Interest Rate Models — Lecture 9 Notes

Sections 3.4–3.5: CIR extensions and the Black–Karasinski model

1 Why the natural CIR extension is not as clean as Hull–White

Hull–White extended Vasicek by making the drift level time dependent. The analogous CIR idea is

$$dr(t) = \big[\vartheta(t) - a(t)r(t)\big]\,dt + \sigma(t)\sqrt{r(t)}\,dW(t). \tag{3.51}$$

The intention is attractive: keep CIR's square-root volatility and positive-rate behavior, while using time-dependent coefficients to fit the initial curve. The problem is that the pleasant CIR tractability mostly disappears.

Even in the simpler case

$$dr(t) = \big[\vartheta(t) - ar(t)\big]\,dt + \sigma\sqrt{r(t)}\,dW(t), \tag{3.52}$$

where only $\vartheta(t)$ is time dependent, there is no simple analytical formula for $\vartheta(t)$ from the observed initial yield curve. More importantly, the positivity condition becomes fragile. In the constant-parameter CIR model, $2a\theta > \sigma^2$ keeps the origin inaccessible. If the mean level is now forced to move with the market curve, a low market curve can imply a low fitted $\vartheta(t)$. Then the square-root process may approach zero, and the diffusion coefficient $\sqrt{r(t)}$ becomes numerically and mathematically delicate.

Bond prices still have an affine-looking form,

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)},$$

but the Riccati equation for B is explicitly solvable only in special cases. Thus the model may require numerical procedures even before pricing options. This is why this direct time-inhomogeneous CIR extension is much less convenient than the Gaussian Hull–White extension.

A related warning is that square-root models often leave the current short rate $r(t)$ explicitly inside long-dated option prices, not only implicitly through market-observable bond prices like $P(t,T)$. For long-dated pricing and hedging, this extra dependence on the instantaneous rate can be undesirable.

2 The Black–Karasinski idea

Black–Karasinski takes a different route to avoid negative short rates. Instead of making $r(t)$ Gaussian or square-root, it makes $\ln r(t)$ Gaussian mean-reverting:

$$d\ln r(t) = \big[\theta(t) - a(t)\ln r(t)\big]\,dt + \sigma(t)dW(t), \qquad r(0) = r_0 > 0. \tag{3.53}$$

The common simplified version fixes a and σ as constants and uses $\theta(t)$ to fit the initial curve:

$$d\ln r(t) = \big[\theta(t) - a\ln r(t)\big]\,dt + \sigma dW(t). \tag{3.54}$$

The key modeling point is

$$r(t) = e^{\ln r(t)} > 0.$$

So even if the log-rate becomes very negative, the short rate itself remains positive. This is the same basic positivity logic as Exponential Vasicek, but now with a time-dependent drift function available for curve fitting.

Applying Ito's lemma gives the short-rate dynamics directly:

$$dr(t) = r(t)\left[\theta(t) + \frac{\sigma^2}{2} - a\ln r(t)\right]dt + \sigma r(t)dW(t).$$

The parameter a controls the speed at which the log-rate is pulled toward its deterministic center. The parameter σ is the instantaneous proportional volatility of the short rate, meaning the volatility of $dr(t)/r(t)$, not the caplet or swaption Black volatility.

3 Conditional moments and why they matter

Solving the log-rate equation gives, for $s \le t$,

$$r(t) = \exp\left(\ln r(s) e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\theta(u)du + \sigma \int_s^t e^{-a(t-u)}dW(u)\right).$$

Conditional on $\mathcal{F}_s$, the exponent is normal, so $r(t)$ is lognormal. The first two conditional moments are

$$E_s[r(t)] = \exp\left(\ln r(s)e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\theta(u)du + \frac{\sigma^2}{4a}\big(1 - e^{-2a(t-s)}\big)\right),$$

$$E_s[r(t)^2] = \exp\left(2\ln r(s)e^{-a(t-s)} + 2\int_s^t e^{-a(t-u)}\theta(u)du + \frac{\sigma^2}{a}\big(1 - e^{-2a(t-s)}\big)\right).$$

These matter because a lognormal distribution is usually understood through the mean and variance of the underlying normal exponent. Once those are known, expectations of powers of $r(t)$, and therefore basic moment checks in a tree or simulation, are controlled.

Define the deterministic center of the log-rate by

$$\alpha(t) = \ln(r_0)e^{-at} + \int_0^t e^{-a(t-u)}\theta(u)du. \tag{3.55}$$

Then

$$r(t) = e^{\alpha(t)+x(t)}, \tag{3.56}$$

where $x(t)$ is the zero-mean stochastic Gaussian part. This decomposition is the bridge to the tree construction.

4 The Black–Karasinski trinomial tree

BK has no closed-form bond price or bond-option formula, so pricing is done numerically. The tree is built in the same spirit as the Hull–White tree, but the final rate at each node is exponential rather than additive.

1. Build the recombining trinomial tree for the normal state variable x, with nodes

   $$x_{i,j} = j\Delta x_i.$$

   The movement probabilities p_u, p_m, p_d and central node choice are inherited from the Hull–White tree construction.
2. At each time layer t_i, find a common deterministic displacement α_i so that the tree reproduces the market zero-coupon bond price $P^M(0, t_{i+1})$. In BK the node short rate is

   $$r_{i,j} = \exp(x_{i,j} + \alpha_i).$$

3. Let $Q_{i,j}$ be the time-0 value of a security paying 1 if node (i, j) is reached and 0 otherwise. Once α_i is known, propagate state prices by discounting with the BK node rate:

   $$Q_{i+1,j} = \sum_h Q_{i,h}q(h,j)\exp\big[-\exp(\alpha_i + h\Delta x_i)\Delta t_i\big].$$

4. Choose α_i by solving

   $$P^M(0, t_{i+1}) - \sum_{j=\underline{j}_i}^{\bar{j}_i} Q_{i,j}\exp\big[-\exp(\alpha_i + j\Delta x_i)\Delta t_i\big] = 0.$$

   This is the exact-fitting step: the deterministic shift is not guessed; it is solved so the tree prices the next market discount factor correctly.

Summary

The direct time-dependent CIR extension tries to preserve square-root positivity but loses much of CIR's analytical tractability and can struggle with positivity when the fitted mean level follows a low market curve. Black–Karasinski avoids negative rates by modeling the logarithm of the short rate. The price is tractability: bonds and options are not closed-form, so the model is implemented through a calibrated trinomial tree. The crucial tree formula is $r_{i,j} = \exp(x_{i,j} + \alpha_i)$: build a normal tree for x, solve the time-layer shifts α_i, then exponentiate to get positive short rates.