IR Simulation Model Validation Notes

Findings Term-Structure Diagnostics

Amine Badii

December 19, 2025

Contents

1 Scope and objective

This document substantiates key validation findings for the interest-rate simulation framework used for PFE, with emphasis on pathwise arbitrage diagnostics and benchmarking against an arbitrage-free reference model (HW1F).

2 Finding 1 — Absence of Pathwise Arbitrage-Free Term-Structure Dynamics

2.1 Finding statement

The IR simulation model does not enforce arbitrage-free term-structure relationships along individual simulation paths. Zero rates at each maturity pillar are simulated independently, with cross-tenor dependence introduced solely through correlations between latent stochastic drivers. As a result, static and dynamic no-arbitrage identities governing discount factors and forward-rate reconstruction may be violated along individual paths, particularly at long horizons and under elevated volatility conditions.

2.2 Model features giving rise to the finding

The finding arises directly from the structural design of the model:

- Each maturity pillar is driven by a dedicated stochastic process calibrated independently to market-implied variance targets.
- Cross-tenor dependence is introduced via correlations between drivers rather than through a unified term-structure factor model.
- No short-rate representation or arbitrage-free drift restriction (e.g. HJM consistency) is imposed.

While this construction ensures marginal calibration accuracy at each tenor, it does not guarantee global coherence of the simulated yield curve along individual paths.

2.3 Risk implications for PFE

The absence of arbitrage-free dynamics may lead to:

- Violations of static discount-factor monotonicity across maturities.
- Reduced cross-tenor smoothness, resulting in locally implausible curve shapes.
- Inconsistencies in discount-factor reconstruction over time, which may accumulate with maturity and affect long-dated exposure profiles.

These effects are expected to be more pronounced for long-horizon PFE metrics and portfolios with significant sensitivity to the shape of the yield curve.

2.4 Validation strategy and diagnostics

Three complementary diagnostics are implemented:

1. **Static discount-factor monotonicity** across maturities.
2. **Cross-tenor smoothness** using a kink index (local curvature proxy).
3. **Pathwise discount-factor multiplicative consistency** via a wedge metric.

All diagnostics are benchmarked against an arbitrage-free Hull–White one-factor (HW1F) model. Together, these diagnostics provide complementary coverage of static, cross-sectional, and dynamic arbitrage constraints.

2.5 Empirical evidence and plots

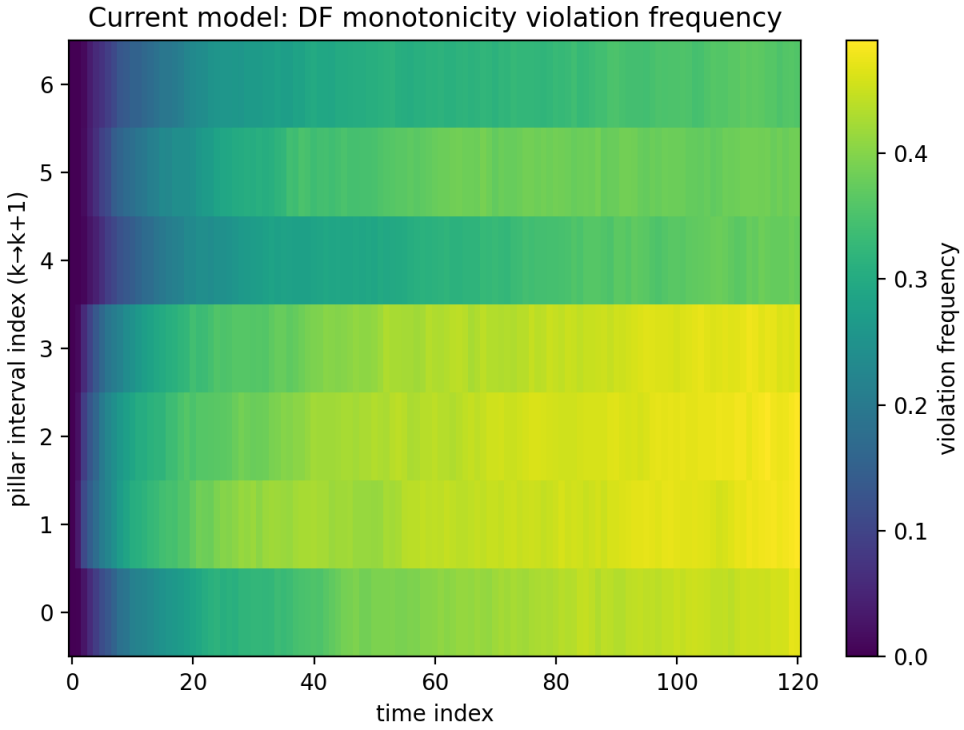

Figure 1: Current model: frequency of discount-factor monotonicity violations across time and pillar intervals.

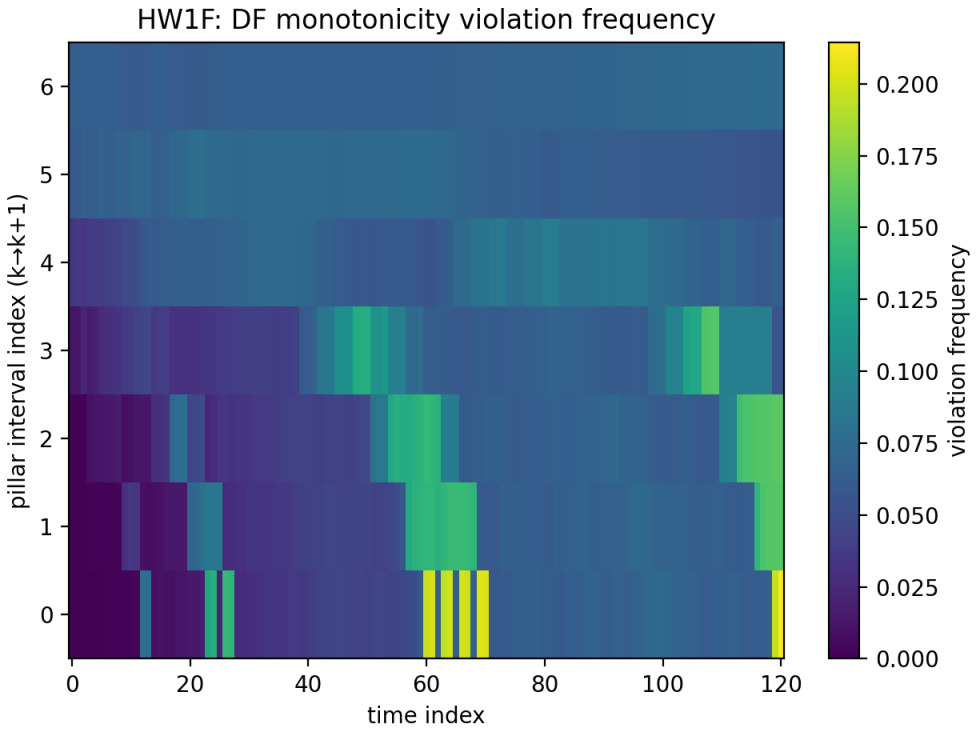

Figure 2: HW1F benchmark: frequency of discount-factor monotonicity violations across time and pillar intervals.

Discount-factor monotonicity violations.

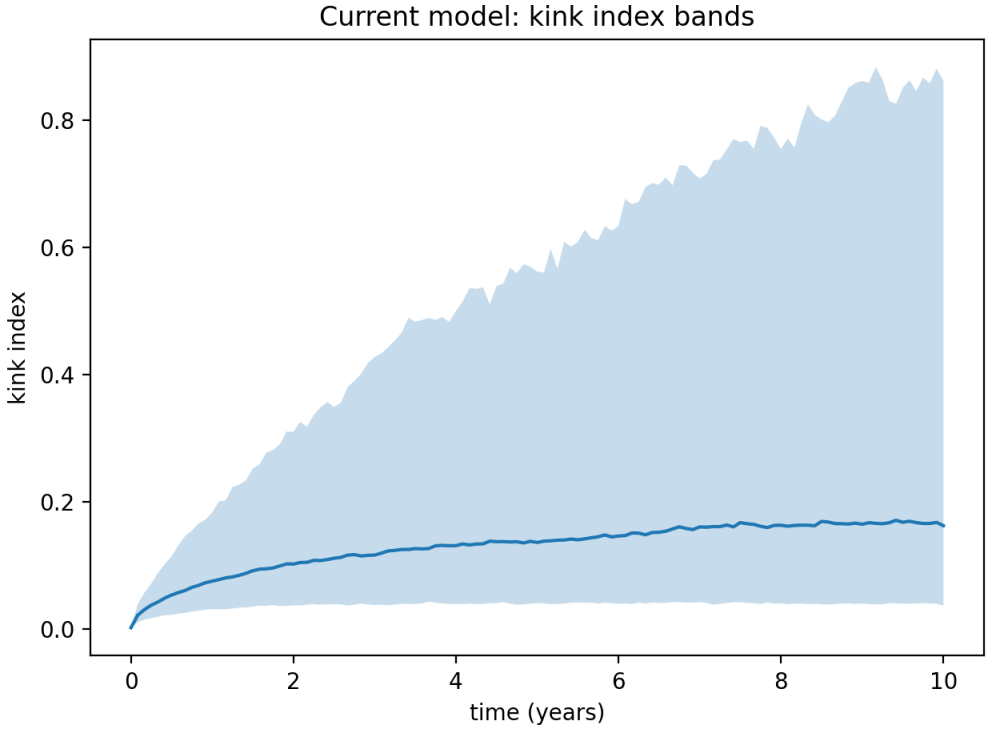

Figure 3: Current model: kink index bands over time (median and dispersion band).

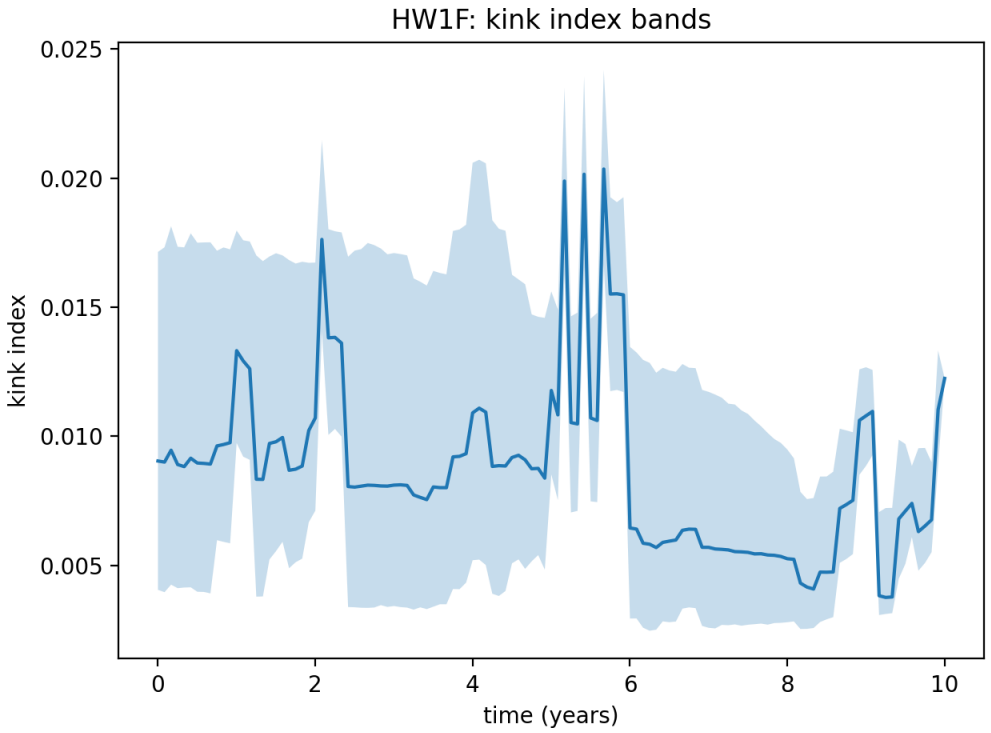

Figure 4: HW1F benchmark: kink index bands over time (median and dispersion band).

Kink index bands (cross-tenor smoothness).

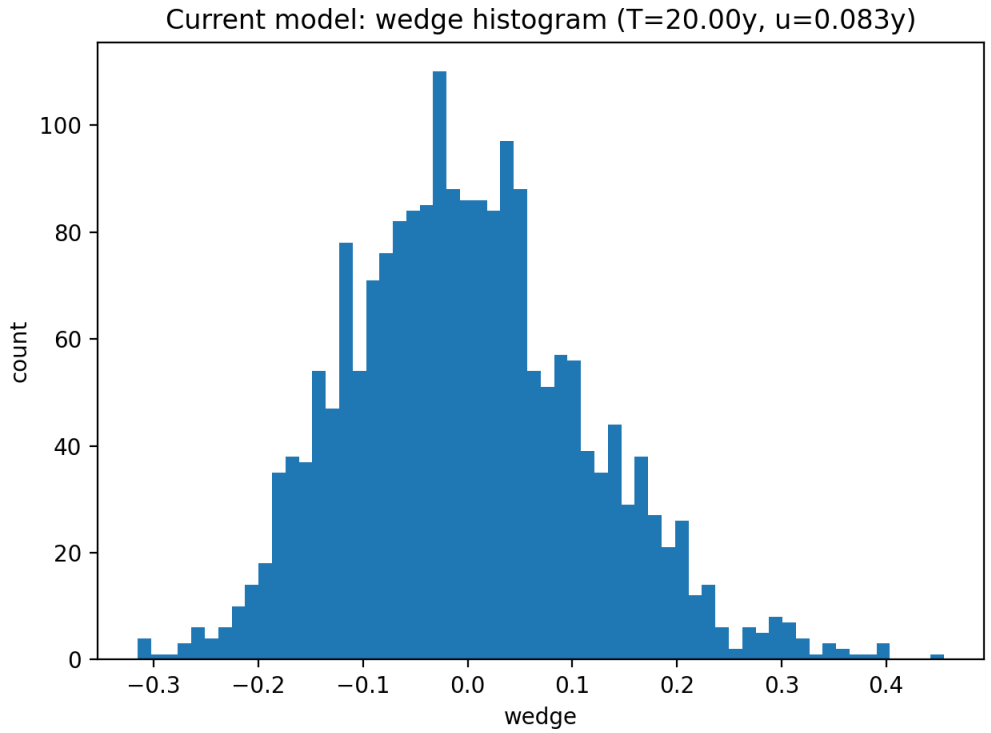

Figure 5: Current model: wedge histogram for a representative (T, u) pair.

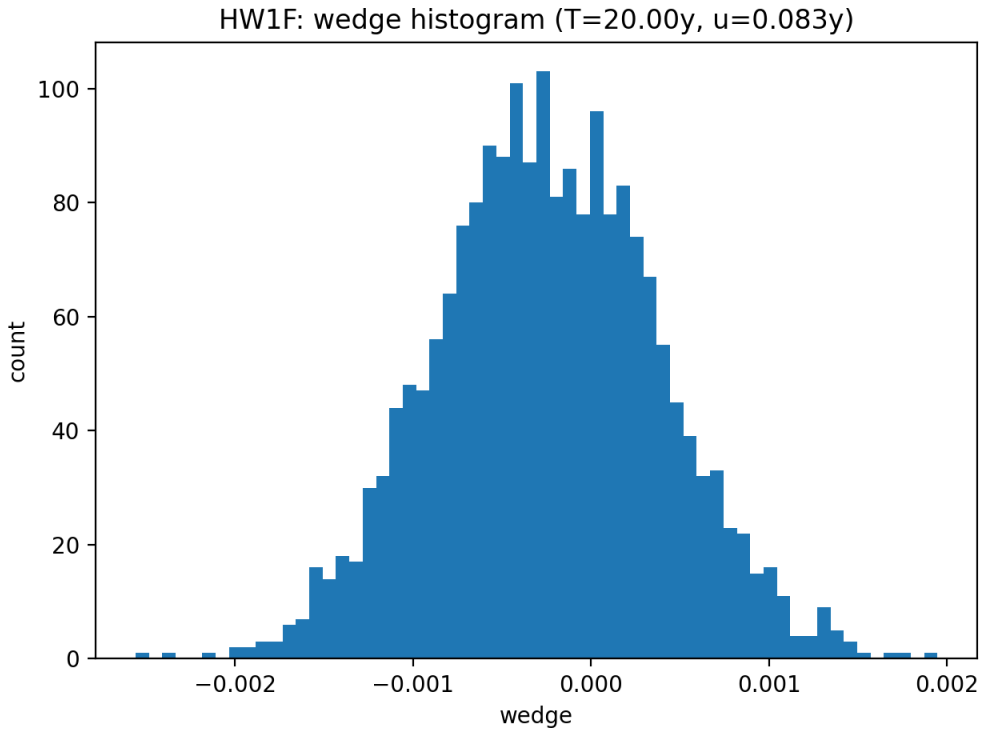

Figure 6: HW1F benchmark: wedge histogram for a representative (T, u) pair.

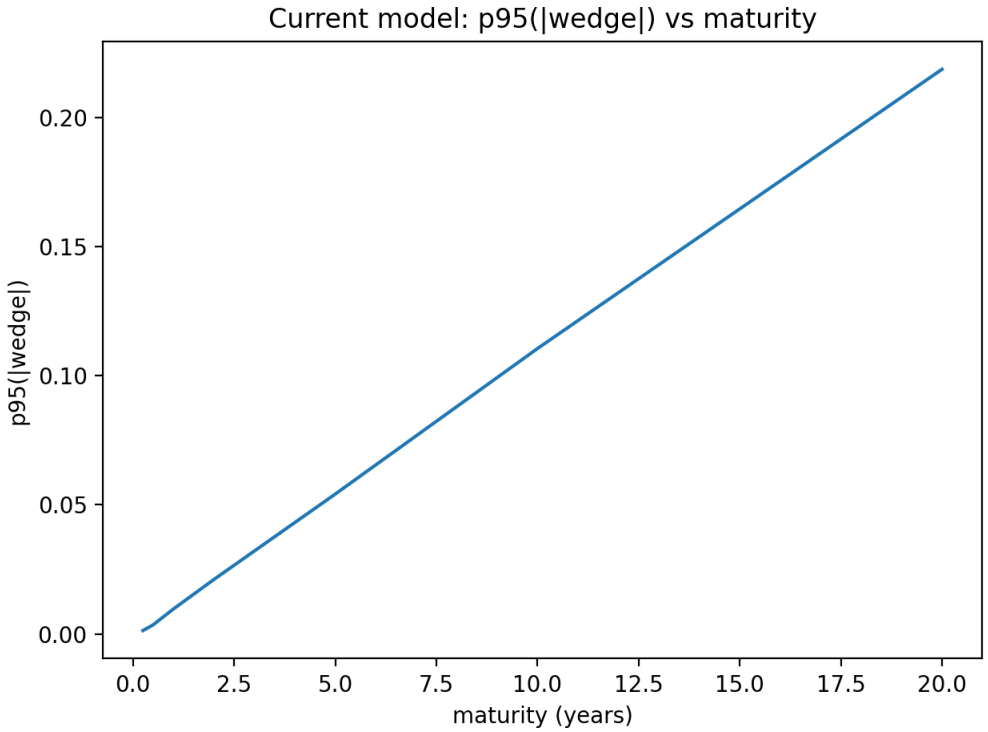

Figure 7: Current model: $p95(|\text{wedge}|)$ vs maturity.

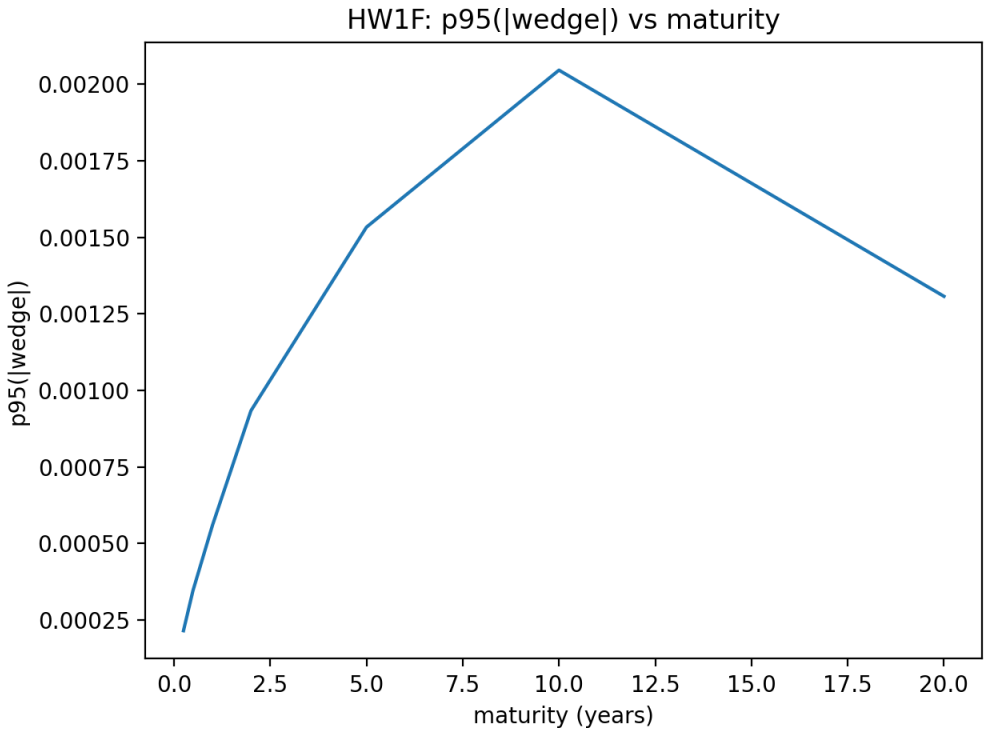

Figure 8: HW1F benchmark: $p95(|\text{wedge}|)$ vs maturity.

Discount-factor wedge (multiplicative consistency).

2.6 Conclusion

The diagnostics demonstrate that the IR simulation model does not enforce pathwise arbitrage-free term-structure dynamics. Violations are systematic and maturity-dependent, and are materially larger than those observed under an arbitrage-free benchmark. This behaviour is a structural consequence of the model design and must be monitored, benchmarked, and appropriately controlled when assessing long-horizon PFE metrics, particularly for long-dated exposure profiles.

3 Finding 2 — Pillar Discretisation and Interpolation Artefacts

3.1 Finding statement

The IR simulation framework represents the yield curve on a finite set of maturity pillars and relies on numerical interpolation to obtain curve values at non-pillar maturities required for pricing, discounting, and exposure computations. This representation introduces structural artefacts that may affect local curve smoothness, implied forward rates, and exposure profiles. In particular, model outputs may exhibit sensitivity to (i) the chosen interpolation scheme and (ii) the pillar grid density.

3.2 Model features giving rise to the finding

The finding arises from the following design features:

- Zero rates are simulated only at discrete maturity pillars $\{T_k\}$.
- Discount factors and forward rates at intermediate maturities are obtained via interpolation.
- Interpolation is applied ex post and is not enforced by the stochastic model dynamics.

Consequently, the interpolation operator imposes additional structure that is not implied by the underlying stochastic specification.

3.3 Risk implications for PFE

Interpolation artefacts may lead to:

- Artificial oscillations in the implied forward curve and local forward-rate spikes.
- Distortions of intermediate discount factors and local curve curvature.
- Sensitivity of exposure profiles (especially long-dated or convexity-sensitive products) to numerical choices rather than economic risk-factor dynamics.

3.4 Validation strategy and diagnostics

Three complementary diagnostics are implemented:

1. **Interpolation sensitivity (T2.1):** compare interpolated curves generated from identical pillar data under alternative interpolation schemes (linear-on-zero vs linear-on-logDF) and report RMS / max deviations.
2. **Implied forward roughness (T2.2):** compute the implied instantaneous forward curve on a dense grid and quantify roughness via an integrated curvature proxy.
3. **Pillar density stress (T2.3):** compare curves reconstructed from the full pillar set to curves reconstructed from a coarse pillar subset and quantify deviations on a common dense grid.

Diagnostics are produced for the current model and for the HW1F benchmark to separate structural interpolation artefacts from numerical noise.

3.5 Empirical evidence

Figure 9: Current model (T2.1): interpolation sensitivity (RMS difference) over time.

Figure 10: HW1F benchmark (T2.1): interpolation sensitivity (RMS difference) over time.

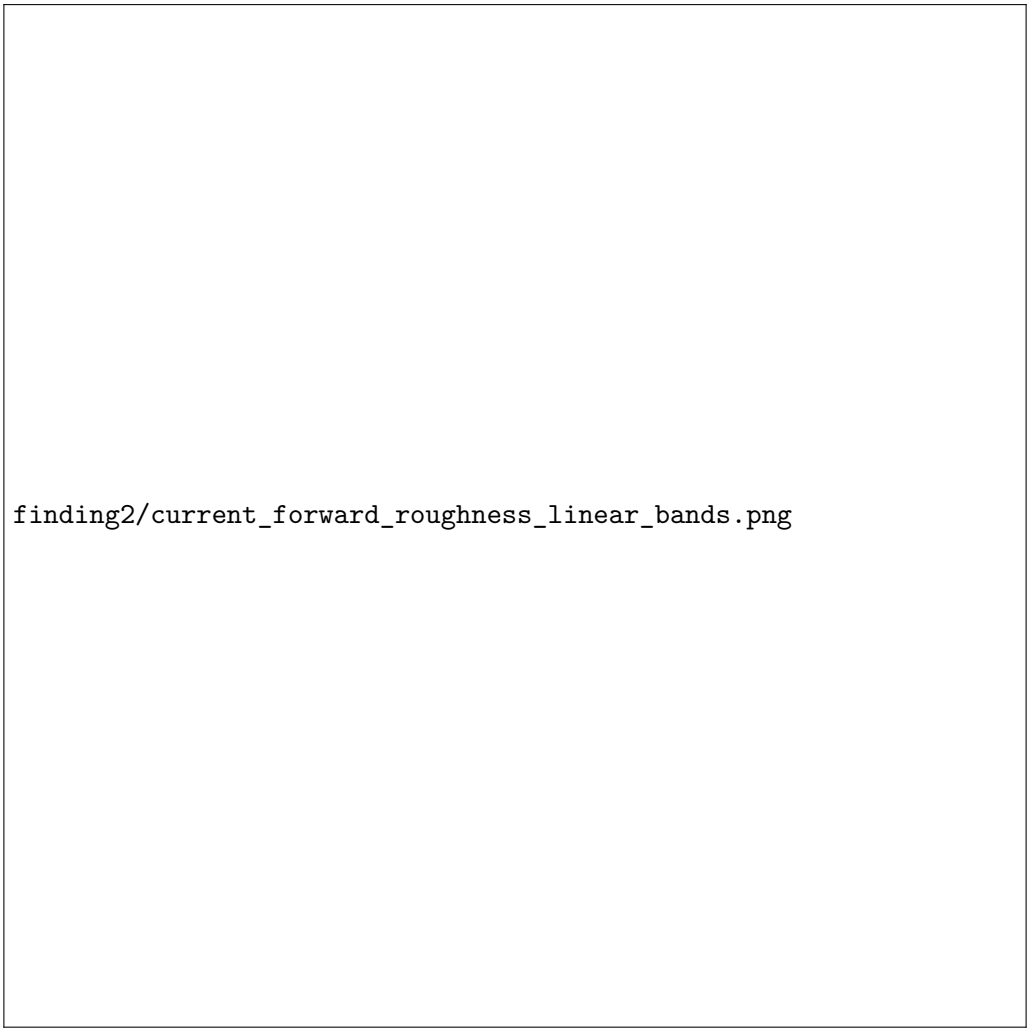

Figure 11: Current model (T2.2): implied forward roughness (linear-on-zero) over time.

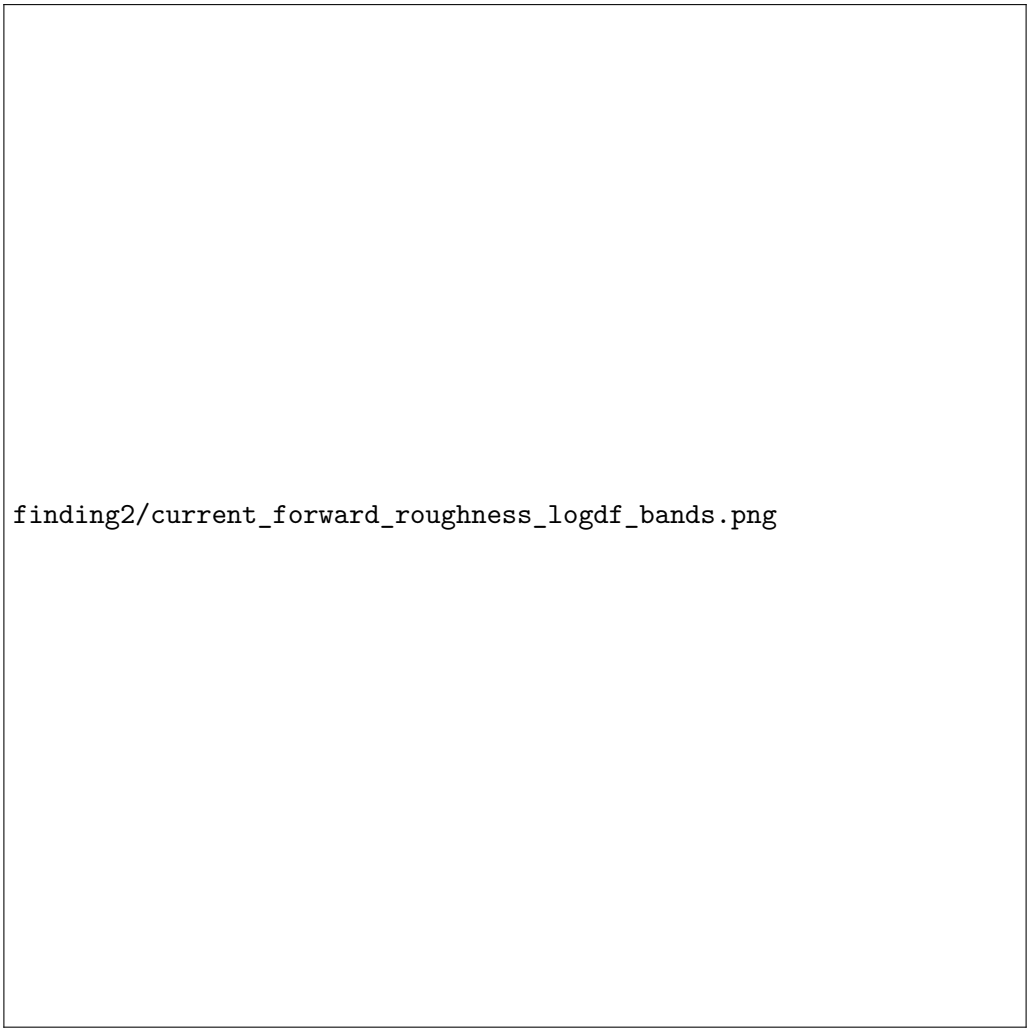

Figure 12: Current model (T2.2): implied forward roughness (linear-on-logDF) over time.

Figure 13: Current model (T2.3): pillar density stress (full vs coarse pillar set), RMS difference over time.

3.6 Conclusion

The discrete pillar representation combined with ex post interpolation introduces structural artefacts that are not implied by the underlying stochastic model. The implemented diagnostics quantify sensitivity to interpolation scheme, identify potential roughness in implied forward curves, and demonstrate grid-dependence under pillar coarsening. These effects are model-design driven and should be controlled via (i) consistent interpolation policy, (ii) sensitivity monitoring, and (iii) appropriate pillar grid specification for long-dated exposure calculations.

A Appendix A — Theory background

A.1 No-arbitrage relationships used in diagnostics

Discount factors and monotonicity. Let $P(t,T)$ denote the discount factor at time t for maturity T. In the absence of negative rates (or under standard admissibility conditions), discount factors are non-increasing in maturity:

$$P(t,T_{k+1}) \le P(t,T_k) \quad \text{for } T_{k+1} > T_k.$$

Equivalently, using zero rates $z(t,T)$ defined by $P(t,T) = \exp(-z(t,T)(T-t))$, monotonicity may be violated when simulated cross-maturity shapes become inconsistent.

Forward-rate reconstruction identity. For $0 \le t < T < T+u$, the forward discount factor satisfies

$$P(t,T,T+u) := \frac{P(t,T+u)}{P(t,T)}.$$

Under coherent term-structure dynamics, pricing identities ensure internal consistency of ratios across maturity intervals.

A.2 Pathwise vs marginal consistency

The model may be *marginally consistent* by construction (each maturity pillar calibrated to variance targets), yet not *pathwise consistent* across maturities because:

- the joint evolution across maturities is not derived from a single arbitrage-free term-structure model;
- correlations are imposed at driver level rather than induced by a coherent curve model;
- no drift restriction enforces HJM-consistency.

A.3 Benchmark rationale: Hull–White one-factor (HW1F)

HW1F provides a standard arbitrage-free reference with an explicit short-rate representation. In the benchmark, discount factors and forward ratios inherit consistency properties from the model's term-structure construction (up to numerical error). Therefore, systematic deviations under the current model can be attributed to structural design rather than plotting or discretization.

B Appendix B — Formal diagnostic definitions

B.1 Static discount-factor monotonicity test

Let $\{T_k\}_{k=0}^{K}$ denote maturity pillars and let $P^{(n)}(t_i,T_k)$ be the simulated discount factor on path n at time t_i.

Define the indicator of a monotonicity violation on path n as:

$$\mathbb{I}_{i,k}^{(n)} = \mathbf{1}\Big\{P^{(n)}(t_i,T_{k+1}) > P^{(n)}(t_i,T_k)\Big\}.$$

The empirical violation frequency is:

$$\widehat{p}_{i,k} = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}_{i,k}^{(n)}.$$

This diagnostic captures static violations of the basic no-arbitrage ordering of discount factors across maturities.

B.2 Kink index: cross-tenor smoothness diagnostic

Let $z^{(n)}(t_i, T_k)$ denote the simulated zero rate at pillar T_k. Define the discrete second difference:

$$\Delta^2 z_{i,k}^{(n)} = z^{(n)}(t_i, T_{k+1}) - 2z^{(n)}(t_i, T_k) + z^{(n)}(t_i, T_{k-1}), \quad k = 1, \ldots, K-1.$$

The *kink index* for path n at time t_i is defined as:

$$\mathrm{Kink}^{(n)}(t_i) = \sum_{k=1}^{K-1} w_k \left| \Delta^2 z_{i,k}^{(n)} \right|,$$

where w_k are optional scaling weights (e.g. based on maturity spacing).

The kink index measures local curvature and detects abrupt slope changes between adjacent pillars. Large values indicate reduced cross-tenor smoothness and economically implausible curve shapes.

B.3 Discount-factor wedge: multiplicative consistency diagnostic

For a fixed horizon T and increment $u > 0$, define the theoretical identity:

$$P(t, T+u) \equiv P(t, T)\, P(t, T, T+u).$$

Let $\widehat{P}^{(n)}(t; T, T+u)$ denote the forward discount factor reconstructed from simulated quantities (e.g. via zero-rate interpolation).

The *discount-factor wedge* is defined as:

$$\mathrm{Wedge}^{(n)}(t; T, u) = \log P^{(n)}(t, T+u) - \log P^{(n)}(t, T) - \log \widehat{P}^{(n)}(t; T, T+u).$$

Under pathwise arbitrage-free dynamics, the wedge should be identically zero (up to numerical error). Persistent dispersion or maturity-dependent growth of the wedge indicates structural inconsistency.

B.4 Interpretation and complementarity

- DF monotonicity tests static arbitrage constraints.
- The kink index captures cross-sectional smoothness and local coherence.
- The wedge diagnostic targets dynamic multiplicative consistency.

Together, these diagnostics provide comprehensive coverage of pathwise arbitrage-free properties.

C Appendix C — Algorithmic implementation and benchmark model

C.1 Hull–White one-factor (HW1F) model

The Hull–White one-factor model specifies the short rate as:

$$r(t) = x(t) + \phi(t),$$

where $x(t)$ follows an Ornstein–Uhlenbeck process:

$$dx(t) = -a\, x(t)\, dt + \sigma\, dW_t.$$

The deterministic shift $\phi(t)$ is chosen to fit the initial term structure exactly.

C.1.1 Zero-coupon bond prices

Under HW1F, the zero-coupon bond price admits the closed form:

$$P(t,T) = A(t,T)\exp(-B(t,T)\,r(t))\,,$$

with:

$$B(t,T) = \frac{1-e^{-a(T-t)}}{a},$$

and $A(t,T)$ determined to match the initial yield curve.

C.1.2 Calibration

Model parameters (a,σ) are calibrated by minimizing the discrepancy between model-implied and market-observed swaption volatilities:

$$\min_{a,\sigma} \sum_i \left(\sigma_{\text{model}}^{\text{swaption}}(i) - \sigma_{\text{market}}^{\text{swaption}}(i)\right)^2.$$

HW1F is arbitrage-free by construction and therefore serves as a suitable benchmark for structural diagnostics.

C.2 Algorithmic computation of diagnostics

C.2.1 DF monotonicity heatmap

For each simulation time and maturity interval:

```
for time i:
for maturity k:
violations = mean(DF[:,i,k+1] > DF[:,i,k])
```

C.2.2 Kink index bands

```
for time i:
for path n:
kink[n] = sum_k |Z[n,i,k+1] - 2 Z[n,i,k] + Z[n,i,k-1]|
compute median and quantiles of kink
```

C.2.3 Wedge statistics

```
for maturity T:
for path n:
wedge = log(P(T+u)) - log(P(T)) - log(P_hat(T,T+u))
compute histogram and p95(|wedge|)
```

C.3 Reproducibility and governance considerations

All diagnostics are computed under fixed random seeds and deterministic configuration files. Output artefacts are versioned by model configuration, ensuring traceability between reported figures and underlying simulations.

C.4 Interpolation and discretisation: theoretical considerations

A continuous yield curve is approximated in practice by values at a finite set of maturity pillars $\{T_k\}_{k=0}^{K}$. Any valuation requiring intermediate maturities $T \notin \{T_k\}$ necessarily introduces an interpolation operator $\mathcal{I}$ such that:

$$\tilde{z}(t,T) = \mathcal{I}\big(\{(T_k, z(t,T_k))\}_{k=0}^{K}\big)(T).$$

This operator imposes additional structure beyond the stochastic model specification. Different interpolation choices correspond to different assumptions about local behaviour of discount factors, zero rates, and implied forward rates. In general, interpolation is not arbitrage-neutral unless it is consistent with an arbitrage-free term-structure model (e.g. constructed from a short-rate or HJM specification). Consequently, discretisation and interpolation can induce artefacts such as: (i) local curvature spikes (kinks), (ii) oscillatory implied forwards, and (iii) grid-dependent valuations.

C.5 RMS difference (interpolation sensitivity)

Given a dense maturity grid $\{T_j\}_{j=1}^{J}$ and two interpolated curves $\tilde{z}^{(1)}(t,T)$ and $\tilde{z}^{(2)}(t,T)$ derived from the same pillar data, define pointwise differences:

$$\Delta z(t,T_j) = \tilde{z}^{(1)}(t,T_j) - \tilde{z}^{(2)}(t,T_j).$$

The root-mean-square (RMS) difference at time t is:

$$\mathrm{RMS}_z(t) = \sqrt{\frac{1}{J}\sum_{j=1}^{J}\big(\Delta z(t,T_j)\big)^2}.$$

RMS provides a stable measure of the typical magnitude of interpolation discrepancies while penalising large local deviations.

C.6 Implied forward curve and roughness proxy

On a dense maturity grid, the instantaneous forward curve is defined by:

$$f(t,T) = \frac{\partial}{\partial T}\big(T\,\tilde{z}(t,T)\big).$$

A roughness proxy used in the diagnostics is the integrated absolute curvature:

$$R_f(t) = \int \left|\frac{\partial^2 f(t,T)}{\partial T^2}\right| dT,$$

approximated numerically using finite differences. Elevated $R_f(t)$ indicates oscillatory or non-smooth forward curves potentially induced by discretisation and interpolation.

C.7 Pillar density stress metric

Let $\tilde{z}^{\mathrm{full}}(t,T)$ denote the interpolated curve reconstructed from the full pillar set and $\tilde{z}^{\mathrm{coarse}}(t,T)$ the curve reconstructed from a coarse pillar subset. Define:

$$\Delta z_{\mathrm{dens}}(t,T_j) = \tilde{z}^{\mathrm{coarse}}(t,T_j) - \tilde{z}^{\mathrm{full}}(t,T_j), \quad \mathrm{RMS}_{\mathrm{dens}}(t) = \sqrt{\frac{1}{J}\sum_{j=1}^{J}\big(\Delta z_{\mathrm{dens}}(t,T_j)\big)^2}.$$

Large values indicate grid dependence of reconstructed curves.

C.8 Finding 2 — Implementation mapping

The diagnostics are implemented under:

- `xva_engine/validation/ir/interpolation/test_interpolation_sensitivity.py`
- `xva_engine/validation/ir/interpolation/test_forward_roughness.py`
- `xva_engine/validation/ir/interpolation/test_pillar_density.py`
- `xva_engine/validation/ir/interpolation/reporting/plots_interpolation.py`

C.9 Interpolation sensitivity (T2.1): pseudo-code

```
grid = dense_maturity_grid(pillars)
z1 = interpolate(pillars, z_pillars, scheme="linear_zero", grid=grid)
z2 = interpolate(pillars, z_pillars, scheme="linear_logdf", grid=grid)
diff = z1 - z2
RMS(t) = sqrt(mean_T(diff^2))
Max(t) = max_T(|diff|)
```

C.10 Forward roughness (T2.2): pseudo-code

```
f(T) = d/dT [T z(T)]  (finite differences on dense grid)
roughness = integral |d^2 f / dT^2| dT
```

C.11 Pillar density stress (T2.3): pseudo-code

```
coarse_pillars = pillars[::2] (+ ensure last pillar)
z_full  = interpolate(full_pillars, z_pillars, scheme, dense_grid)
z_coarse = interpolate(coarse_pillars, z_pillars_coarse, scheme, dense_grid)
diff = z_coarse - z_full
report RMS and Max of diff
```