GEM - International Journal on Geomathematics (2025) 16:16
https://doi.org/10.1007/s13137-025-00273-4

RESEARCH

A parallel adaptive space-time discontinuous Galerkin method for transport in porous media

Daniele Corallo[1] · **Christian Wieners**[1]

Received: 31 March 2025 / Accepted: 15 September 2025

Abstract
We introduce a parallel adaptive space-time discontinuous Galerkin method for the linear transport equation, where the transport vector is determined from Darcy's law for porous media flow. Given the permeability distribution, in the first step the pressure head and the flux is computed by a mixed approximation of the linear porous media problem. Then, for a given initial pollution distribution the linear transport is approximated by an adaptive DG space-time discretization on a truncated space-time cylinder which turns out to be very efficient since the adaptively refined region is transported with the pollution distribution. The full linear system in space and time is solved with a parallel multigrid method where the stopping criterium for the linear solver is controlled by the convergence of a linear goal functional. Finally we apply this method to solve the inverse problem to reconstruct the initial pollution distribution from measurements of the outflow.

Keywords Space-time methods · Discontinuous Galerkin discretization · Linear transport in porous media

Mathematics Subject Classification 65M60

1 Introduction

The numerical approximation of transport in porous media is a very classical application in geosciences which has been studied for many decades. The first challenge is the appropriate approximation of the flux determined by Darcy's law for porous media flow, which yields a second order linear elliptic equation for the pressure head depending on the permeability of the material. Here, the simple approximation with standard finite elements is not appropriate, since in this case the corresponding discontinuous

✉ Christian Wieners
christian.wieners@kit.edu

Daniele Corallo
daniele.corallo@kit.edu

[1] Institute of Applied and Numerical Mathematics, KIT, Kaiserstr. 12, 76128 Karlsruhe, Germany

Published online: 22 September 2025

flux approximation is not volume preserving. So, in the first step it is required, e.g., to use a mixed method for the pressure head and the flux, so that one gets a conforming flux approximation in the Hilbert space H(div). This is then appropriate for the next step, the linear transport of a pollution density along this flux. The linear transport problem is hyperbolic, and finite volume and discontinuous Galerkin methods in space combined with suitable time-stepping methods are now well established (see, e.g., Bear and Bachmat (2012) for an overview of modeling aspects, the textbooks (Di Pietro and Ern 2011; Ern and Guermond 2021; Hesthaven and Warburton 2007) for numerical methods, and Flemisch et al. (2011); Bastian et al. (2021) for parallel software).

Recently, space-time discretizations for the transport problem were developed for two reasons. In standard porous media applications the pollution distribution is locally in space but transported in time, so that reasonable adaptivity strategies also require to be time-dependent. More complex are inverse problems which aim to recover the initial distribution of the pollution density from measurements at the outflow (see, e.g., the overview in Becker and Rannacher (2001)). The established solution method by gradient descent requires solving the adjoint problem backward in time, where the right-hand side of the adjoint problem relies on the solution of the forward problem in space and time. Adaptive space-time methods for parabolic optimization problems where introduced in Meidner and Vexler (2007), and meanwhile this is also applied to hyperbolic applications (see, e.g., Besier and Rannacher (2012)).

In this work we summerize our recent results on space-time methods and we demonstrate how this applies to classical problems in porous media. The article is organized as follows. We will start with summarizing the standard approach realized in our parallel finite element software system M++ (Baumgarten and Wieners 2021), where the Darcy equation is solved with Raviart-Thomas finite elements of lowest order and then using DG in space (without adaptivity) together with different established time stepping methods for the linear transport problem. Then, in Sect. 3 we derive a weak variational formulation in space and time which builds the basis for the discontinuous Galerkin method introduced in Sect. 4, where we present the extension to the adaptive space-time DG method introduced in Corallo et al. (2023) for general linear hyperbolic Friedrichs systems and specified in Wieners (2023) for linear transport. Now we combine this with the space-time multigrid method introduced first for a Petrov-Galerkin method in time in Dörfler et al. (2016) and extended to a DG method in time in Dörfler et al. (2019). In addition, we show that this can be realized on a truncated space-time mesh analyzed in Ernesti and Wieners (2019), which turns out to be very efficient for practical applications where the initial conditions and the measurements are local in space.

The efficiency of the adaptive method is demonstrated for a typical porous media setting with impermeable obstacles in Sect. 5, and the parallel performance of the adaptive multigrid method is demonstrated for a large-scale example. Finally, in Sect. 6 we show that the space-time discretization forward and backward in time is well suited for an inverse problem where we recover the position of the pollution source by measuring the outflow.

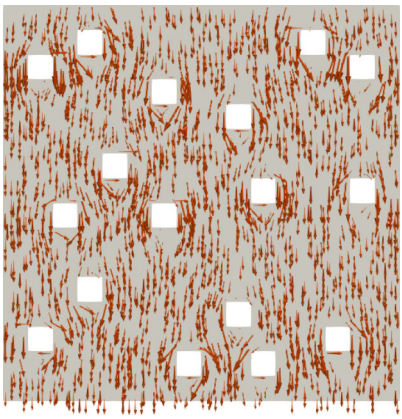


Fig. 1 Flux vector $\mathbf{q} = -\kappa \nabla p$ for water seeping into the groundwater

2 An application for transport in porous media

We start with an illustrative example in 2D. Let $\Omega \subset (0,1)^2$ be a simplified configuration (cf. Fig. 1) modeling top earth layers of sand with permeability distribution

$$\kappa : \overline{\Omega} \longrightarrow (\kappa_{\min}, \kappa_{\max}) \subset (0, \infty) .$$

In this configuration $(0,1)^2 \setminus \Omega$ are impermeable stones and rocks. The surface where it is raining is denoted by $\Gamma_{\text{top}} = (0,1) \times \{1\} \subset \partial\Omega$, and let $\Gamma_{\text{bottom}} = [0,1] \times \{0\} \subset \partial\Omega$ be the groundwater level. In the first step, we compute the flux vector $\mathbf{q} : \overline{\Omega} \longrightarrow \mathbb{R}^2$ by solving the Darcy equation, i.e., depending on Dirichlet and Neumann data p_{D} and g_{N} we have to determine the pressure head $p : \overline{\Omega} \longrightarrow \mathbb{R}$

$$-\nabla \cdot \kappa \nabla p = 0 \ \text{ in } \ \Omega , \quad p = p_{\text{D}} \ \text{ on } \Gamma_{\text{D}} = \Gamma_{\text{bottom}} , \quad -\kappa \nabla p = g_{\text{N}} \ \text{ on } \ \Gamma_{\text{N}} = \partial\Omega \setminus \Gamma_{\text{D}} .$$

This defines the flux vector $\mathbf{q} = -\kappa \nabla p$, and in direction of the outer normal vector $\mathbf{n}$ the inflow boundary $\Gamma_{\text{in}} = \big\{ \mathbf{x} \in \partial\Omega : \mathbf{n}(\mathbf{x}) \cdot \mathbf{q}(\mathbf{x}) < 0 \big\}$.

In the second step, the pollution density $u : (0,T) \times \Omega \longrightarrow \mathbb{R}$ is transported along the flux vector $\mathbf{q}$ solving

$$\partial_t u(t, \mathbf{x}) + \operatorname{div}(u(t, \mathbf{x}) \mathbf{q}(\mathbf{x})) = 0 , \qquad (t, \mathbf{x}) \in (0,T) \times \Omega$$

with initial distribution $u(0, \mathbf{x}) = u_0(\mathbf{x})$ for $\mathbf{x} \in \Omega$ at $t = 0$ and inflow boundary condition $u(t, \mathbf{x}) = u_{\text{in}}(t, \mathbf{x})$ for $(t, \mathbf{x}) \in (0,T) \times \Gamma_{\text{in}}$.

In our example we simply use $\kappa \equiv 1$, $p_{\text{D}} = 0$ on Γ_{bottom}, $g_{\text{N}} = 1$ on Γ_{top} and $g_{\text{N}} = 0$ on $\partial\Omega \setminus (\Gamma_{\text{top}} \cup \Gamma_{\text{bottom}})$, and we only consider $u_{\text{in}} = 0$. We obtain $p \geq 0$ in Ω by monotonicity principle (Knabner and Angermann 2021, Thm. 8.21) and thus $(\mathbf{n} \cdot \mathbf{q}) u \geq 0$ on $(0,T) \times \Gamma_{\text{bottom}}$.

We start with $u_0(\mathbf{x}) = B(\mathbf{x} - \mathbf{z})$ for a bubble function with $B(\mathbf{d}) = 0$ for $|\mathbf{d}| > r_0$ and the source position $\mathbf{z} \in \Omega$, where $\mathbf{d} = \mathbf{x} - \mathbf{z}$. The solution depending on the

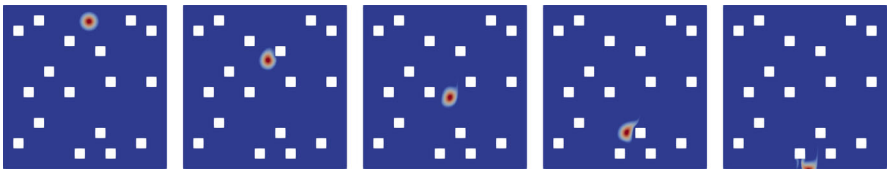


Fig. 2 Distribution of the pollution density $u_{\mathbf{z}}(t)$ at $t = 0$ and then along the transport for $t > 0$. This example is included in the tutorial of our parallel finite element software system M++ (Baumgarten and Wieners 2021)

source position is denoted by $u_{\mathbf{z}}$. For the simulation in Fig. 2 the source position is $\mathbf{z} = (0.5, 0.8025)^\top$ and $B(\mathbf{d}) = \max\{0, 100 - 1600|\mathbf{d}|\}$.

3 Modeling transport in porous media

Let $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3) be a bounded domain in space with Lipschitz boundary $\partial\Omega$ and outer normal vector $\mathbf{n}$. In general, the permeability can be anisotropic, i.e., it is a symmetric tensor, and we assume that $\boldsymbol{\kappa}$ is uniformly positive definite and bounded. The flux vector $\mathbf{q}$ is determined by the difference of the pressure head p at two points $\mathbf{x}_0, \mathbf{x}_1 \in \overline{\Omega}$: for any smooth curve $\gamma : [0, 1] \longrightarrow \overline{\Omega}$ from $\mathbf{x}_0 = \gamma(0)$ to $\mathbf{x}_1 = \gamma(1)$ we have

$$p(\mathbf{x}_1) - p(\mathbf{x}_0) = -\int_0^1 \boldsymbol{\kappa}(\gamma(s))^{-1}\mathbf{q}(\gamma(s)) \cdot \dot{\gamma}(s)\mathrm{d}s\,.$$

If the pressure head is smooth, this results for the limit $\mathbf{x}_1 \longrightarrow \mathbf{x}_0$ into the material law $\mathbf{q} = -\boldsymbol{\kappa}\nabla p$. In the following, our discretization is based on the weak formulation: testing with smooth vector fields $\boldsymbol{\phi} : \overline{\Omega} \longrightarrow \mathbb{R}^d$ with $\mathbf{n} \cdot \boldsymbol{\phi} = 0$ on Γ_{N}, integrating by parts and inserting the Dirichlet boundary values results into

$$\int_\Omega \left(p \operatorname{div} \boldsymbol{\phi} - \boldsymbol{\kappa}^{-1}\mathbf{q} \cdot \boldsymbol{\phi}\right) \mathrm{d}\mathbf{x} = \int_{\Gamma_{\mathrm{D}}} p_{\mathrm{D}}\mathbf{n} \cdot \boldsymbol{\phi}\, \mathrm{d}\mathbf{a}\,. \tag{1a}$$

Flux preserving requires for all inner subdomains $K \subset \Omega$ where the outer normal $\mathbf{n}_K$ exists almost everywhere on ∂K

$$\int_{\partial K} \mathbf{q} \cdot \mathbf{n}_K \, \mathrm{d}\mathbf{a} = 0\,,$$

so that we obtain, integrating by parts, for all smooth test functions $\psi : \overline{\Omega} \longrightarrow \mathbb{R}$ with $\psi = 0$ on Γ_{D} and inserting the Neumann data

$$\int_\Omega \operatorname{div} \mathbf{q}\, \psi \, \mathrm{d}\mathbf{x} = \int_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}}\psi \, \mathrm{d}\mathbf{a}\,. \tag{1b}$$

Together, (1) yields a saddle point formulation for $(p, \mathbf{q}) \in \mathrm{L}_2(\Omega) \times \mathrm{H}(\mathrm{div}, \Omega)$ of the elliptic Darcy equation.

For the transport of the pollution density u we define the time interval $I = (0, T)$ and the space-time cylinder $Q = (0, T) \times \Omega$. Flux preserving for the linear flux function $\mathbf{f}(u) = u\mathbf{q}$ requires for all convex subdomains $K \subset \Omega$ and all time intervals $(t_1, t_2) \subset I$

$$\int_K u(t_1)\,\mathrm{d}\mathbf{x} = \int_K u(t_2)\,\mathrm{d}\mathbf{x} + \int_{t_1}^{t_2} \int_{\partial K} \mathbf{f}(u) \cdot \mathbf{n}_K \,\mathrm{d}\mathbf{a}\,\mathrm{d}t\,.$$

If the pollution density and the flux are smooth, this yields $\partial_t u + \mathrm{div}(u\mathbf{q}) = 0$ in Q.

In the following, our discretization is based on the weak space-time formulation. Using smooth test functions $\varphi\colon \overline{Q} \longrightarrow \mathbb{R}$ with $\varphi(T) = 0$ and $\varphi = 0$ on $I \times \partial\Omega \setminus \Gamma_{\mathrm{in}}$, integrating by parts and inserting the initial data and the inflow boundary values results into

$$\int_Q u\big(\partial_t \varphi + \mathbf{q} \cdot \nabla\varphi\big)\,\mathrm{d}(t, \mathbf{x}) = \int_0^T \int_{\Gamma_{\mathrm{in}}} \mathbf{n} \cdot \mathbf{q}\, u_{\mathrm{in}} \varphi \,\mathrm{d}\mathbf{a}\,\mathrm{d}t - \int_\Omega u_0 \varphi(0)\,\mathrm{d}\mathbf{x}\,. \tag{2}$$

4 Approximating transport in porous media

For $h \in \mathcal{H} \subset (0, h_0)$ let $\Omega_h = \bigcup_{K \in \mathcal{K}_h} K$ be meshes where the elements $K \subset \Omega$, $K \in \mathcal{K}_h$ are open triangles/tetrahedra. Let $F \in \mathcal{F}_K$ be the faces of the element K, and we set $\mathcal{F}_h = \bigcup_{K \in \mathcal{K}_h} \mathcal{F}_K$, so that $\partial\Omega_h = \overline{\bigcup_{F \in \mathcal{F}_h} F}$ is the skeleton in space and $\overline{\Omega} = \Omega_h \cup \partial\Omega_h$.

We assume that the meshes are compatible with the different boundary parts of the models, so that $\Gamma_{\mathrm{D}} = \overline{\bigcup_{F \in \mathcal{F}_h \cap \Gamma_{\mathrm{D}}} F}$. For every cell $K \in \mathcal{K}_h$ and inner face $F \in \mathcal{F}_K \cap \Omega$ the neighboring cell is denoted by $K_F \in \mathcal{K}_h$ and is determined by $\overline{F} = \partial K \cap \partial K_F$.

Approximating the Darcy equation we use Raviart-Thomas elements

$$W_h = \big\{(p_h, \mathbf{q}_h) \in \mathrm{L}_2(\Omega) \times \mathrm{H}(\mathrm{div}, \Omega)\colon\ p_h|_K \in \mathbb{P}_0(K) \text{ for all } K \in \mathcal{K}_h \text{ and } \mathbf{q}_h|_K \in \mathbb{P}_1(K)^d \text{ such that } \mathbf{n}_F \cdot \mathbf{q}_h|_F \in \mathbb{P}_0(F) \text{ for all } F \in \mathcal{F}_K\big\}\,,$$

and compute $(p_h, \mathbf{q}_h) \in W_h$ with $\int_F \mathbf{n} \cdot \mathbf{q}_h \,\mathrm{d}\mathbf{a} = \int_F g_{\mathrm{N}} \,\mathrm{d}\mathbf{a}$ for $F \in \mathcal{F}_h \cap \Gamma_{\mathrm{N}}$ solving

$$\int_\Omega \Big(\boldsymbol{\kappa}^{-1}\mathbf{q}_h \cdot \boldsymbol{\phi}_h - p_h \,\mathrm{div}\,\boldsymbol{\phi}_h - \mathrm{div}\,\mathbf{q}_h\,\psi_h\Big)\,\mathrm{d}\mathbf{x} = -\int_{\Gamma_{\mathrm{D}}} p_{\mathrm{D}} \mathbf{n} \cdot \boldsymbol{\phi}_h \,\mathrm{d}\mathbf{a}$$

for all $(\psi_h, \boldsymbol{\phi}_h) \in W_h$ with $\mathbf{n} \cdot \boldsymbol{\phi}_h = 0$ on Γ_{N}.

For the discontinuous Galerkin method in space approximating the transport equation we define the discrete flux $\mathbf{f}_h(v_h) = v_h\,\mathbf{q}_h$ with the conforming approximation of the flux vector $\mathbf{q}_h$ in $\mathrm{H}(\mathrm{div}, \Omega)$. For the construction of the upwind flux, we observe that $\mathbf{q}_h \cdot \mathbf{n}_K$ is constant on every face F (since we use lowest-order Raviart-Thomas

elements). For discontinuous functions in space

$$\mathcal{S}_h = \left\{ v_h \in \mathrm{L}_2(\Omega) \colon v_{h,K} = v_h|_K \in \mathbb{P}(K) \text{ for all } K \in \mathcal{K}_h \right\}$$

which are polynomials in every cell K we observe for $v_h, w_h \in \mathcal{S}_h$

$$\left(\operatorname{div} \mathbf{f}_h(v_h), w_h \right)_{\Omega_h} = \sum_{K \in \mathcal{K}_h} \Big(-\left(\mathbf{f}_h(v_{h,K}), \nabla w_{h,K} \right)_K + \sum_{F \in \mathcal{F}_K} \left(\mathbf{f}_h(v_{h,K}) \cdot \mathbf{n}_K, w_{h,K} \right)_F \Big).$$

This is approximated with the bilinear form

$$a_h(v_h, w_h) = \sum_{K \in \mathcal{K}_h} \Big(- \left(\mathbf{f}_h(v_{h,K}), \nabla w_{h,K} \right)_K + \sum_{F \in \mathcal{F}_K} \left(\mathbf{f}^{\mathrm{up}}_{K,F}(v_h) \cdot \mathbf{n}_K, w_{h,K} \right)_F \Big)$$

by replacing $\mathbf{f}_h(v_{h,K}) \cdot \mathbf{n}_K$ with the upwind flux

$$\mathbf{f}^{\mathrm{up}}_{K,F}(v_h) = \begin{cases} \mathbf{f}_h(v_{h,K}), & F \in \mathcal{F}^{\mathrm{out}}_K \\ \mathbf{f}_h(v_{h,K_F}), & F \in \mathcal{F}^{\mathrm{in}}_K \setminus \Gamma_{\mathrm{in}} \\ \mathbf{0}, & F \in \mathcal{F}^{\mathrm{in}}_K \cap \Gamma_{\mathrm{in}} \end{cases}$$

depending of the in- or outgoing flux on the faces, i.e.,

$$\mathcal{F}^{\mathrm{out}}_K = \left\{ F \in \mathcal{F}_K \colon \mathbf{q}_h \cdot \mathbf{n}_K \geq 0 \text{ on } F \right\}, \qquad \mathcal{F}^{\mathrm{in}}_K = \left\{ F \in \mathcal{F}_K \colon \mathbf{q}_h \cdot \mathbf{n}_K < 0 \text{ on } F \right\}.$$

The discontinuous Galerkin method in space is complemented by the full-upwind discretization in time. For $0 = t_0 < t_1 < \cdots < t_N = T$, we define time intervals $I_{n,h} = (t_{n-1}, t_n)$ and

$$I_h = (t_0, t_1) \cup \cdots \cup (t_{N-1}, t_N) \subset I = (0, T), \qquad \partial I_h = \{t_0, t_1, \ldots, t_{N-1}, t_N\}.$$

Combined with the mesh in space, we obtain a decomposition into space-time cells $R \in \mathcal{R}_h = \left\{ I_{n,h} \times K \colon K \in \mathcal{K}_h \text{ and } n = 1, \ldots, N \right\}$ and

$$Q_h = I_h \times \Omega_h = \bigcup_{n=1}^{N} Q_{n,h} = \bigcup_{R \in \mathcal{R}_h} R, \qquad Q_{n,h} = \bigcup_{K \in \mathcal{K}_h} I_{n,h} \times K \ \subset \ I_{n,h} \times \Omega.$$

For $v_h, w_h \in \mathrm{H}^1(Q_h)$ we obtain after integration by parts in all intervals $I_{n,h}$

$$\begin{aligned} \left(\partial_t v_h, w_h \right)_{Q_h} = \sum_{n=1}^{N} \Big(& - \left(v_{n,h}, \partial_t w_{n,h} \right)_{Q_{n,h}} \\ & + \left(v_{n,h}(t_n), w_{n,h}(t_n) \right)_{\Omega} - \left(v_{n,h}(t_{n-1}), w_{n,h}(t_{n-1}) \right)_{\Omega} \Big). \end{aligned}$$

With $[w_h]_n = w_{n+1,h}(t_n) - w_{n,h}(t_n)$, $n = 1, \dots, N-1$ and $[w_h]_N = -w_{N,h}(t_N)$ set

$$m_h(v_h, w_h) = \sum_{n=1}^{N} \Big(- \big(v_{n,h}, \partial_t w_{n,h}\big)_{Q_{n,h}} - \big(v_{n,h}(t_n), [w_h]_n\big)_{\Omega} \Big) .$$

Together, we obtain in

$$\mathcal{V}_h = \big\{ v_h \in \mathrm{L}_2(Q) \colon v_{n,h,K} = v_{n,h}|_K \in \mathbb{P}(I_{n,h} \times K) \text{ for } K \in \mathcal{K}_h \text{ and } n = 1, \dots, N \big\}$$

the discrete bilinear form

$$b_h(v_h, w_h) = m_h(v_h, w_h) + \int_0^T a_h(v_h(t), w_h(t)) \, \mathrm{dt} , \qquad v_h, w_h \in \mathcal{V}_h .$$

For every discrete finite element space $V_h \subset \mathcal{V}_h$ the space-time approximation $u_h \in V_h$ is defined by solving the linear system

$$b_h(u_h, \varphi_h) = \int_0^T \int_{\Gamma_{\mathrm{in}}} \mathbf{n} \cdot \mathbf{q}_h \, u_{h,\mathrm{in}} \varphi_h \, \mathrm{d}(t, \mathbf{a}) - \int_{\Omega} u_0 \varphi_h(0) \, \mathrm{d}\mathbf{x} , \qquad \varphi_h \in V_h . \tag{3}$$

Well-posedness and convergence of the discrete approximation (3) are analyzed in Wieners (2023). In particular, up to the difference $\mathbf{q}_h - \mathbf{q}$ the bilinear form is consistent with the weak formulation (2): using smooth test functions $\varphi \colon \overline{Q} \longrightarrow \mathbb{R}$ with $\varphi(T) = 0$ and $\varphi = 0$ on $I \times \partial\Omega \setminus \Gamma_{\mathrm{in}}$ we have for all $v_h \in \mathcal{V}_h$

$$b_h(v_h, \varphi) = \int_Q v_h \big(\partial_t \varphi + \mathbf{q} \cdot \nabla \varphi \big) \, \mathrm{d}(t, \mathbf{x}) \quad + \text{data error approximating flux vector } \mathbf{q}.$$

In order to calibrate the accuracy in space and time, we assume

$$c_{\mathrm{ref}} \triangle t \le h , \qquad \triangle t = \max(t_n - t_{n-1}) , \qquad h = \max \operatorname{diam}(K) ,$$

where $c_{\mathrm{ref}} > 0$ is a reference velocity depending on the flux vector $\mathbf{q}$.

5 Convergence tests for the space-time DG method

Now we test the approximation for the example in Sect. 1 with $\Omega \subset (0,1)^2$ and the time interval $(0, T) = (0, 2)$. In Fig. 1 the flux vector $\mathbf{q}_h$ approximated with Raviart-Thomas elements is shown. In this application, we have a bounded transport velocity $\|\mathbf{q}\|_\infty \le c < \infty$, and we start with a local pollution source, so that $u_0(\mathbf{x}) = 0$ for $|\mathbf{x} - \mathbf{z}| \ge r_0$. Thus, we get for the solution $u_{\mathbf{z}}(t, \mathbf{x}) = 0$ for $|\mathbf{x} - \mathbf{z}| \ge r_0 + ct$.

In our application we only observe the space-time solution at the outflow boundary $(0, T) \times \Gamma_{\mathrm{bottom}}$, so we do not need to compute $u_{\mathbf{z}}(t, \mathbf{x})$ for $\operatorname{dist}(\mathbf{x}, \Gamma_{\mathrm{bottom}}) > c(T - t)$. Also for the reconstruction of the point source from the outflow data only the results

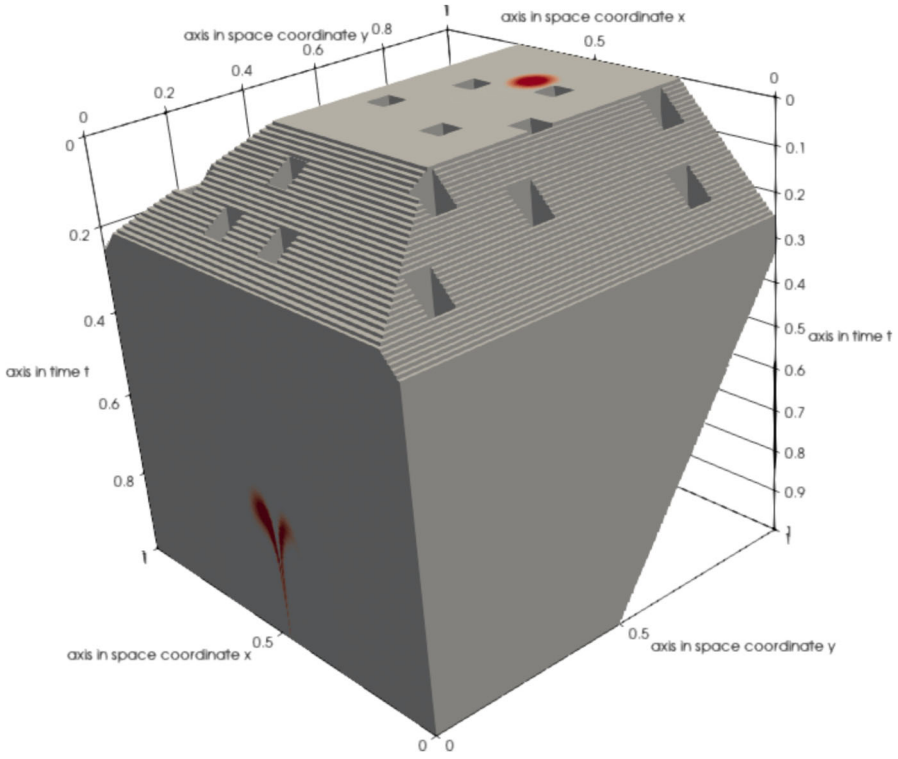


Fig. 3 Space-time approximation u_h on the truncated space-time cylinder Q^{tr} of the solution $u_{\mathbf{z}}$ for $\mathbf{z} = (0.5, 0.8015)^\top$ (with the initial value on top and the outflow on the left) shown in Fig. 2 at different time steps

on the outflow boundary are required. So it is sufficient to evaluate $u_{\mathbf{z}}$ on a truncated space-time domain $Q^{\mathrm{tr}} \subset (0, T) \times \Omega$, cf. Fig. 3.

Depending on the pollution density we define the goal functional

$$G_{\mathrm{out}}(u)(t) = \int_{\Gamma_{\mathrm{bottom}}} \mathbf{n} \cdot \mathbf{f}(u(t)) \, \mathrm{d}\mathbf{a} \,,$$

and we observe for the solution of the transport problem

$$\int_{\Omega} u_0 \, \mathrm{d}\mathbf{x} = \int_0^{\infty} G_{\mathrm{out}}(u)(t) \, \mathrm{dt} \,,$$

so that this quantity is a standard choice to test the reliability of the approximation. For our convergence tests we now compare $G_{\mathrm{out}}(u) \colon [0, T] \longrightarrow \mathbb{R}$ for different polynomial degrees in space and time. For the tests we use fixed time steps $\triangle t = T/N$ for $N \in \mathbb{N}_0$ and $t_n = n \triangle t$.

We start with the classical lowest-order approximation using finite volumes in space

$$S_h^{\mathrm{p}=0} = \left\{ v_h \in \mathrm{L}_2(\Omega) \colon v_{h,K} = v_h|_K \in \mathbb{P}_0(K) \text{ for all } K \in \mathcal{K}_h \right\}$$

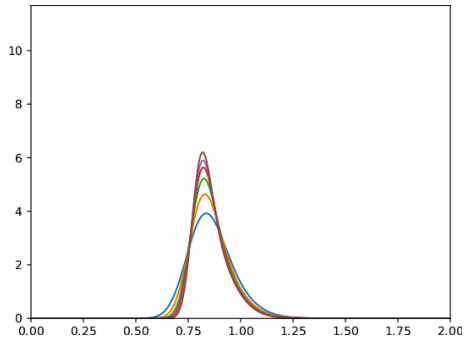
and a discrete approximation of the initial value $u_h^0 \in S_h^{\mathrm{p}=0}$. Then, we determine by the implicit Euler method for $n = 1, \ldots, N$

$$u_h^n \in S_h^{\mathrm{p}=0} \colon \qquad (u_h^n, v_h)_\Omega + \triangle t \, a_h(u_h^n, v_h) = (u_h^{n-1}, v_h)_\Omega \,, \qquad v_h \in S_h^{\mathrm{p}=0} \,,$$

since the implicit Euler method is preserving positivity and mass. Note that this is equivalent to the space-time approximation (3) in

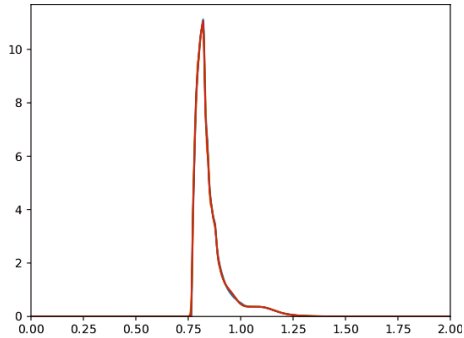
$$V_h^{\mathrm{p}=0} = \left\{ v_h \in \mathrm{L}_2(Q) \colon v_h|_R \in \mathbb{P}_0(R) \text{ for all } R \in \mathcal{R}_h \right\},$$

so that it can be computed parallel in time and on the truncated space-time cylinder.



Goal functional $G_{\mathrm{out}}(u_h)$ for $t \in [0, 2]$ and finite volume approximation in $S_h^{\mathrm{p}=0}$ with $\#\mathcal{K}_h = 122\,368$ triangles in space and $N_m = 160 \cdot 2^m$ time steps for $m = 1, \ldots, 7$ using the implicit Euler method.

We observe convergence in time, but the resolution of the mesh is not sufficient: comparing the finite volume method with the DG approximation of polynomial degree $\mathrm{p} = 2$ in space and implicit midpoint rule in time results in a far better approximation of the solution, but the approximation is not strictly positive. Nevertheless, the error in preserving positivity and mass of the implicit midpoint rule is very small for fine discretizations.



Goal functional $G_{\mathrm{out}}(u_h)$ for $t \in [0, 2]$ and DG approximation in $S_h^{\mathrm{p}=2}$ with $\#\mathcal{K}_h = 122\,368$ triangles in space and $N_m = 160 \cdot 2^m$ time steps for $m = 4, \ldots, 7$ using implicit midpoint rule.

This clearly shows, that finite volume in space and implicit Euler in time are not appropriate for approximating transport in porous media. Nevertheless, this approach with higher order approximations is numerically quite expensive. So this is compared with our space-time adaptive strategy. Therefore, we use the error indicator

$$\eta_{\text{res},h} = \Big(\sum_{R \in \mathcal{R}_h} \eta^2_{\text{res},R} \Big)^{1/2}$$

for $R = (t_{n-1}, t_n) \times K$, $n = 1, \dots, N$, with

$$\begin{aligned}
&\eta_{\text{res},R}(u_h)^2 \\
&\quad = \eta_{\text{res},n,K}(u_h)^2 + 2h_K \left\| \partial_t u_h + \operatorname{div} \mathbf{f}_h(u_h) \right\|^2_R \\
&\qquad + \frac{1}{4} \left\| |\mathbf{q}_h \cdot \mathbf{n}_K|^{1/2} [u_h]_{K,F} \right\|^2_{I_{n,h} \times \partial K \cap \Omega} + \left\| |\mathbf{q}_h \cdot \mathbf{n}|^{-1/2} (g_{\text{in}} - \mathbf{f}_h(u_h) \cdot \mathbf{n}) \right\| I_{n,h} \\
&\qquad \times \partial K \cap \Gamma^2_{\text{in}}
\end{aligned}$$

and the local contributions

$$\begin{aligned}
\eta_{\text{res},1,K}(u_h)^2 &= \frac{1}{2} \left\| u_0 - u_h(0) \right\|^2_K + \frac{1}{4} \left\| [u_h]_1 \right\|^2_K , && R = (0, t_1) \times K , \\
\eta_{\text{res},n,K}(u_h)^2 &= \frac{1}{4} \left\| [u_h]_{n-1} \right\|^2_K + \frac{1}{4} \left\| [u_h]_n \right\|^2_K , && R = (t_{n-1}, t_n) \times K , \ 1 < n < N , \\
\eta_{\text{res},N,K}(u_h)^2 &= \frac{1}{4} \left\| [u_h]_{N-1} \right\|^2_K , && R = (t_{N-1}, T) \times K .
\end{aligned}$$

The error indicator approximates the error $u - u_h$ in a suitable mesh-dependent DG norm, see Wieners (2023) for details.

For the following test, we use the notation $\eta_{\text{res,max}}(u_h) = \max_R \eta_{\text{res},R}(u_h)$ and p_R for the polynomial degree in the space-time cell R. We start with the solution $u_h^0 \in V_h^{\mathrm{p}=0}$, i.e., $\mathrm{p}_R = 0$ for all space-time cells R. Then, for $k = 1, 2, 3$ depending on parameters $0 < \theta_0 < \theta_1 < 1$ we increase the polynomial degree to $\mathrm{p}_R + 1$ for all space-time cells R with $\eta_{\text{res},R}(u_h) > \theta_1 \eta_{\text{res,max}}(u_h)$. The polynomial degree is decreased to $\mathrm{p}_R - 1$, if $\eta_{\text{res},R}(u_h) < \theta_0 \eta_{\text{res,max}}(u_h)$ and $\mathrm{p}_R > 0$.

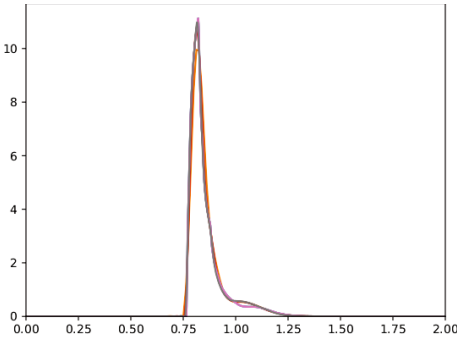


Table 1 Comparison of the adaptive space-time solution u_h^k with a reference solution u_h^{ref} using DG in space for p = 2 on 489 472 triangles and $N_7 = 160 \cdot 2^7$ time steps with implicit midpoint rule

k	0	1	2	3
$\dim V_h$	37,944,289	68,398,396	127,598,522	247,760,863
$\left\|G(u_h^m) - G(u_h^{\text{ref}})\right\|_{\text{L}_1(0,T)}$	0.1262	0.0570	0.0236	0.0146

Here, the goal functional $G_{\text{out}}(u_h)$ for the adaptive space-time approximations in Table 1 and the reference solution is presented. This clearly shows that the space-time adaptive approximation on the truncated space-time cylinder is very accurate with far less degrees of freedom in space and time compared with a classical time-stepping method. Nevertheless, one has to solve a very large linear system, where we use a parallel full space-time multigrid method with adaptivity. This extends our parallel space-time multigrid methods introduced in Dörfler et al. (2016, 2019) to the p-adaptive DG discretization for the refinement steps $k = 0, 1, 2, \ldots$ together with an integrated error control. This stops the multigrid iteration $u_{h\ell_{\max}}^{k,0}, u_{h\ell_{\max}}^{k,1}, u_{h\ell_{\max}}^{k,2}, u_{h\ell_{\max}}^{k,3}, u_{h\ell_{\max}}^{k,4}, \ldots$ when the finite element error of the discretization is achieved, since it is not required to solve the linear system with full arithmetic accuracy. Here this is tested heuristically, i.e., we choose a relative tolerance $\varepsilon_{\text{rel}} > 0$ and stop the iteration if no further progress in the approximation of the goal functional is achieved, i.e., if

$$\left\| G_{\text{out}}(u_{h\ell_{\max}}^{k,j}) - G_{\text{out}}(u_{h\ell_{\max}}^{k,j-1}) \right\|_{\text{L}_1(0,T)} \leq \varepsilon_{\text{rel}} \left\| G_{\text{out}}(u_{h\ell_{\max}}^{k,j}) \right\|_{\text{L}_1(0,T)}. \tag{4}$$

Together, the parallel solution algorithm works as follows (see Table 2 for examples):

- refine the mesh from level $\ell = 0$ up to $\ell_{\min}$, build the truncated space-time cylinder Q_h^{tr} on level $\ell_{\min}$ and distribute Q_h^{tr} non-overlapping onto the parallel process;
- determine and then distribute the parallel overlap required for the upwind flux;
- for $\ell = \ell_{\min}, \ldots, \ell_{\max}$ approximate the flux vector $\mathbf{q}_{h_\ell}$ and then assemble in parallel the finite element matrices for the linear system (3) for $\ell = \ell_{\min}, \ldots, \ell_{\max}$;
- approximate $u_{h\ell_{\max}}^0 \in V_{h\ell_{\max}}^{\text{p=0}}$ using the parallel space-time linear multigrid method with GMRES solver on the coarse level $\ell_{\min}$ and using the stopping criterion (4);
- starting adaptivity, set $u_{h\ell_{\max}}^{1,0} = u_{h\ell_{\max}}^0$ in $V_{h\ell_{\max}}^0 = V_{h\ell_{\max}}^{\text{p=0}}$;
- for $k = 1, \ldots, k_{\max}$ adapt the discretization by increasing and decreasing the polynomial degree depending on the space-time error indicator, prolongate $u_{h\ell_{\max}}^{k-1}$ to the new discretization and then approximate $u_{h\ell_{\max}}^k \in V_{h\ell_{\max}}^k$ using parallel space-time multigrid solver starting with the prologation and using the stopping criterion (4).

Table 2 Two examples for the parallel adaptive multigrid solver on 128 processes for $\ell_{\max} = 6$ and on 512 processes for $\ell_{\max} = 7$ with $2^{1+\ell_{\max}}$ time steps: we use $\ell_{\min} = 4$, $k_{\max} = 2$, and using N_{mg} multigrid V-cycles with block-Gauss-Seidel smoothing on every processor (where all degrees of freedom per space-time cell build a block). Here we use 10 smoothing steps on the levels $\ell < \ell_{\max}$ and more smoothing on the finest level $\ell = \ell_{\max}$. The Euclidean norm of the residual in $\mathbb{R}^{\dim V_h}$ is denoted by $|r_h|_2$. For the finest discretization with 3 886 260 361 DoFs the solution is obtained in approximately one hour using six adaptive multigrid steps on 512 parallel processes

$\ell_{\max}$	k	$\dim W_h$	$\dim V_h$	N_{mg}	$\lvert r_h\rvert_2$	$\lVert G_{\mathrm{out}}\rVert_{\mathrm{L}_1}$	$\lVert G_{\mathrm{out}}\rVert_{\mathrm{L}_\infty}$
6	0	2,941,056	228,288,512	7	0.0000002	0.99993	3.916
6	1		461,590,312	5	0.0000006	1.00880	10.593
6	2		616,774,443	4	0.0000092	1.00206	10.956
7	0	11,755,776	1,826,308,096	6	0.0000007	0.99985	5.071
7	1		3,188,987,351	7	0.0000365	1.00241	10.838
7	2		3,886,260,361	6	0.0000003	1.00074	11.052

6 Localization of the pollution source position

The reconstruction of the source position for the example in Sect. 1 is an established inverse problem. For a given source position $\mathbf{z} \in \Omega$, the outflow on the bottom is denoted by $g_{\mathbf{z}} = u_{\mathbf{z}}\mathbf{n} \cdot \mathbf{q} : (0, T) \times \Gamma_{\mathrm{bottom}} \longrightarrow \mathbb{R}$.

Now we assume that we do not know the source position which will be reconstructed from data $g_{\mathrm{data}} : (0, T) \times \Gamma_{\mathrm{bottom}} \longrightarrow \mathbb{R}$ for the outflow. We assume that the initial distribution is given by the bubble function B which is calibrated by mass conservation, i.e.,

$$\int_{\mathbb{R}^2} B(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_0^T \int_{\Gamma_{\mathrm{bottom}}} g_{\mathrm{data}}(t, \mathbf{x})\,\mathrm{d}\mathbf{a}\,\mathrm{d}t\,.$$

The source position $\mathbf{z}$ is located by minimizing the quadratic functional

$$J(\mathbf{z}) = \frac{1}{2}\int_0^T \int_{\Gamma_{\mathrm{bottom}}} \big(u_{\mathbf{z}}(t, \mathbf{x})\mathbf{n}(\mathbf{x}) \cdot \mathbf{q}(\mathbf{x}) - g_{\mathrm{data}}(t, \mathbf{x})\big)^2\,\mathrm{d}\mathbf{a}\,\mathrm{d}t \tag{5}$$

subject to the solution of the transport equation

$$\begin{aligned} \partial_t u_{\mathbf{z}} + \mathrm{div}(u_{\mathbf{z}}\mathbf{q}) = 0 \ \text{ in } \ (0, T) \times \Omega\,, \qquad u_{\mathbf{z}}(0, \mathbf{x}) = B(\mathbf{x} - \mathbf{z})\,, \quad \mathbf{x} \in \Omega\,, \\ u_{\mathbf{z}}(t, \mathbf{x}) = 0 \ \text{ on } \ (0, T) \times \Gamma_{\mathrm{in}}\,. \end{aligned} \tag{6}$$

Then we obtain for the derivative of the quadratic functional J at $\mathbf{z}$ in direction $\mathbf{y}$

$$\mathrm{DJ}(\mathbf{z})[\mathbf{y}] = \lim_{\delta \longrightarrow 0} \frac{1}{\delta} \int_\Omega u_{\mathbf{z}}^{\mathrm{ad}}(0, \mathbf{x})\big(B(\mathbf{x} - \mathbf{z} - \delta\mathbf{y}) - B(\mathbf{x} - \mathbf{z})\big)\,\mathrm{d}\mathbf{x}\,,$$

where the gradient is evaluated by solving the linear adjoint problem backward in time

$$\begin{aligned} \partial_t u_{\mathbf{z}}^{\mathrm{ad}} + \mathbf{q} \cdot \nabla u_{\mathbf{z}}^{\mathrm{ad}} = 0 \qquad & \text{in} \quad (0,T) \times \Omega\,, \qquad u_{\mathbf{z}}^{\mathrm{ad}}(T,\mathbf{x}) = 0\,, \quad \mathbf{x} \in \Omega\,, \\ u_{\mathbf{z}}^{\mathrm{ad}} \mathbf{n} \cdot \mathbf{q} = u_{\mathbf{z}} \mathbf{n} \cdot \mathbf{q} - g_{\mathrm{data}} \quad & \text{on} \quad (0,T) \times \Gamma_{\mathrm{bottom}}\,, \\ u_{\mathbf{z}}^{\mathrm{ad}} \mathbf{n} \cdot \mathbf{q} = 0 \qquad & \text{on} \quad (0,T) \times \partial\Omega \setminus (\Gamma_{\mathrm{bottom}} \cup \Gamma_{\mathrm{in}}) \end{aligned} \tag{7}$$

(see App. A for details).

To minimize $J(\cdot)$ we use a gradient descent method. Starting with a guess for the source point $\mathbf{z}_0 \in \Omega$, the source $\mathbf{z}$ is approximated by a gradient descent minimizing J. Therefore, we select a fixed mesh in space and compute the approximation of the flux vector first. Then we select a suitable time steps size and construct and distribute in parallel the corresponding truncated space-time cylinder. Then we select the space-time finite element space V_h for the transport equation, we assemble the two sparse matrices A_h and A_h^{ad}, and we construct the preconditioner for the forward and the adjoint problem, since in every step of the gradient descent only the right-hand is updated and we can use the same matrix.
Then, for $k = 0, 1, 2, 3, \ldots$

- compute the approximation of the forward problem $u_{h,\mathbf{z}_k}$ by assembling the right-hand side depending on the $\mathbf{z}_k$ and solving the linear system with matrix A_h (starting for $k > 0$ the iterative linear solver with $u_{h,\mathbf{z}_{k-1}}$);
- compute the approximation of the adjoint problem $u_{h,\mathbf{z}_k}^{\mathrm{ad}}$ by assembling the right-hand side depending on the forward problem $u_{h,\mathbf{z}_k}$ and solving the linear system with matrix A_h^{ad} (starting for $k > 0$ the iterative linear solver with $u_{h,\mathbf{z}_{k-1}^{\mathrm{ad}}}$);
- approximate the gradient of the goal functional by finite differences with $\delta_k = \frac{0.01}{2^k}$, i.e., for $\mathbf{y} = (1,0)^\top$ and $\mathbf{y} = (0,1)^\top$ we evaluate

$$\mathrm{DJ}_{\delta_k}(\mathbf{z}_k)[\mathbf{y}] = \frac{1}{\delta_k} \int_\Omega u_{h,\mathbf{z}_k}^{\mathrm{ad}}(0,\mathbf{x}) \big(B(\mathbf{x} - \mathbf{z}_k - \delta_k \mathbf{y}) - B(\mathbf{x} - \mathbf{z}_k) \big)\, \mathrm{d}\mathbf{x}\,,$$

  and then update the source location $\mathbf{z}_{k+1} = \mathbf{z}_k - \alpha_k \mathrm{DJ}_{\delta_k}(\mathbf{z}_k)$ with a suitable damping parameter $\alpha_k > 0$;
- stop the iteration if the goal functional is sufficiently small.

In our numerical tests the data are generated with $\mathbf{z} = (0.5, 0.8025)^\top$, and we start the iteration with the guess $\mathbf{z}_0 = (0.5, 0.6875)^\top$. The convergence of the method is shown in Table 3, and the solutions $u_{\mathbf{z}_k}$ and the adjoints $u_{\mathbf{z}_k}^{\mathrm{ad}}$ are illustrated in Fig. 4.

Conclusion and outlook

We have shown that our p-adaptive space-time discontinuous Galerkin method approximates classical applications in porous media very accurately, and that the parallel distribution of the truncated space-time cylinder allows for an efficient load balancing on many processors. Moreover, the space-time setting performs well for an inverse

Table 3 Convergence of the gradient descent method on level $\ell = 3$ and $\ell = 4$.

$\dim V_h^{\mathrm{p}=1} = 8\,702\,208$		$\dim V_h^{\mathrm{p}=1} = 70\,553\,040$	
$J(\mathbf{z}_0) = 27.924125$	$\mathbf{z}_0 = (0.5000, 0.6875)^\top$	$J(\mathbf{z}_0) = 31.735728$	$\mathbf{z}_0 = (0.5000, 0.6875)^\top$
$J(\mathbf{z}_1) = 11.484516$	$\mathbf{z}_1 = (0.4859, 0.7580)^\top$	$J(\mathbf{z}_1) = 6.886961$	$\mathbf{z}_1 = (0.4859, 0.7580)^\top$
$J(\mathbf{z}_2) = 3.274998$	$\mathbf{z}_2 = (0.4973, 0.7806)^\top$	$J(\mathbf{z}_2) = 1.380585$	$\mathbf{z}_2 = (0.4973, 0.7806)^\top$
$J(\mathbf{z}_3) = 0.679856$	$\mathbf{z}_3 = (0.5001, 0.7929)^\top$	$J(\mathbf{z}_3) = 0.327434$	$\mathbf{z}_3 = (0.5001, 0.7929)^\top$
$J(\mathbf{z}_4) = 0.102828$	$\mathbf{z}_4 = (0.5003, 0.7990)^\top$	$J(\mathbf{z}_4) = 0.085922$	$\mathbf{z}_4 = (0.5003, 0.7990)^\top$
$J(\mathbf{z}_5) = 0.011519$	$\mathbf{z}_5 = (0.5001, 0.8014)^\top$	$J(\mathbf{z}_5) = 0.024657$	$\mathbf{z}_5 = (0.5001, 0.8014)^\top$
$J(\mathbf{z}_6) = 0.000959$	$\mathbf{z}_6 = (0.5000, 0.8022)^\top$	$J(\mathbf{z}_6) = 0.007495$	$\mathbf{z}_6 = (0.5000, 0.8022)^\top$
$J(\mathbf{z}_7) = 0.000039$	$\mathbf{z}_7 = (0.5000, 0.8024)^\top$	$J(\mathbf{z}_7) = 0.002277$	$\mathbf{z}_7 = (0.5000, 0.8024)^\top$
$J(\mathbf{z}_8) = 0.000001$	$\mathbf{z}_8 = (0.5000, 0.8025)^\top$	$J(\mathbf{z}_8) = 0.000666$	$\mathbf{z}_8 = (0.5000, 0.8025)^\top$

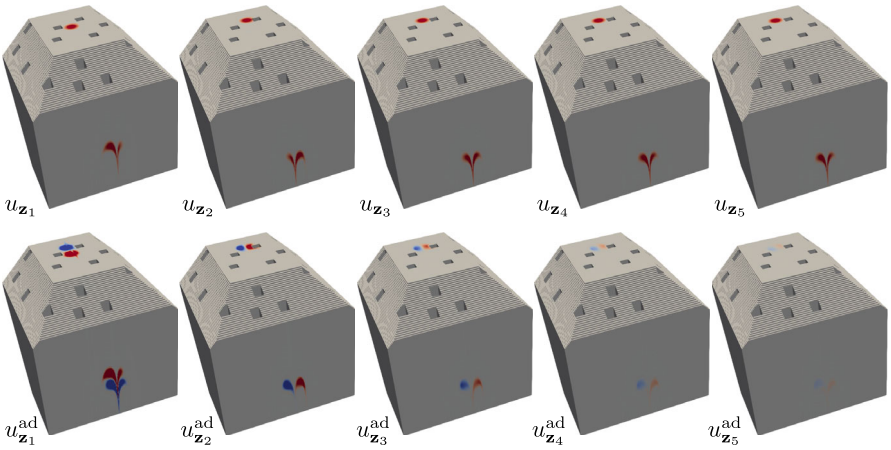


Fig. 4 Approximation of the solution and its adjoint on the truncated space-time cylinder for the first 5 steps of the gradient descent method

problem, where the adjoint solution backward in time depends on the forward problem in the full time interval. Here this is tested for a simplified configuration where we identify only the position of the initial pollution distribution. For the more complex problem to identify further parameters describing the initial pollution distribution the same procedure can be used, but in general an additional regularization term is required.

For the next steps in our research the adaptive space-time discretization will be extended to applications in three space dimensions, and it will be combined with our budgeted multilevel Monte Carlo method (Baumgarten et al. 2024) for the probabilistic evaluation of the flux vector $\mathbf{q}$, so that we can also approximate porous media applications with uncertain permeability.

A Evaluation of the gradient DJ

Depending on the pollution source position $\mathbf{z} \in \Omega$ we define $B_{\mathbf{z}}(\mathbf{x}) = B(\mathbf{x} - \mathbf{z})$ and the forward operator $\mathcal{F}\colon \Omega \longrightarrow \mathrm{L}_2(Q)$ solving the transport equation (6) by

$$\int_Q \mathcal{F}(\mathbf{z})\big(\partial_t \varphi + \mathbf{q} \cdot \nabla \varphi\big)\,\mathrm{d}(t, \mathbf{x}) = -\int_\Omega B_{\mathbf{z}} \varphi(0)\,\mathrm{d}\mathbf{x}$$

for all smooth test functions $\varphi\colon \overline{Q} \longrightarrow \mathbb{R}$ with $\varphi(T) = 0$ and $\varphi = 0$ on $(0, T) \times \partial\Omega \setminus \Gamma_{\mathrm{in}}$. Then, we have $\mathcal{F}(\mathbf{z}) = u_{\mathbf{z}}$, which yields for the quadratic functional (5)

$$J(\mathbf{z}) = \frac{1}{2}\left\| \mathcal{F}(\mathbf{z})\mathbf{n} \cdot \mathbf{q} - g_{\mathrm{data}} \right\|^2_{(0,T)\times\Gamma_{\mathrm{bottom}}}$$

the derivative

$$\mathrm{DJ}(\mathbf{z})[\mathbf{y}] = \big(\mathcal{F}(\mathbf{z})\mathbf{n}\cdot\mathbf{q} - g_{\mathrm{data}}, \mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\mathbf{n}\cdot\mathbf{q}\big)_{(0,T)\times\Gamma_{\mathrm{bottom}}}$$

by inserting the derivative $\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]$ of the forward operator $\mathcal{F}(\mathbf{z})$ which is given by the solution of

$$\int_Q \mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\big(\partial_t\varphi + \mathbf{q}\cdot\nabla\varphi\big)\,\mathrm{d}(t,\mathbf{x}) = -\int_\Omega \mathrm{D}B_{\mathbf{z}}[\mathbf{y}]\varphi(0)\,\mathrm{d}\mathbf{x}$$

for all test functions φ. Inserting $u_{\mathbf{z}}^{\mathrm{ad}}$ solving (7) yields

$$\begin{aligned}
\mathrm{DJ}(\mathbf{z})[\mathbf{y}] &= \big(u_{\mathbf{z}}^{\mathrm{ad}}, \mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\mathbf{n}\cdot\mathbf{q}\big)_{(0,T)\times\Gamma_{\mathrm{bottom}}}\\
&= \big(u_{\mathbf{z}}^{\mathrm{ad}}, \mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\mathbf{n}\cdot\mathbf{q}\big)_{(0,T)\times\partial\Omega}\\
&= \int_Q \nabla\cdot\big(u_{\mathbf{z}}^{\mathrm{ad}}\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\mathbf{q}\big)\,\mathrm{d}(t,\mathbf{x})\\
&= \int_Q \nabla\cdot\big(u_{\mathbf{z}}^{\mathrm{ad}}\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\mathbf{q}\big)\,\mathrm{d}(t,\mathbf{x}) + \int_Q \partial_t\big(u_{\mathbf{z}}^{\mathrm{ad}}\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\big)\,\mathrm{d}(t,\mathbf{x})\\
&\quad - \int_\Omega \big(u_{\mathbf{z}}^{\mathrm{ad}}(T)\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}](T) - u_{\mathbf{z}}^{\mathrm{ad}}(0)\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}](0)\big)\,\mathrm{d}\mathbf{x}\\
&= \int_Q u_{\mathbf{z}}^{\mathrm{ad}}\big(\partial_t\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}] + \mathbf{q}\cdot\nabla\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\big)\,\mathrm{d}(t,\mathbf{x})\\
&\quad + \int_Q \big(\partial_t u_{\mathbf{z}}^{\mathrm{ad}} + \mathbf{q}\cdot\nabla u_{\mathbf{z}}^{\mathrm{ad}}\big)\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]\,\mathrm{d}(t,\mathbf{x}) + \int_\Omega u_{\mathbf{z}}^{\mathrm{ad}}(0)\mathrm{D}B_{\mathbf{z}}[\mathbf{y}]\,\mathrm{d}\mathbf{x}\\
&= \int_\Omega u_{\mathbf{z}}^{\mathrm{ad}}(0)\mathrm{D}B_{\mathbf{z}}[\mathbf{y}]\,\mathrm{d}\mathbf{x}
\end{aligned}$$

if $u_{\mathbf{z}}^{\mathrm{ad}}$ and $\mathrm{D}\mathcal{F}(\mathbf{z})[\mathbf{y}]$ are sufficiently regular. For the gradient decent this is simply approximated by

$$\mathrm{D}_\delta J(\mathbf{z})[\mathbf{y}] = \frac{1}{\delta}\int_\Omega u_{\mathbf{z}}^{\mathrm{ad}}(0)\big(B_{\mathbf{z}+\delta\mathbf{y}} - B_{\mathbf{z}}\big)\,\mathrm{d}\mathbf{x}\,.$$

Author contributions Both authors contributed equally.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded by the CRC 1173 "Wave Phenomena", Deutsche Forschungsgemeinschaft (DFG, Project-ID 258734477). The authors gratefully acknowledge the computing time provided on the high-performance computer HoreKa by the National High-Performance Computing Center at KIT (NHR@KIT). This center is jointly supported by the Federal Ministry of Education and Research and the Ministry of Science, Research and the Arts of Baden-Württemberg, as part of the National High-Performance Computing (NHR) joint funding program (https://www.nhr-verein.de/en/our-partners). HoreKa is partly funded by the German Research Foundation (DFG).

Data availability The code used for the numerical experiments are available at https://gitlab.kit.edu/kit/mpp/mpp/-/tree/c4b6ec68122cf4addd4e09b71c5a5ce01f50468b, for further references and a tutorial we refer to https://www.math.kit.edu/ianm3/page/mplusplus.

Declarations

Competing interests The authors declare no competing interests.

References

Bastian, P., Blatt, M., Dedner, A., Dreier, N.-A., Engwer, C., Fritze, R., Gräser, C., Grüninger, C., Kempf, D., Klöfkorn, R., et al.: The DUNE framework: Basic concepts and recent developments. Computers & Mathematics with Applications **81**, 75–112 (2021). https://doi.org/10.1016/j.camwa.2020.06.007

Baumgarten, N., Wieners, C.: The parallel finite element system M++ with integrated multilevel preconditioning and multilevel Monte Carlo methods. Computers & Mathematics with Applications **81**, 391–406 (2021). https://doi.org/10.1016/j.camwa.2020.03.004

Baumgarten, N., Krumscheid, S., Wieners, C.: A fully parallelized and budgeted multilevel monte carlo method and the application to acoustic waves. SIAM/ASA Journal on Uncertainty Quantification **12**(3), 901–931 (2024). https://doi.org/10.1137/23M1588354

Bear, J., Bachmat, Y.: Introduction to Modeling of Transport Phenomena in Porous Media vol. 4. Springer (2012). https://doi.org/10.1007/978-94-009-1926-6

Becker, R., Rannacher, R.: An optimal control approach to a posteriori error estimation in finite element methods. Acta Numerica **10**, 1–102 (2001). https://doi.org/10.1017/S0962492901000010

Besier, M., Rannacher, R.: Goal-oriented space-time adaptivity in the finite element Galerkin method for the computation of nonstationary incompressible flow. Int. J. Numer. Meth. Fluids **70**(9), 1139–1166 (2012). https://doi.org/10.1002/fld.2735

Corallo, D., Dörfler, W., Wieners, C.: Space-time discontinuous Galerkin methods for weak solutions of hyperbolic linear symmetric Friedrichs systems. J. Sci. Comput. **94**(1), 27 (2023). https://doi.org/10.1007/s10915-022-02076-3

Di Pietro, D.A., Ern, A.: Mathematical Aspects of Discontinuous Galerkin Methods vol. 69. Springer (2011). https://doi.org/10.1007/978-3-642-22980-0

Dörfler, W., Findeisen, S., Wieners, C., Ziegler, D.: Parallel adaptive discontinuous Galerkin discretizations in space and time for linear elastic and acoustic waves. In: Langer, U., Steinbach, O. (eds.) Space-Time Methods. Applications to Partial Differential Equations. Radon Series on Computational and Applied Mathematics, vol. 25, pp. 97–127. Walter de Gruyter (2019). https://doi.org/10.1515/9783110548488-002

Dörfler, W., Findeisen, S., Wieners, C.: Space-time discontinuous Galerkin discretizations for linear first-order hyperbolic evolution systems. Comput. Methods Appl. Math. **16**(3), 409–428 (2016). https://doi.org/10.1515/cmam-2016-0015

Ern, A., Guermond, J.-L.: Finite Elements III: First-Order and Time-Dependent PDEs vol. 74. Springer (2021). https://doi.org/10.1007/978-3-030-57348-5

Ernesti, J., Wieners, C.: Space-time discontinuous Petrov-Galerkin methods for linear wave equations in heterogeneous media. Comput. Methods Appl. Math. **19**(3), 465–481 (2019). https://doi.org/10.1515/cmam-2018-0190

Flemisch, B., Darcis, M., Erbertseder, K., Faigle, B., Lauser, A., Mosthaf, K., Müthing, S., Nuske, P., Tatomir, A., Wolff, M., et al.: Dumux: dune for multi-{phase, component, scale, physics,...} flow and

transport in porous media. Adv. Water Resour. **34**(9), 1102–1112 (2011). https://doi.org/10.1016/j.advwatres.2011.03.007

Hesthaven, J.S., Warburton, T.: Nodal discontinuous Galerkin methods: algorithms, analysis, and applications. Springer (2007). https://doi.org/10.1007/978-0-387-72067-8

Knabner, P., Angermann, L.: Numerical Methods for Elliptic and Parabolic Partial Differential Equations. Texts in Applied Mathematics, vol. 44, p. 802. Springer (2021). https://doi.org/10.1007/978-3-030-79385-2

Meidner, D., Vexler, B.: Adaptive space?time finite element methods for parabolic optimization problems. SIAM J. Control. Optim. **46**(1), 116–142 (2007). https://doi.org/10.1137/060648994

Wieners, C.: A space-time discontinuous Galerkin discretization for the linear transport equation. Computers & Mathematics with Applications **152**, 294–307 (2023). https://doi.org/10.1016/j.camwa.2023.10.031