Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 21/14

Anna Clevenhaus, Matthias Ehrhardt, Michael Günther

The Parareal Algorithm and the Sparse Grid Combination Technique in the Application of the Heston Model

May 4, 2021

http://www.imacm.uni-wuppertal.de

The Parareal Algorithm and the Sparse Grid Combination Technique in the Application of the Heston Model

A. Clevenhaus, M. Ehrhardt, and M. Günther

Abstract The sparse grid combination technique is an efficient method to reduce the course of dimensionality for high-dimensional problems, since it uses only selected sparse grids for spatial discretization. To further reduce the computational complexity in the temporal dimension, we choose the Parareal algorithm, a parallel-in-time algorithm. For the coarse and fine solvers in time, we use an efficient implementation of the Alternating Direction Implicit (ADI) method, which is an unusual choice due to the larger computational cost compared to the usual choice of one-step or Runge-Kutta methods. We analyze the Heston model with correlation as an example to illustrate this advantageous combination of the sparse grid with the Parareal algorithm. Finally, we present further ideas to improve this advantageous combination of methods.

1 American Option Pricing under the Heston model

The payoff function for a Put option with a predefined strike K and the price for the underlying asset S is given by

$$\phi(S) = \max(K-S,0).$$

To price an American put option we have to solve a free boundary value problem, where we seek for $\big(P(S,t),S_f(t)\big)$ in $t \in [0,T]$, where $S_f(t)$ is the free boundary value at time t and $P(S,t)$ fulfills

$$P(S,t) = \phi(S) \quad \text{for} \quad S \leq S_f(t), \quad P(S,t) > \phi(S) \quad \text{for} \quad S > S_f(t).$$

A. Clevenhaus · M. Ehrhardt · M. Günther
Bergische Universität Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany
e-mail: {clevenhaus, ehrhardt, guenther}@uni-wuppertal.de

The dynamics of the price S and the volatility ν are described by the Heston model with correlation [4]. The differential operator for $P(S,\nu,t)$ is given by

$$\mathscr{L}[P] = \frac{1}{2}\nu S^2\frac{\partial^2 P}{\partial S^2} + \rho_{S\nu}\sigma_\nu S\nu\frac{\partial^2 P}{\partial S\nu} + \frac{1}{2}\sigma_\nu^2\nu\frac{\partial^2 P}{\partial \nu^2} + rS\frac{\partial P}{\partial S} + \kappa_\nu(\nu-\mu_\nu)\frac{\partial P}{\partial \nu} - rP,$$

where r is the interest rate, $\nu > 0$ is the square of the volatility of the underlying, κ_ν is the mean-reversion rate and μ_ν is the long-term mean of the volatility ν and σ_ν is the volatility-of-variance. The correlation between S and ν is denoted by $\rho_{S\nu} \in [-1,1]$. If the Feller condition $2\kappa_\nu\mu_\nu > \sigma_\nu$ is fulfilled, $\nu > 0$ holds. After time reversal $\tau = T - t$, the differential operator has to fulfill the inequality

$$\frac{\partial P}{\partial \tau} - \mathscr{L}[P] \leq 0$$

and the initial condition

$$P(S,\nu,0) = \phi(S), \qquad S > S_f(0).$$

To avoid an explicit computation of the free boundary value problem, we apply an operator splitting and recast the problem into a *linear complementarity problem* with an auxiliary variable λ [7]

$$\begin{cases} \mathscr{L}[P] - \frac{\partial P}{\partial \tau} = \lambda, \\ \lambda \geq 0, P - \phi(S) \geq 0, \quad \big(P-\phi(S)\big)\lambda = 0. \end{cases} \tag{LCPλ}$$

In this mixed formulation of the LCP problem, λ plays the role of a Lagrange multiplier.

2 The Sparse Grid Combination Technique

The sparse grid idea is motivated to reduce the course of dimensionality for solving PDEs [1]. Let $\mathbf{x} \in \Omega_2 = [0,1]^2$ be defined by the multi-indices

$$\mathbf{l} = (l_1,l_2) \in \mathrm{N}_0^2, \quad \mathbf{j} = (j_1,j_2) \in \mathrm{N}_0^2, \quad \mathbf{N} = (N_1,N_2) = (2^{l_1},2^{l_2}). \tag{1}$$

such that we can define a tensor based grid $\Omega_\mathbf{l}$ whose grid nodes are given by

$$\mathbf{x}_{\mathbf{l},j} = (x_{l_1,j_1}, x_{l_2,j_2}) \quad \text{for} \quad j_1 = 0,\ldots,\mathrm{N}_1 \quad \text{and} \quad j_2 = 0,\ldots,\mathrm{N}_2.$$

The mesh width defined by this grid is $h = (2^{-l_1}, 2^{-l_2})$. To avoid errors due to sensitivities to disordered grid, we set a minimum for $l_i > l_{\min} = 3$, s.t. each spacial direction has at least 9 grid points. Let u be the continuous solution on Ω_2 and $u_\mathbf{l}$ the discrete solution on $\Omega_\mathbf{l}$ with $l = (l_1,l_2)$. The hierarchical surplus of $u_\mathbf{l}$

$$\delta(u_{\mathbf{l}}) = u_{\mathbf{l}} - u_{\mathbf{l}-e_1} - u_{\mathbf{l}-e_1} + u_{\mathbf{l}-e_1-e_2} \text{ with } e_1 = (1,0)^\top, e_2 = (0,1)^\top$$

Based on the error splitting

$$u - u_{\mathbf{l}} = h_1^2 w_1(h_1) + h_2^2 w_2(h_2) + h_1^2 h_2^2 w_{1,2}(h_1,h_2)$$

we derive the error spitting of the hierarchical surplus

$$\delta(u - u_{\mathbf{l}}) = \mathscr{O}\left(h_1^2 h_2^2\right) = \mathscr{O}\left(2^{-2|\mathbf{l}|_1}\right).$$

For the highest information gain for the sparse grid solution u_n^s of level $n = |\mathbf{l}|_1$, we use the hierarchical surplus and the error splitting and derive the sparse grid combination technique

$$u_n^s = \sum_{|\mathbf{l}|_1 \leq n} \delta(u_{\mathbf{l}}) = \sum_{|\mathbf{l}|_1 = n} u_{\mathbf{l}} - \sum_{|\mathbf{l}|_1 = n-1} u_{\mathbf{l}}$$

Since the sparse grid combination technique is developed on Ω_2, we define $x = (y,z) \in [0,1]^2$ and obtain $S \in [S_{\min}, S_{\max}]$ and $\nu \in [\nu_{\min}, \nu_{\max}]$ by using the following transformation

$$\psi^{-1}(y) = S_0 + \alpha \cdot \sinh(y \cdot (c_2 - c_1) + c_1),$$
$$c_1 = \sinh^{-1}\left(\frac{S_{\min} - S_0}{\alpha}\right), \qquad c_2 = \sinh^{-1}\left(\frac{S_{\max} - S_0}{\alpha}\right),$$

where α describes the non-uniformity of the grid. If α is small, we obtain a highly non-uniform grid and else wise a uniform grid. For z we use the transformation analogously. Using finite difference stencils of second order, the semi-discrete *partial differential complementarity problem (PDCPλ)*

$$\frac{\partial P}{\partial \tau} = FP(\tau) + \lambda(\tau), \quad P(\tau) \geq \phi\left(\psi^{-1}(y)\right), \quad \left(P(\tau) - \phi\left(\psi^{-1}(y)\right)\right)^\top \lambda(\tau) = 0,$$

is derived.

3 Temporal Discretization and the Parareal Algorithm

We discretize the time uniformly, using $\Delta_\tau = T/N_t$ we obtain the temporal time points $\tau_k = k \cdot \Delta_\tau$ with $k = 0, \ldots, N_t$. With u^k describing the discrete solution at time step τ_k and g describing the discrete payoff value, we gain the fully *discrete linear complementarity problem*, cf. [7]

$$\mathscr{T}(u^k, \lambda^k, \tau^k) = \begin{cases} u^{k+1} = Au^k + \Delta_\tau \lambda^k, \\ \lambda^{k+1} \geq 0, u^{k+1} \geq g, (\lambda^{k+1})^\top (u^{k+1} - g). \end{cases} \quad \text{(DLCP}\lambda\text{)}$$

Within this problem, we have to solve two separate problems. In the first step a system of linear equations has to be solved and in the second one a variable update is done. The system of equations is solved by the modified Craig-Sneyd scheme with the additional parameter λ

$$\begin{cases} Y_0 = u^k + \Delta_\tau \mathscr{A}(\tau^k, u^k) \boxed{+\Delta_\tau \lambda^k}, \\ Y_i = Y_{i-1} + \theta \Delta_\tau \left(\mathscr{A}_i(\tau^k, Y_i) - \mathscr{A}_i(\tau^k, u^k) \right), \quad i = 1,2, \\ \hat{Y}_0 = Y_0 + \theta \Delta_\tau \left(\mathscr{A}_0(\tau^k, Y_0) - \mathscr{A}_0(\tau^k, u^k) \right) \\ \tilde{Y}_0 = \hat{Y}_0 + (\frac{1}{2} - \theta) \Delta_\tau \left(\mathscr{A}(\tau^k, \hat{Y}_0) - \mathscr{A}(\tau^k, u^k) \right) \\ \tilde{Y}_i = \tilde{Y}_{i-1} + \theta \Delta_\tau \left(\mathscr{A}_i(\tau^k, Y_i) - \mathscr{A}_i(\tau^k, u^k) \right), \quad i = 1,2, \\ \tilde{u}^{k+1} = \tilde{Y}_2, \end{cases}$$

where $\mathscr{A}_0$ is the operator for the mixed derivatives, $\mathscr{A}_1$ the operator of the derivatives of the first coordinate direction, $\mathscr{A}_2$ the operator of the derivative of the second direction and $\mathscr{A}$ the sum of all operators. Further a improved way of implementation of the ADI schemes is used [9]. Since numerical results show $N_1 - 2N_2 = 2^{l_1} - 2 \cdot 2^{l_2} = 0$ is a feasible choice [5], we apply additional restrictions to $\mathbf{l}$ [2]. The restrictions can variate from the strict condition $l_1 > l_2$ being fulfilled for every single sparse grid to a softer condition where $\max l_1 > \max l_2$. The second step, the variable update can be done component wise by applying

$$\begin{cases} u^{k+1} & = \max(\tilde{u}^{k+1} - \Delta_\tau \lambda^k, u^0), \\ \lambda^{k+1} & = \max(0, \lambda^k + (u^0 - \tilde{u}^{k+1})/\Delta_\tau) \end{cases}$$

Since we have an initial condition, we set λ^0 as the zero vector.
The Parareal algorithm is an iterative parallel-in-time method and can be viewed as either a multigrid method or a multiple shot method [8]. For both solvers, we consider the temporal operators as previously described. The difference between the fine and coarse solvers is based on the spatial domain considered, the fine solver $\mathscr{F}$ solves the problem on u_n^s with $N_{\mathscr{F}}$ time steps and the coarse solver $\mathscr{G}$ on u_{n-1}^s with $N_{\mathscr{G}}$ time steps. We initialize the algorithm by introducing N_τ equal time slices, s.t. $\tilde{\tau}_p = [\tau_0 + (p-1) \cdot \Delta_{\tilde{\tau}}, \tau_0 + p \cdot \Delta_{\tilde{\tau}}]$, where $\Delta_{\tilde{\tau}} = \frac{T}{N_\tau}$. The initial value for each time slice is calculated by the coarse solver. The initial value for the first time slice is always given by the initial condition. Since the fast solver solves one time slice in each iteration, the maximum number of iterations J must be much smaller than N_τ. After initialization, the iterative procedure begins. First, the fine solver computes in parallel the solution of each time slice with the initial values. Let u_j^i be the discrete solution to the time slice $\tilde{\tau}_i$ at the j-th iteration. A serial correction step over all time slices follows

$$u_{i+1}^{j+1} = \mathscr{G}(u_i^{j+1}, t_i, t_{i+1}) + \mathscr{F}(u_i^j, t_i, t_{i+1}) - \mathscr{G}(u_i^j, t_i, t_{i+1})$$

4 Numerical Results

In this section, we analyze the effect of reducing the grid resolution in the volatility direction on the accuracy as well as the application of the Parareal algorithm to the run time. We consider the following set of parameters

$$T = 0.25,\ K = 10,\ \rho = 0.1,\ r = 0.1,\ \kappa = 5,\ \mu = 0.16,\ \sigma_v = 0.9,\ J = 3$$
$$|l|_1 = 12, l_{\min} = 3, S \in [0, 3K], v \in [0, 3], \alpha_S = \alpha_v = 2, N_\tau = 16,\ N_F = 100,\ N_G = 25.$$

Table 1 contains the accuracy results for different grid resolutions, we get that even for very small volatility values and a high reduction in resolution the error is comparable to the full sparse grid solution, which requires almost twice the amount of sparse grid and thus twice the computation time.

		$\tilde{v} = 0.0625$					
$\tilde{S}$		8	9	10	11	12	
[3]		2.0000	1.1081	0.5204	0.2143	0.0827	Grids
Reduced Resolution	0	2.0000	1.1078	0.5202	0.2138	0.0821	13
	1	2.0000	1.1078	0.5202	0.2138	0.0821	11
	2	2.0000	1.1075	0.5202	0.2138	0.0821	9
	3	2.0000	1.1076	0.5201	0.2137	0.0821	7

Table 1 Solution values for the different spot asset prices and spot volatilities for the parameter sets compared to reference values.

Figure 1 shows the run time results for different parallel processors. We observe that the sparse grid technique is more efficient than the combination with the Parareal algorithm, due to increased communication time. To underline this fact, we observe that the runtime increases almost linearly with the number of processors.

5 Conclusion and Outlook

The numerical results show that even the strong additional restriction $l_1 > l_2$, which leads to a high resolution reduction in the volatility direction is feasible and leads to a large reduction of the computational effort and thus to a shorter runtime. To obtain better results for using the Parareal algorithm in combination with the sparse grid approach, we need to further improve the resulting algorithm. Fortunately, there are two ways to reduce the computational cost. The first idea is based on the structure of the sparse grid combination technique. Since in the presented approach all sparse grids of level $n-1$ have to be computed by the fine and the coarse solver, we can easily reduce the overhead by reusing the results. The second is based on paralleliz-

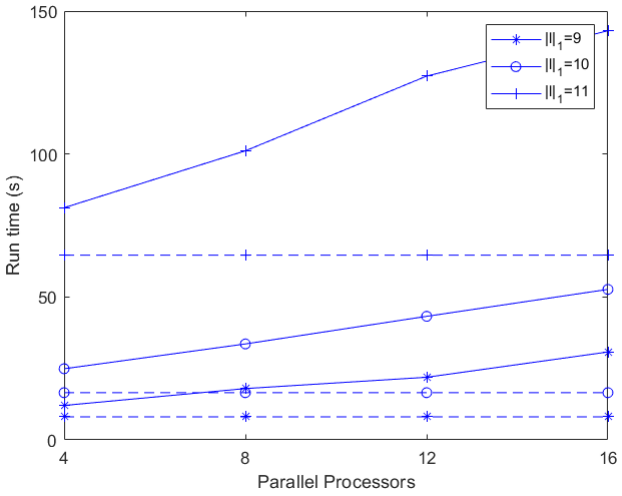

Fig. 1 The dashed line corresponds to the constant serial run time and the solid line represents the run time for the Parareal Algorithm with 4, 8, 12 and 16 parallel processors.

ing the computation of the sparse grids within the coarse solver, since they can each be computed independently.

References

1. H. Bungartz and M. Griebel, Sparse Grids, Cambridge University Press, 2004, 1–123.
2. A. Clevenhaus, M. Ehrhardt, M. Günther, An ADI Sparse Grid method for pricing efficiently American Options under the Heston model, to appear: Adv. App. Math. Mech., (2021)
3. T. Haentjens and K. J. in't Hout, ADI schemes for pricing American Options under the Heston model, Appl. Math. Fin., 22 (2013), 207-237.
4. S. L. Heston, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, In: Review of Financial Studies 6 (2) (1993), 327-343.
5. K. J. in't Hout and S. Foulon, ADI Finite Difference Schemes for Option Pricing in the Heston Model with Correlation, Int. J. Numer. Anal. Mod., 7 (2010), 303-320.
6. K. J. in't Hout and B. Welfert, Unconditional stability of second-order ADI schemes applied to multi-dimensional diffusion equations with mixed derivative terms, Appl. Numer. Math., 59(3-4) (2009), 677-692.
7. S. Ikonen and J. Toivanen, Operator splitting methods for pricing American options under stochastic volatility, Numer. Math., 113(2) (2009), 299-324.
8. J.-L. Lions, Y. Maday, and G. Turinici, Résolution d'EDP par un schéma en temps "pararéel", C.R.A.S. Sér. I Math., 332(7) (2000), 661–668.
9. L. Teng and A. Clevenhaus. Accelerated implementation of the ADI schemes for the Heston model with stochastic correlation. *J. Comput. Sci.*, 36 (2019), 101022
10. C. Zenger, Sparse Grids, Technical Report, Institut für Informatik, Technische Universität München, October 1990.