Exam: Solutions

1. **a)** The Heaviside function

$$H(x) := \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

is an example of a function in $L^2(G) \setminus H^1(G)$.

b) A positive refracting time is necessary in situations of energy markets, where a positive amount of time is needed to restore capacity needed for the next exercise.

c) We expect to observe a convergence rate of less than 2 due to the low regularity of the initial data ($\notin H^2$). We could use a graded mesh in time to overcome this difficulty.

d) The Heston model is a stochastic volatility model, thus the market is incomplete (financial difference) and there is a range of arbitrage-free prices. Numerically we have to solve a two-dimensional problem to calculate prices in a Heston model whereas in the CEV model the problem is one-dimensional.

e) For $\frac{1}{2} \leq \theta \leq 1$, the θ timestepping scheme is unconditionally stable.
For $0 \leq \theta < \frac{1}{2}$ there is a condition on the step size $k > 0$ for it to be stable: there exists a constant $C > 0$ independent of h and θ such that

$$k \leq C \frac{h_{\text{min}}^2}{1 - 2\theta}.$$

f) The corresponding covariance matrix $\mathcal{Q}$ is positive definite, i.e., there exists a constant $\gamma > 0$ such that $x^\top \mathcal{Q} x \geq \gamma x^\top x$, $\forall x \in \mathbb{R}^d$.

g) i) The condition $\beta_+ > 1$ is required to control the truncation error in the case of call options, due to linear growth of the payoff. For put options the value $\beta_+ = 1$ the condition that

$$\int_{|z|>1} e^z k(z) \mathrm{d}z < \infty$$

is also satisfied using that $\alpha > 0$.

ii) The case $k(z) = 0$, $z \in \mathbb{R}$, means that the Lévy measure is zero. This implies that the stiffness matrix is sparse, because the infinitesimal generator of X is just a differential operator of second order. Numerically, linear systems are fast to solve. In the case $k(z) > 0$, $z \in \mathbb{R}$, the infinitesimal generator of X has an integration part, which implies that the stiffness matrix has in general no

Please turn over.

sparsity. Numerically, this means that linear systems are expensive to solve, which limits the space discretization, where speed and needed memory to store the matrix becomes challenging.

iii) For $\alpha = 2$, we cannot conclude under the made assumption that

$$\int_{\mathbb{R}} \min\{1, z^2\} k(z) \mathrm{d}z < \infty.$$

The assumptions $k(z) \leq C/|z|^{1+\alpha}$ only implies that

$$\int_{|z|<1} z^2 k(z) \mathrm{d}z \leq C \int_{|z|<1} \frac{1}{|z|} \mathrm{d}z = 2C \int_0^1 \frac{1}{z} \mathrm{d}z = -2C \lim_{z \to 0} \log(z) = \infty.$$

iv) The process $\{e^{X_t} : t \geq 0\}$ is a martingale if and only if

$$\frac{\sigma^2}{2} + \gamma + \int_{\mathbb{R}} (e^z - 1 - z) \nu(\mathrm{d}z) = 0.$$

Continued overleaf.

2. **a)** The interest rate and volatilities r, Σ are assumed to be constant in time, which is not necessarily the case. In fact, market behaviour suggests that the volatilities behave like a so-called skew, i.e. they are convex functions of the spot price. Another argument is that the BS market model is a pure diffusion model and does not incorporate jumps, where jumps are obviously observed in markets.

b) According to the Feynman-Kac procedure in Section 8.1 of the lecture notes, a process given by the system of SDEs (1) has an infinitesimal generator

$$\mathcal{A}[f] := \frac{1}{2}\operatorname{tr}\left[\Sigma(\mathbf{s})\Sigma(\mathbf{s})^\top D^2 f(\mathbf{s})\right] + b(\mathbf{s})^\top \nabla f(\mathbf{s}) \,.$$

Hence, the pricing PDE is given as

$$\begin{aligned} \partial_t V(t,\mathbf{s}) + \mathcal{A}[V](t,\mathbf{s}) - rV(t,\mathbf{s}) &= 0 \quad \text{in } J \times \mathbb{R}^d \,, \\ V(T,\mathbf{s}) &= g(\mathbf{s}) \quad \text{in } G \,. \end{aligned}$$

c) Let us fix a time $t \in J$. We multiply with a test function $v \in C_0^1(G)$ and integrate over G. The test functions are taken with compact support because of the homogeneous Dirichlet boundary conditions. We obtain

$$(\partial_t u(t,\cdot), v) - \frac{1}{2}\int_G v(\mathbf{x})\, \nabla\cdot(\mathcal{Q}\nabla u(t,\mathbf{x}))\, d\mathbf{x} + \int_G \mu^\top \nabla u(t,\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x} + r\int_G u(t,\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x} = 0$$

and by application of a corollary of the Divergence Theorem, this is equivalent to

$$\begin{aligned} (\partial_t u(t,\cdot), v) + \frac{1}{2}\int_G \nabla v(\mathbf{x})^\top\, \mathcal{Q}\nabla u(t,\mathbf{x})\, d\mathbf{x} + \int_G \mu^\top \nabla u(t,\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x} + r\int_G u(t,\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x} \\ = \int_{\partial G} \boldsymbol{\nu}(\mathbf{x})^\top \nabla u(t,\mathbf{x})\, v(\mathbf{x})\, d\Gamma(\mathbf{x}) = 0 \,, \quad (1) \end{aligned}$$

where $\boldsymbol{\nu}$ denotes the outward pointing unit normal vector and Γ is a parametrization of ∂G. The last equality is implied by the compact support of v.
We define the bilinear form $a(w,v) := \frac{1}{2}\int_G \nabla w(\mathbf{x})^\top\, \mathcal{Q}\nabla v(\mathbf{x})\, d\mathbf{x} + \int_G \mu^\top \nabla w(\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x} + r\int_G w(\mathbf{x})\, v(\mathbf{x})\, d\mathbf{x}$.
We define $H^{-1}(G) := (H_0^1(G))^*$ to be the dual of $H_0^1(G)$. The calculations above motivate to state the following weak form:

Find $u \in H^1(J; H^{-1}(G)) \cap L^2(J; H_0^1(G))$ s. t. for all $v \in H_0^1(G)$ and for a.e. $t \in J$:

$$\begin{aligned} \langle \partial_t u(t,\cdot), v\rangle_{H^{-1}(G)\times H_0^1(G)} + a(u(t,\cdot), v) &= 0 \\ (u(0,\cdot), v) &= (g(\exp(\cdot), \ldots, \exp(\cdot)), v) \,. \end{aligned}$$

This equation is well-posed, and the initial condition is well-defined since by the Theorem in Chapter 4, $u \in C^0(\bar{J}; L^2(G))$.

d) We claim that if $\mathcal{Q}$ is positive definite, $a(\cdot,\cdot)$ satisfies the Gårding inequality. A matrix $\mathcal{Q}$ is positive definite if and only if there is a constant $\gamma > 0$ such that for

Please turn over.

all $0 \neq x \in \mathbb{R}^d$, $x^\top \boldsymbol{Q} x \geq \gamma x^\top x$.
First of all, the Dirichlet boundary conditions imply that

$$\int_G \nabla v(\mathbf{x}) v(\mathbf{x})\, d\mathbf{x} = \frac{1}{2}\int_G \nabla(v(\mathbf{x})^2) = 0 \,.$$

Hence we can go on and compute for $v \in H_0^1(G)$:

$$\begin{aligned}
a(v,v) &= \frac{1}{2}\int_G \nabla v(\mathbf{x})^\top \boldsymbol{Q} \nabla v(\mathbf{x})\, d\mathbf{x} + \int_G \mu^\top \nabla v(\mathbf{x})^2\, d\mathbf{x} + r\int_G v(\mathbf{x})^2\, d\mathbf{x} \\
&\geq \frac{\gamma}{2}|\nabla v|^2_{H^1(G)} + r\|v\|^2_{L^2(G)} \\
&= \frac{\gamma}{2}\|\nabla v\|^2_{H^1(G)} + (r - \gamma/2)\,\|v\|^2_{L^2(G)} \\
&\geq \frac{\gamma}{2}\|\nabla v\|^2_{H^1(G)} - |r - \gamma/2|\,\|v\|^2_{L^2(G)} \,.
\end{aligned}$$

This holds with no further restriction to r.

e) Let N be the number of total degrees of freedom in a FE approximation of the d-dimensional problem. It is proportional to the total computational work. As seen in class, the convergence rate is given by $\mathcal{O}(N^{-1/d})$, hence for $d \geq 3$, Monte Carlo yields a better approximation rate (independent of the dimension).
A related topic is that the matrices get larger with the dimension. On a d-dimensional product mesh with one-dimensional meshwidth h, we have that $N \sim h^{-d}$. If d is large, matrices of size $\sim h^{-d}$ can get too large to store and to do efficient numerical computations on them.

f) We are looking for a matrix $\mathbf{A}$ with entries $A_{i,i'} := a(\phi_{i'}, \phi_i)$.
We plug in the formulae for $\phi_i, \phi_{i'}$:

$$\begin{aligned}
A_{i,i'} &= a(\phi_{i'}, \phi_i) = a(b_{i_1'} b_{i_2'}, b_{i_1} b_{i_2}) \\
&= \frac{Q_{11}}{2}\int_G \partial_{x_1}(b_{i_1'} b_{i_2'})\partial_{x_1}(b_{i_1} b_{i_2})\mathrm{d}x + \frac{Q_{12}}{2}\int_G \partial_{x_1}(b_{i_1'} b_{i_2'})\partial_{x_2}(b_{i_1} b_{i_2})\mathrm{d}x \\
&\quad + \frac{Q_{21}}{2}\int_G \partial_{x_2}(b_{i_1'} b_{i_2'})\partial_{x_1}(b_{i_1} b_{i_2})\mathrm{d}x + \frac{Q_{22}}{2}\int_G \partial_{x_2}(b_{i_1'} b_{i_2'})\partial_{x_2}(b_{i_1} b_{i_2})\mathrm{d}x \\
&\quad + \mu_1 \int_G \partial_{x_1}(b_{i_1'} b_{i_2'}) b_{i_1} b_{i_2}\mathrm{d}x + \mu_2 \int_G \partial_{x_2}(b_{i_1'} b_{i_2'}) b_{i_1} b_{i_2}\mathrm{d}x + r\int_G b_{i_1'} b_{i_2'} b_{i_1} b_{i_2}\mathrm{d}x \\
&= \frac{Q_{11}}{2}\int_{-R}^{R} b_{i_1'}' b_{i_1}'\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'} b_{i_2}\mathrm{d}x_2 + \frac{Q_{12}}{2}\int_{-R}^{R} b_{i_1'}' b_{i_1}\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'} b_{i_2}'\mathrm{d}x_2 \\
&\quad + \frac{Q_{21}}{2}\int_{-R}^{R} b_{i_1'} b_{i_1}'\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'}' b_{i_2}\mathrm{d}x_2 + \frac{Q_{22}}{2}\int_{-R}^{R} b_{i_1'} b_{i_1}\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'}' b_{i_2}'\mathrm{d}x_2 \\
&\quad + \mu_1 \int_{-R}^{R} b_{i_1'}' b_{i_1}\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'} b_{i_2}\mathrm{d}x_2 + \mu_2 \int_{-R}^{R} b_{i_1'} b_{i_1}\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'}' b_{i_2}\mathrm{d}x_2 \\
&\quad + r\int_{-R}^{R} b_{i_1'} b_{i_1}\mathrm{d}x_1 \int_{-R}^{R} b_{i_2'} b_{i_2}\mathrm{d}x_2.
\end{aligned}$$

Continued overleaf.

Using that the one-dimensional matrices $\mathbf{S}$, $\mathbf{B}$ and $\mathbf{M}$ are given by $\mathbf{S}_{i,i'} = \int b'_{i'} b'_{i}$, $\mathbf{B}_{i,i'} = \int b'_{i'} b_i$ and $\mathbf{M}_{i,i'} = \int b_{i'} b_i$, and noting that $\int b_{i'} b'_i = -\int b'_{i'} b_i = -\mathbf{B}_{i,i'}$, we obtain

$$
\begin{aligned}
A_{j,j'} = {} & \frac{Q_{11}}{2}\mathbf{S}_{i_1,i'_1}\mathbf{M}_{i_2,i'_2} + \frac{Q_{12}}{2}\mathbf{B}_{i_1,i'_1}(-\mathbf{B})_{i_2,i'_2} + \frac{Q_{21}}{2}(-\mathbf{B})_{i_1,i'_1}\mathbf{B}_{i_2,i'_2} + \frac{Q_{22}}{2}\mathbf{M}_{i_1,i'_1}\mathbf{S}_{i_2,i'_2} \\
& + \mu_1 \mathbf{B}_{i_1,i'_1}\mathbf{M}_{i_2,i'_2} + \mu_2 \mathbf{M}_{i_1,i'_1}\mathbf{B}_{i_2,i'_2} + r\mathbf{M}_{i_1,i'_1}\mathbf{M}_{i_2,i'_2}.
\end{aligned}
$$

Denote by $\mathbf{S}^k$ the matrix with entries $\mathbf{S}_{i_k,i'_k}$, $k = 1, 2,$, and analogously for $\mathbf{B}^k$, $\mathbf{M}^k$. By our anterior calculations, the stiffness matrix $\mathbf{A}$ to the bilinear form $a(\cdot,\cdot)$ is given by

$$
\begin{aligned}
\mathbf{A} = {} & \frac{Q_{11}}{2}\mathbf{S}^1 \otimes \mathbf{M}^2 - Q_{12}\mathbf{B}^1 \otimes \mathbf{B}^2 + \frac{Q_{22}}{2}\mathbf{M}^1 \otimes \mathbf{S}^2 \\
& + \mu_1 \mathbf{B}^1 \otimes \mathbf{M}^2 + \mu_2 \mathbf{M}^1 \otimes \mathbf{B}^2 + r\mathbf{M}^1 \otimes \mathbf{M}^2.
\end{aligned}
$$

Since the Kronecker product is distributive, we can reduce the number of products by

$$
\mathbf{A} = \left(\frac{Q_{11}}{2}\mathbf{S}^1 + \mu_1\mathbf{B}^1 + r\mathbf{M}^1\right) \otimes \mathbf{M}^2 + \big(-Q_{12}\mathbf{B}^1 + \mu_2\mathbf{M}^1\big) \otimes \mathbf{B}^2 + \frac{Q_{22}}{2}\mathbf{M}^1 \otimes \mathbf{S}^2.
$$

g)

```
L = 5;
N = 2.^(L+1)-1;

sigma = [.1; .3];
rho = -.6;
Q=[sigma(1)^2, sigma(1)*sigma(2)*rho; sigma(1)*sigma
   (2)*rho, sigma(2)^2];

% one dimensional matrices
M = matrices(@(x)(0), @(x)(0), @(x)(1));
S = matrices(@(x)(1), @(x)(0), @(x)(0));
B = matrices(@(x)(0), @(x)(1), @(x)(0));


% tensor products
A = 0.5*Q(1,1)*kron(S,M) + 0.5*Q(2,2)*kron(M,S) ...
   - Q(1,2)*kron(B,B) + (0.5*Q(1,1)-r)*kron(B,M) ...
   + (0.5*Q(2,2)-r)*kron(M,B) + r*kron(M,M);
```

h)

```
function gvec = L2proj(g, N)
  g_rhs = int2dbasis(g, N);
  M_oned = matrices(@(x)(0), @(x)(0), @(x)(1));

  M = kron(M_oned, M_oned);

```

Please turn over.

```
    gvec = M\g_rhs;
end
```

i) We consider the worst-case convergence rate w.r.t. to N_{2D}. The quasi-optimal convergence rate is $\mathcal{O}(N_{2D}^{-1/2})$ since we solve the equation in two price dimensions. However, since the initial condition is not in $H^1(G)$, we expect a lower convergence rate than this. As we have seen in the Problem Sheets, a β-graded selection of timesteps $\{t_m := (m/M)^\beta\}_{m=1,\dots,M}$, graded towards $t = 0$ with a parameter $\beta > 1$ big enough results in the optimal rate again.
Another approach is to use a uniform time-grid, the implicit Euler scheme for a low number of steps with timestep k^2, and to continue with Crank-Nicolson.

Continued overleaf.

3. **a)** Define the matrices

$$\mathbf{M}_{ij} = \int_G b_i b_j, \qquad \mathbf{A}^{\mathrm{BS}}_{ij} = a^{\mathrm{BS}}(b_j, b_i),$$

and the vector

$$(\underline{f})_j = a^{\mathrm{BS}}(g, b_j).$$

The algebraic formulation is given by

$$\begin{aligned} &\mathbf{B}\underline{u}_N^{m+1} \geq \underline{F}^m \\ &\underline{u}_N^{m+1} \geq \underline{0} \\ &(\underline{u}_N^{m+1})^\top (\mathbf{B}\underline{u}_N^{m+1} - \underline{F}^m) = 0, \end{aligned}$$

where $\mathbf{B} := \mathbf{M} + \theta \Delta t \mathbf{A}^{\mathrm{BS}}$ and $\underline{F}^m := \Delta t \underline{f} + (\mathbf{M} - \Delta t(1-\theta)\mathbf{A})\,\underline{u}_N^m$, where Δt is the time step.

b) Implicit Euler

c) i) The long butterfly payoff is given by, for $0 < K_0 < K_1 < K_2$, with $K_0 - 2K_1 + K_2 = 0$:

$$g(s) = \max(K_0 - s, 0) - 2\max(K_1 - s, 0) + \max(K_2 - s, 0). \tag{2}$$

or by

$$g(s) = \max(s - K_2, 0) - 2\max(s - K_1, 0) + \max(s - K_0, 0).$$

Note that $g(s) = 0$ for all $s < K_0$ and for all $s > K_2$.

ii) A long butterfly strategy is profitable is the future volatility of the underlying assets is expected to be low (i.e., the investor makes a profit if the future volatility is lower than the implied volatility).

iii) Change

```
payoff = max(K-S, 0);
```

to

```
payoff = max(K0-S, 0) - 2*max(K1-S, 0) + max(K2-S
    ,0);
```

and replace lines 22–26 with

```
Ks = [K0, K1, K2];
fs = zeros(n+2, 3)
for i=1:3
  K = Ks(i);
  j = find(x<=log(K), 1, 'last');
  fs(2:j-1, i) = - r*K*h;
  fs(j, i) = sigma^2/2*K*(x(j+1)-log(K))/h - r*K
    /2*(2*h-1/h*(x(j+1)-log(K))*(x(j+1)-log(K)));
```

```
  fs(j+1, i) = sigma^2/2*K*(x(j)<log(K))*(-x(j)+log(K))/h - r*K/2*(1/h*(log(K)-x(j))*(log(K)-x(j)));
end
f = fs(:, 1) - 2*fs(:, 2) + fs(:, 3);
```

where we have exploited (2) and the fact that the listing provided a computation of the right hand side for a single put payoff.

d) There holds, for all $s \in \mathbb{R}_+$,

$$|\sigma(s)s| \leq 1 + 2|s|$$

and, for all $s, t \in \mathbb{R}_+$

$$|\sigma(s)s - \sigma(t)t| \leq 2|s - t|.$$

Hence, by Theorem 1.2.6 in the book, the SDE admits a unique solution.

e) The function $v(t, s)$ satisfies the following system of inequalities

$$\begin{array}{rcll} \partial_t v - \mathcal{A}^{\mathrm{LV}} v + rv & \geq & 0 & \text{in } J \times \mathbb{R}_+, \\ v(t,s) & \geq & g(s) & \text{in } J \times \mathbb{R}_+, \\ (\partial_t v - \mathcal{A}^{\mathrm{LV}} v + rv)(g - v) & = & 0 & \text{in } J \times \mathbb{R}_+, \\ v(0,s) & = & g(s) & \text{in } \mathbb{R}_+, \end{array} \tag{3}$$

where

$$(\mathcal{A}^{\mathrm{LV}} f)(s) = \frac{1}{2} s^2 \sigma(s)^2 \partial_{ss} f(s) + rs\partial_s f(s).$$

f) The variational formulation reads

Find $u_R \in L^2(J; H^1_0(G)) \cap H^1(J; L^2(G))$ such that $u_R(t, \cdot) \in \mathcal{K}_{0,R}$ and

$$(\partial_t u_R, v - u_R) + a^{\mathrm{LV}}(u_R, v - u_R) \geq -a^{\mathrm{LV}}(g \circ \exp, v - u_R), \;\; \forall v \in \mathcal{K}_{0,R}, \tag{4}$$

$u_R(0) = 0$.

We shall now derive the bilinear form $a^{\mathrm{LV}}(\cdot, \cdot)$. Let

$$\widetilde{\sigma} = \sigma \circ \exp\,.$$

The bilinear form associated to $f(x) \mapsto [-\frac{1}{2}\widetilde{\sigma}^2 \partial_{xx} f(x) - (r - \frac{\widetilde{\sigma}^2}{2})\partial_x f(x) + rf(x)]$: for $\phi, \varphi \in C_0^\infty(G)$,

$$a^{\mathrm{LV}}(\phi, \varphi) = -\int_{-R}^{R} \frac{1}{2}\widetilde{\sigma}^2 \partial_{xx}\phi\varphi \mathrm{d}x + \int_{-R}^{R} \left(\frac{\widetilde{\sigma}^2}{2} - r\right) \partial_x\phi\varphi + r\phi\varphi \mathrm{d}x.$$

Note that by integration by parts

$$\int_{-R}^{R} \frac{1}{2}\widetilde{\sigma}^2 \partial_{xx}\phi\varphi \mathrm{d}x = \frac{1}{2}\widetilde{\sigma}^2 \partial_x\phi\varphi\Big|_{-R}^{R} - \int_{-R}^{R} \frac{1}{2}\partial_x\phi\partial_x(\widetilde{\sigma}^2\varphi)\mathrm{d}x = -\int_{-R}^{R} \frac{1}{2}\partial_x\phi(\widetilde{\sigma}^2\partial_x\varphi + 2\widetilde{\sigma}\partial_x\widetilde{\sigma}\varphi)\mathrm{d}x.$$

Thus,

$$a^{\mathrm{LV}}(\phi, \varphi) = \int_{-R}^{R} \frac{\widetilde{\sigma}^2}{2}\partial_x\phi\partial_x\varphi \mathrm{d}x + \int_{-R}^{R} \left(\widetilde{\sigma}\partial_x\widetilde{\sigma} + \frac{\widetilde{\sigma}^2}{2} - r\right) \partial_x\phi\varphi + r\phi\varphi \mathrm{d}x.$$

The testfunction space is in this case $V = H^1_0(G)$, since we have Dirichlet boundary conditions.

Continued overleaf.

g) We provide the full listing:

```
function [u, fb] = amput_LV(n, R, T, K, r)

  h = 2*R/(n+1);
  x = linspace(-R, R, n+2)';
  S = exp(x);
  M = ceil(T/h);
  k = T/M;


  sigma = @(x) min(1+exp(x), 2);
  sigma_p = @(x) exp(x)*(x<=0);

  alpha = @(x)0.5 * sigma(x).^2;
  beta = @(x)sigma(x).*sigma_p(x) + 0.5*sigma(x).^2 - r;

  gamma = @(x)r;

  Am = matrices(x, @(x)0, @(x)0, @(x)1);
  A = matrices(x, alpha, beta, gamma);

  payoff = max(K-S, 0);

  u = zeros(n+2,1);

  fb =zeros(M+1,1);
  fb(1) = K;

  dof = 2:n+1;

  f = -A*payoff(dof);

  B = Am+k*A;
  C = Am;

  % loop over time points
  for i=1:m
      u(dof) = psor(B, k*f + C*u(dof), zeros(n, 1));

      J1 = find(u(dof) > 1.e-6);
      fb(i+1) = S(J1(1));
  end

  u = u + payoff;
end
```

Please turn over.