PDSim: A Shiny App for Polynomial Diffusion Model Simulation and Estimation

Peilun He [1*], **Nino Kordzakhia** [2*], **Gareth W. Peters** [3*], and **Pavel V. Shevchenko** [1*]

1 Department of Actuarial Studies and Business Analytics, Macquarie University, Australia **2** Department of Mathematics and Statistics, Macquarie University, Australia **3** Department of Statistics and Applied Probability, University of California Santa Barbara, USA * These authors contributed equally.

Summary

The Schwartz-Smith two-factor model (Schwartz & Smith, 2000) was commonly used in the pricing of crude oil futures and some other futures in the last two decades. In 2016, (Filipovic & Larsson, 2016) introduced a new polynomial diffusion framework which allows a more complicated structure of spot price. This framework has been applied to electricity forwards (Kleisinger-Yu et al., 2020), in which the spot price is modelled in a quadratic form of two factors. PDSim aims to estimate futures prices as well as the latent state variables, and provides well-designed visualisations. This application is available at https://github.com/peilun-he/polynomial-diffusion-model-simulation-and-estimation.

Schwartz-Smith two-factor model

Under the Schwartz-Smith framework, the spot price S_t is modelled as the sum of two factors χ_t and ξ_t,

$$\log(S_t) = \chi_t + \xi_t, \tag{1}$$

where χ_t represents the short-term fluctuation and ξ_t is the long-term equilibrium price level. Additionally, we assume both χ_t and ξ_t follow a risk-neutral Ornstein-Uhlenbeck process,

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dW_t^\chi, \tag{2}$$

and

$$d\xi_t = (\mu_\xi - \gamma\xi_t - \lambda_\xi)dt + \sigma_\xi dW_t^\xi, \tag{3}$$

where $\kappa, \gamma \in \mathbb{R}^+$ are the speed of mean-reversion parameters, $\mu_\xi \in \mathbb{R}$ is the mean level of the long-term factor, $\sigma_\chi, \sigma_\xi \in \mathbb{R}^+$ are the volatilities, and $\lambda_\chi, \lambda_\xi \in \mathbb{R}$ are risk premiums. The processes $(W_t^\chi)_{t\geq 0}$ and $(W_t^\xi)_{t\geq 0}$ are correlated standard Brownian Motions with

$$\mathbb{E}\left(dW_t^\chi dW_t^\xi\right) = \rho dt.$$

We set $\lambda_\chi = \lambda_\xi = 0$ in Equation 2 and Equation 3 to get the real processes. We use the risk-neutral processes for futures pricing, and real processes for modelling state variables.

In discrete time, χ_t and ξ_t are jointly normally distributed. Therefore, the spot price, which is equal to the sum of χ_t and ξ_t, is log-normally distributed. Moreover, under the arbitrage-free assumption, the futures price $(F_{t,T})$ at current time t must be equal to the expected value of spot price at maturity time T,

$$\log(F_{t,T}) = \log(\mathbb{E}^*(S_T|\mathcal{F}_t)),$$

where $\mathcal{F}_t$ is the filtration and $\mathbb{E}^*(\cdot)$ is the expectation under the risk-neutral processes Equation 2 and Equation 3. Then we can get the linear Gaussian state space model:

$$x_t = c + Ex_{t-1} + w_t, \tag{4}$$

$$y_t = d_t + F_t x_t + v_t, \tag{5}$$

where $x_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix}$, $c = \begin{bmatrix} 0 \\ \frac{\mu_\xi}{\gamma}\left(1 - e^{-\gamma\Delta t}\right) \end{bmatrix}$, $E = \begin{bmatrix} e^{-\kappa\Delta t} & 0 \\ 0 & e^{-\gamma\Delta t} \end{bmatrix}$, $y_t = \left(\log\left(F_{t,T_1}\right), \dots, \log\left(F_{t,T_m}\right)\right)^\top$

$(A(T_1 - t), \dots, A(T_m - t))^\top$, $F_t = \begin{bmatrix} e^{-\kappa(T_1-t)}, \dots, e^{-\kappa(T_m-t)} \\ e^{-\gamma(T_1-t)}, \dots, e^{-\gamma(T_m-t)} \end{bmatrix}^\top$ and m is the number of

futures contracts. The function $A(\cdot)$ is given by

$$A(t) = -\frac{\lambda_\chi}{\kappa}(1 - e^{-\kappa t}) + \frac{\mu_\xi - \lambda_\xi}{\gamma}(1 - e^{-\gamma t})$$
$$+ \frac{1}{2}\left(\frac{1 - e^{-2\kappa t}}{2\kappa}\sigma_\chi^2 + \frac{1 - e^{-2\gamma t}}{2\gamma}\sigma_\xi^2 + 2\frac{1 - e^{-(\kappa+\gamma)t}}{\kappa + \gamma}\sigma_\chi\sigma_\xi\rho\right).$$

w_t and v_t are multivariate Gaussian noises with mean $\mathbf{0}$ and covariance matrix Σ_w and Σ_v respectively, where

$$\Sigma_w = \begin{bmatrix} \frac{1 - e^{-2\kappa\Delta t}}{2\kappa}\sigma_\chi^2 & \frac{1 - e^{-(\kappa+\gamma)\Delta t}}{\kappa+\gamma}\sigma_\chi\sigma_\xi\rho \\ \frac{1 - e^{-(\kappa+\gamma)\Delta t}}{\kappa+\gamma}\sigma_\chi\sigma_\xi\rho & \frac{1 - e^{-2\gamma\Delta t}}{2\gamma}\sigma_\xi^2 \end{bmatrix},$$

and we assume Σ_v is diagonal, $\Sigma_v = diag(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2)$. Under this framework, c, E, Σ_w and Σ_v are deterministic but d_t and F_t are time-variant.

Polynomial diffusion model

In this section, we present a general framework of the polynomial diffusion model first, and then we give the application in the two-factor model. The mathematical foundations and applications of polynomial diffusion model in finance are provided in (Filipovic & Larsson, 2016).

Consider the stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{6}$$

where W_t is a d-dimensional standard Brownian motion and map $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is continuous. Define $a := \sigma\sigma^\top$. For maps $a : \mathbb{R}^d \to \mathbb{S}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$, suppose we have $a_{ij} \in Pol_2$ and $b_i \in Pol_1$. $\mathbb{S}^d$ is the set of all real symmetric $d \times d$ matrices and Pol_n is the set of all polynomials of degree at most n. Then the solution of Equation 6 is a polynomial diffusion.

Moreover, we define the generator $\mathcal{G}$ associated to the polynomial diffusion X_t as

$$\mathcal{G}f(x) = \frac{1}{2}Tr\left(a(x)\nabla^2 f(x)\right) + b(x)^\top\nabla f(x) \tag{7}$$

for $x \in \mathbb{R}^d$ and any C^2 function f. Let N be the dimension of Pol_n, and $H : \mathbb{R}^d \to \mathbb{R}^N$ be a function whose components form a basis of Pol_n. Then for any $p \in Pol_n$, there exists a unique vector $\vec{p} \in \mathbb{R}^N$ such that

$$p(x) = H(x)^\top\vec{p} \tag{8}$$

and $\vec{p}$ is the coordinate representation of $p(x)$. Moreover, there exists a unique matrix representation $G \in \mathbb{R}^{N \times N}$ of the generator $\mathcal{G}$, such that $G\vec{p}$ is the coordinate vector of $\mathcal{G}p$. So we have

$$\mathcal{G}p(x) = H(x)^\top G\vec{p}. \tag{9}$$

59 Theorem 1: Let $p(x) \in Pol_n$ be a polynomial with coordinate representation $\vec{p} \in \mathbb{R}^N$ satisfying
60 Equation 8, $G \in \mathbb{R}^{N \times N}$ be a matrix representation of generator $\mathcal{G}$ satisfying Equation 9, and
61 $X_t \in \mathbb{R}^d$ satisfies Equation 6. Then for $0 \le t \le T$, we have

$$\mathbb{E}\left(p(X_T)|\mathcal{F}_t\right) = H(X_t)^\top e^{(T-t)G}\vec{p},$$

62 where $\mathcal{F}_t$ is a σ-algebra generated up tp time t.

63 The proof of Theorem 1 is given in (Filipovic & Larsson, 2016).

64 Next, we apply this theorem to the two-factor model. Assume the spot price S_t is modelled as

$$S_t = p_n(x_t), \tag{10}$$

65 where $x_t = (\chi_t, \xi_t)^\top$ is a vector of state variables and $p_n(\cdot)$ is a polynomial function with a
66 degree at most n. χ_t and ξ_t are the short-term and long-term factors defined in Equation 2
67 and Equation 3 for risk-neutral processes. x_t satisfies the stochastic differential equation
68 Equation 6, with

$$b(x_t) = \begin{bmatrix} -\kappa\chi_t - \lambda_\chi \\ \mu_\xi - \gamma\xi_t - \lambda_\xi \end{bmatrix}, \sigma(x_t) = \begin{bmatrix} \sigma_\chi & 0 \\ 0 & \sigma_\xi \end{bmatrix}, a(x_t) = \sigma(x_t)\sigma(x_t)^\top = \begin{bmatrix} \sigma_\chi^2 & 0 \\ 0 & \sigma_\xi^2 \end{bmatrix}.$$

69 For any basis $H_n(x_t)$, the polynomial $p_n(x_t)$ can be uniquely represented as

$$p_n(x_t) = H_n(x_t)^\top \vec{p}.$$

70 The generator $\mathcal{G}$ is given by

$$\mathcal{G}f(x) = \frac{1}{2}Tr\left(\begin{bmatrix} \sigma_\chi^2 & 0 \\ 0 & \sigma_\xi^2 \end{bmatrix}\nabla^2 f(x)\right) + \begin{bmatrix} -\kappa\chi_t - \lambda_\chi \\ \mu_\xi - \gamma\xi_t - \lambda_\xi \end{bmatrix}^\top \nabla f(x).$$

71 By applying $\mathcal{G}$ to each element of $H_n(x_t)$, we get the matrix representation G. Then, by
72 Theorem 1, the futures price $F_{t,T}$ is given by

$$F_{t,T} = \mathbb{E}^*(S_T|\mathcal{F}_t) = H(x_t)^\top e^{(T-t)G}\vec{p}. \tag{11}$$

73 Therefore, we have the non-linear state-space model

$$x_t = c + Ex_{t-1} + w_t, w_t \sim N(\mathbf{0}, \Sigma_w), \tag{12}$$

74 and

$$y_t = H_n(x_t)^\top e^{(T-t)G}\vec{p} + v_t, v_t \sim N(\mathbf{0}, \Sigma_v). \tag{13}$$

75 In this application, we assume the spot price is a polynomial of state variables with degree
76 $n = 2$, $S_t = \alpha_1 + \alpha_2\chi_t + \alpha_3\xi_t + \alpha_4\chi_t^2 + \alpha_5\chi_t\xi_t + \alpha_6\xi_t^2$, and the dimension of Pol_2 is $N = 6$.
77 The coordinate representation $\vec{p} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)^\top$. $\alpha_i, i = 1, \ldots, 6$ are parameters
78 which can be chosen by users.

Filtering methods

80 In this section, we use the notation

$$a_{t|t-1} := \mathbb{E}(x_t|\mathcal{F}_{t-1}), \qquad\qquad P_{t|t-1} := Cov(x_t|\mathcal{F}_{t-1}),$$
$$a_t := \mathbb{E}(x_t|\mathcal{F}_t), \qquad\qquad P_t := Cov(x_t|\mathcal{F}_t).$$
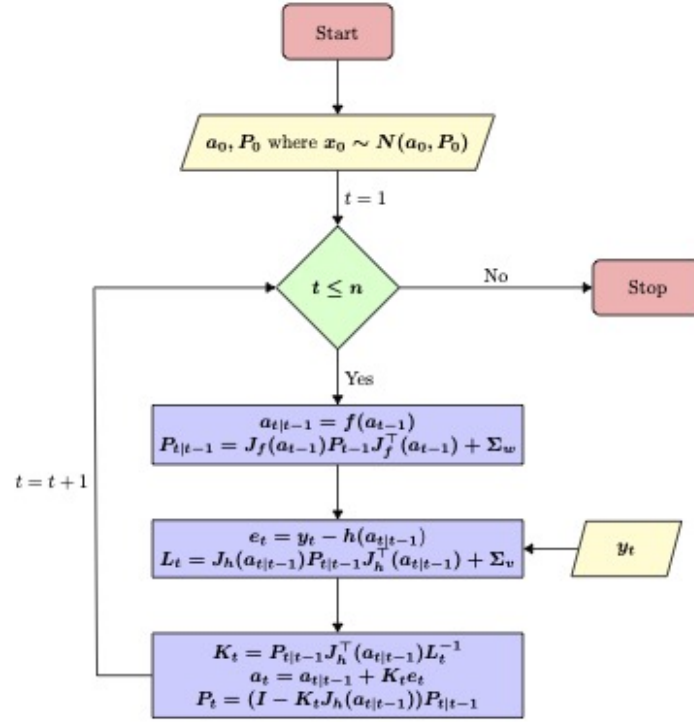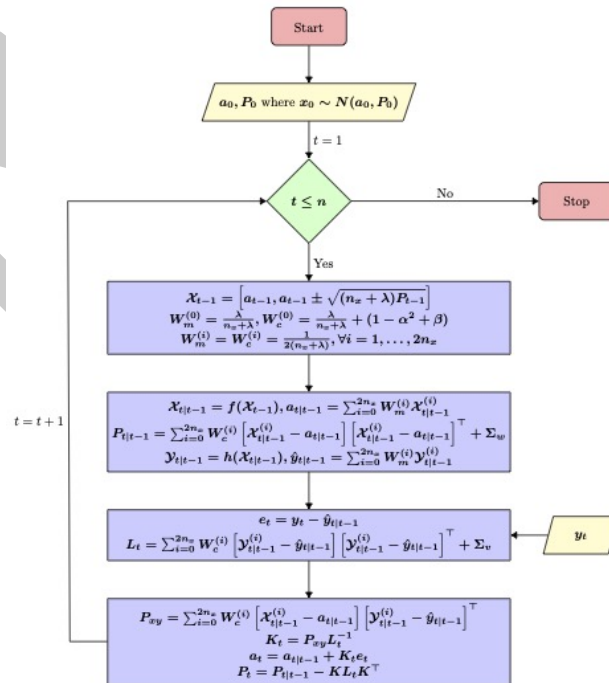
Figure 1: Flowcharts of EKF

Figure 2: Flowcharts of UKF

The Kalman Filter (KF) (Harvey, 1990) is a commonly used filtering method in estimating hidden state variables. However, KF can only deal with the linear Gaussian state model. To capture the non-linear dynamics in the PD model, we use Extended Kalman Filter (EKF)

84 (Julier & Uhlmann, 1997) and Unscented Kalman Filter (UKF) (Julier & Uhlmann, 2004; Wan
85 & Van Der Merwe, 2000).
Suppose we have a non-linear state-space model

$$x_t = f(x_{t-1}) + w_t, w_t \sim N(\mathbf{0}, \Sigma_w),$$
$$y_t = h(x_t) + v_t, v_t \sim N(\mathbf{0}, \Sigma_v).$$

87 The EKF linearises the state and measurement equations through the first-order Taylor series.
88 To run KF, we just replace J_f with E and replace J_h with F_t, where J_f and J_h are the
89 Jacobian of $f(\cdot)$ and $h(\cdot)$ respectively. In contrast, the UKF uses a set of carefully chosen
90 points, called sigma points, to represent the true distributions of state variables. Then, these
91 sigma points are propagated through the state equation. The flowchats of EKF and UKF are
92 given in Figure 1 and Figure 2. In this application, we use KF for the Schwartz-Smith model,
93 and EKF/UKF for the polynomial diffusion model.

Statement of need

95 This application is aimed at researchers who are pricing commodity futures by Schwartz-Smith
96 model or PD model. It has been designed with the following goals:

97 1. To provide a simulation tool for the polynomial diffusion model. Users can declare all
98   model specifications and parameters. The generated data is downloadable.

99 2. To provide two filtering methods, EKF and UKF, to estimate the futures prices and
100   hidden state variables. Currently, there is no filtering toolbox for the polynomial diffusion
101   model.

102 3. To provide well-designed visualisations. That includes the futures prices, the state
103   variables, the estimates of futures prices and state variables, and some downloadable
104   tables. Moreover, all these plots are interactive. Users can zoom in/out, highlight a
105   specific curve, download these plots, and so on.

106 4. To provide the estimation errors including root mean squared error (RMSE), mean
107   absolute error (MSE) and mean relative error (MRE). These errors are presented in
108   tables and plots.

109 5. To provide all functions listed above for the Schwartz-Smith model as a comparison.

Comparison with existing libraries

111 The R package "NFCP" (Aspinall et al., 2022) was developed for multi-factor pricing of
112 commodity futures, which is a generalisation of the Schwartz-Smith model. However, this
113 package doesn't accommodate the polynomial diffusion model. There are no R packages
114 available for PD models currently.

115 There are many packages in R for KF, for example, "dse", "FKF", "sspir", "dlm", "KFAS".
116 However, all of these packages are limited. "dse" can only take time-invariant state and
117 measurement transition matrices. "FKF" emphasizes computation speed but cannot run
118 smoother. "sspir", "dlm" and "KFAS" have no deterministic inputs in state and measurement
119 equations. For the non-linear state-space model, the functions "ukf" and "ekf" in package
120 "bssm" run the EKF and UKF respectively. However, this package was designed for Bayesian
121 inference where a prior distribution of unknown parameters is required. To achieve the best
122 collaboration of filters and models, we use our functions of KF, EKF and UKF.

References

124 Aspinall, T., Gepp, A., Harris, G., Kelly, S., Southam, C., & Vanstone, B. (2022). NFCP:
125   N-factor commodity pricing through term structure estimation. In *The Comprehensive R*

Archive Network. https://cran.rstudio.com/web/packages/NFCP/index.html; Comprehensive R Archive Network (CRAN).

Filipovic, D., & Larsson, M. (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics*, *20*(4), 931–972.

Harvey, A. C. (1990). *Forecasting, structural time series models and the kalman filter*. Cambridge University Press.

Julier, S. J., & Uhlmann, J. K. (1997). New extension of the kalman filter to nonlinear systems. *Signal Processing, Sensor Fusion, and Target Recognition VI*, *3068*, 182–193.

Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, *92*(3), 401–422.

Kleisinger-Yu, X., Komaric, V., Larsson, M., & Regez, M. (2020). A multifactor polynomial framework for long-term electricity forwards with delivery period. *SIAM Journal on Financial Mathematics*, *11*(3), 928–957.

Schwartz, E. S., & Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, *46*(7), 893–911.

Wan, E. A., & Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, 153–158.