Short-rate models with smile and applications to Valuation Adjustments

Utrecht University & Rabobank, the Netherlands

T. van der Zwaard

ICCF24 - Amsterdam - April 5, 2024

Acknowledgements & Disclaimer

Based on:
T. van der Zwaard, L.A. Grzelak, C.W. Oosterlee. *"On the Hull-White model with volatility smile for Valuation Adjustments"*. Preprint submitted to arXiv (2403.14841) [8].

Acknowledgements
This work has been financially supported by Rabobank.

Disclaimer
The views expressed in this work are the personal views of the authors and do not necessarily reflect the views or policies of their current or past employers.

Outline

Goal: incorporate smiles in Valuation Adjustments (xVAs).

Steps:

1. Introduction.
2. Our contribution.
3. SDE with state-dependent drift / diffusion.
4. Randomized Affine Diffusion (RAnD).
5. Calibration, simulation and exposures.
6. Conclusions.

Introduction

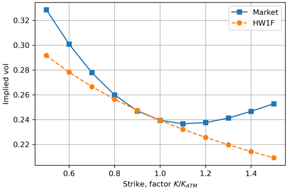
1. Background on xVAs:
    a. Economic value = risk-neutral value − xVA.
    b. Valuation Adjustments (xVAs), e.g., CVA, DVA, FVA, MVA, KVA.
    c. Computational challenges.
2. Focus on xVAs for IR derivatives.
3. Common xVA modeling setup in a Monte Carlo framework:
    a. Use one-factor short-rate model in Affine Diffusion class.
    b. Analytic tractability motivates use for xVA purposes.
    c. Example: Hull-White one-factor model (HW).

HW model

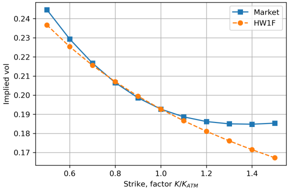
1. Impossible to fit to the whole market volatility surface (expiry $\times$ tenor $\times$ strike).
2. Time-dependent piece-wise constant volatility parameter used to calibrate the model to a strip of ATM co-terminal swaptions.
3. Forward rate under HW is shifted-lognormal: there is skew but it cannot be controlled.
4. The model does not generate volatility smile.
5. HW dynamics in the G1++ form:

$$r(t) = x(t) + b(t), \quad \mathrm{d}x(t) = -a_x x(t)\mathrm{d}t + \sigma_x(t)\mathrm{d}W(t).$$

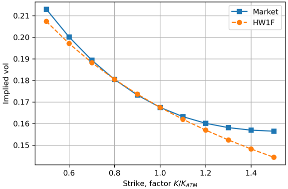
Smile and skew: the market vs HW



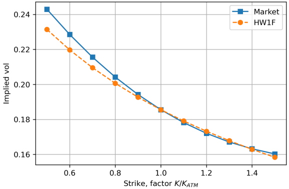
(a) 1Y expiry, 29Y tenor.



(b) 5Y expiry, 25Y tenor.



(c) 10Y expiry, 20Y tenor.



(d) 25Y expiry, 5Y tenor.

Figure: USD 30Y co-terminal swaption volatility strips (02/12/2022).

Smile and skew: xVA

1. Smile and skew typically absent from xVA calculations.
2. Challenge: find a model that captures smile and skew, but also allows for efficient calibration and pricing.
3. Smile and skew can be relevant for xVA:
    a. Obvious case: derivatives that take into account smile.
    b. Also for linear derivatives: legacy trades that are off-market and not primarily driven by ATM vols.
    c. Larger effect expected on PFE as this is a tail metric.
4. Cheyette-type examples in literature, e.g., Andreasen [1], Hoencamp *et al.* [5].
   Downsides:
    a. Only the smile curvature of one strip can be included: curvature of all smiles have to be roughly equivalent to have a sensible model.
    b. Calibration to European swaptions requires swap rate approximations.

Our contribution

1. Find SDE with state-dependent drift / diffusion that is consistent with the convex combination of N different HW models, where one model parameter is varied.
2. This model allows to capture market smile and skew.
3. Profit from the analytic tractability of Affine Diffusion dynamics.
4. The model allows for fast and semi-analytic swaption calibration.
5. Monte Carlo pricing using regression methods.
6. Use the idea of the RAnD method to parameterize the model: this results in the Randomized Hull-White (rHW) model, which has one additional degree of freedom w.r.t. HW.
7. Demonstrate the effect of smile on exposures of IR derivatives and the corresponding xVA metrics.

SDE with state-dependent drift / diffusion

1. General dynamics for $r(t)$ for which we try to find the (potentially) state-dependent drift and diffusion:

$$\mathrm{d}r(t) = \mu_r^{\mathbb{Q}_r}(t, r(t))\mathrm{d}t + \eta_r(t, r(t))\mathrm{d}W^{\mathbb{Q}_r}(t). \tag{1}$$

2. We want to find $\mu_r^{\mathbb{Q}_r}(t, r(t))$ and $\eta_r(t, r(t))$ s.t. $\forall t$ the density is consistent with the convex combination of N densities of analytically tractable models $r_n(t)$:

$$f_{r(t)}^{\mathbb{Q}_r}(y) := \sum_{n=1}^{N} \omega_n f_{r_n(t)}^{\mathbb{Q}_r}(y), \tag{2}$$

$$\mathrm{d}r_n(t) = \mu_{r_n}^{\mathbb{Q}_r}(t, r_n(t))\mathrm{d}t + \eta_{r_n}(t, r_n(t))\mathrm{d}W^{\mathbb{Q}_r}(t). \tag{3}$$

3. Eq. (2) holds for all measures and $\forall t$.
4. $\sum_{n=1}^{N} \omega_n = 1$ and $\omega_n > 0 \ \forall n$.
5. All dynamics are driven by the same Brownian motion $W^{\mathbb{Q}_r}(t)$.

Fokker-Planck: applied to our case

We derive $\mathrm{d}r(t)$ using the FP equation for both $r(t)$ and $r_n(t)$. Using

$$f^{\mathbb{Q}_r}_{r(t)}(y) := \sum_{n=1}^{N} \omega_n f^{\mathbb{Q}_r}_{r_n(t)}(y), \tag{4}$$

and linearity of the derivative operator we obtain:

$$\mathrm{d}r(t) = \mu^{\mathbb{Q}_r}_r(t, r(t))\mathrm{d}t + \eta_r(t, r(t))\mathrm{d}W^{\mathbb{Q}_r}(t), \tag{5}$$

$$\mu^{\mathbb{Q}_r}_r(t, y) = \sum_{n=1}^{N} \mu^{\mathbb{Q}_r}_{r_n}(t, y)\Lambda^{\mathbb{Q}_r}_n(t, y), \tag{6}$$

$$\eta^2_r(t, y) = \sum_{n=1}^{N} \eta^2_{r_n}(t, y)\Lambda^{\mathbb{Q}_r}_n(t, y), \tag{7}$$

$$\Lambda^{\mathbb{Q}_r}_n(t, y) = \frac{\omega_n f^{\mathbb{Q}_r}_{r_n(t)}(y)}{\sum_{i=1}^{N} \omega_i f^{\mathbb{Q}_r}_{r_i(t)}(y)}. \tag{8}$$

So an SDE with state-dependent drift and diffusion.

The $r_n(t)$ dynamics

- We work with the HW model in the G1++ formulation, where each $r_n(t)$ has a different mean-reversion $a_x = \theta_n$:

$$r_n(t) = x_n(t) + b_n(t), \tag{9}$$

$$\mathrm{d}x_n(t) = -\theta_n x_n(t)\mathrm{d}t + \sigma_x \mathrm{d}W(t), \tag{10}$$

$$b_n(t) = f^{\mathrm{M}}(0, t) - x_n(0)\mathrm{e}^{-\theta_n t} + \frac{1}{2}\sigma_x^2 {B_n}^2(0, T), \tag{11}$$

$$B_n(s, t) = \frac{1}{\theta_n}\left(1 - \mathrm{e}^{-\theta_n(t-s)}\right). \tag{12}$$

- $r_n(t) \sim \mathcal{N}\left(\mathbb{E}_s\left[x_n(t)\right] + b_n(t), \mathbb{V}\mathrm{ar}_s\left(x_n(t)\right)\right)$ conditional on $\mathcal{F}_s$.
- So $f_{r_n(t)}(y)$ is a normal pdf.

The $r(t)$ dynamics

For the underlying HW dynamics we obtain the following SDE:

$$\mathrm{d}r(t) = \mu_r^{\mathbb{Q}_r}(t, r(t))\mathrm{d}t + \eta_r(t, r(t))\mathrm{d}W^{\mathbb{Q}_r}(t), \tag{13}$$

$$\mu_r^{\mathbb{Q}_r}(t, r(t)) = \sum_{n=1}^{N} \left[\frac{\mathrm{d}\, f^{\mathrm{M}}(0,t)}{\mathrm{d}\, t} + \theta_n f^{\mathrm{M}}(0,t) - \theta_n r(t) + \mathbb{V}\mathrm{ar}_0\left(r_n(t)\right) \right] \cdot \Lambda_n^{\mathbb{Q}_r}(t, r(t)), \tag{14}$$

$$\eta_r(t, r(t)) = \sqrt{\sum_{n=1}^{N} \sigma_x^2 \cdot \Lambda_n^{\mathbb{Q}_r}(t, r(t))} = \sigma_x, \tag{15}$$

as $\sum_{n=1}^{N} \Lambda_n^{\mathbb{Q}_r}(t, y) = 1 \ \forall y$.
This means that the diffusion component $\eta_r(t, r(t))$ is unchanged, whereas the drift $\mu_r^{\mathbb{Q}_r}(t, r(t))$ is state-dependent.

Fast pricing equation for calibration

- Main result:

$$V_r(t;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(t;T). \tag{16}$$

- Both $V_r(t;T)$ and $V_{r_n}(t;T)\ \forall n$ are arbitrage-free, but only the former prices back the market.
- Eq. (16) only holds for non-path-dependent derivatives.
- For more complex derivatives, derive state-dependent (local-vol type) dynamics as before.
- We use it at $t = 0$ for calibration purposes.
- Under the HW model, $V_{r_n}(t;T)$ semi-analytic using Jamshidian decomposition.

Randomized Affine Diffusion

Randomized Affine Diffusion (RAnD) method [3, 4]:

1. Take an Affine Diffusion (AD) model.
2. Pick model parameter ϑ to randomize.
3. The r.v. ϑ is defined on domain $D_\vartheta := [a, b]$ with PDF $f_\vartheta(x)$ and CDF $F_\vartheta(x)$, and realization θ, $\vartheta(\omega) = \theta$, with finite moments.
4. For valuation, we use Gauss-quadrature weights $\{\omega_n, \theta_n\}_{n=1}^N$ where the nodes θ_n are based on $F_\vartheta(x)$, see [4, Appendix A.2]. Then, for the valuation:

$$V_{r(t;\vartheta)}(t;T) = \int_{[a,b]} V_{r(t;\theta)}(t;T)\mathrm{d}F_\vartheta(\theta) \approx \sum_{n=1}^{N} \omega_n V_{r(t;\theta_n)}(t;T).$$

5. Compare with the result we derived before:

$$V_r(t;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(t;T). \tag{17}$$

RAnD for model parametrization

1. Use the idea of the RAnD method to reduce dimensionality of our model parameters.
2. We do not suffer from the quadrature error when pricing Europeans.
3. We work with the HW dynamics.
4. We choose $\vartheta = a_x$, i.e., the mean-reversion parameter.
5. Impose $\mathcal{N}\left(\mu_\vartheta, \sigma_\vartheta^2\right)$ as randomizer (constant over time).
6. We call this the Randomized Hull-White (rHW) model.
7. $N = 5$ suitable when ϑ follows a normal (or uniform) distribution.
8. Key advantage: one additional degree of freedom w.r.t. HW.

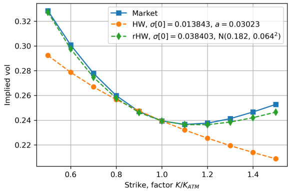
Calibration of the rHW dynamics $r(t)$

1. Calibration of the $r_n(t)$ HW dynamics in the usual way.
2. Mean-reversion parameterized as $a_x \sim \mathcal{N}\left(\mu_\vartheta, \sigma_\vartheta^2\right)$.
   For each choice of μ_ϑ and σ_ϑ^2:
   a. Compute collocation points (Gauss-quad weights) $\{\omega_n, \theta_n\}_{n=1}^N$.
   b. Initialize N HW models with mean-reversion parameter $a_x = \theta_n$.
3. Use fast valuation

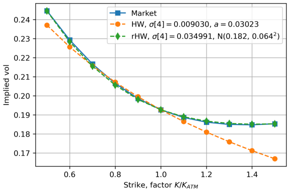
$$V_r(0;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(0;T).$$

4. Calibrate the parametrization of the mean-reversion $a_x \sim \mathcal{N}\left(\mu_\vartheta, \sigma_\vartheta^2\right)$ according to the desired strategy.
5. Bootstrap calibration of piece-wise constant model volatility to get a good ATM fit to the coterminal swaption strip.

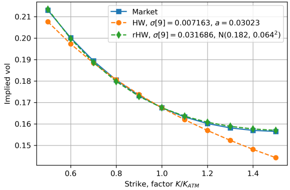
Calibration results



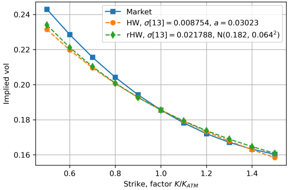
(a) 1Y expiry, 29Y tenor.



(b) 5Y expiry, 25Y tenor.



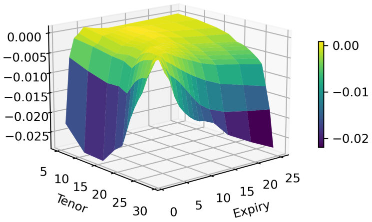
(c) 10Y expiry, 20Y tenor.



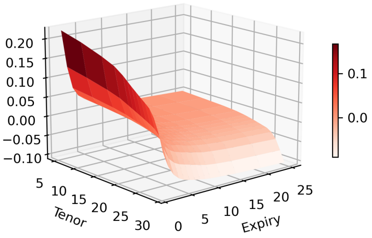
(d) 25Y expiry, 5Y tenor.

Figure: Market and model swaption implied volatilities.

Calibration results



(a) HW model.



(b) rHW model.

Figure: Implied volatility calibration error for all ATM points when calibrating to all coterminal smiles. USD market data from 02/12/2022.

Simulation of the rHW dynamics $r(t)$

1. Euler-Maruyama discretization always works:

$$r(t_{i+1}) = r(t_i) + \mu_r(t_i, r(t_i))\Delta t + \eta_r(t, r(t_i))\sqrt{\Delta t} Z, \qquad (18)$$

where $Z \sim \mathcal{N}(0,1)$.

2. Ideally we make large time steps. Hence, we integrate $\mathrm{d}r(t)$ to obtain an expression for $r(t)$ conditional on $r(s)$ for $s < t$, i.e.,

$$r(t) = r(s) + \int_s^t \mu_r(u, r(u))\mathrm{d}u + \int_s^t \eta_r(u, r(u))\mathrm{d}W(u). \quad (19)$$

3. The integrated drift is difficult to compute:

$$\int_s^t \mu_r(u, {\color{red}r(u)})\mathrm{d}u = f^{\mathrm{M}}(0,t) - f^{\mathrm{M}}(0,s)$$
$$+ \int_s^t \sum_{n=1}^{N} \left[\theta_n f^{\mathrm{M}}(0,u) - \theta_n {\color{red}r(u)} + \mathbb{V}\mathrm{ar}_0\left(r_n(u)\right)\right] \Lambda_n(u, {\color{red}r(u)})\mathrm{d}u.$$

4. Alternatively: machine learning, e.g., Seven-League scheme [6].

Simulation of the rHW dynamics $r(t)$

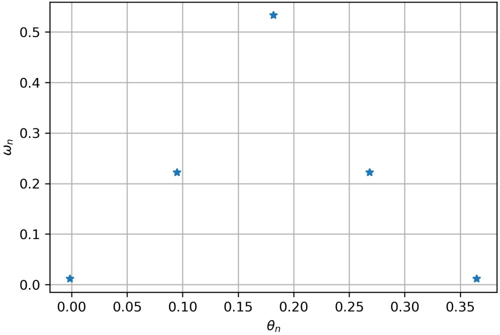


Figure: Example of quadrature points $\{\omega_n, \theta_n\}_{n=1}^{N}$ for $N = 5$ and $\mathcal{N}(\hat{a}, \hat{b}^2)$ with $\hat{a} = 0.181711$ and $\hat{b} = 0.064055$.

Simulation of the rHW dynamics $r(t)$

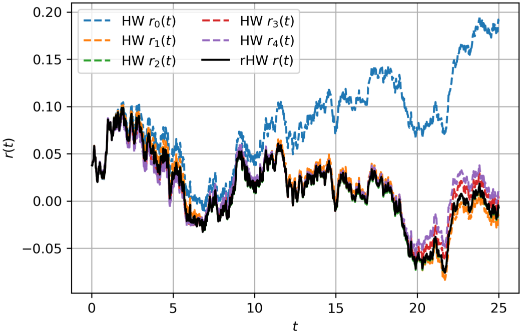


Figure: Path example: regular paths.

Simulation of the rHW dynamics $r(t)$

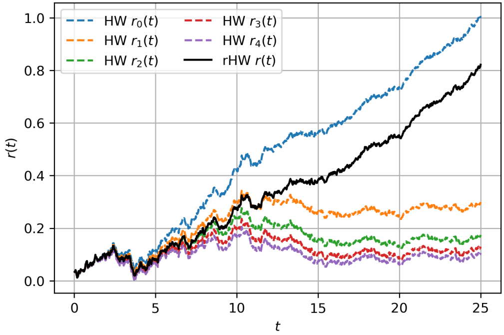


Figure: Path example: path ending high.

Simulation of the rHW dynamics $r(t)$

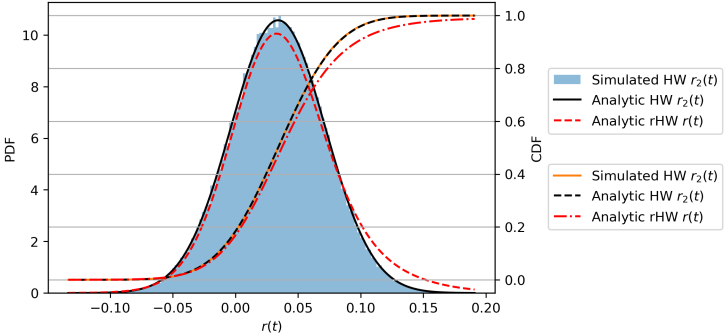


Figure: Comparing the rHW PDF and CDF at $t = 25$ with the HW process $r_2(t)$.

Simulation of the rHW dynamics $r(t)$

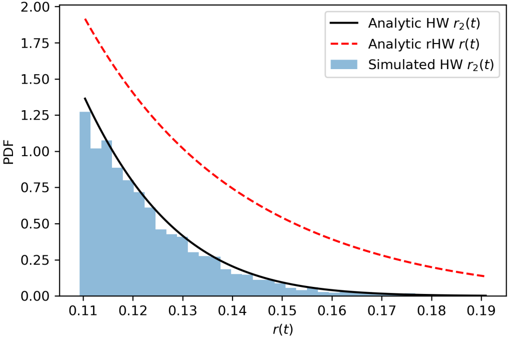


Figure: Right tail of the PDF.

Pricing under the rHW dynamics $r(t)$

1. Valuation as convex combination of underlying prices only for Europeans.
2. In general, we can use Monte Carlo with regression:
   a. For example, we simulate r from t_0 to t and at this point we want to compute $P_r(t, T) = \mathbb{E}_t\left[\mathrm{e}^{-\int_t^T r(s)\mathrm{d}s}\right]$.
   b. For each $P_r(t, T)$ we need for pricing, it is regressed on $r(t)$.
   c. For example, an n-th order polynomial can be used as regression function, or something of exponential form.
3. This step of ZCB calibration can be done with a separate, independent simulation before looking at the pricing.

Swap exposures

$$V^{\text{Swap}}(t) = A(t)\left[K - S(t)\right],$$

$$S(t) := \frac{P(t,T_0) - P(t,T_m)}{A(t)}, \quad A(t) := \sum_{k=1}^{m} (T_k - T_{k-1}) P(t, T_k).$$

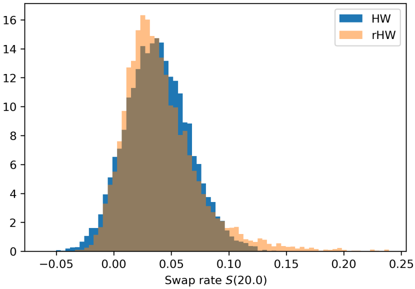
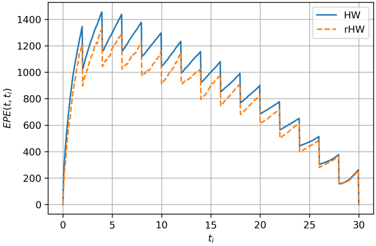
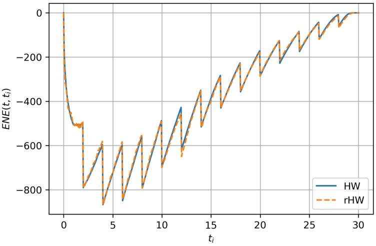


Figure: Swap rate distribution at $t = 20$ for a receiver swap, starting at $T_0 = 0$, ending at $T_m = 30$, with payments every two years.

Swap exposures



(a) $\mathrm{EPE}(t, t_i)$



(b) $\mathrm{ENE}(t, t_i)$

Figure: Comparing average exposures for an ATM receiver swap ($K = K_{\mathrm{ATM}}$). Runtime HW exposures 2.03 sec, runtime rHW exposure 1.90 sec (averages over 20 runs).

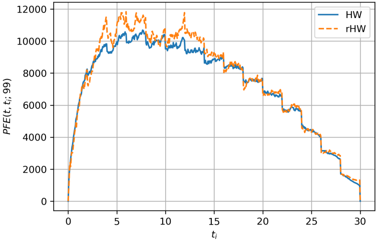
Swap BCVA

Model	K	Moneyness	BCVA(t_0)
HW	K_{ATM}	ATM	289.612
rHW			249.831
HW	$1.5 \cdot K_{\text{ATM}}$	ITM	862.803
rHW			820.243
HW	$0.5 \cdot K_{\text{ATM}}$	OTM	-165.661
rHW			-196.049

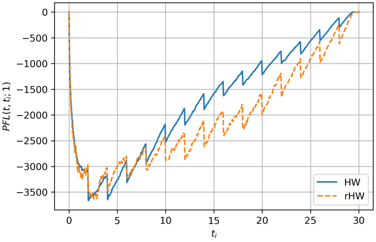
Table: BCVA metrics for the receiver swap example, for various strikes.

- Significant smile impact for all moneyness types.
- Impact relatively the smallest and absolutely the largest in the ITM case.
- Impact relatively the largest and absolutely the smallest in the OTM case.

Swap tail exposures



(a) $\text{PFE}(t, t_i; 99)$



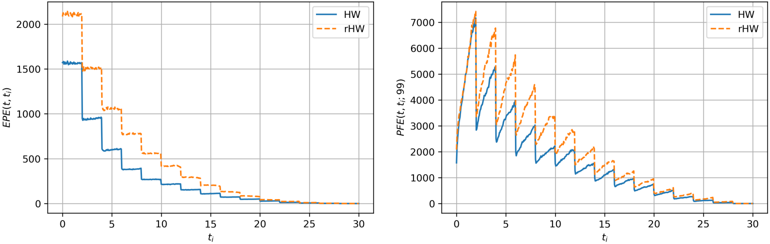
(b) $\text{PFL}(t, t_i; 1)$

Figure: Comparing tail exposures for an ATM receiver swap ($K = K_{\text{ATM}}$).

Bermudan swaption exposures

1. Receiver Bermudan swaption.
2. Cash-settled: so zero exposure after exercise.
3. The underlying swap starts at $T_0 = 0$, ends at $T_m = 30$, has payments every two years and early-exercise dates at every swap payment date until the swap maturity.

Bermudan swaption exposures



(a) $\text{EPE}(t, t_i)$ (b) $\text{PFE}(t, t_i; 99)$

Figure: Comparing exposures for a receiver Bermudan swaption on an ATM swap ($K = K_{\text{ATM}}$).

Bermudan swaption CVA

Model	K	Moneyness	CVA(t_0)
HW	K_{ATM}	ATM	159.156
rHW			258.035
HW	$0.5 \cdot K_{\text{ATM}}$	OTM	106.292
rHW			157.305
HW	$1.5 \cdot K_{\text{ATM}}$	ITM	190.855
rHW			320.358

Table: CVA metrics for a receiver Bermudan swaption on a swap, for various strikes.

Conclusions

1. Find SDE with state-dependent drift / diffusion that is consistent with the convex combination of N different HW models, where one model parameter is varied.
2. This model allows to capture market smile and skew.
3. Profit from the analytic tractability of Affine Diffusion dynamics.
4. The model allows for fast and semi-analytic swaption calibration.
5. Monte Carlo pricing using regression methods.
6. Use the idea of the RAnD method to parameterize the model: this results in the Randomized Hull-White (rHW) model, which has one additional degree of freedom w.r.t. HW.
7. Demonstrate the effect of smile on exposures of IR derivatives and the corresponding xVA metrics.

Short-rate models with smile and applications to Valuation Adjustments

Utrecht University & Rabobank, the Netherlands

T. van der Zwaard

ICCF24 - Amsterdam - April 5, 2024

References I

[1] J. Andreasen.
CVA on an iPad Mini - Part 2: The Beast.
Aarhus Kwant Factory PhD Course, January 2020.

[2] D. Brigo and F. Mercurio.
A mixed-up smile.
Risk, pages 123–126, September 2000.

[3] L.A. Grzelak.
On Randomization of Affine Diffusion Processes with Application to Pricing of Options on VIX and S&P 500.
arXiv Electronic Journal, August 2022.

[4] L.A. Grzelak.
Randomization of Short-Rate Models, Analytic Pricing and Flexibility in Controlling Implied Volatilities.
arXiv Electronic Journal, November 2022.

[5] J.H. Hoencamp, J.P. de Kort, and B.D. Kandhai.
The Impact of Stochastic Volatility on Initial Margin and MVA for Interest Rate Derivatives.
Applied Mathematical Finance, 29(2):141–179, December 2022.

[6] S. Liu, L.A. Grzelak, and C.W. Oosterlee.
The Seven-League Scheme: Deep Learning for Large Time Step Monte Carlo Simulations of Stochastic Differential Equations.
Risks, 10(47), February 2022.

[7] V. Piterbarg.
Mixture of Models: A Simple Recipe for a ... Hangover?
SSRN Electronic Journal, July 2003.

References II

[8] T. van der Zwaard, L.A. Grzelak, and C.W. Oosterlee.
On the Hull-White model with volatility smile for Valuation Adjustments.
arXiv Electronic Journal, March 2024.