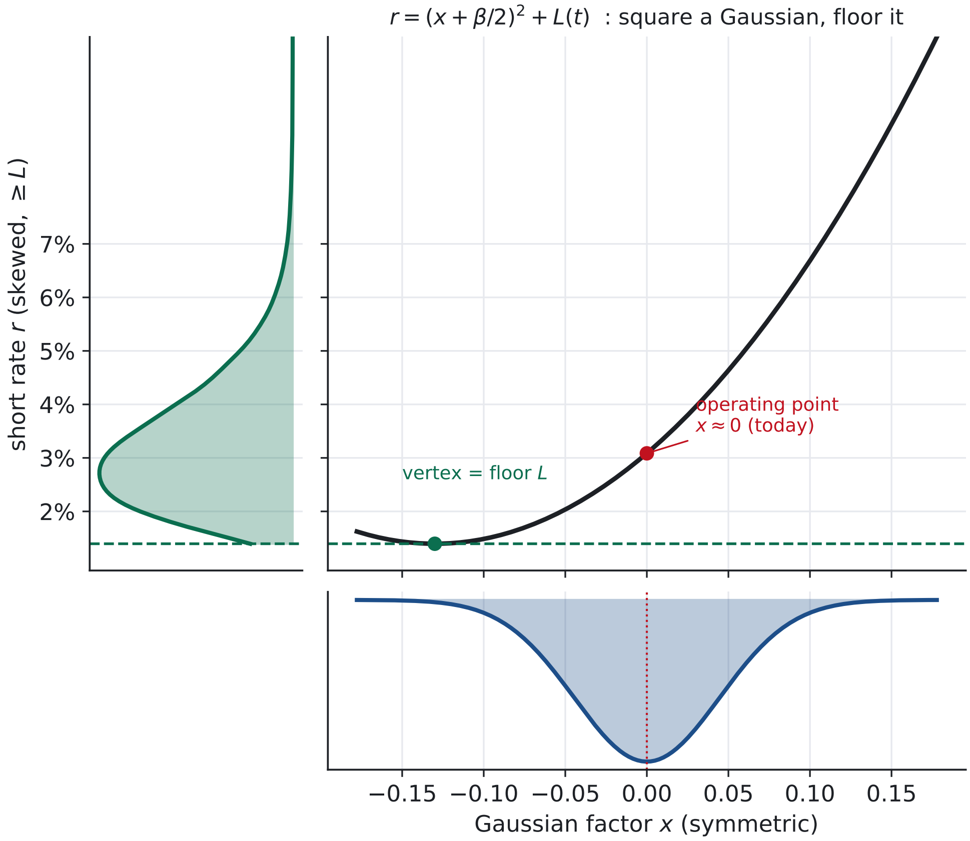

Figure 2: The quadratic map. A symmetric Gaussian factor x (bottom) is squared and shifted by the parabola $r = (x + \beta/2)^2 + L$ (centre) into a short rate (left) that is bounded below by the floor L (the parabola's vertex) and is right-skewed. Today the factor starts at $x = 0$ (red), the operating point part-way up the parabola; the floor sits below it. This single change of variable is the entire model.

The skew, for free. Squaring does something to the rate's volatility that a linear map cannot. Apply Itô's lemma to (3):

$$\begin{aligned} \mathrm{d}r_t &= \big(2x_t + \beta\big)\,\mathrm{d}x_t + \tfrac{1}{2}\cdot 2\,(\mathrm{d}x_t)^2 + \phi'(t)\,\mathrm{d}t \\ &= \underbrace{\Big[\sigma^2 - \kappa\, x_t(2x_t + \beta) + \phi'(t)\Big]}_{\text{drift}}\,\mathrm{d}t \;+\; \underbrace{(2x_t + \beta)\,\sigma}_{\text{diffusion}}\,\mathrm{d}W_t. \end{aligned} \tag{5}$$

The instantaneous volatility of the short rate is $(2x_t + \beta)\sigma$, which depends on the *state*. Using (4) to write $2x + \beta = \pm 2\sqrt{r - L}$, this is

$$\text{local vol of } r_t \;=\; 2\sigma\sqrt{r_t - L(t)}\,. \tag{6}$$

The rate's own volatility is a *square-root* function of how far the rate sits above its floor. This is the same square-root local volatility that the CIR model is famous for, but obtained here while keeping a tractable Gaussian factor underneath. Volatility that rises with the level of the rate is, by definition, a skew: high-strike options—struck where the rate is more volatile—carry more volatility than low-strike ones. The flat smile of Hull–White is gone, replaced by a shape the model generates on its own. We will see and price it in Section 6; for now the point is that *both* of Hull–White's gaps are closed by the same parabola: the floor in (4), the skew in (6).

3 The model, precisely

Collecting the pieces, the single-factor quadratic Gaussian model is

$$\boxed{\mathrm{d}x_t = -\kappa\, x_t\, \mathrm{d}t + \sigma\, \mathrm{d}W_t, \qquad r_t = x_t^2 + \beta\, x_t + \phi(t)} \tag{7}$$

with $x_0 = 0$. It carries the same two-layer spine as Hull–White, and it is worth naming the layers because everything downstream respects the split.

Layer one: the drift that fits the curve. The deterministic function $\phi(t)$ is free, an entire curve's worth of freedom. It does exactly one job: make the model reprice today's discount curve $P(0, T)$ *exactly*, for every maturity. We derive its explicit form in Section 4. As in Hull–White, this layer is not where the modelling lives; it is a calibration device that absorbs the shape of the initial curve so the dynamics need not.

Layer two: the parameters that carry the dynamics. Three constants remain: κ, σ, and β. The first two are inherited wholesale from Hull–White—mean reversion and factor volatility—and do the same work: they set the overall level and the term structure of volatility. The third, β, is new, and it is the price of admission to the smile. It positions the *operating point* $x \approx 0$ along the parabola relative to the vertex: a larger β pushes the vertex (the floor) further below the current rate, a smaller β brings the floor up toward it. Through (6) that distance controls how much the rate's volatility varies across levels—the slope of the skew—and, through (4), how low the floor sits. There is a genuine tension here, and it is worth stating early because it is structural, not a defect:

> *Volatility and the floor compete.* The rate's volatility, $2\sigma\sqrt{r - L}$, comes from the *slope* of the parabola at the operating point. A steeper slope (more volatility) requires the operating point to sit further from the vertex—i.e. a lower floor. To buy more volatility you must accept a lower floor; to raise the floor toward the current rate you must give up volatility.

This is the one-factor model's real limit, and a second factor (Section 10) relaxes it. For now we keep a single factor and a floor pinned just above zero.

A note on $\beta = 0$. The purest version of the model drops the linear term: $r = x^2 + \phi$, a rate that is the bare square of the factor. It has the cleanest floor ($L = \phi$, tracking the forward curve) but a degenerate volatility: at the operating point $x = 0$ the slope $2x + \beta = 0$ vanishes, so the at-the-money rate has almost no diffusion and the smile collapses to a near-singular valley. The linear term exists to move the operating point off the vertex and give the rate a sensible volatility. We keep $\beta > 0$ throughout, and treat $\beta = 0$ as the instructive limiting case it is.

4 From the factor to bond prices

In Hull–White the bond price was *exponential-affine*, $P = A\, e^{-Br}$: the log bond price was linear in the state. With a quadratic rate that affine form cannot hold; the right generalisation is the one the algebra forces on us—*exponential-quadratic.*

The ansatz and why it closes. A zero-coupon bond is the risk-neutral expectation of the discount factor,

$$P(t, T) = \mathbb{E}^Q\Big[\exp\Big(-\int_t^T r_s\, \mathrm{d}s\Big) \,\Big|\, \mathcal{F}_t\Big]. \tag{8}$$

The integrand $\int r\,\mathrm{d}s$ is a quadratic functional of the Gaussian path x, and the expectation of the exponential of a quadratic form in a Gaussian is itself the exponential of a quadratic. So we try

$$P(t,T) = \exp\big(-A(t,T)\,x_t^2 - B(t,T)\,x_t - C(t,T)\big), \tag{9}$$

and check it against the pricing PDE. By Feynman–Kac, P solves

$$\partial_t P - \kappa x\,\partial_x P + \tfrac{1}{2}\sigma^2\partial_{xx}P - r\,P = 0, \qquad P(T,T) = 1, \tag{10}$$

with $r = x^2 + \beta x + \phi$. Substituting (9), dividing through by P, and matching the coefficients of x^2, x^1 and x^0 separately (they must each vanish for the equation to hold at all x) gives three ordinary differential equations, solved backward from $A(T,T) = B(T,T) = C(T,T) = 0$:

$$\begin{aligned} \partial_t A &= 2\sigma^2 A^2 + 2\kappa A - 1, \\ \partial_t B &= \big(\kappa + 2\sigma^2 A\big)B - \beta, \\ \partial_t C &= \tfrac{1}{2}\sigma^2 B^2 - \sigma^2 A - \phi(t). \end{aligned} \tag{11}$$

The structure is exactly that of Hull–White's affine ODEs, with one decisive new term. The equation for A is a *Riccati* equation—it has the quadratic term $2\sigma^2 A^2$. That single nonlinearity is the whole difference between affine and quadratic models: it is what makes the bond price quadratic in the state, and it is the algebraic fingerprint of the squared rate. The equation for B is linear in B (driven by β), and C is a plain integral once A and B are known. The derivation is laid out term by term in Appendix A.

Constant coefficients, closed forms. Because κ, σ, β are constant, A, B and C depend on t and T only through the time-to-maturity $\tau = T - t$. Writing $A(t,T) = \hat{A}(\tau)$ and likewise for B, the Riccati equation $\hat{A}'(\tau) = 1 - 2\kappa\hat{A} - 2\sigma^2\hat{A}^2$ has the standard tanh solution, and in particular a finite long-maturity limit where $\hat{A}' = 0$:

$$A_\infty = \frac{-\kappa + \sqrt{\kappa^2 + 2\sigma^2}}{2\sigma^2}, \qquad B_\infty = \frac{\beta}{\sqrt{\kappa^2 + 2\sigma^2}}. \tag{12}$$

Figure 3 plots $A(0,T)$ and $B(0,T)$ against maturity: both rise from zero and saturate at (12), the quadratic coefficient A playing the role the affine $B(0,T) \to 1/a$ played in Hull–White.

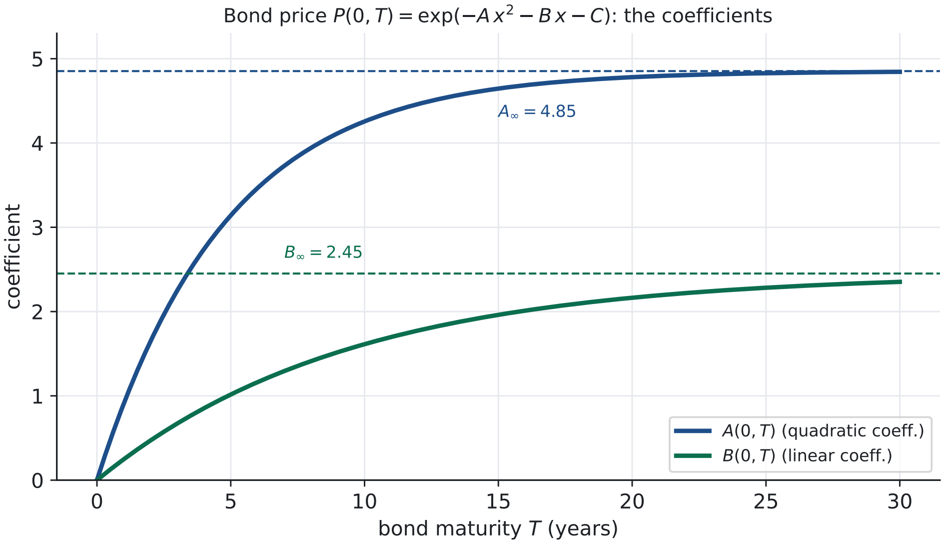

Figure 3: The exponential-quadratic bond coefficients $A(0,T)$ (quadratic) and $B(0,T)$ (linear), as functions of maturity, for $(\kappa, \sigma, \beta) = (0.10, 0.025, 0.26)$. Both grow from zero at $T = 0$ and saturate at the long-maturity limits (12), $A_\infty \approx 4.85$ and $B_\infty \approx 2.45$. The Riccati term $2\sigma^2 A^2$ is what bends A to a finite asymptote.

Fitting today's curve, exactly. Layer one now does its job. At $t = 0$ the factor starts at $x_0 = 0$, so (9) collapses to $P(0, T) = e^{-C(0,T)}$: only C matters for the initial curve. Setting this equal to the market bond $P(0, T) = \exp(-\int_0^T f(0, s)\, ds)$ and differentiating in T pins ϕ down completely. Because of the constant-coefficient structure the result is a clean, closed expression (derived in Appendix B):

$$\phi(T) \;=\; f(0,T) \;-\; \sigma^2 A(0,T) \;+\; \tfrac{1}{2}\sigma^2 B(0,T)^2. \tag{13}$$

Read it as Hull–White's $\alpha(t)$ = forward + convexity, generalised. The floor itself, $L(T) = \phi(T) - \beta^2/4$, is therefore the forward curve minus the model's own convexity terms minus $\beta^2/4$. Two checks confirm the construction. First, with (13) the model reprices the whole curve to within 3×10^{-4} basis points—machine-exact, as it must be. Second, for the parameters above the floor $L(t)$ stays positive for every maturity out to thirty years, dipping no lower than about 0.21%: the model's rates are guaranteed non-negative, the very thing Figure 1 showed Hull–White cannot promise.

5 The lower bound, and what it buys

The floor is not a soft tendency—it is a hard, deterministic barrier the rate provably never crosses. Figure 4 drives the contrast home with simulated paths. A Hull–White short rate, calibrated to the same curve, scatters symmetrically and sends a visible fraction of its paths below zero. The quadratic Gaussian rate, by construction (4), rides above the floor $L(t)$ at all times: the squared factor simply cannot reach down past the vertex.

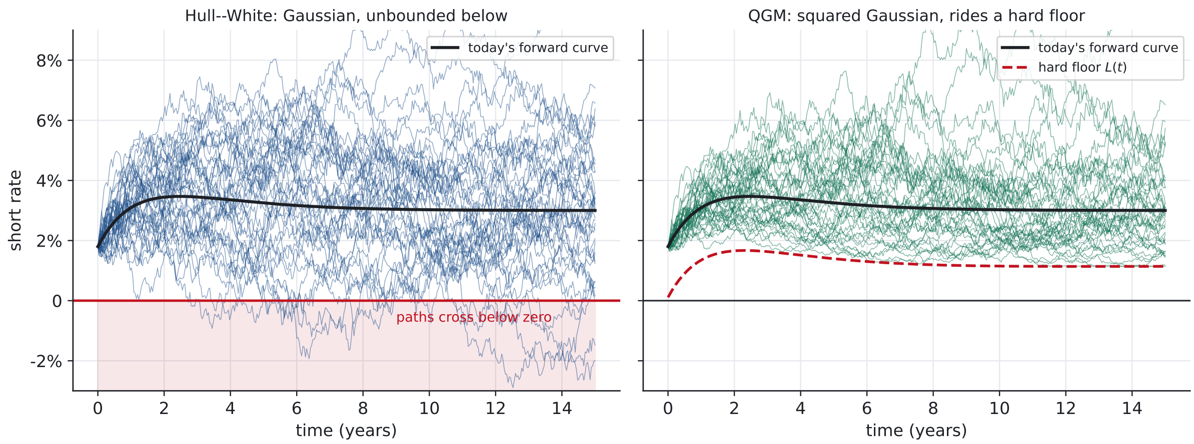

Figure 4: Simulated short-rate paths over fifteen years on the same initial curve. *Left:* Hull–White ($a = 0.10, \sigma = 1\%$). Being Gaussian, paths cross below zero (shaded). *Right:* the quadratic Gaussian model $(\kappa, \sigma, \beta) = (0.10, 0.025, 0.26)$. Every path stays above the hard, time-dependent floor $L(t)$ (red dashed); the rate is a squared Gaussian and cannot go lower.

The distributional picture in Figure 1 (left) is the same fact seen at a single horizon: the quadratic Gaussian density is the law of a (shifted, scaled) non-central χ^2 variable—squared Gaussian—which lives entirely on $[L, \infty)$ and is right-skewed, while the Hull–White density is the symmetric normal that spills below zero. In a normal-rate environment the difference is a modest left tail; in a low-rate environment, where the curve sits near the floor, the Gaussian tail below zero swells and the two models part company sharply. Any product whose payoff is sensitive to that tail—a coupon with a guaranteed minimum, a note that must return at least par—is precisely where the floor earns its keep (Section 9).

6 The smile the model makes

We met the mechanism in (6): the rate's local volatility, $2\sigma\sqrt{r-L}$, rises with the level of the rate. Figure 5 draws it. At the floor the volatility is zero—there is nowhere left to fall, so the rate goes quiet—and it climbs as a square root from there. A Hull–White rate, by contrast, has a constant volatility at every level: a flat line. State-dependent volatility is the necessary and sufficient condition for a non-flat smile, and the quadratic Gaussian model has it built in.

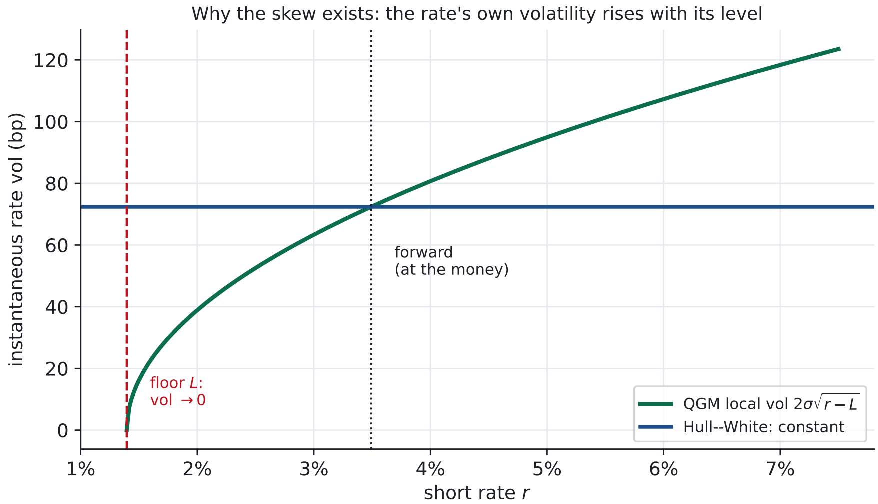

Figure 5: Why the skew exists. The instantaneous volatility of the short rate, $2\sigma\sqrt{r-L}$ (green), vanishes at the floor and grows with the rate's level; a Hull–White rate has constant volatility (blue). The two agree, by calibration, at the forward (at the money), but everywhere else the quadratic model assigns different volatility to different levels—which is a skew.

Turning the local volatility into an implied-volatility smile requires pricing caplets across strikes (Section 7), but the result is worth seeing first. Figure 6 (left) shows the implied normal volatility of a two-year caplet as a function of strike: the quadratic Gaussian model traces a clean, upward-sloping skew—low strikes (where the rate is near the floor and quiet) carry less volatility, high strikes more—while Hull–White is the flat line through the at-the-money point. This is the rates analogue of the equity volatility skew, and it is produced by the model's own dynamics, not bolted on.

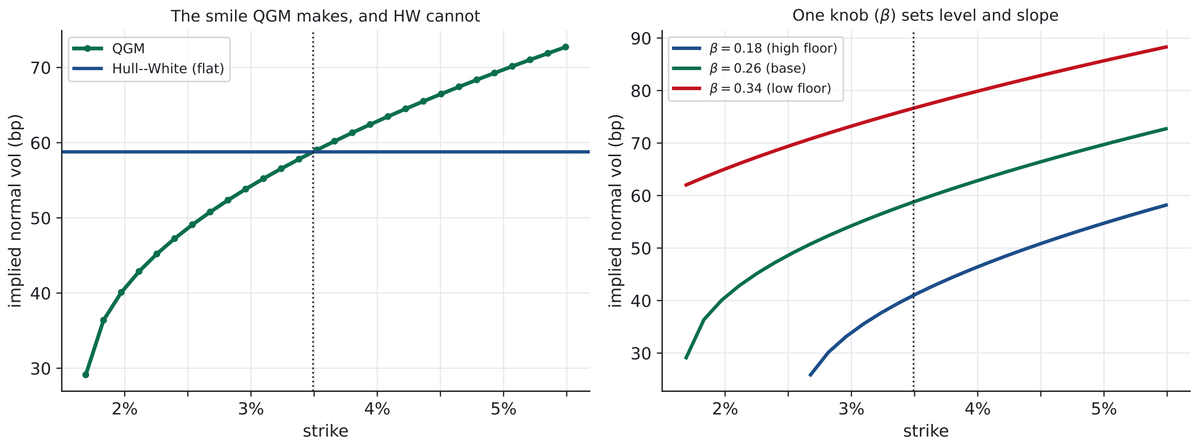

Figure 6: *Left:* the implied normal-volatility smile of a two-year caplet. The quadratic Gaussian model (green) makes a genuine skew; Hull–White (blue) is flat, matching only the at-the-money level. *Right:* the knob. Lowering β raises the floor toward the forward and steepens the skew while dropping its level; raising β lowers the floor, lifts the level and flattens the slope. One parameter controls both where the smile sits and how it tilts.

Figure 6 (right) shows what β does. Because β sets the distance from the operating point to the floor, it moves the whole smile up or down (a lower floor means a more volatile rate at the money) and changes its slope (a floor nearer the forward makes volatility vary more steeply across the accessible range). This is precisely the strike-dependence Hull–White lacked. It is, in spirit, the same device the Hull–White note flagged as the cheapest production fix—a displaced, $(r + s)$-style diffusion that interpolates between normal and lognormal dynamics—except that here it falls out of the squared-Gaussian construction rather than being added by hand, and it comes bundled with the floor. With a single factor there is one skew knob; a multi-factor version (Section 10) adds curvature and term-structure control over the smile.

7 Pricing the vanillas

The reason the smile is computable, and not merely describable, is that the Gaussian factor stays Gaussian under the right measure—so option prices collapse to one-dimensional integrals against a normal density. This section sets that up and prices a caplet and a swaption; the closed-form details are in Appendices C–D.

The factor is Gaussian under the forward measure. As in Hull–White, the clean way to price an option expiring at T_O is to switch to the T_O-forward measure, under which the numeraire is the bond $P(\cdot, T_O)$ and an option price is its forward value times a forward-measure expectation. The key fact, and the reason everything stays tractable, is:

> Under the T_O-forward measure, the factor x_{T_O} is *Gaussian*, $x_{T_O} \sim \mathcal{N}(m_f,\, v_f)$.

This is not obvious—the change of measure reweights paths by the bond $P(\cdot, T_O) = \exp(-Ax^2 - Bx - C)$, an exponential-quadratic in x—but a Gaussian reweighted by the exponential of a quadratic is still Gaussian, only with a shifted mean and a different variance. Appendix C derives the moments explicitly:

$$\begin{aligned} v_f &= \sigma^2 \int_0^{T_O} E(s)^2 \, \mathrm{d}s, \\ m_f &= -\sigma^2 \int_0^{T_O} E(s)\, B(s, T_O)\, \mathrm{d}s, \\ E(s) &= \exp\Big(-\int_s^{T_O} \big(\kappa + 2\sigma^2 A(u, T_O)\big) \mathrm{d}u\Big). \end{aligned} \tag{14}$$

The mean m_f is non-zero precisely because $\beta \neq 0$ (through B): the asymmetry the linear term injects is what makes the smile a skew rather than a symmetric smile.

A caplet. Recall from the Hull–White note that a caplet on the simple rate over $[T_O, T_B]$ with strike K is a put on the zero-coupon bond $P(T_O, T_B)$ struck at $X = 1/(1 + K\tau)$, scaled by $(1 + K\tau)$, where $\tau = T_B - T_O$. Its price is the forward bond value times a forward-measure expectation of the put payoff, and since x_{T_O} is Gaussian and $P(T_O, T_B)$ is an explicit function of it, this is a single integral:

$$\text{Caplet} = (1 + K\tau)\, P(0, T_O)\, \mathbb{E}_f\Big[\big(X - P(T_O, T_B \mid x_{T_O})\big)^+\Big], \qquad x_{T_O} \sim \mathcal{N}(m_f, v_f). \tag{15}$$

The integral has a closed form in terms of the normal CDF (Appendix D), but its *geometry* is the instructive part, and it differs from Hull–White in a way that is pure quadratic Gaussian.

The exercise set is a band. In Hull–White the bond price is monotone in the state—higher rate, cheaper bond—so a bond option is in the money on a half-line: one threshold, which is what Jamshidian's trick exploits. Here the rate is *quadratic* in x, so the bond price $\exp(-Ax^2 - Bx - C)$ is a bell-shaped function of x: it is cheap when $|x|$ is large in *either* direction, because rates are high in both wings of the parabola. The option is therefore in the money in two wings, separated by two thresholds (Figure 7). Algebraically, exercise requires $Ax^2 + Bx + (C + \ln X) > 0$, a quadratic inequality whose two roots are the thresholds. This costs nothing—a quadratic has a closed-form root—but it is the visible signature of the squared rate, and it is why the single-threshold Jamshidian decomposition does not transfer unchanged.

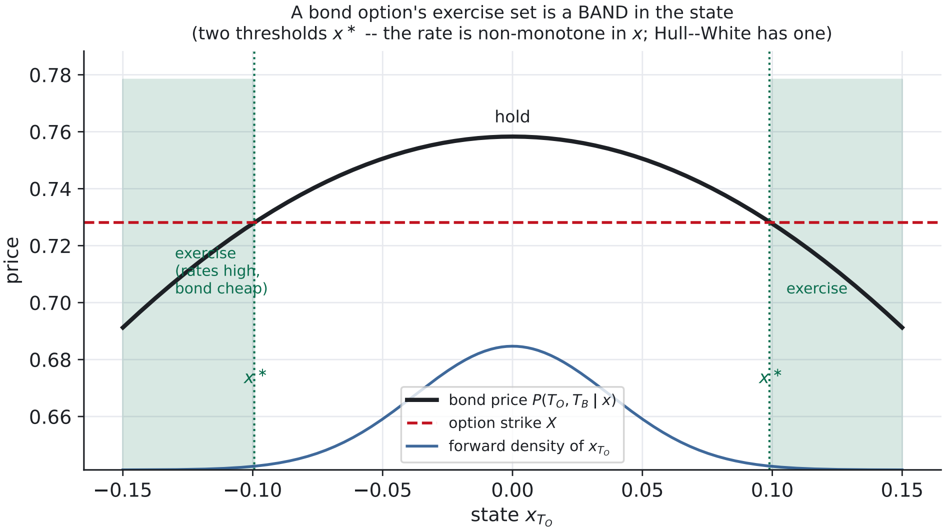

Figure 7: The exercise set of a bond option (the building block of a caplet) is a *band* in the state. Because the rate is quadratic in x, the bond price (black) is a bell: it falls below the strike X (red) in two wings, where rates are high. Exercise is therefore the region outside an interval—two thresholds x^* (green dotted), against Hull–White's single one. The forward density of x_{T_O} (blue) sets how much each wing contributes. Drawn for the symmetric $\beta = 0$ case so the bell is centred; a non-zero β merely shifts it sideways.

It works, and it is right. For $(\kappa, \sigma, \beta) = (0.10, 0.025, 0.26)$ on the note's curve, the at-the-money two-year caplet on the $2y \to 2.5y$ rate prices to 15.37 bp by the forward-measure integral, against 15.41 ± 0.09 bp from an independent Monte Carlo simulation of (7)—agreement within the Monte Carlo error. Sweeping the strike and inverting each price to an implied normal volatility produces the skew of Figure 6: from about 40 bp at a strike 1.5% below the forward, through 59 bp at the money, to 70 bp 1.5% above. That is a 30-bp range of implied volatility across strike—exactly the dimension Hull–White cannot move.

Swaptions. A European swaption is an option on the fixed-leg coupon bond, struck at par. Under the forward measure the coupon bond $\sum_i c_i\, P(T_O, T_i)$ is a sum of exponential-quadratics in the single Gaussian x_{T_O}—a smooth, bell-shaped function of one variable—so the swaption is again a one-dimensional Gaussian integral, evaluated by quadrature in milliseconds. As a check, a five-into-five at-the-money payer swaption prices to 143.8 bp by quadrature against 142.8 ± 0.7 bp by Monte Carlo. The exercise region is once more a band in x rather than a half-line, for the same reason as the caplet. (For genuinely high-dimensional baskets one falls back on the same semi-analytic approximations used for G2++ in the Hull–White note; with one factor the integral is exact.)

8 Calibration

The two-layer split organises calibration exactly as it did for Hull–White, with one extra parameter and one extra target.

Layer one is free. The curve fit is not an optimisation. Equation (13) hands $\phi(t)$ in closed form from today's forwards and the already-known coefficients A, B; the curve is reproduced to machine precision for *any* choice of the dynamic parameters. So ϕ is never something one searches for—it is recomputed, instantly, whenever (κ, σ, β) move.

Layer two: three parameters, two jobs. The dynamic parameters (κ, σ, β) are fit to option prices. Their division of labour is the clean part of the story:

- σ sets the overall *level* of volatility—how big the rate's moves are;
- κ shapes the *term structure* of volatility—how that level decays with option expiry, through the mean-reversion of the factor, exactly as in Hull–White;
- β controls the *smile*—its slope across strike and, jointly with the others, the floor.

The contrast with Hull–White is the whole point. There, (a, σ) were spent entirely on the at-the-money term structure, leaving nothing for the smile; the smile came out flat whether the market liked it or not. Here, (κ, σ) do that same at-the-money job and β is left over to match the skew. The natural calibration target is therefore not a single at-the-money strip but a *slice of the smile*: at-the-money volatilities across expiry to set (κ, σ), plus a away-from-the-money volatility (a risk-reversal, or a single out-of-the-money strike) to set β. One disciplines the level and term structure; the other disciplines the shape. As always, the fit is staged—(κ, σ) first on the at-the-money surface, then β on the skew—because the at-the-money level is nearly independent of β, which keeps the optimisation well conditioned.

What one factor cannot reach. A single β buys a single, roughly maturity-stable skew. It cannot make the skew steepen or flatten with expiry, nor add curvature (a true convex smile, with both wings up) independently of the slope. Those need more shape in the volatility function—more factors, or a state-dependent or stochastic volatility on top—which is the subject of Section 10 and the closing comparison.

9 What it prices that one factor cannot

The abstract gains—a floor and a skew—turn into concrete prices on exactly the products where Hull–White was silently wrong. In each case below the quadratic Gaussian model and a one-factor Hull–White model are calibrated to the *same* at-the-money volatility, so they agree on at-the-money vanillas and differ only where the skew or the floor bites. All numbers use $(\kappa, \sigma, \beta) = (0.10, 0.025, 0.26)$ on the note's curve and are computed by the accompanying pricer.

1. Out-of-the-money caps and floors. This is the cleanest case, because it is the smile itself. A cap or floor struck away from the money is priced off the wing of the smile, and Hull–White's wings are wrong by construction—it has only the at-the-money level to offer. Pricing the two-year caplet at strikes 1% either side of the forward, against a flat-volatility (Hull–White) model matched at the money:

	quadratic Gaussian	flat Gaussian (HW)	difference
caplet, $K = F - 1\%$ (in the money)	47.36 bp	48.48 bp	−1.12 bp
caplet, $K = F + 1\%$ (out of the money)	3.20 bp	2.15 bp	+1.05 bp

The skew makes the out-of-the-money caplet almost 50% richer than the flat model says, and the in-the-money one cheaper. A book of out-of-the-money caps and floors marked on a flat Hull–White surface carries exactly this error on every line—small per option, but one-signed and cumulative across a portfolio.

2. The swaption skew. The same effect, scaled up to swaptions, is what a desk trading the swaption cube actually faces: receiver volatilities richer than payer volatilities (or vice versa), and wings that bow away from the at-the-money level. Hull–White can be re-struck to any single point on that cube but cannot hold the whole strike dimension at once; the quadratic Gaussian model fits a skewed slice with its one β. Mark a portfolio of low-strike receivers and high-strike payers, and the flat model misprices the two wings in opposite directions—a structural skew risk that does not net out.

3. Convexity and CMS. A constant-maturity-swap payment depends on a swap rate paid on a date that does not match the rate's own natural payment schedule, and the resulting *convexity adjustment* is, by static-replication, an integral of swaption prices *across all strikes*. It therefore depends on the entire smile, not just its at-the-money value. A flat-smile model gets the at-the-money point right and the integral wrong; the sign and size of the CMS convexity adjustment are a direct read-out of the skew. The wing differences in the table above are the building blocks of that integral: a model that misprices each out-of-the-money option by a basis point misprices the convexity adjustment that sums them. CMS caps, floors and spread coupons inherit the error. The quadratic Gaussian model, carrying a calibrated skew, prices the replication integral consistently with the vanilla wings it was fit to.

4. Floored and capital-guaranteed structures. Here it is the lower bound, not the skew, that does the work. Consider a note whose coupon is a rate with a guaranteed minimum, or a structure that promises return of at least par—both are long a floor at or near zero. Under Hull–White that floor has value coming partly from the Gaussian tail of rates below zero, an artefact of the model rather than a feature of the world; the guarantee is mispriced, and in a low-rate environment materially so. The quadratic Gaussian model places *no* mass below its floor $L(t)$, so a guarantee at or below that level is valued on economically meaningful dynamics. The cleanest case is exact: a floorlet struck at zero is worth *precisely nothing* in the quadratic Gaussian model—the short rate never falls below $L(t) > 0$, so the realised forward rate is always positive and the floor never pays—whereas Hull–White assigns it a strictly positive value drawn entirely from its sub-zero tail. The closer the curve sits to zero, the larger that spurious value grows—and the more the choice of model is really a choice about whether negative rates are a thing your product should pay for.

The common thread is that each product reaches into a part of the distribution Hull–White gets wrong: the wings (the skew) or the floor (positivity). Where a product lives at the money and away from the floor, Hull–White remains an excellent, cheaper choice—the quadratic Gaussian model earns its extra complexity precisely on the structures above.

10 More factors: smile and twist together

A single factor, we saw, has a genuine limit: one skew knob, and a volatility–floor trade-off. It also inherits Hull–White's other one-factor limitation untouched—every rate on the curve is driven by the one Brownian motion, so distant forward rates move in lockstep and the curve cannot genuinely twist. The Hull–White note fixed that with a second Gaussian factor (G2++). The quadratic Gaussian model generalises in exactly the same direction, and the generalisation is what makes the family genuinely useful in production.

Let $x_t \in \mathbb{R}^n$ be a vector Ornstein–Uhlenbeck process and let the short rate be a quadratic *form*:

$$\mathrm{d}x_t = -K x_t \,\mathrm{d}t + \Sigma \,\mathrm{d}W_t, \qquad r_t = x_t^\top \Gamma \, x_t + \psi^\top x_t + \phi(t), \tag{16}$$

with Γ a symmetric (positive-semidefinite) matrix. Every structural feature survives the move to many dimensions. Bond prices stay exponential-quadratic, $P = \exp(-x^\top \mathbf{A}\, x - \mathbf{b}^\top x - C)$, where the matrix $\mathbf{A}(t,T)$ now solves a *matrix* Riccati equation and $\mathbf{b}$ a linear system—the same (11) with scalars promoted to matrices. The factor vector stays Gaussian under the forward measure, so vanillas remain low-dimensional integrals. And the curve is still fit by the deterministic ϕ.

What the extra factors buy is two things at once, which is the point. The quadratic form still delivers the *floor and the skew* of the single-factor model. The *multiple* factors deliver *decorrelation and twist*, exactly as the second Gaussian factor did in G2++: rates at different maturities can now move by different amounts *and* in different directions, the volatility term structure can hump, and the skew can acquire its own term structure and curvature. A two- or three-factor quadratic Gaussian model is thus, roughly, "G2++ with a smile"—it prices CMS spread options and steepeners (which need decorrelation) *and* their skew (which needs state-dependent volatility) in one consistent, arbitrage-free, curve-fitting model. The cost is the bookkeeping of matrix Riccati equations and a higher-dimensional calibration; the reward is a single model that covers the products of both this note and Part II of the Hull–White note. We leave the matrix algebra to Appendix A and the references.

11 Choosing a model, and where this one sits

It helps to place the quadratic Gaussian model on the same map as the others. All of these share the two-layer spine—a drift slaved to the curve, volatility parameters for the dynamics—and differ in what their volatility layer can express.

model	rates	floor?	smile/skew?	decorr.?
1-factor Hull–White	Gaussian	no	flat	no
G2++ (2-factor HW)	Gaussian	no	flat	yes
1-factor quadratic Gaussian	squared Gaussian	yes	skew	no
n-factor quadratic Gaussian	squared Gaussian	yes	skew + term struct.	yes
local-/stochastic-vol	general	varies	full, dynamic	yes

Read across the rows and the quadratic Gaussian model is the smallest step from Hull–White that closes *both* of its gaps: it keeps the Gaussian factor, the exact curve fit, the analytic bonds and the integral-form vanillas, and pays for the floor and the skew only with a Riccati equation in place of a linear one. It is not the end of the road—a genuinely dynamic smile, or independent control of slope and curvature across expiries, still wants a richer volatility specification (local-volatility Cheyette models, stochastic-volatility extensions, full market models). But as the tractable core that those models extend, the quadratic Gaussian model occupies the same place on the rates side that a displaced or CEV diffusion occupies on the equity side: the simplest thing past Black that has a skew and respects a boundary. If Hull–White is the Black–Scholes of rates, the quadratic Gaussian model is its displaced-diffusion cousin—and, in its multi-factor form, a complete and practical pricing model in its own right.

A The bond ODEs

We derive (11). With $r = x^2 + \beta x + \phi$, Feynman–Kac gives the bond PDE (10). Write the ansatz $P = \exp(-Ax^2 - Bx - C)$ with A, B, C functions of t (suppressing T). The derivatives are

$$\begin{aligned} \partial_x P &= -(2Ax + B)\,P, \\ \partial_{xx} P &= \left[(2Ax + B)^2 - 2A\right] P, \\ \partial_t P &= -(A'x^2 + B'x + C')\,P, \end{aligned} \tag{17}$$

where $' = \partial_t$. Substituting into (10) and dividing by P,

$$-(A'x^2 + B'x + C') - \kappa x \left[-(2Ax + B)\right] + \tfrac{1}{2}\sigma^2\left[(2Ax + B)^2 - 2A\right] - (x^2 + \beta x + \phi) = 0. \tag{18}$$

Expanding $-\kappa x(-(2Ax+B)) = 2\kappa Ax^2 + \kappa Bx$ and $\frac{1}{2}\sigma^2(2Ax+B)^2 = 2\sigma^2A^2x^2 + 2\sigma^2ABx + \frac{1}{2}\sigma^2B^2$, and collecting powers of x:

$$\begin{aligned} x^2 :&\quad -A' + 2\kappa A + 2\sigma^2A^2 - 1 = 0, \\ x^1 :&\quad -B' + \kappa B + 2\sigma^2AB - \beta = 0, \\ x^0 :&\quad -C' + \tfrac{1}{2}\sigma^2B^2 - \sigma^2A - \phi = 0. \end{aligned} \tag{19}$$

Rearranging gives (11), with terminal conditions $A = B = C = 0$ at $t = T$ from $P(T,T) = 1$. The x^2 equation is the Riccati equation; its quadratic term $2\sigma^2A^2$ is the sole nonlinearity and the origin of the quadratic (rather than affine) bond price. For the vector model (16) the identical bookkeeping with $A \to \mathbf{A}$ (matrix), $B \to \mathbf{b}$ (vector) yields a matrix Riccati equation $\mathbf{A}' = 2\mathbf{A}\Sigma\Sigma^\top\mathbf{A} + \mathbf{A}K + K^\top\mathbf{A} - \Gamma$ and a linear equation for $\mathbf{b}$.

B The exact curve fit

At $t = 0$, $x_0 = 0$, so $P(0,T) = e^{-C(0,T)}$. From (11), integrating C' backward with $C(T,T) = 0$,

$$C(0,T) = \int_0^T \Big(\sigma^2A(s,T) - \tfrac{1}{2}\sigma^2B(s,T)^2 + \phi(s)\Big)\mathrm{d}s. \tag{20}$$

Setting $C(0,T) = \int_0^T f(0,s)\,\mathrm{d}s$ (the market curve) and differentiating in T,

$$f(0,T) = \phi(T) + \frac{\partial}{\partial T}\int_0^T \Big(\sigma^2A(s,T) - \tfrac{1}{2}\sigma^2B(s,T)^2\Big)\mathrm{d}s, \tag{21}$$

the boundary terms $A(T,T) = B(T,T) = 0$ dropping out. Because the coefficients are constant, $A(s,T) = \hat{A}(T-s)$ and $B(s,T) = \hat{B}(T-s)$, so $\int_0^T A(s,T)\,\mathrm{d}s = \int_0^T \hat{A}(u)\,\mathrm{d}u = \int_0^T A(0,s)\,\mathrm{d}s$, and likewise for B^2; differentiating these in T returns the integrand at $u = T$, namely $A(0,T)$ and $B(0,T)^2$. Hence

$$\phi(T) = f(0,T) - \sigma^2A(0,T) + \tfrac{1}{2}\sigma^2B(0,T)^2, \tag{22}$$

which is (13). The floor follows as $L(T) = \phi(T) - \beta^2/4$.

C The forward-measure law of the factor

We show x_{T_O} is Gaussian under the T_O-forward measure and derive (14). Consider the discounted moment-generating object

$$\psi(t,x;p) = \mathbb{E}^Q\Big[\exp\Big(-\int_t^{T_O} r_s\,\mathrm{d}s + p\,x_{T_O}\Big)\,\Big|\,x_t = x\Big]. \tag{23}$$

It solves the same PDE (10) as the bond but with terminal value e^{px}, so it is exponential-quadratic, $\psi = \exp(-\mathcal{A}x^2 - \mathcal{B}x - \mathcal{C})$, with the *same* $\mathcal{A}, \mathcal{B}, \mathcal{C}$ ODEs and terminal conditions $\mathcal{A}(T_O) = 0$, $\mathcal{B}(T_O) = -p$, $\mathcal{C}(T_O) = 0$. Since the $\mathcal{B}$ equation is linear, $\mathcal{B}(s;p) = B(s,T_O) - p\,E(s)$ where E solves the homogeneous part with $E(T_O) = 1$, giving $E(s) = \exp(-\int_s^{T_O}(\kappa + 2\sigma^2 A)\,\mathrm{d}u)$. The forward-measure MGF is $M_f(p) = \psi(0,0;p)/P(0,T_O)$, and evaluating $\mathcal{C}(0;p) - \mathcal{C}(0;0)$ (the only p-dependence at $x_0 = 0$) gives

$$\ln M_f(p) = -\sigma^2\Big(\int_0^{T_O} EB\,\mathrm{d}s\Big)p + \tfrac{1}{2}\sigma^2\Big(\int_0^{T_O} E^2\,\mathrm{d}s\Big)p^2. \tag{24}$$

This is quadratic in p, so x_{T_O} is Gaussian, with mean and variance the coefficients of p and $\frac{1}{2}p^2$—which is (14). With $\beta = 0$, $B \equiv 0$ and the mean vanishes: the smile is then symmetric.

D The caplet in closed form

With $\beta = 0$ (so $B(T_O, T_B) = 0$ and $m_f = 0$) the integral (15) is fully explicit; the general case adds the shift m_f and a linear term and is the same calculation. Write $A = A(T_O, T_B)$, $C = C(T_O, T_B)$, $v = v_f$, and the bond at expiry as $P = e^{-Ax^2 - C}$. Exercise ($P < X$) requires $x^2 > q := (-C - \ln X)/A$; if $q \le 0$ the option is always in the money, otherwise exercise is $|x| > x^* = \sqrt{q}$. Then

$$\mathbb{E}_f\big[(X - P)^+\big] = X \cdot 2\,\Phi\Big(-\tfrac{x^*}{\sqrt{v}}\Big) \;-\; \frac{e^{-C}}{\sqrt{2v\,a_2}}\,\mathrm{erfc}\big(x^*\sqrt{a_2}\big), \qquad a_2 = A + \tfrac{1}{2v}, \tag{25}$$

the first term being the strike times the probability of landing in the two exercise wings, the second the forward-measure value of the bond over the same region (a Gaussian times an exponential-quadratic is an unnormalised Gaussian, integrated to a complementary error function). Multiplying by $(1 + K\tau)P(0, T_O)$ gives the caplet. The two thresholds $\pm x^*$ are the band of Figure 7; with $\beta \ne 0$ they become the two roots of $Ax^2 + Bx + (C + \ln X) = 0$.

References

[1] D. Beaglehole, M. Tenney. General solutions of some interest rate-contingent claim pricing equations. *Journal of Fixed Income*, 1991.

[2] F. Jamshidian. Bond, futures and option evaluation in the quadratic interest rate model. *Applied Mathematical Finance*, 3:93–115, 1996.

[3] N. El Karoui, A. Frachot, H. Geman. A note on the behaviour of long-term zero-coupon rates in a no-arbitrage framework. Working paper, 1992 (squared Gaussian / quadratic models).

[4] L. C. G. Rogers. The potential approach to the term structure of interest rates and foreign exchange rates. *Mathematical Finance*, 7:157–176, 1997.

[5] D.-H. Ahn, R. F. Dittmar, A. R. Gallant. Quadratic term structure models: theory and evidence. *Review of Financial Studies*, 15:243–288, 2002.

[6] M. Leippold, L. Wu. Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis*, 37:271–295, 2002.

[7] A. Pelsser. A tractable yield-curve model that guarantees positive interest rates. *Review of Derivatives Research*, 1:269–284, 1997.

[8] L. B. G. Andersen, V. V. Piterbarg. *Interest Rate Modeling*. Atlantic Financial Press, 2010.

[9] D. Brigo, F. Mercurio. *Interest Rate Models—Theory and Practice*. Springer, 2nd ed., 2006.

[10] D. Filipović. *Term-Structure Models: A Graduate Course*. Springer, 2009.

[11] J. Hull, A. White. Pricing interest-rate-derivative securities. *Review of Financial Studies*, 3:573–592, 1990.