# Valuation Adjustments with an Affine-Diffusion-based Interest Rate Smile

Utrecht University & Rabobank, the Netherlands

T. van der Zwaard

21st Winter School on Math. Finance - Soesterberg - January 23, 2024

# Acknowledgements & Disclaimer

Joint work with L. Grzelak (UU, Rabobank) & C. Oosterlee (UU).

**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 pricing.
6. Conclusions.

# Introduction

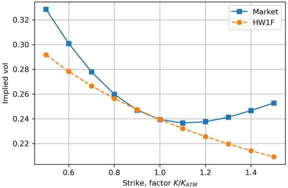
1. Background on xVAs:
   a. Economic value = risk-neutral value − xVA.
   b. Valuation Adjustments (xVAs), e.g., CVA, DVA, FVA, MVA, KVA.
   c. Computational challenges.
2. Focus on xVAs for IR derivatives.
3. Common xVA modeling setup in a Monte Carlo framework:
   a. Use one-factor short-rate model in Affine Diffusion class.
   b. Analytic tractability motivates use for xVA purposes.
   c. Example: Hull-White one-factor model (HW1F).

# HW1F model

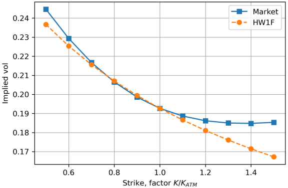
1. Impossible to fit to the whole market volatility surface (expiry $\times$ tenor $\times$ strike).
2. Time-dependent piece-wise constant volatility parameter used to calibrate the model to a strip of ATM co-terminal swaptions.
3. Forward rate under HW1F is shifted-lognormal: there is skew but it cannot be controlled.
4. The model does not generate volatility smile.
5. HW1F dynamics in the G1++ form:

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

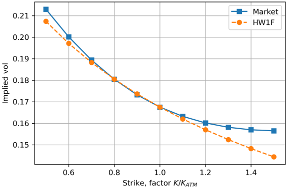
# Smile and skew: the market vs HW1F



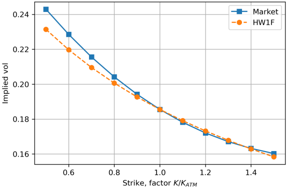

(a) 1Y expiry, 29Y tenor.




(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. Examples in literature:
   a. Andreasen used a four-factor Cheyette model with local and stochastic volatility [1].
   b. Quadratic Gaussian models (quadratic form for the short rate) also allow smile control [3, Section 16.3.2].

# Our contribution

1. Find SDE with state-dependent drift / diffusion that is consistent with the convex combination of $N$ different HW1F 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: one additional degree of freedom for HW1F.

# 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:

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

2. We want to find $\mu_r^{\mathbb{M}}(t, r(t))$ 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 $r_n(t)$:

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

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

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

# 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

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

and linearity of the derivative operator we obtain:

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

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

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

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

So an SDE with state-dependent drift and diffusion.

# The $r_n(t)$ dynamics

- We work with the HW1F 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 HW1F dynamics we obtain the following SDE:

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

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

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

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

# Fast pricing equation for calibration

- We start from martingale pricing under each of the individual underlying affine models $r_n(t)$, i.e.,

$$\begin{aligned} V_{r_n}(t;T) &= \mathbb{E}_t^{\mathbb{Q}_{r_n}}\left[\frac{B_{r_n}(t)}{B_{r_n}(T)}V_{r_n}(T;T)\right] \\ &= \mathbb{E}_t^{\mathbb{Q}_{r_n}}\left[\frac{B_{r_n}(t)}{B_{r_n}(T)}H(T;r_n(T))\right] \\ &= P_{r_n}(t,T)\mathbb{E}_t^{\mathbb{Q}_{r_n}^T}\left[H(T;r_n(T))\right], \end{aligned} \tag{16}$$

where $V_{r_n}(t;T)$ denotes the time $t$ value of the derivative that has payoff $H(T;r_n(T))$ at time $T$. $B_{r_n}(t)$ is the risk-neutral bank account and $P_{r_n}(t,T)$ is the Zero Coupon Bond.

# Fast pricing equation for calibration

$$\begin{aligned}
\sum_{n=1}^{N} \omega_n V_{r_n}(t;T) &= \sum_{n=1}^{N} \omega_n P_{r_n}(t,T)\mathbb{E}_t^{\mathbb{Q}_{r_n}^T}\left[H(T;r_n(T))\right] \\
&= P_r(t,T)\sum_{n=1}^{N} \omega_n \mathbb{E}_t^{\mathbb{Q}_r^T}\left[H(T;r_n(T))\right] \\
&= P_r(t,T)\sum_{n=1}^{N} \omega_n \int_{\mathbb{R}} H(T;x) f_{r_n(T)}^{\mathbb{Q}_r^T}(x)\mathrm{d}x \\
&= P_r(t,T)\int_{\mathbb{R}} H(T;x) \sum_{n=1}^{N} \omega_n f_{r_n(T)}^{\mathbb{Q}_r^T}(x)\mathrm{d}x \\
&= P_r(t,T)\int_{\mathbb{R}} H(T;x) f_{r(T)}^{\mathbb{Q}_r^T}(x)\mathrm{d}x \\
&= P_r(t,T)\mathbb{E}_t^{\mathbb{Q}_r^T}\left[H(T;r(T))\right] = V_r(t;T). \qquad (17)
\end{aligned}$$

# Fast pricing equation for calibration

- Main result:

$$V_r(t;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(t;T). \tag{18}$$

- 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. (18) 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 HW1F model, $V_{r_n}(t;T)$ semi-analytic using Jamshidian decomposition.

# Randomized Affine Diffusion

Randomized Affine Diffusion (RAnD) method [4, 5]:

1. Take an Affine Diffusion (AD) model.
2. Pick model parameter $\vartheta$ to randomize.
3. The r.v. $\vartheta$ is defined on domain $D_\vartheta := [a, b]$ with PDF $f_\vartheta(x)$ and CDF $F_\vartheta(x)$, and realization $\theta$, $\vartheta(\omega) = \theta$, with finite moments.
4. For valuation, we use Gauss-quadrature weights $\{\omega_n, \theta_n\}_{n=1}^N$ where the nodes $\theta_n$ are based on $F_\vartheta(x)$, see [5, 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:

$$V_r(t;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(t;T). \tag{19}$$

# 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 HW1F 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. $N = 5$ suitable when $\vartheta$ follows a normal (or uniform) distribution.
7. Key advantage: one additional degree of freedom w.r.t. HW1F.

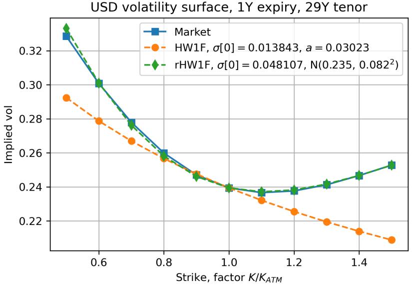
# Calibration of the $r(t)$ dynamics

1. Calibration of the $r_n(t)$ HW1F 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 $\mu_\vartheta$ and $\sigma_\vartheta^2$:
   a. Compute collocation points (Gauss-quad weights) $\{\omega_n, \theta_n\}_{n=1}^{N}$.
   b. Initialize $N$ HW1F models with mean-reversion parameter $a_x = \theta_n$.
3. Use fast valuation

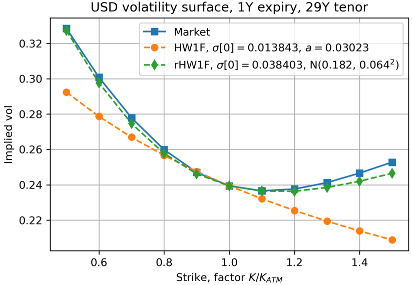
$$V_r(0;T) = \sum_{n=1}^{N} \omega_n V_{r_n}(0;T).$$

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:
   a. Fit the initial coterminal smile.
   b. Fit all coterminal smiles.
5. Bootstrap calibration of piece-wise constant model volatility to get a good ATM fit to the coterminal swaption strip.

# Calibration results




(a) Fit initial coterminal smile.




(b) Fit all coterminal smiles.

Figure: Initial coterminal smile. USD market data from 02/12/2022.

# Calibration results

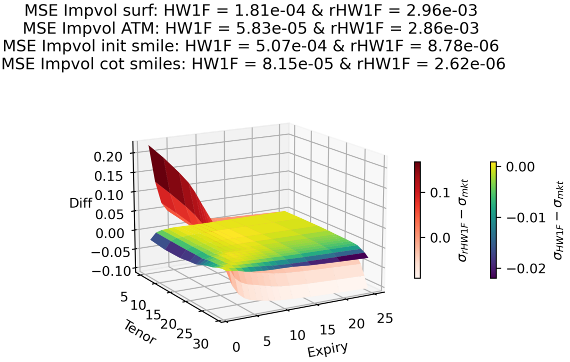



Figure: Difference in ATM implied vols when calibrating to all coterminal smiles. USD market data from 02/12/2022.

# Simulation of the $r(t)$ dynamics

1. Euler-Maruyama discretization always works:

$$r(t_{i+1}) = r(t_i) + \mu_r(t_i, r(t_i))\Delta t + \eta_r(t, r(t_i))\sqrt{\Delta t} Z, \qquad (20)$$

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.,

$$r(t) = r(s) + \int_s^t \mu_r(u, r(u))\mathrm{d}u + \int_s^t \eta_r(u, r(u))\mathrm{d}W(u). \qquad (21)$$

3. The integrated drift is difficult to compute:

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

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

# Simulation of the $r(t)$ dynamics

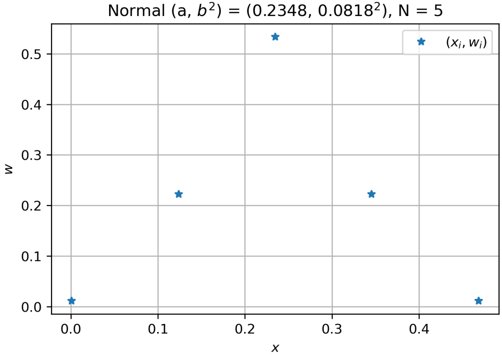



Figure: Example of quadrature points.

# Simulation of the $r(t)$ dynamics

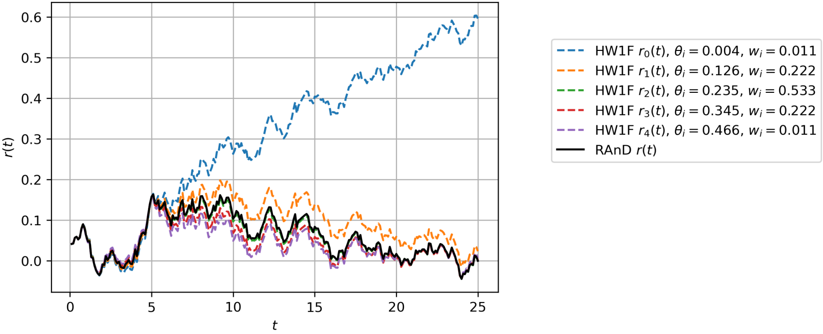



Figure: Path example: regular paths.

# Simulation of the $r(t)$ dynamics

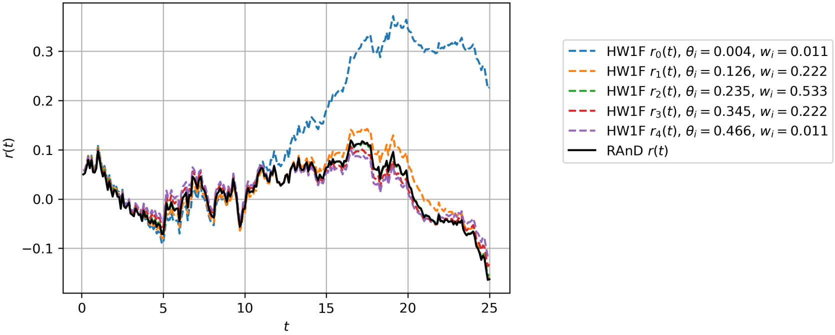



Figure: Path example: path ending low.

# Simulation of the $r(t)$ dynamics

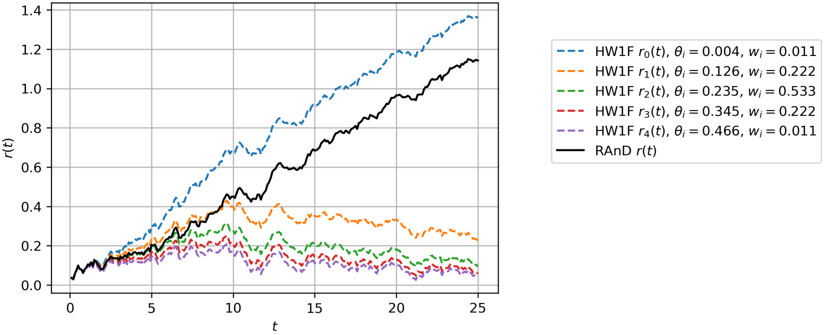



Figure: Path example: path ending high.

# Simulation of the $r(t)$ dynamics

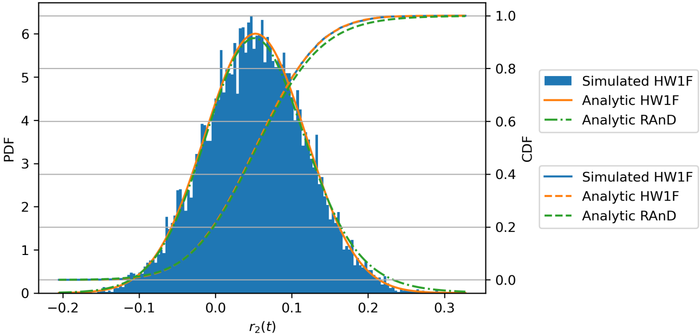
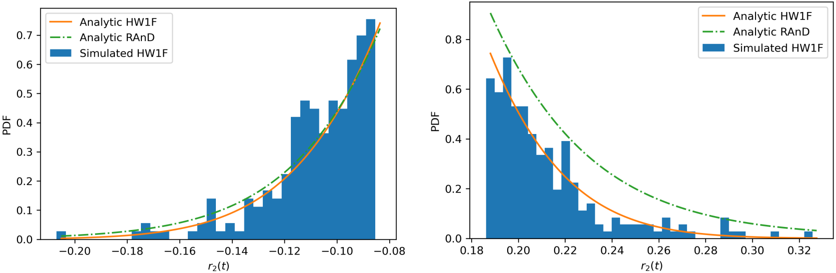



Figure: $r(t)$ vs $r_2(t)$.

# Simulation of the $r(t)$ dynamics




(a) Left tail. (b) Right tail.

Figure: $r(t)$ vs $r_2(t)$.

# Pricing under the $r(t)$ dynamics

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. These regression-based methods lend themselves naturally for xVA calculations, also known as American Monte Carlo.

# Pricing a swaption under the $r(t)$ dynamics

1. Swaption with 10k notional, 5y expiry, on a 5y payer swap with annual payments.
2. Use $10^5$ MC paths (antithetic variates turned on) and 100 simulation dates per year.
3. Polynomial regression of degree 1 or 2.

| | Value | Imp.vol |
|---|---|---|
| HW1F: analytic | 328.63814 | 0.221863 |
| Fast pricing eq: analytic | 581.86990 | 0.401957 |
| Fast pricing eq: MC regressed ZCB | 581.72496 | 0.401850 |
| RAnD dynamics: MC regressed ZCB | 582.26208 | 0.402247 |
| Abs diff | 0.53711 | 0.000397 |

Table: Results for all coterminal smiles calibration. Diff between fast pricing eq. and RAnD dynamics values using the MC with regressed ZCB. RAnD 95% conf.int.: $(580.09, 584.43)$.

# Conclusions

1. Find SDE with state-dependent drift / diffusion that is consistent with the convex combination of $N$ different HW1F 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: one additional degree of freedom for HW1F.

# Valuation Adjustments with an Affine-Diffusion-based Interest Rate Smile


## Utrecht University & Rabobank, the Netherlands

T. van der Zwaard


21st Winter School on Math. Finance - Soesterberg - January 23, 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] A. Green.
*XVA: Credit, Funding and Capital Valuation Adjustments*.
John Wiley & Sons, first edition, November 2015.
ISBN 978-1-118-55678-8.

# References II


[4] 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.

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

# References III


[7] C.W. Oosterlee and L.A. Grzelak.
*Mathematical Modeling and Computation in Finance*.
World Scientific, first edition, November 2019.
ISBN 978-1-78634-794-7.

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