Numerische Mathematik (2019) 141:21–62
https://doi.org/10.1007/s00211-018-0994-y

Numerische
Mathematik

CrossMark

Numerical analysis of a two-phase flow discrete fracture matrix model

Jérôme Droniou[1] · **Julian Hennicker**[2,3] · **Roland Masson**[2]

Received: 23 January 2017 / Revised: 25 May 2018 / Published online: 5 September 2018

Abstract
We present a new model for two phase Darcy flows in fractured media, in which fractures are modelled as submanifolds of codimension one with respect to the surrounding domain (matrix). Fractures can act as drains or as barriers, since pressure discontinuities at the matrix-fracture interfaces are permitted. Additionally, a layer of damaged rock at the matrix-fracture interfaces is accounted for. The numerical analysis is carried out in the general framework of the Gradient Discretisation Method. Compactness techniques are used to establish convergence results for a wide range of possible numerical schemes; the existence of a solution for the two phase flow model is obtained as a byproduct of the convergence analysis. A series of numerical experiments conclude the paper, with a study of the influence of the damaged layer on the numerical solution.

Mathematics Subject Classification 65M08 · 65M60 · 76M12 · 76M10 · 76S05 · 65M12 · 35K51 · 35K55 · 46N40

1 Introduction

Flow and transport in fractured porous media are of paramount importance for many applications such as petroleum exploration and production, geological storage of

✉ Julian Hennicker
julian.hennicker@unice.fr

Jérôme Droniou
jerome.droniou@monash.edu

Roland Masson
roland.masson@unice.fr

1 School of Mathematical Sciences, Monash University, Victoria 3800, Australia

2 Université Côte d’Azur, CNRS, Inria team COFFEE, LJAD, Nice, France

3 Total SA, Centre scientifique et technique Jean-Féger, Avenue Larribau, 64018 Pau, France

carbon dioxide, hydrogeology, or geothermal energy. Two classes of models, dual continuum and discrete fracture matrix models, are typically employed and possibly coupled to simulate flow and transport in fractured porous media. Dual continuum models assume that the fracture network is well connected and can be homogenised as a continuum coupled to the matrix continuum using transfer functions. On the other hand, discrete fracture matrix models (DFM), on which this paper focuses, represent explicitly the fractures as co-dimension one surfaces immersed in the surrounding matrix domain. The use of lower dimensional rather than equi-dimensional entities to represent the fractures has been introduced in [4,8,31,36,37] to facilitate the grid generation and to reduce the number of degrees of freedom of the discretised model. The reduction of dimension in the fracture network is obtained from the equi-dimensional model by integration and averaging along the width of each fracture. The resulting so called hybrid-dimensional model couple the 3D model in the matrix with a 2D model in the fracture network taking into account the jump of the normal fluxes as well as additional transmission conditions at the matrix-fracture interfaces. These transmission conditions depend on the mathematical nature of the equi-dimensional model and on additional physical assumptions. They are typically derived for a single phase Darcy flow for which they specify either the continuity of the pressure in the case of fractures acting as drains [4,9] or Robin type conditions in order to take into account the discontinuity of the pressure for fractures acting either as drains or barriers [5,11,31,37].

Fewer works deal with the extension of hybrid-dimensional models to two-phase Darcy flows. Most of them build directly the model at the discrete level as in [8,34,40] or are limited to the case of continuous pressures at the matrix-fracture interfaces as in [8,10,40]. In [35], an hybrid-dimensional two-phase flow model with discontinuous pressures at the matrix-fracture interfaces is proposed using a global pressure formulation. However, the transmission conditions at the interface do not take into account correctly the transport from the matrix to the fracture.

In this paper, a new hybrid-dimensional two-phase Darcy flow model is proposed accounting for complex networks of fractures acting either as drains or barriers. The model takes into account discontinuous capillary pressure curves at the matrix-fracture interfaces. It also includes a layer of damaged rock at the matrix-fracture interface with its own mobility and capillary pressure functions. This additional layer is not only a modelling tool. It also plays a major role in the convergence analysis of the model by giving time estimates on the approximate interfacial saturations, which yield their compactness (see Remark 4.7) and enables the identification of their limit. Moreover, when solving the discrete equations with a Newton-Raphson method, a non-zero distribution of volume at the interfacial unknowns is in general required for the Jacobian not to be degenerate. The sensitivity of the discrete solution as well as of the computational performance on interfacial parameters is studied in the test case section. The results suggest that the model converges with vanishing interfacial volume. However, this is still an open question.

The discretisation of hybrid-dimensional Darcy flow models has been the object of many works using cell-centred Finite Volume schemes with either Two Point or Multi Point Flux Approximations (TPFA and MPFA) [2,3,5,33,36,41,43], Mixed or Mixed Hybrid Finite Element methods (MFE and MHFE) [4,34,37], Hybrid Mimetic Mixed Methods (HMM, which contains Mixed/Hybrid Finite Volume and Mimetic Finite Difference schemes [22]) [6,9,11,30], Control Volume Finite Element Methods (CVFE) [8,33,38–40], and the Vertex Approximate Gradient (VAG) scheme [9–11,44,45]. Let us also mention that non-matching discretisations of the fracture and matrix meshes are studied for single phase Darcy flows in [7,14,32,42]. The convergence analysis for single-phase flow models with a single fracture is established in [4,37] for MFE methods, in [14] for non matching MFE discretisations, and in [5] for TPFA discretisations. The case of single-phase flows with complex fracture networks is studied in the general framework of the gradient discretisation method in [9] for continuous pressure models and in [11] for discontinuous pressure models. For hybrid-dimensional two-phase flow models, the only convergence analysis is to our knowledge done in [10] for the VAG discretisation of the continuous pressure model with fractures acting only as drains. Let us recall that the gradient discretisation method (GDM) enables convergence analysis of both conforming and non conforming discretisations for linear and non-linear second order elliptic and parabolic problems. It accounts for various discretisations such as conforming Finite Element methods, MFE and MHFE methods, some TPFA and symmetric MPFA schemes, and the VAG and HHM schemes [24]. The main advantage of this framework is to provide, for a given model, a convergence proof for all schemes satisfying some abstract conditions, at the reduced cost of a single convergence analysis; see e.g. [19,20,23,28,29]. We refer to the monograph [21] for a detailed presentation of the GDM.

The main purpose of this paper is to propose an extension of the gradient discretisation method to our hybrid-dimensional two-phase Darcy flow model. This provides, in an abstract framework, the convergence of the approximate solution to a weak solution of the model; as a by-product, this proves the existence of a solution to this continuous model. The numerical analysis is partially based on the previous work [29] dealing with the gradient discretisation method for single medium two-phase Darcy flows. The main new difficulty addressed in this work compared with the analysis of [29] and [10] comes from the transmission conditions at the matrix-fracture interfaces; these conditions involve an upwinding between the fracture phase pressures and the traces of the matrix phase pressures. Note that, as in [29] and [10], the convergence analysis assumes that the phase mobilities do not vanish.

The outline of this paper is as follows. Section 2 introduces the geometry of the fracture network, the function spaces, the strong and weak formulations of the model as well as the assumptions on the data. Section 3 details the gradient discretisation method, including the definition of the abstract reconstruction operators, of the discrete variational formulation (gradient scheme), and of the coercivity, consistency, limit conformity and compactness properties. Section 4 proves the main result of this paper which is the convergence of the gradient scheme solution to a weak solution of the model. This convergence is established using compactness arguments, and requires us to establish various compactness results on the approximation solutions: averaged in time and space, uniform-in-time and weak-in-space, etc. The Minty monotonicity trick is used to identify the limit of the non-linear term resulting from the the upwinding between the fracture and matrix phase pressures. Section 5 studies on a 2D numerical example the influence of the additional layer of damaged rock at the matrix-fracture interface on the solution of the model. The discretisation used in this test case is based on the VAG scheme which can be shown from [11] to satisfy the assumptions of our

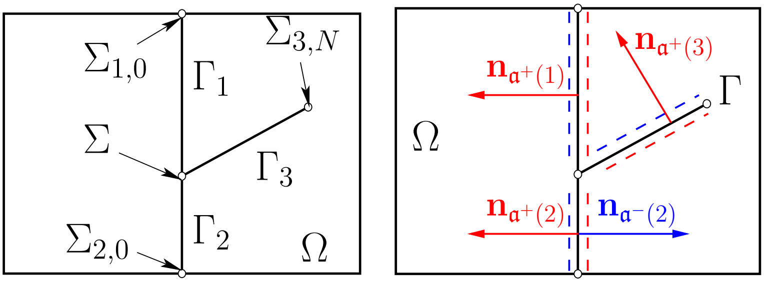

Fig. 1 Example of a 2D domain Ω and 3 intersecting fractures Γ_i, $i = 1, 2, 3$. We define the fracture plane orientations by $\mathfrak{a}^{\pm}(i) \in \chi$ for Γ_i, $i \in I$

gradient discretisation method. Note that numerical comparisons of our model with the equi-dimensional model as well as with the continuous pressure model of [10] can be found in [1,12] without the accumulation term in the interfacial layer, which plays a minor role in the numerical tests when this layer is thin with respect to the fracture (see Sect. 5). It is shown that the discontinuous pressure model analysed in this paper is more accurate than the continuous pressure model of [10] even in the case of fractures acting only as drains; this improved accuracy is due to more accurate transmission conditions at the matrix-fracture interfaces.

2 Notation and model

2.1 Geometry

Let Ω denote a bounded domain of $\mathbb{R}^d$ $(d = 2, 3)$, polyhedral for $d = 3$ and polygonal for $d = 2$. To fix ideas the dimension will be fixed to $d = 3$ when it needs to be specified, for instance in the naming of the geometrical objects or for the space discretisation in the next section. The adaptations to the case $d = 2$ are straightforward.

Let $\overline{\Gamma} = \bigcup_{i\in I} \overline{\Gamma}_i$ and its interior $\Gamma = \overline{\Gamma} \backslash \partial\overline{\Gamma}$ denote the network of fractures $\Gamma_i \subset \Omega$, $i \in I$. Each Γ_i is a planar polygonal simply connected open domain included in a plane $\mathcal{P}_i$ of $\mathbb{R}^d$. It is assumed that the angles of Γ_i are strictly smaller than 2π, and that $\Gamma_i \cap \overline{\Gamma}_j = \emptyset$ for all $i \neq j$. For all $i \in I$, let us set $\Sigma_i = \partial\Gamma_i$, with $\mathbf{n}_{\Sigma_i}$ as unit vector in $\mathcal{P}_i$, normal to Σ_i and outward to Γ_i. Further $\Sigma_{i,j} = \Sigma_i \cap \Sigma_j$ for $i \neq j$, $\Sigma_{i,0} = \Sigma_i \cap \partial\Omega$, $\Sigma_{i,N} = \Sigma_i \backslash (\bigcup_{j\in I\backslash\{i\}} \Sigma_{i,j} \cup \Sigma_{i,0})$, $\Sigma = \bigcup_{(i,j)\in I\times I, i\neq j} (\Sigma_{i,j} \backslash \Sigma_{i,0})$ and $\Sigma_0 = \bigcup_{i\in I} \Sigma_{i,0}$. It is assumed that $\Sigma_{i,0} = \overline{\Gamma}_i \cap \partial\Omega$.

We define the two unit normal vectors $\mathbf{n}_{\mathfrak{a}^{\pm}(i)}$ at each planar fracture Γ_i, such that $\mathbf{n}_{\mathfrak{a}^+(i)} + \mathbf{n}_{\mathfrak{a}^-(i)} = 0$ and oriented outward to the matrix side $\mathfrak{a}^{\pm}(i)$ (cf. Fig. 1). We define the set of indices $\chi = \{\mathfrak{a}^+(i), \mathfrak{a}^-(i) \mid i \in I\}$, such that $\#\chi = 2\#I$. For ease of notation, we use the convention $\Gamma_{\mathfrak{a}^+(i)} = \Gamma_{\mathfrak{a}^-(i)} = \Gamma_i$.

For $\mathfrak{a} = \mathfrak{a}^{\pm}(i) \in \chi$, we denote by $\gamma_{\mathfrak{a}}$ the one-sided trace operator on $\Gamma_{\mathfrak{a}}$. It satisfies the condition $\gamma_{\mathfrak{a}}(h) = \gamma_{\mathfrak{a}}(h \restriction_{\omega_{\mathfrak{a}}})$, where $\omega_{\mathfrak{a}} = \{\mathbf{x} \in \Omega \mid (\mathbf{x} - \mathbf{y}) \cdot \mathbf{n}_{\mathfrak{a}} < 0,\ \forall \mathbf{y} \in \Gamma_i\}$.

On the fracture network Γ, the tangential gradient is denoted by ∇_τ, and is such that

$$\nabla_\tau v = (\nabla_{\tau_i} v_i)_{i\in I},$$

where, for each $i \in I$, the tangential gradient ∇_{τ_i} is defined by fixing a reference Cartesian coordinate system of the plane $\mathcal{P}_i$ containing Γ_i. In the same manner, we denote by $\operatorname{div}_\tau \mathbf{q} = (\operatorname{div}_{\tau_i} \mathbf{q}_i)_{i\in I}$ the tangential divergence operator.

2.2 Continuous model and hypotheses

We describe here the continuous model and assumptions that are implicitly made throughout the paper. In the matrix domain $\Omega\backslash\overline{\Gamma}$, let us denote by $\Lambda_m \in L^\infty(\Omega)^{d\times d}$ the symmetric permeability tensor, chosen such that there exist $\overline{\lambda}_m \geq \underline{\lambda}_m > 0$ with

$$\underline{\lambda}_m|\zeta|^2 \leq \Lambda_m(\mathbf{x})\zeta\cdot\zeta \leq \overline{\lambda}_m|\zeta|^2 \text{ for all } \zeta \in \mathbb{R}^d, \mathbf{x} \in \Omega.$$

Analogously, in the fracture network Γ, we denote by $\Lambda_f \in L^\infty(\Gamma)^{(d-1)\times(d-1)}$ the symmetric tangential permeability tensor, and assume that there exist $\overline{\lambda}_f \geq \underline{\lambda}_f > 0$, such that

$$\underline{\lambda}_f|\zeta|^2 \leq \Lambda_f(\mathbf{x})\zeta\cdot\zeta \leq \overline{\lambda}_f|\zeta|^2 \text{ for all } \zeta \in \mathbb{R}^{d-1}, \mathbf{x} \in \Gamma.$$

On the fracture network Γ, we introduce an orthonormal system $(\boldsymbol{\tau}_1(\mathbf{x}), \boldsymbol{\tau}_2(\mathbf{x}), \mathbf{n}(\mathbf{x}))$, defined a.e. on Γ. Inside the fractures, the normal direction is assumed to be a permeability principal direction. The normal permeability $\lambda_{f,\mathbf{n}} \in L^\infty(\Gamma)$ is such that $\underline{\lambda}_{f,\mathbf{n}} \leq \lambda_{f,\mathbf{n}}(\mathbf{x}) \leq \overline{\lambda}_{f,\mathbf{n}}$ for a.e. $\mathbf{x} \in \Gamma$ with $0 < \underline{\lambda}_{f,\mathbf{n}} \leq \overline{\lambda}_{f,\mathbf{n}}$. We also denote by $d_f \in L^\infty(\Gamma)$ the width of the fractures, assumed to be such that there exist $\overline{d}_f \geq \underline{d}_f > 0$ with $\underline{d}_f \leq d_f(\mathbf{x}) \leq \overline{d}_f$ for a.e. $\mathbf{x} \in \Gamma$. The half normal transmissibility in the fracture network is denoted by

$$T_f = \frac{2\lambda_{f,\mathbf{n}}}{d_f}.$$

Furthermore, ϕ_m and ϕ_f are the matrix and fracture porosities, respectively, $\rho^\alpha \in \mathbb{R}^+$ denotes the density of phase α (with $\alpha = 1$ the non-wetting and $\alpha = 2$ the wetting phase) and $\mathbf{g} \in \mathbb{R}^d$ is the gravitational vector field. We assume that $\underline{\phi}_{m,f} \leq \phi_{m,f} \leq \overline{\phi}_{m,f}$, for some $\underline{\phi}_{m,f}, \overline{\phi}_{m,f} > 0$. (k_m^α, k_f^α) and (S_m^α, S_f^α) are the matrix and fracture phase mobilities and saturations, respectively. Hypothesis on these functions are stated below.

The PDEs model writes: find phase pressures $(\overline{u}_m^\alpha, \overline{u}_f^\alpha)$ and velocities $(\mathbf{q}_m^\alpha, \mathbf{q}_f^\alpha)$ $(\alpha = 1, 2)$, such that

$$\left\{\begin{array}{rll}
\phi_m \partial_t S_m^\alpha(\overline{p}_m) + \operatorname{div}(\mathbf{q}_m^\alpha) &= h_m^\alpha & \text{on } (0,T)\times\Omega\backslash\overline{\Gamma}\\
\mathbf{q}_m^\alpha &= -[kS]_m^\alpha(\overline{p}_m)\,\Lambda_m\nabla\overline{u}_m^\alpha & \text{on } (0,T)\times\Omega\backslash\overline{\Gamma}\\
\phi_f d_f \partial_t S_f^\alpha(\overline{p}_f) + \operatorname{div}_\tau(\mathbf{q}_f^\alpha) - \sum_{\mathfrak{a}\in\chi} Q_{f,\mathfrak{a}}^\alpha &= d_f h_f^\alpha & \text{on } (0,T)\times\Gamma\\
\mathbf{q}_f^\alpha &= -d_f[kS]_f^\alpha(\overline{p}_f)\,\Lambda_f\nabla_\tau\overline{u}_f & \text{on } (0,T)\times\Gamma\\
(\overline{p}_m, \overline{p}_f)|_{t=0} &= (\overline{p}_{m,0}, \overline{p}_{f,0}) & \text{on } (\Omega\setminus\overline{\Gamma})\times\Gamma.
\end{array}\right. \quad (1a)$$

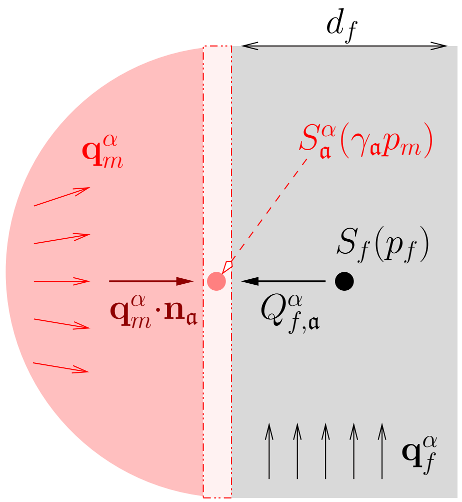

Fig. 2 Illustration of the coupling condition. It can be seen as an upwind two point approximation of $Q^\alpha_{f,\mathfrak{a}}$. The upwinding takes into account the damaged rock type at the matrix-fracture interfaces. The arrows show the positive orientation of the normal fluxes $\mathbf{q}^\alpha_m \cdot \mathbf{n}_\mathfrak{a}$ and $Q^\alpha_{f,\mathfrak{a}}$

The matrix-fracture coupling condition on $(0, T) \times \Gamma_\mathfrak{a}$ (for all $\mathfrak{a} \in \chi$) are

$$\begin{cases} \mathbf{q}^\alpha_m \cdot \mathbf{n}_\mathfrak{a} + Q^\alpha_{f,\mathfrak{a}} = \eta \partial_t S^\alpha_\mathfrak{a}(\gamma_\mathfrak{a} \overline{p}_m) \\ \qquad\qquad\;\; Q^\alpha_{f,\mathfrak{a}} = [kS]^\alpha_f(\overline{p}_f) T_f [\![\overline{u}^\alpha]\!]^-_\mathfrak{a} - [kS]^\alpha_\mathfrak{a}(\gamma_\mathfrak{a} \overline{p}_m) T_f [\![\overline{u}^\alpha]\!]^+_\mathfrak{a}, \end{cases} \tag{1b}$$

where $\eta = d_\mathfrak{a} \phi_\mathfrak{a}$, with $d_\mathfrak{a} \in (0, \frac{d_f}{2})$ representing the interfacial width and $\phi_\mathfrak{a} \in (0, 1]$ the interfacial porosity (cf. Fig. 2). We assume that each of these parameters is uniformly bounded below. In these equations, we have

$$S^2_\mu = 1 - S^1_\mu \text{ for } \mu \in \{m, f\} \cup \chi, \text{ and } (\overline{p}_m, \overline{p}_f) = (\overline{u}^1_m - \overline{u}^2_m, \overline{u}^1_f - \overline{u}^2_f). \tag{1c}$$

In the above, we used the shorthand notations

$$[\![\overline{u}^\alpha]\!]_\mathfrak{a} = \gamma_\mathfrak{a} \overline{u}^\alpha_m - \overline{u}^\alpha_f\,, \quad [\![\overline{u}^\alpha]\!]^+_\mathfrak{a} = \max(0, [\![\overline{u}^\alpha]\!]_\mathfrak{a}) \quad \text{and} \quad [\![\overline{u}^\alpha]\!]^-_\mathfrak{a} = [\![-\overline{u}^\alpha]\!]^+_\mathfrak{a}$$

as well as, for $\mu \in \{m, f\} \cup \chi$, $\varphi_\mu \in L^2((0, T) \times M_\mu)$ and a.e. $(t, \mathbf{x}) \in (0, T) \times M_\mu$,

$$S^\alpha_\mu(\varphi_\mu)(t, \mathbf{x}) = S^\alpha_\mu(\mathbf{x}, \varphi_\mu(t, \mathbf{x})) \quad \text{and} \quad [kS]^\alpha_\mu(\varphi_\mu)(t, \mathbf{x}) = k^\alpha_\mu(\mathbf{x}, S^\alpha_\mu(\mathbf{x}, \varphi_\mu(t, \mathbf{x}))).$$

Here and in the following, M_μ is defined by

$$M_\mu = \begin{cases} \Omega & \text{if } \mu = m \\ \Gamma & \text{if } \mu = f \\ \Gamma_\mathfrak{a} & \text{if } \mu = \mathfrak{a} \in \chi\,. \end{cases}$$

The various boundary conditions imposed on the domain are: homogeneous Dirichlet conditions at the boundary of the domain, pressure continuity and flux conservation at the fracture-fracture intersections, and zero normal flux at the immersed fracture tips. In other words,

$$\gamma_{\partial\Omega\setminus\partial\Gamma}\bar{u}_m = 0 \text{ on } \partial\Omega \setminus \partial\Gamma, \qquad \gamma_{\partial\Omega\cap\partial\Gamma}\bar{u}_f = 0 \text{ on} \partial\Omega \cap \partial\Gamma$$
$$\gamma_{\Sigma_i}\bar{u}_{f,i} = \gamma_{\Sigma_j}\bar{u}_{f,j} \text{ on } \Sigma_{i,j} \text{ for all } i \neq j \text{ such that } \Sigma_{i,j}$$
$$\text{has a non zero} d-2 \text{ Lebesgue measure}$$
$$\sum_{i\in I}\mathbf{q}_{f,i}\cdot\mathbf{n}_{\Sigma_i} = 0 \text{ on } \Sigma, \qquad \mathbf{q}_{f,i}\cdot\mathbf{n}_{\Sigma_i} = 0 \text{ on } \Sigma_{i,N},\ i \in I$$

Let us define $L^2(\Gamma) = \{v = (v_i)_{i\in I}, v_i \in L^2(\Gamma_i), i \in I\}$. The assumptions under which the model is considered are:

- $\overline{p}_{m,0} \in H^1(\Omega\backslash\overline{\Gamma})$ and $\overline{p}_{f,0} \in L^2(\Gamma)$,
- For $\mu \in \{m, f\}$ and $\alpha = 1, 2$, $h^\alpha_\mu \in L^2((0, T) \times M_\mu)$,
- For $\mu \in \{m, f\} \cup \chi$: $S^1_\mu : M_\mu \times \mathbb{R} \to [0, 1]$ is a Caratheodory function; for a.e. $\mathbf{x} \in M_\mu$, $S^1_\mu(\mathbf{x}, \cdot)$ is a non-decreasing Lipschitz continuous function on $\mathbb{R}$; for all $q \in \mathbb{R}$, $S^1_\mu(\cdot, q)$ is piecewise constant on a finite partition $(M^j_\mu)_{j\in J_\mu}$ of polytopal subsets of M_μ.
- For $\alpha = 1, 2$ and $\mu \in \{m, f\} \cup \chi$: there exist constants $\underline{k}_\mu, \overline{k}_\mu > 0$, such that $k^\alpha_\mu : M_\mu \times [0, 1] \to [\underline{k}_\mu, \overline{k}_\mu]$ is a Caratheodory function.

Recall that a Caratheodory function is measurable w.r.t. its first argument and continuous w.r.t. its second argument.

2.3 Weak formulation

The subspace $H^1(\Gamma)$ of $L^2(\Gamma)$ consists in functions $v = (v_i)_{i\in I}$ such that $v_i \in H^1(\Gamma_i)$ for all $i \in I$, with continuous traces at the fracture intersections $\Sigma_{i,j}$ for all $i \neq j$. Its subspace of functions with vanishing traces on Σ_0 is denoted by $H^1_{\Sigma_0}(\Gamma)$.

Let us now define the hybrid-dimensional function spaces that are used as variational spaces for the Darcy flow model. Starting from

$$V = H^1(\Omega\backslash\overline{\Gamma}) \times H^1(\Gamma),$$

consider the subspace

$$V^0 = V^0_m \times V^0_f$$

where (with $\gamma_{\partial\Omega}: H^1(\Omega\backslash\overline{\Gamma}) \to L^2(\partial\Omega)$ the trace operator on $\partial\Omega$)

$$V^0_m = \{v \in H^1(\Omega\backslash\overline{\Gamma}) \mid \gamma_{\partial\Omega}v = 0 \text{ on } \partial\Omega\} \quad \text{and} \quad V^0_f = H^1_{\Sigma_0}(\Gamma).$$

The weak formulation of (1) amounts to finding $(\overline{u}_m^\alpha, \overline{u}_f^\alpha)_{\alpha=1,2} \in [L^2(0,T;V_m^0) \times L^2(0,T;V_f^0)]^2$ satisfying the following variational equalities, for any $\alpha = 1, 2$ and any $(\overline{\varphi}_m^\alpha, \overline{\varphi}_f^\alpha) \in C_0^\infty([0,T)\times\Omega) \times C_0^\infty([0,T)\times\Gamma)$:

$$
\begin{aligned}
&\sum_{\mu\in\{m,f\}} \Big(-\int_0^T \int_{M_\mu} \phi_\mu S_\mu^\alpha(\overline{p}_\mu)\partial_t \overline{\varphi}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t + \int_0^T \int_{M_\mu} [kS]_\mu^\alpha(\overline{p}_\mu)\ \Lambda_\mu \nabla \overline{u}_\mu^\alpha \cdot \nabla \overline{\varphi}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \\
&\quad - \int_{M_\mu} \phi_\mu S_\mu^\alpha(\overline{p}_{\mu,0})\overline{\varphi}_\mu^\alpha(0,\cdot)\mathrm{d}\tau_\mu \Big) \\
&+ \sum_{\mathfrak{a}\in\chi} \Big(\int_0^T \int_{\Gamma_\mathfrak{a}} T_f \Big([kS]_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_m)[\![\overline{u}^\alpha]\!]_\mathfrak{a}^+ - [kS]_f^\alpha(\overline{p}_f)[\![\overline{u}^\alpha]\!]_\mathfrak{a}^- \Big) [\![\overline{\varphi}^\alpha]\!]_\mathfrak{a} \mathrm{d}\tau \mathrm{d}t \\
&\quad - \int_0^T \int_{\Gamma_\mathfrak{a}} \eta S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_m)\partial_t \gamma_\mathfrak{a}\overline{\varphi}_m^\alpha \mathrm{d}\tau \mathrm{d}t - \int_{\Gamma_\mathfrak{a}} \eta S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_{m,0})\gamma_\mathfrak{a}\overline{\varphi}_m^\alpha(0,\cdot)\mathrm{d}\tau \Big) \\
&= \sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} h_\mu^\alpha \overline{\varphi}_\mu^\alpha \mathrm{d}\tau_\mu .
\end{aligned}
\tag{2}
$$

Here,

$$
\mathrm{d}\tau_\mu(\mathbf{x}) = \begin{cases} \mathrm{d}\mathbf{x} & \text{if } \mu = m \\ \mathrm{d}\tau_f(\mathbf{x}) = d_f(\mathbf{x})\mathrm{d}\tau(\mathbf{x}) & \text{if } \mu = f \end{cases}
$$

with $\mathrm{d}\tau(x)$ the $d-1$ dimensional Lebesgue measure on Γ.

3 The gradient discretisation method

The gradient discretisation method consists in selecting a set (called a gradient discretisation) of a finite-dimensional space and reconstruction operators on this space, and in substituting them for their continuous counterpart in the weak formulation of the model. The scheme thus obtained is called a gradient scheme. Let us first define the set of discrete elements that make up a gradient discretisation.

Definition 3.1 (*Gradient discretisation (GD)*) A spatial gradient discretisation for a discrete fracture matrix model is $\mathcal{D}_S = (X^0, (\Pi_{\mathcal{D}_S}^\mu, \nabla_{\mathcal{D}_S}^\mu)_{\mu\in\{m,f\}}, ([\![\cdot]\!]_{\mathfrak{a},\mathcal{D}_S})_{\mathfrak{a}\in\chi}, (\mathbb{T}_{\mathcal{D}_S}^\mathfrak{a})_{\mathfrak{a}\in\chi})$, where

- X^0 is a finite-dimensional space of degrees of freedom (DOFs),
- For $\mu \in \{m, f\}$, $\Pi_{\mathcal{D}_S}^\mu : X^0 \to L^2(M_\mu)$ reconstructs a function on M_μ from the DOFs,
- For $\mu \in \{m, f\}$, $\nabla_{\mathcal{D}_S}^\mu : X^0 \to L^2(M_\mu)^{\dim M_\mu}$ reconstructs a gradient on M_μ from the DOFs,
- For $\mathfrak{a} \in \chi$, $[\![\cdot]\!]_{\mathfrak{a},\mathcal{D}_S} : X^0 \to L^2(\Gamma_\mathfrak{a})$ reconstructs, from the DOFs, a jump on $\Gamma_\mathfrak{a}$ between the matrix and fracture,

- For $\mathfrak{a} \in \chi$, $\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S} : X^0 \to L^2(\Gamma_{\mathfrak{a}})$ reconstructs, from the DOFs, a trace on $\Gamma_{\mathfrak{a}}$ from the matrix.

These operators must be chosen such that the following expression defines a norm on X^0:

$$\|w\|_{\mathcal{D}_S} = \Big(\|\nabla^m_{\mathcal{D}_S} w\|^2_{L^2(\Omega)^d} + \|\nabla^f_{\mathcal{D}_S} w\|^2_{L^2(\Gamma)^{d-1}} + \sum_{\mathfrak{a}\in\chi} \|[\![w]\!]_{\mathfrak{a},\mathcal{D}_S}\|^2_{L^2(\Gamma_{\mathfrak{a}})}\Big)^{1/2}.$$

The spatial gradient discretisation $\mathcal{D}_S$ is extended to a space-time gradient discretisation by setting $\mathcal{D} = (\mathcal{D}_S, \mathrm{I}_{\mathcal{D}}, (t_n)_{n=0,\ldots,N})$ with

- $0 = t_0 < t_1 < \cdots < t_N = T$ a discretisation of the time interval $[0, T]$,
- $\mathrm{I}_{\mathcal{D}} \colon H^1(\Omega\backslash\overline{\Gamma}) \times L^2(\Gamma) \to X^0$ an operator designed to interpolate the initial condition.

The space-time operators act on a family $u = (u_n)_{n=0,\ldots,N} \in (X^0)^{N+1}$ the following way: for all $n = 0, \ldots, N-1$ and all $t \in (t_n, t_{n+1}]$,

$$\begin{aligned} &\Pi^\mu_{\mathcal{D}} u(t,\cdot) = \Pi^\mu_{\mathcal{D}_S} u_{n+1}, \qquad &&\nabla^\mu_{\mathcal{D}} u(t,\cdot) = \nabla^\mu_{\mathcal{D}_S} u_{n+1}, \\ &\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} u(t,\cdot) = \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S} u_{n+1}, \qquad &&[\![u]\!]_{\mathfrak{a},\mathcal{D}}(t,\cdot) = [\![u_{n+1}]\!]_{\mathfrak{a},\mathcal{D}_S}. \end{aligned} \tag{3}$$

We extend these functions at $t = 0$ by considering the corresponding spatial operators on u_0.

If $w = (w_n)_{n=0,\ldots,N}$ is a family in X^0, the discrete time derivatives $\delta_t w : (0, T] \to X^0$ are defined such that, for all $n = 0, \ldots, N-1$ and all $t \in (t_n, t_{n+1}]$, with $\Delta t_{n+\frac{1}{2}} = t_{n+1} - t_n$,

$$\delta_t w(t) = \frac{w_{n+1} - w_n}{\Delta t_{n+\frac{1}{2}}} \in X^0.$$

Let $(e_\nu)_{\nu\in\mathrm{DOF}_{\mathcal{D}}}$ be a basis of X^0. If $w \in X^0$, we write $w = \sum_{\nu\in\mathrm{DOF}_{\mathcal{D}}} w_\nu e_\nu$. Then, for $g \in C(\mathbb{R})$, we define $g(w) \in X^0$ by $g(w) = \sum_{\nu\in\mathrm{DOF}_{\mathcal{D}}} g(w_\nu) e_\nu$. In other words, $g(w)$ is defined by applying g to each degree of freedom of w. Although this definition depends on the choice of basis $(e_\nu)_{\nu\in\mathrm{DOF}_{\mathcal{D}}}$, we do not explicitly indicate this dependency. This definition of $g(w)$ is particularly meaningful in the context of piecewise constant reconstructions, see Remark 3.3 below.

The gradient scheme for (1) consists in writing the weak formulation (2) with continuous spaces and operators replaced by their discrete counterparts, after a formal integration-by-parts in time. In other words, the gradient scheme is: find $(u^\alpha)_{\alpha=1,2} \in [(X^0)^{N+1}]^2$ such that, with $p = u^1 - u^2$,

$$p_0 = \mathrm{I}_{\mathcal{D}}(\overline{p}_{m,0}, \overline{p}_{f,0}) \tag{4}$$

and, for any $\alpha = 1, 2$ and $v^\alpha \in (X^0)^{N+1}$,

$$\begin{aligned}
&\sum_{\mu\in\{m,f\}} \Big(\int_0^T \int_{M_\mu} \phi_\mu \Pi_{\mathcal{D}}^\mu \Big[\delta_t S_\mu^\alpha(p) \Big] \Pi_{\mathcal{D}}^\mu v^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \\
&\quad + \int_0^T \int_{M_\mu} [kS]_\mu^\alpha (\Pi_{\mathcal{D}}^\mu p)\, \Lambda_\mu \nabla_{\mathcal{D}}^\mu u^\alpha \cdot \nabla_{\mathcal{D}}^\mu v^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \Big) \\
&\quad + \sum_{\mathfrak{a}\in\chi} \Big(\int_0^T \int_{\Gamma_\mathfrak{a}} \Big([kS]_\mathfrak{a}^\alpha (\mathbb{T}_{\mathcal{D}}^\mathfrak{a} p) T_f [\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^+ - [kS]_f^\alpha (\Pi_{\mathcal{D}}^f p) T_f [\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^- \Big) [\![v^\alpha]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau \mathrm{d}t \\
&\quad + \int_0^T \int_{\Gamma_\mathfrak{a}} \eta \mathbb{T}_{\mathcal{D}}^\mathfrak{a} \Big[\delta_t S_\mathfrak{a}^\alpha(p) \Big] \mathbb{T}_{\mathcal{D}}^\mathfrak{a} v^\alpha \mathrm{d}\tau \mathrm{d}t \Big) = \sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} h_\mu^\alpha \Pi_{\mathcal{D}}^\mu v^\alpha \mathrm{d}\tau_\mu \mathrm{d}t.
\end{aligned} \tag{5}$$

3.1 Properties of gradient discretisations

The convergence analysis of the GDM is based on a few properties that sequences of GDs must satisfy.

Definition 3.2 (*Piecewise constant reconstruction operator*) Let $(e_\nu)_{\nu\in\mathrm{DOF}_\mathcal{D}}$ be the basis of X^0 chosen in Sect. 3. For $\mu \in \{m, f\} \cup \chi$, an operator $\Pi : X^0 \to L^2(M_\mu)$ is called piecewise constant if it has the representation

$$\Pi u = \sum_{\nu\in\mathrm{DOF}_\mathcal{D}} u_\nu \mathbb{1}_{\omega_\nu^\mu} \qquad \text{for all } u = \sum_{\nu\in\mathrm{DOF}_\mathcal{D}} u_\nu e_\nu \in X^0,$$

where $(\omega_\nu^\mu)_{\nu\in\mathrm{DOF}_\mathcal{D}}$ is a partition of M_μ up to a set of zero measure, and $\mathbb{1}_{\omega_\nu^\mu}$ is the characteristic function of ω_ν^μ.

In the following, all considered function reconstruction operators $\Pi_{\mathcal{D}}^\mu$ and $\mathbb{T}_{\mathcal{D}}^\mathfrak{a}$ are assumed to be piecewise constant.

Remark 3.3 Recall that, if $g \in C^0(\mathbb{R})$ and $u \in X^0$, then $g(u) \in X^0$ is defined by the degrees of freedom $(g(u_\nu))_{\nu\in\mathrm{DOF}_\mathcal{D}}$. Then, any piecewise constant reconstruction operator Π commutes with g in the sense that $g(\Pi u) = \Pi g(u)$.

The coercivity property enables us to control the functions and trace reconstruction by the norm on X^0. This is a combination of a discrete Poincaré inequality and a discrete trace inequality.

Definition 3.4 (*Coercivity of spatial GD*) Let

$$C_{\mathcal{D}_S} = \max_{0\neq v\in X^0} \frac{\|\Pi_{\mathcal{D}_S}^m v\|_{L^2(\Omega)} + \|\Pi_{\mathcal{D}_S}^f v\|_{L^2(\Gamma)} + \sum_{\mathfrak{a}\in\chi} \|\mathbb{T}_{\mathcal{D}_S}^\mathfrak{a} v\|_{L^2(\Gamma_\mathfrak{a})}}{\|v\|_{\mathcal{D}_S}}.$$

A sequence $(\mathcal{D}_S^l)_{l\in\mathbb{N}}$ of gradient discretisations is coercive if there exists $\mathcal{C}_P > 0$ such that

$$\mathcal{C}_{\mathcal{D}_S^l} \le \mathcal{C}_P \text{ for all } l \in \mathbb{N}. \tag{6}$$

This consists in choosing a family

The consistency ensures that a certain interpolation error goes to zero along sequences of GDs.

Definition 3.5 (*Consistency of spatial GD*) For $\overline{u} = (\overline{u}_m, \overline{u}_f) \in V^0$ and $v \in X^0$, define

$$\begin{aligned} s_{\mathcal{D}_S}(v, \overline{u}) = {} & \|\nabla^m_{\mathcal{D}_S} v - \nabla \overline{u}_m\|_{L^2(\Omega)^d} + \|\nabla^f_{\mathcal{D}_S} v - \nabla_\tau \overline{u}_f\|_{L^2(\Gamma)^{d-1}} \\ & + \|\Pi^m_{\mathcal{D}_S} v - \overline{u}_m\|_{L^2(\Omega)} + \|\Pi^f_{\mathcal{D}_S} v - \overline{u}_f\|_{L^2(\Gamma)} \\ & + \sum_{\mathfrak{a}\in\chi} \Big(\|[\![v]\!]_{\mathfrak{a},\mathcal{D}_S} - [\![\overline{u}]\!]_{\mathfrak{a}}\|_{L^2(\Gamma_a)} + \|\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S} v - \gamma_{\mathfrak{a}} \overline{u}_m\|_{L^2(\Gamma_a)} \Big), \end{aligned}$$

and $\mathcal{S}_{\mathcal{D}_S}(\overline{u}) = \min_{v\in X^0} s_{\mathcal{D}_S}(v, \overline{u})$. A sequence $(\mathcal{D}_S^l)_{l\in\mathbb{N}}$ of gradient discretisations is GD-consistent (or consistent for short) if, for all $\overline{u} = (\overline{u}_m, \overline{u}_f) \in V^0$,

$$\lim_{l\to\infty} \mathcal{S}_{\mathcal{D}_S^l}(\overline{u}) = 0. \tag{7}$$

To define the notion of limit-conformity, we need the following two spaces:

$$\begin{aligned} \mathbf{C}^\infty_\Omega &= C^\infty_b(\Omega \setminus \overline{\Gamma})^d, \\ \mathbf{C}^\infty_\Gamma &= \Big\{ \mathbf{q}_f = (\mathbf{q}_{f,i})_{i\in I} \mid \mathbf{q}_{f,i} \in C^\infty(\overline{\Gamma}_i)^{d-1},\ \textstyle\sum_{i\in I} \mathbf{q}_{f,i} \cdot \mathbf{n}_{\Sigma_i} = 0 \text{ on} \Sigma, \\ & \qquad \mathbf{q}_{f,i} \cdot \mathbf{n}_{\Sigma_i} = 0 \text{ on} \Sigma_{i,N},\ i \in I \Big\}, \end{aligned}$$

where $C^\infty_b(\Omega \setminus \overline{\Gamma}) \subset C^\infty(\Omega \setminus \overline{\Gamma})$ is the set of functions φ, such that for all $\mathbf{x} \in \Omega$ there exists $r > 0$, such that for all connected components ω of $\{\mathbf{x}+\mathbf{y} \in \mathbb{R}^d \mid |\mathbf{y}| < r\} \cap (\Omega \setminus \overline{\Gamma})$ one has $\varphi \in C^\infty(\overline{\omega})$, and such that all derivatives of φ are bounded. The limit-conformity imposes that, in the limit, the discrete gradient and function reconstructions satisfy a natural integration-by-part formula (Stokes' theorem).

Definition 3.6 (*Limit-conformity of spatial GD*) For all $\mathbf{q} = (\mathbf{q}_m, \mathbf{q}_f) \in \mathbf{C}^\infty_\Omega \times \mathbf{C}^\infty_\Gamma$, $\varphi_{\mathfrak{a}} \in C^\infty_0(\Gamma_{\mathfrak{a}})$ and $v \in X^0$, define

$$\begin{aligned} w_{\mathcal{D}_S}(v, \mathbf{q}, \varphi_{\mathfrak{a}}) = {} & \int_\Omega \Big(\nabla^m_{\mathcal{D}_S} v \cdot \mathbf{q}_m + (\Pi^m_{\mathcal{D}_S} v) \mathrm{div} \mathbf{q}_m \Big) \mathrm{d}\mathbf{x} \\ & + \int_\Gamma \Big(\nabla^f_{\mathcal{D}_S} v \cdot \mathbf{q}_f + (\Pi^f_{\mathcal{D}_S} v) \mathrm{div}_\tau \mathbf{q}_f \Big) \mathrm{d}\tau(\mathbf{x}) \\ & - \sum_{\mathfrak{a}\in\chi} \int_{\Gamma_{\mathfrak{a}}} \mathbf{q}_m \cdot \mathbf{n}_{\mathfrak{a}} \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S} v \, \mathrm{d}\tau(\mathbf{x}) \end{aligned}$$

$$+\sum_{\mathfrak{a}\in\chi}\int_{\Gamma_{\mathfrak{a}}}\varphi_{\mathfrak{a}}\Big(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S}v-\Pi^f_{\mathcal{D}_S}v-[\![v]\!]_{\mathfrak{a},\mathcal{D}_S}\Big)\mathrm{d}\tau(\mathbf{x})$$

and $\mathcal{W}_{\mathcal{D}_S}(\mathbf{q},\varphi_{\mathfrak{a}})=\max_{0\neq v\in X^0}\frac{1}{\|v\|_{\mathcal{D}_S}}|w_{\mathcal{D}_S}(v,\mathbf{q},\varphi_{\mathfrak{a}})|$. A sequence $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ of gradient discretisations is limit-conforming if, for all $\mathbf{q}=(\mathbf{q}_m,\mathbf{q}_f)\in\mathbf{C}^\infty_\Omega\times\mathbf{C}^\infty_\Gamma$ and all $\varphi_{\mathfrak{a}}\in C^\infty_0(\Gamma_{\mathfrak{a}})$,

$$\lim_{l\to\infty}\mathcal{W}_{\mathcal{D}^l_S}(\mathbf{q},\varphi_{\mathfrak{a}})=0. \tag{8}$$

Remark 3.7 (Domain of $\mathcal{W}_{\mathcal{D}_S}$) Usually, the measure $\mathcal{W}_{\mathcal{D}_S}$ of limit-conformity is defined on spaces in which the Darcy velocities of solutions to the model are expected to be, not smooth spaces as $\mathbf{C}^\infty_\Omega\times\mathbf{C}^\infty_\Gamma$ [21, Definition 2.6]. However, if we do not aim at obtaining error estimates (which is the case here, given that such estimates would require unrealistic regularity assumptions on the data and the solution), $\mathcal{W}_{\mathcal{D}_S}$ only needs to be defined and to converge to 0 on spaces of smooth functions—see Lemma A.2.

For any space-dependent function f, define $\mathrm{T}_{\boldsymbol{\xi}}f(\mathbf{x})=f(\mathbf{x}+\boldsymbol{\xi})$. Likewise, for any time-dependent function g, let $\mathrm{T}_h g(t)=g(t+h)$. The compactness property ensures a sort of discrete Rellich theorem (compact embedding of H^1_0 into L^2). By the Kolmogorov theorem, this compactness is equivalent to a uniform control of the translates of the functions.

Definition 3.8 (*Compactness of spatial GD*) For all $v\in X^0$ and $\boldsymbol{\xi}=(\boldsymbol{\xi}_m,\boldsymbol{\xi}_f)$, with $\boldsymbol{\xi}_m\in\mathbb{R}^d$ and $\boldsymbol{\xi}_f=(\boldsymbol{\xi}^i_f)_{i\in I}\in\bigoplus_{i\in I}\tau(\mathcal{P}_i)$, where $\tau(\mathcal{P}_i)$ is the (constant) tangent space of $\mathcal{P}_i$, define

$$\begin{aligned}\tau_{\mathcal{D}_S}(v,\boldsymbol{\xi})&=\|\mathrm{T}_{\boldsymbol{\xi}_m}\Pi^m_{\mathcal{D}_S}v-\Pi^m_{\mathcal{D}_S}v\|_{L^2(\mathbb{R}^d)}\\&\quad+\sum_{i\in I}\Big(\|\mathrm{T}_{\boldsymbol{\xi}^i_f}\Pi^f_{\mathcal{D}_S}v-\Pi^f_{\mathcal{D}_S}v\|_{L^2(\mathcal{P}_i)}+\sum_{\mathfrak{a}=\mathfrak{a}^\pm(i)}\|\mathrm{T}_{\boldsymbol{\xi}^i_f}\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S}v-\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S}v\|_{L^2(\mathcal{P}_i)}\Big),\end{aligned}$$

where all the functions on Ω (resp. Γ_i) have been extended to $\mathbb{R}^d$ (resp. $\mathcal{P}_i$) by 0 outside their initial domain. Let $\mathcal{T}_{\mathcal{D}_S}(\boldsymbol{\xi})=\max_{0\neq v\in X^0}\frac{1}{\|v\|_{\mathcal{D}_S}}\tau_{\mathcal{D}_S}(v,\boldsymbol{\xi})$. A sequence $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ of gradient discretisations is compact if

$$\lim_{|\boldsymbol{\xi}|\to 0}\sup_{l\in\mathbb{N}}\mathcal{T}_{\mathcal{D}^l_S}(\boldsymbol{\xi})=0. \tag{9}$$

All these properties for spatial GDs naturally extend to space–time GDs with, for the consistency, additional requirements on the time steps and on the interpolants of the initial conditions.

Definition 3.9 (*Properties of space-time gradient discretisations*) A sequence of space-time gradient discretisations $(\mathcal{D}^l)_{l\in\mathbb{N}}$ is

1. Coercive if $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ is coercive.

2. Consistent if

   (i) $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ is consistent,
   (ii) $\Delta t^l = \max_{n=0,\ldots,N-1} \Delta t^l_{n+\frac{1}{2}} \to 0$ as $l \to \infty$, and
   (iii) For all $\overline{\varphi} = (\overline{\varphi}_m, \overline{\varphi}_f) \in H^1(\Omega\backslash\overline{\Gamma}) \times L^2(\Gamma)$, letting $\varphi^l = \mathrm{I}_{\mathcal{D}^l}(\overline{\varphi}_m, \overline{\varphi}_f)$ we have, as $l \to \infty$,

$$\|\overline{\varphi}_m - \Pi^m_{\mathcal{D}^l_S}\varphi^l\|_{L^2(\Omega)} \to 0,$$
$$\|\gamma_{\mathfrak{a}}\overline{\varphi}_m - \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l_S}\varphi^l\|_{L^2(\Gamma_{\mathfrak{a}})} \to 0 \quad \forall \mathfrak{a} \in \chi,$$
$$\|\overline{\varphi}_f - \Pi^f_{\mathcal{D}^l_S}\varphi^l\|_{L^2(\Gamma)} \to 0.$$

3. Limit-conforming if $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ is limit-conforming.
4. Compact if $(\mathcal{D}^l_S)_{l\in\mathbb{N}}$ is compact.

Elements of $(X^0)^{N+1}$ are identified with functions $(0, T] \to X^0$ by setting, for $u \in (X^0)^{N+1}$ with $u = (u_n)_{n=0,\ldots,N}$,

$$\forall n = 0, \ldots, N-1\,,\ \forall t \in (t_n, t_{n+1}]\,,\ u(t) = u_{n+1}. \tag{10}$$

This definition is compatible with the choices of space-time operators made in Definition 3.1, in the sense that, for any $t \in (0, T]$, $\Pi^\mu_{\mathcal{D}} u(t, \mathbf{x}) = \Pi^\mu_{\mathcal{D}_S}(u(t))(\mathbf{x})$ (and similarly for the other reconstruction operators). With the identification (10), the norm on $(X^0)^{N+1}$ is

$$\|u\|^2_{\mathcal{D}} = \int_0^T \|u(t)\|^2_{\mathcal{D}_S} \mathrm{d}t.$$

4 Convergence analysis

In the rest of this paper, when the phase parameter α is absent this implicitly means that it is equal to 1. For example, we write S_μ for S^1_μ. The main convergence result is the following.

Theorem 4.1 (Convergence Theorem) *Let $(\mathcal{D}^l)_{l\in\mathbb{N}}$ be a coercive, consistent, limit-conforming and compact sequence of space-time gradient discretisations, with piecewise constant reconstructions. Then for any $l \in \mathbb{N}$ there is a solution $(u^{\alpha,l})_{\alpha=1,2}$ of (5) with $\mathcal{D} = \mathcal{D}^l$.*

Moreover, there exists $(\overline{u}^\alpha)_{\alpha=1,2} = (\overline{u}^\alpha_m, \overline{u}^\alpha_f)_{\alpha=1,2} \in [L^2(0, T; V^0_m) \times L^2(0, T; V^0_f)]^2$ solution of (2) such that, up to a subsequence as $l \to \infty$,

1. The following weak convergences hold, for $\alpha = 1, 2$,

$$
\begin{cases}
\Pi^{\mu}_{\mathcal{D}^l} u^{\alpha,l} \rightharpoonup \overline{u}^{\alpha}_{\mu} & \textit{weakly in } L^2((0,T)\times M_\mu)\,,\ \textit{for } \mu\in\{m,f\},\\
\nabla^{\mu}_{\mathcal{D}^l} u^{\alpha,l} \rightharpoonup \nabla\overline{u}^{\alpha}_{\mu} & \textit{weakly in } L^2((0,T)\times M_\mu)^{\dim_{M_\mu}}\,,\ \textit{for } \mu\in\{m,f\},\\
\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l} u^{\alpha,l} \rightharpoonup \gamma_{\mathfrak{a}}\overline{u}^{\alpha}_{m} & \textit{weakly in } L^2((0,T)\times \Gamma_{\mathfrak{a}})\,,\ \textit{for all} \mathfrak{a}\in\chi,\\
[\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l} \rightharpoonup [\![\overline{u}^{\alpha}]\!]_{\mathfrak{a}} & \textit{weakly in } L^2((0,T)\times \Gamma_{\mathfrak{a}})\,,\ \textit{for all} \mathfrak{a}\in\chi.
\end{cases}
\tag{11}
$$

2. *The following strong convergences hold, with $p = u^1 - u^2$ and $\overline{p}_\mu = \overline{u}^1_\mu - \overline{u}^2_\mu$:*

$$
\begin{cases}
\Pi^{\mu}_{\mathcal{D}^l} S_\mu(p^l) \to S_\mu(\overline{p}_\mu) & \textit{in } L^2((0,T)\times M_\mu)\,,\ \textit{for } \mu\in\{m,f\},\\
\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l} S_{\mathfrak{a}}(p^l) \to S_{\mathfrak{a}}(\gamma_{\mathfrak{a}}\overline{p}_m) & \textit{in } L^2((0,T)\times \Gamma_{\mathfrak{a}})\,,\ \textit{for all } \mathfrak{a}\in\chi.
\end{cases}
\tag{12}
$$

Remark 4.2 (Uniform-in-time strong-in-space convergence) It is additionally proved in [26] that the saturations converge uniformly-in-time strongly in L^2 (that is, in $L^\infty(0,T;L^2(\Omega))$).

Remark 4.3 (Discretisation spaces varying with the time step) As mentioned in [13, Remark 3.5] for a different model, it is also possible to consider gradient schemes in which the gradient discretisation changes at each time step. This consists in choosing a *family* $\widetilde{\mathcal{D}}_S = (\mathcal{D}_{S,n})_{n=0,\ldots,N_l}$ of spatial gradient discretisations $\mathcal{D}_{S,n}$ (as in Definition 3.1), in considering unknowns $u = (u_n)_{n=0,\ldots,N} \in \prod_{n=0}^N X^0_{S,n}$ and in defining the space-time operators (3) with $\Pi^\mu_{\mathcal{D}_S} u_{n+1}$, $\nabla^\mu_{\mathcal{D}_S} u_{n+1}$, $\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_S} u_{n+1}$ and $[\![u_{n+1}]\!]_{\mathfrak{a},\mathcal{D}_S}$ respectively replaced by $\Pi^\mu_{\mathcal{D}_{S,n+1}} u_{n+1}$, $\nabla^\mu_{\mathcal{D}_{S,n+1}} u_{n+1}$, $\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}_{S,n+1}} u_{n+1}$ and $[\![u_{n+1}]\!]_{\mathfrak{a},\mathcal{D}_{S,n+1}}$. The gradient scheme is then written as in (5). For a sequence $(\widetilde{\mathcal{D}}^l_S)_{l\in\mathbb{N}}$ of such families of spatial GDs, the notions coercivity, consistency, limit-conformity and compactness are defined by writing the bound and convergences in (6), (7), (8) and (9) with $\mathcal{C}_{\mathcal{D}^l_S}$, $\mathcal{S}_{\mathcal{D}^l_S}(\overline{u})$, $\mathcal{W}_{\mathcal{D}^l_S}(\mathbf{q},\varphi_{\mathfrak{a}})$ and $\mathcal{T}_{\mathcal{D}^l_S}(\boldsymbol{\xi})$ replaced by

$$
\sup_{n=0,\ldots,N_l} \mathcal{C}_{\mathcal{D}^l_{S,n}}\,,\quad \sup_{n=0,\ldots,N_l} \mathcal{S}_{\mathcal{D}^l_{S,n}}(\overline{u})\,,\quad \sup_{n=0,\ldots,N_l} \mathcal{W}_{\mathcal{D}^l_{S,n}}(\mathbf{q},\varphi_{\mathfrak{a}}) \quad\text{and}\quad \sup_{n=0,\ldots,N_l} \mathcal{T}_{\mathcal{D}^l_{S,n}}(\boldsymbol{\xi}).
$$

With these notions, Theorem 4.1 still holds.

By using spatial GDs that change at each time step, one can represent in the GDM framework numerical methods with moving or dynamically refined meshes, or whose gradient reconstruction involves time-dependent parameters (as in $\mathbb{RT}_k$ Mixed Finite Elements with a diffusion tensor that depends on some unknown of the system; see [13, Section 4.1]).

Before delving into the proof of the theorem, let us give an overview of the strategy. The convergence of the solutions to the gradient schemes (5) is established by a compactness technique, as briefly described in [18, Section 1.2]: (i) prove *a priori* estimates on the solutions to the scheme, (ii) using discrete compactness theorems, deduce from these estimates that the (reconstructions of the) approximate solutions are compact in appropriate spaces, (iii) prove that any limit, in these spaces, of the approximate solutions is a solution to the continuous model (2).

(i) *A priori estimates*. The first *a priori* estimates are classically obtained by using the approximate solution u^α itself as a test function in the scheme (5). After summing the two equations corresponding to each phase, the diffusion terms then directly yield an estimate on $\nabla^\mu_{\mathcal{D}} u^\alpha$. The time derivative term form the discrete counterpart of $S_\mu(\overline{p}_\mu)\partial_t \overline{p}_\mu$ which, after integration in time, would yield an estimate on $\mathfrak{S}_\mu(\overline{p}_\mu)$ with $(\mathfrak{S}_\mu)' = S_\mu$. To make explicit that this estimate is actually an estimate on the saturation, we re-write $\mathfrak{S}_\mu$ as $B_\mu(S_\mu)$ for a well-chosen B_μ. These *a priori* estimates are stated in Lemma 4.4. These initial estimates only concern spatial derivatives of the approximate solution (they are a discrete equivalent of $L^2(0,T;H^1_0)$ estimates). Since this solution depends on both time and space, estimates are also required on its (discrete) time derivative to establish the compactness in an appropriate space. These time derivative estimates are the purpose of Lemma 4.6 and, classically for parabolic PDEs, they are obtained in a weak spatial norm (a sort of discrete H^{-1} norm). They are obtained on $\delta_t S_\mu(p)$ and, thanks to the modelling of the damaged rock type at the matrix-fracture interface (term $\eta\partial_t S^\alpha_{\mathfrak{a}}(\gamma_{\mathfrak{a}}\overline{p}_m)$ in (1b)), also on $\delta_t S_{\mathfrak{a}}(p)$. These estimates are instrumental to obtain the compactness in time and space of all the saturations in the model.

(ii) *Compactness*. The estimates on the discrete spatial and temporal derivatives, together with the compactness property of the gradient discretisations, yield estimates on the spatial and temporal translates of the saturations (Lemmas 4.8 and 4.9). A use of the Kolmogorov theorem and of the consistency of the gradient discretisations (to identify, through Lemma A.2, weak limits of reconstructed gradients and traces as the gradient and trace of the limit of the approximate solutions) then give the convergences (11) and (12); this is stated in Theorem 4.11. A discontinuous Ascoli-Arzela theorem (Theorem A.1) is then applied in Theorem 4.13 to obtain the convergence of the saturations uniformly-in-time and weakly in $L^2(\Omega)$. This uniform-in-time convergence is essential to pass to the limit, in (iii) below, in the energy estimate (16) (which involves pointwise-in-time values of the saturations).

(iii) *The limit is a solution of the model*. The conclusion, presented in Sect. 4.3, consists in proving that the limit of the approximate solutions is a solution to the continuous model. As we do not have strong convergence of the phase pressures u^α, the main challenge in analysing this limit arises from the non-linear upwinding terms $[kS]^\alpha_{\mathfrak{a}}(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p) T_f [\![u^\alpha]\!]^+_{\mathfrak{a},\mathcal{D}} - [kS]^\alpha_f(\Pi^f_{\mathcal{D}} p) T_f [\![u^\alpha]\!]^-_{\mathfrak{a},\mathcal{D}}$. The limit of this term is obtained by using the monotony properties of this upwinding, a Minty trick, and the discrete energy estimate (16).

4.1 Preliminary estimates

Let us introduce some useful auxiliary functions. These functions are the same as in [19,25], with adjustments to account for the fact that the saturations depends on $\mathbf{x}$ and might not vanish at $p = 0$. For $\mu \in \{m, f\} \cup \chi$, let $R_{S_\mu(\mathbf{x},\cdot)}$ be the range of $S_\mu(\mathbf{x},\cdot)$. The pseudo-inverse of $S_\mu(\mathbf{x},\cdot)$ is the mapping $[S_\mu(\mathbf{x},\cdot)]^i : R_{S_\mu(\mathbf{x},\cdot)} \to \mathbb{R}$ defined by

$$[S_\mu(\mathbf{x},\cdot)]^i(q) = \begin{cases} \inf\{z\in\mathbb{R}\mid S_\mu(\mathbf{x},z)=q\} & \text{if } q > S_\mu(\mathbf{x},0)\,, \\ 0 & \text{if } q = S_\mu(\mathbf{x},0)\,, \\ \sup\{z\in\mathbb{R}\mid S_\mu(\mathbf{x},z)=q\} & \text{if } q < S_\mu(\mathbf{x},0). \end{cases}$$

That is, $[S_\mu(\mathbf{x},\cdot)]^i(q)$ is the point z in $R_{S_\mu(\mathbf{x},\cdot)}$ that is the closest to $S_\mu(\mathbf{x},0)$ and such that $S_\mu(\mathbf{x},z)=q$. The function $B_\mu(\mathbf{x},\cdot):\mathbb{R}\to[0,\infty]$ is given by

$$B_\mu(\mathbf{x},q) = \begin{cases} \displaystyle\int_{S_\mu(\mathbf{x},0)}^q [S_\mu(\mathbf{x},\cdot)]^i(\tau)\mathrm{d}\tau & \text{if } q\in R_{S_\mu(\mathbf{x},\cdot)}\,, \\ \infty & \text{else.} \end{cases}$$

$B_\mu(\mathbf{x},\cdot)$ is convex lower semi-continuous (l.s.c.) and satisfies the following properties [25]

$$B_\mu(\mathbf{x},S_\mu(\mathbf{x},r)) = \int_0^r \tau\frac{\partial S_\mu}{\partial q}(\mathbf{x},\tau)\mathrm{d}\tau, \tag{13}$$

$$\forall a,b\in\mathbb{R}\,,\; a(S_\mu(\mathbf{x},b)-S_\mu(\mathbf{x},a)) \le B_\mu(\mathbf{x},S_\mu(\mathbf{x},b)) - B_\mu(\mathbf{x},S_\mu(\mathbf{x},a)) \tag{14}$$

and, for some K_0, K_1 and K_2 not depending on $\mathbf{x}$ or r,

$$K_0 S_\mu(\mathbf{x},r)^2 - K_1 \le B_\mu(\mathbf{x},S_\mu(\mathbf{x},r)) \le K_2 r^2. \tag{15}$$

In the following, we write $A\lesssim B$ for "$A\le MB$ for a constant M depending only on an upper bound of $C_\mathcal{D}$ and on the data in the assumptions of Sect. 2.2".

Lemma 4.4 (Energy estimates) *Under the assumptions of Sect. 2.2, let $\mathcal{D}$ be a gradient discretisation with piecewise constant reconstructions $\Pi_\mathcal{D}^\mu$, $\mathbb{T}_\mathcal{D}^\mathfrak{a}$. Let $(u^\alpha)_{\alpha=1,2}\in[(X^0)^{N+1}]^2$ be a solution of the gradient scheme of (5). Take $T_0\in(0,T]$ and $k\in\{0,\ldots,N-1\}$ such that $T_0\in(t_k,t_{k+1}]$. Then*

$$\begin{aligned}
&\sum_{\mu\in\{m,f\}}\int_{M_\mu}\phi_\mu\left[B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p(T_0))) - B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p_0))\right]\mathrm{d}\tau_\mu \\
&\quad+\sum_{\alpha=1}^2\sum_{\mu\in\{m,f\}}\int_0^{T_0}\int_{M_\mu}[kS]_\mu^\alpha(\Pi_\mathcal{D}^\mu p)\Lambda_\mu\nabla_\mathcal{D}^\mu u^\alpha\cdot\nabla_\mathcal{D}^\mu u^\alpha\mathrm{d}\tau_\mu\mathrm{d}t \\
&\quad+\sum_{\mathfrak{a}\in\chi}\int_{\Gamma_\mathfrak{a}}\eta\left[B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}_S}^\mathfrak{a} p(T_0))) - B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}_S}^\mathfrak{a} p_0))\right]\mathrm{d}\tau \\
&\quad+\sum_{\alpha=1}^2\sum_{\mathfrak{a}\in\chi}\int_0^{T_0}\int_{\Gamma_\mathfrak{a}}\Big([kS]_\mathfrak{a}^\alpha(\mathbb{T}_\mathcal{D}^\mathfrak{a} p)T_f[\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^+ \\
&\quad-[kS]_f^\alpha(\Pi_\mathcal{D}^f p)T_f[\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^-\Big)[\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}\mathrm{d}\tau\mathrm{d}t \\
&\qquad\le\sum_{\alpha=1}^2\sum_{\mu\in\{m,f\}}\int_0^{t_{k+1}}\int_{M_\mu}h_\mu^\alpha\Pi_\mathcal{D}^\mu u^\alpha\mathrm{d}\tau_\mu\mathrm{d}t.
\end{aligned} \tag{16}$$

As a consequence,

$$\sum_{\alpha=1,2} \|u^\alpha\|_{\mathcal{D}}^2 \lesssim 1 + \sum_{\mu\in\{m,f\}} \|\Pi_{\mathcal{D}}^\mu p_0\|_{L^2(M_\mu)}^2 + \sum_{\mathfrak{a}\in\chi} \|\mathbb{T}_{\mathcal{D}_S}^{\mathfrak{a}} p_0\|_{L^2(\Gamma_\mathfrak{a})}^2. \tag{17}$$

Proof We remove the spatial coordinate $\mathbf{x}$ in the arguments, when not needed. Reasoning as in [19, Lemma 4.1], Property (14) gives

$$\begin{aligned}
&\sum_{\mu\in\{m,f\}} \int_0^{t_{k+1}} \int_{M_\mu} \phi_\mu \Pi_{\mathcal{D}}^\mu \Big[\delta_t S_\mu(p)\Big] \Pi_{\mathcal{D}}^\mu p \,\mathrm{d}\tau_\mu \mathrm{d}t \\
&\quad = \sum_{\mu\in\{m,f\}} \sum_{n=0}^{k} \int_{M_\mu} \phi_\mu \left[S_\mu(\Pi_{\mathcal{D}_S}^\mu p_{n+1}) - S_\mu(\Pi_{\mathcal{D}_S}^\mu p_n)\right] \Pi_{\mathcal{D}_S}^\mu p_{n+1} \mathrm{d}\tau_\mu \\
&\quad \geq \sum_{\mu\in\{m,f\}} \sum_{n=0}^{k} \int_{M_\mu} \phi_\mu \left[B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p_{n+1})) - B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p_n))\right] \mathrm{d}\tau_\mu \\
&\quad = \sum_{\mu\in\{m,f\}} \int_{M_\mu} \phi_\mu \left[B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p(T_0))) - B_\mu(S_\mu(\Pi_{\mathcal{D}_S}^\mu p_0))\right] \mathrm{d}\tau_\mu
\end{aligned} \tag{18}$$

where we have used, by definition, $\Pi_{\mathcal{D}_S}^\mu p(T_0) = \Pi_{\mathcal{D}_S}^\mu p_{k+1}$. Similarly,

$$\begin{aligned}
\int_0^{t_{k+1}} \int_{\Gamma_\mathfrak{a}} \eta \mathbb{T}_{\mathcal{D}}^{\mathfrak{a}} \Big[\delta_t S_\mathfrak{a}(p)\Big] \mathbb{T}_{\mathcal{D}}^{\mathfrak{a}} p \,\mathrm{d}\tau \mathrm{d}t \geq \int_{\Gamma_\mathfrak{a}} \eta \left[B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}_S}^{\mathfrak{a}} p(T_0)))\right. \\
\left. - B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}_S}^{\mathfrak{a}} p_0))\right] \mathrm{d}\tau.
\end{aligned} \tag{19}$$

Equation (16) is then obtained by taking $v^\alpha = (u_0^\alpha, \ldots, u_{k+1}^\alpha, 0, \ldots, 0)$ (for $\alpha = 1, 2$) in the gradient scheme (5), by summing the resulting equations over $\alpha = 1, 2$, by using (18) and (19), and by reducing the time integrals in the left-hand side from $[0, t_{k+1}]$ to $[0, T_0]$, due to the non-negativity of the integrands.

The inequality (17) is the consequence of a few simple estimates on the terms of (16) with $T_0 = T$. For the symmetric diffusion terms (for $\alpha = 1, 2$ and $\mu \in \{m, f\}$), we write

$$\int_0^T \int_{M_\mu} [kS]_\mu^\alpha(\Pi_{\mathcal{D}}^\mu p) \Lambda_\mu \nabla_{\mathcal{D}}^\mu u^\alpha \cdot \nabla_{\mathcal{D}}^\mu u^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \geq \underline{d}_\mu \underline{k}_\mu \underline{\lambda}_\mu \|\nabla_{\mathcal{D}}^\mu u^\alpha\|_{L^2((0,T)\times M_\mu)}^2 \tag{20}$$

where $\underline{d}_\mu = 1$ if $\mu = m$. The matrix–fracture coupling terms are handled by noticing that, for any $s \in \mathbb{R}$, $s^+ s = (s^+)^2$ and $s^- s = -(s^-)^2$, so that for $\alpha = 1, 2$ and $\mathfrak{a} \in \chi$,

$$\int_0^T \int_{\Gamma_\mathfrak{a}} \Big([kS]_\mathfrak{a}^\alpha(\mathbb{T}_{\mathcal{D}}^{\mathfrak{a}} p) T_f [\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^+ - [kS]_f^\alpha(\Pi_{\mathcal{D}}^f p) T_f [\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}}^-\Big) [\![u^\alpha]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau \mathrm{d}t$$

$$= \int_0^T \int_{\Gamma_\mathfrak{a}} \Big([kS]^\alpha_\mathfrak{a}(\mathbb{T}^\mathfrak{a}_\mathcal{D} p) T_f(\llbracket u^\alpha \rrbracket^+_{\mathfrak{a},\mathcal{D}})^2 + [kS]^\alpha_f(\Pi^f_\mathcal{D} p) T_f(\llbracket u^\alpha \rrbracket^-_{\mathfrak{a},\mathcal{D}})^2 \mathrm{d}\tau \mathrm{d}t \Big)$$
$$\gtrsim \| \llbracket u^\alpha \rrbracket_{\mathfrak{a},\mathcal{D}} \|^2_{L^2((0,T)\times\Gamma_\mathfrak{a})}. \tag{21}$$

Here, we have used $[kS]^\alpha_\mathfrak{a}(\mathbb{T}^\mathfrak{a}_\mathcal{D} p) \geq \underline{k}_\mathfrak{a}$, $[kS]^\alpha_f(\Pi^f_\mathcal{D} p) \geq \underline{k}_f$ and $|s|^2 = (s^+)^2 + (s^-)^2$. Plugging estimates (15), (20) and (21) in (16) (with $T_0 = T$) and invoking Cauchy–Schwarz inequalities leads to

$$\sum_{\alpha=1}^2 \Big[\|\nabla^m_\mathcal{D} u^\alpha\|^2_{L^2((0,T)\times\Omega)^d} + \|\nabla^f_\mathcal{D} u^\alpha\|^2_{L^2((0,T)\times\Gamma)^{d-1}} + \sum_{\mathfrak{a}\in\chi} \|\llbracket u^\alpha \rrbracket_{\mathfrak{a},\mathcal{D}}\|^2_{L^2((0,T)\times\Gamma)} \Big]$$
$$\lesssim \sum_{\mu\in\{m,f\}} \Big[\sum_{\alpha=1}^2 \|h^\alpha_\mu\|_{L^2((0,T)\times M_\mu)} \|\Pi^\mu_\mathcal{D} u^\alpha\|_{L^2((0,T)\times M_\mu)} + \|\Pi^\mu_\mathcal{D} p_0\|^2_{L^2(M_\mu)} \Big]$$
$$+ \|\mathbb{T}^\mathfrak{a}_{\mathcal{D}_S} p_0\|^2_{L^2(M_\mu)}.$$

The proof of (17) is complete by noticing that the left-hand side is equal to $\sum_{\alpha=1}^2 \|u^\alpha\|^2_\mathcal{D}$, and by using Young's inequality and the definition of $C_\mathcal{D}$ in the right-hand side. $\square$

The existence of a solution to the gradient scheme follows by a standard fixed point argument based on the Leray–Schauder topological degree, see e.g. [10, proof of Lemma 3.2] or [23, Step 1 in the proof of Theorem 3.1].

Corollary 4.5 *Under the assumptions of Lemma 4.4, there exists a solution to the gradient scheme* (5).

We now want to obtain estimates on the discrete time derivatives. Let the dual norm of $W = [w_m, w_f, (w_\mathfrak{a})_{\mathfrak{a}\in\chi}] \in (X^0)^{2+\sharp\chi}$ be defined by

$$|W|_{\mathcal{D}_S,*} = \sup \Big\{ \sum_{\mu\in\{m,f\}} \int_{M_\mu} \phi_\mu \Pi^\mu_{\mathcal{D}_S} w_\mu \Pi^\mu_{\mathcal{D}_S} v \mathrm{d}\tau_\mu$$
$$+ \sum_{\mathfrak{a}\in\chi} \int_{\Gamma_\mathfrak{a}} \eta \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S} w_\mathfrak{a} \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S} v \mathrm{d}\tau \;:\; v \in X^0 ,\; \|v\|_{\mathcal{D}_S} \leq 1 \Big\}. \tag{22}$$

Lemma 4.6 (Weak estimate on time derivatives) *Under the assumptions of Sect. 2.2, let $\mathcal{D}$ be a gradient discretisation with piecewise constant reconstructions $\Pi^\mu_\mathcal{D}$, $\mathbb{T}^\mathfrak{a}_\mathcal{D}$. Let $(u^\alpha)_{\alpha=1,2} \in [(X^0)^{N+1}]^2$ be a solution of the gradient scheme of* (5). *Then,*

$$\int_0^T \Big| \Big[\delta_t S_m(p)(t), \delta_t S_f(p)(t), (\delta_t S_\mathfrak{a}(p)(t))_{\mathfrak{a}\in\chi} \Big] \Big|^2_{\mathcal{D}_S,*} \mathrm{d}t \lesssim 1 + \sum_{\alpha=1,2} \|u^\alpha\|^2_\mathcal{D}.$$

Remark 4.7 (Damaged rock modelling) The modelling of the damaged rock type (term $\eta\partial_t S^\alpha_{\mathfrak{a}}(\gamma_{\mathfrak{a}}\overline{p}_m)$ in (1b)) is essential to obtain the estimate on $\delta_t S_{\mathfrak{a}}(p)$ above. These estimates are required to obtain the compactness of this discrete saturation (see Theorems 4.11 and 4.13).

Proof Take $v \in X^0$ and apply (5) with $\alpha = 1$ to the test function $(0, \ldots, 0, v, 0, \ldots, 0)$, where v is at an arbitrary position n. This shows that, for all $n = 0, \ldots, N$ and $t \in (t_n, t_{n+1}]$

$$\sum_{\mu\in\{m,f\}} \int_{M_\mu} \phi_\mu \Pi^\mu_{\mathcal{D}}\Big[\delta_t S_\mu(p)\Big](t)\Pi^\mu_{\mathcal{D}} v \,\mathrm{d}\tau_\mu + \sum_{\mathfrak{a}\in\chi} \int_{\Gamma_{\mathfrak{a}}} \eta \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}\Big[\delta_t S_{\mathfrak{a}}(p)\Big](t)\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} v \,\mathrm{d}\tau$$
$$= \sum_{\mu\in\{m,f\}} \Big(\int_{M_\mu} \Big[\frac{1}{\Delta t_{n+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} h_\mu(s)\mathrm{d}s\Big]\Pi^\mu_{\mathcal{D}} v \,\mathrm{d}\tau_\mu$$
$$- \int_{M_\mu} [kS]_\mu(\Pi^\mu_{\mathcal{D}} p)(t)\ \Lambda_\mu \nabla^\mu_{\mathcal{D}} u(t) \cdot \nabla^\mu_{\mathcal{D}} v \,\mathrm{d}\tau_\mu\Big)$$
$$- \sum_{\mathfrak{a}\in\chi} \int_{\Gamma_{\mathfrak{a}}} \Big([kS]_{\mathfrak{a}}(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p)(t) T_f [\![u(t)]\!]^+_{\mathfrak{a},\mathcal{D}}$$
$$- [kS]_f(\Pi^f_{\mathcal{D}} p)(t) T_f [\![u(t)]\!]^-_{\mathfrak{a},\mathcal{D}}\Big)[\![v]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau$$
$$\lesssim \left\| \frac{1}{\Delta t_{n+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} h_\mu(s)\mathrm{d}s \right\|_{L^2(M_\mu)} \|v\|_{\mathcal{D}_S} + \|u(t)\|_{\mathcal{D}_S}\|v\|_{\mathcal{D}_S},$$

where we have used the definition of $C_{\mathcal{D}}$ in the last step. Taking the supremum over all v such that $\|v\|_{\mathcal{D}_S} \le 1$ shows that

$$\Big\|\Big[\delta_t S_m(p)(t), \delta_t S_f(p)(t), (\delta_t S_{\mathfrak{a}}(p)(t))_{\mathfrak{a}\in\chi}\Big]\Big\|_{\mathcal{D}_S,*} \lesssim \frac{1}{\Delta t_{n+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} \|h_\mu(s)\|_{L^2(M_\mu)}\mathrm{d}s$$
$$+ \|u(t)\|_{\mathcal{D}_S}. \tag{23}$$

Take the square of this relation, use $(a+b)^2 \le 2a^2 + 2b^2$, and apply Jensen's inequality to introduce the square inside the time integral. Multiply then by $\Delta t_{n+\frac{1}{2}}$ and sum over n to conclude. □

Lemma 4.8 (Estimate on time translates) *Under the assumptions of Sect. 2.2, let $\mathcal{D}$ be a gradient discretisation with piecewise constant reconstructions $\Pi^\mu_{\mathcal{D}}$, $\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}$. For any $h > 0$ and any solution $(u^\alpha)_{\alpha=1,2} \in [(X^0)^{N+1}]^2$ of (5),*

$$\sum_{\mu\in\{m,f\}} \|S_\mu(\mathrm{T}_h\Pi^\mu_{\mathcal{D}} p) - S_\mu(\Pi^\mu_{\mathcal{D}} p)\|^2_{L^2((0,T)\times M_\mu)} + \sum_{\mathfrak{a}\in\chi} \|S_{\mathfrak{a}}(\mathrm{T}_h\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p)$$
$$- S_{\mathfrak{a}}(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p)\|^2_{L^2((0,T)\times\Gamma_{\mathfrak{a}})} \lesssim (h + \Delta t)\Big(1 + \sum_{\alpha=1}^{2} \|u^\alpha\|^2_{\mathcal{D}}\Big), \tag{24}$$

where we recall that $\mathrm{T}_h g(s) = g(s+h)$ *and* $\Delta t = \max\{\Delta t_{n+\frac{1}{2}} : n = 0, \ldots, N-1\}$*, and where all functions of time have been extended by* 0 *outside* $(0, T)$.

Proof Let us start by assuming that $h \in (0, T)$, and let us consider integrals over $(0, T-h)$ (we therefore do not use extensions outside $(0, T)$ yet). By the Lipschitz continuity and monotonicity of the saturations $S_\mu = S_\mu^1$ we have $|S_\mu(b) - S_\mu(a)|^2 \lesssim (S_\mu(b) - S_\mu(a))(b-a)$. Thus, setting $n(s) = \min\{k = 1, \ldots, N \mid t_k \geq s\}$ for all $s \in \mathbb{R}$,

$$\begin{aligned}
&\sum_{\mu\in\{m,f\}} \int_0^{T-h} \int_{M_\mu} |S_\mu(\mathrm{T}_h \Pi_{\mathcal{D}}^\mu p) - S_\mu(\Pi_{\mathcal{D}}^\mu p)|^2 \mathrm{d}\tau_\mu \mathrm{d}s \\
&\qquad + \sum_{\mathfrak{a}\in\chi} \int_0^{T-h} \int_{\Gamma_\mathfrak{a}} |S_\mathfrak{a}(\mathrm{T}_h \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p) - S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}}^\mathfrak{a} p)|^2 \mathrm{d}\tau \mathrm{d}s \\
&\lesssim \sum_{\mu\in\{m,f\}} \int_0^{T-h} \int_{M_\mu} \phi_\mu \Big(S_\mu(\mathrm{T}_h \Pi_{\mathcal{D}}^\mu p) - S_\mu(\Pi_{\mathcal{D}}^\mu p) \Big)(s) (\mathrm{T}_h \Pi_{\mathcal{D}}^\mu p \\
&\qquad - \Pi_{\mathcal{D}}^\mu p)(s) \mathrm{d}\tau_\mu \mathrm{d}s \\
&\qquad + \sum_{\mathfrak{a}\in\chi} \int_0^{T-h} \int_{\Gamma_\mathfrak{a}} \eta \Big(S_\mathfrak{a}(\mathrm{T}_h \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p) - S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}}^\mathfrak{a} p) \Big)(s) (\mathrm{T}_h \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p \\
&\qquad - \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p)(s) \mathrm{d}\tau \mathrm{d}s \\
&\lesssim \int_0^{T-h} \Big[\sum_{\mu\in\{m,f\}} \int_{M_\mu} \int_{t_{n(s)}}^{t_{n(s+h)}} \phi_\mu \Pi_{\mathcal{D}}^\mu \Big[\delta_t S_\mu(p) \Big](t) (\mathrm{T}_h \Pi_{\mathcal{D}}^\mu p \\
&\qquad - \Pi_{\mathcal{D}}^\mu p)(s) \mathrm{d}t \mathrm{d}\tau_\mu \\
&\qquad + \sum_{\mathfrak{a}\in\chi} \int_{\Gamma_\mathfrak{a}} \int_{t_{n(s)}}^{t_{n(s+h)}} \eta \mathbb{T}_{\mathcal{D}}^\mathfrak{a} \Big[\delta_t S_\mathfrak{a}(p) \Big](t) (\mathrm{T}_h \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p \\
&\qquad - \mathbb{T}_{\mathcal{D}}^\mathfrak{a} p)(s) \mathrm{d}t \mathrm{d}\tau \Big] \mathrm{d}s.
\end{aligned} \tag{25}$$

In the last line, we simply wrote $S_\mu(\mathrm{T}_h \Pi_{\mathcal{D}}^\mu p)(s) - S_\mu(\Pi_{\mathcal{D}}^\mu p)(s) = S_\mu(\Pi_{\mathcal{D}}^\mu p)(s+h) - S_\mu(\Pi_{\mathcal{D}}^\mu p)(s)$ as the sum of the jumps if $S_\mu(\Pi_{\mathcal{D}}^\mu p)$ between s and $s+h$ (likewise for $S_\mathfrak{a}(\mathbb{T}_{\mathcal{D}}^\mathfrak{a} p)$).

For a fixed s, define $v \in (X^0)^{N+1}$ by

$$v_k = \begin{cases} p_{n(s+h)} - p_{n(s)} & \text{if } n(s)+1 \leq k \leq n(s+h) \\ 0 & \text{else.} \end{cases}$$

With this choice,

$$
\begin{aligned}
\Pi^{\mu}_{\mathcal{D}} v(t,\mathbf{x}) &= \mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\ (\mathrm{T}_h \Pi^{\mu}_{\mathcal{D}} p - \Pi^{\mu}_{\mathcal{D}} p)(s,\mathbf{x}),\\
\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} v(t,\mathbf{x}) &= \mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\ (\mathrm{T}_h \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p - \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} p)(s,\mathbf{x}),\\
\nabla^{\mu}_{\mathcal{D}} v(t,\mathbf{x}) &= \mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\ (\mathrm{T}_h \nabla^{\mu}_{\mathcal{D}} p - \nabla^{\mu}_{\mathcal{D}} p)(s,\mathbf{x})\ ,\ \text{ and}\\
[\![v]\!]_{\mathfrak{a},\mathcal{D}}(t,\mathbf{x}) &= \mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\ (\mathrm{T}_h [\![p]\!]_{\mathfrak{a},\mathcal{D}} - [\![p]\!]_{\mathfrak{a},\mathcal{D}})(s,\mathbf{x}).
\end{aligned}
\tag{26}
$$

We keep s fixed and concentrate on the integrand of the outer integral in the right-hand side of (25). Estimate (23), the definition (22) of $|\cdot|_{\mathcal{D}_S,*}$, and Young's inequality yield

$$
\begin{aligned}
&\sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} \phi_\mu \Pi^{\mu}_{\mathcal{D}}\Big[\delta_t S_\mu(p)\Big]\Pi^{\mu}_{\mathcal{D}} v \mathrm{d}\tau_\mu \mathrm{d}t + \sum_{\mathfrak{a}\in\chi} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \eta \mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}\Big[\delta_t S_{\mathfrak{a}}(p)\Big]\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}} v \mathrm{d}\tau \mathrm{d}t\\
&\quad\lesssim \int_0^T (\|h_\mu(t)\|_{L^2(M_\mu)} + \|u(t)\|_{\mathcal{D}_S})\|v\|_{\mathcal{D}_S}\mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\mathrm{d}t\\
&\quad\lesssim \int_0^T (\|h_\mu(t)\|_{L^2(M_\mu)} + \|u(t)\|_{\mathcal{D}_S})^2\mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\mathrm{d}t + (t_{n(s+h)} - t_{n(s)})\|v\|^2_{\mathcal{D}_S}.
\end{aligned}
$$

Returning to (25), integrate the previous estimate over $s\in(0,T-h)$. In this step, it is crucial to realise that

$$
t_{n(s+h)} - t_{n(s)} \le h + \Delta t \ \text{ and }\ \int_0^{T-h} \mathbb{1}_{(t_{n(s)},t_{n(s+h)}]}(t)\mathrm{d}s \le \int_0^T \mathbb{1}_{[t-h-\Delta t,t]}(s)\mathrm{d}s \le h+\Delta t.
$$

Hence, recalling the definition of v,

$$
\begin{aligned}
\mathrm{RHS}(25) &\lesssim (h+\Delta t)\Bigg[\int_0^T (\|h_\mu(t)\|_{L^2(M_\mu)} + \|u(t)\|_{\mathcal{D}_S})^2\mathrm{d}t\\
&\qquad + \int_0^{T-h} \|p_{n(s+h)}\|^2_{\mathcal{D}_S}\mathrm{d}s + \int_0^{T-h} \|p_{n(s)}\|^2_{\mathcal{D}_S}\mathrm{d}s\Bigg]\\
&\lesssim (h+\Delta t)\Big(1 + \|u\|^2_{\mathcal{D}} + \|p\|^2_{\mathcal{D}}\Big).
\end{aligned}
$$

Since $p = u^1 - u^2$, this proves (24) with $L^2(0,T-h)$ norms in the left-hand side, instead of $L^2(0,T)$ norms. The complete form of (24) follows by recalling that $0\le S_\mu \le 1$, so that $\|S_\mu(\Pi^{\mu}_{\mathcal{D}} p)\|^2_{L^2((T-h,T)\times M_\mu)} \lesssim h$ (and similarly for other saturation terms). □

Lemma 4.9 (Estimate on space translates) *Under the assumptions of Sect. 2.2, let $\mathcal{D}$ be a gradient discretisation with piecewise constant reconstructions $\Pi^{\mu}_{\mathcal{D}}$, $\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}$. Let $(u^\alpha)_{\alpha=1,2} \in [(X^0)^{N+1}]^2$ be a solution of (5), and let $\boldsymbol{\xi} = (\boldsymbol{\xi}_m, \boldsymbol{\xi}_f)$, with $\boldsymbol{\xi}_m \in \mathbb{R}^d$ and $\boldsymbol{\xi}_f = (\boldsymbol{\xi}^i_f)_{i\in I} \in \bigoplus_{i\in I}\tau(\mathcal{P}_i)$, where $\tau(\mathcal{P}_i)$ is the (const.) tangent space of $\mathcal{P}_i$.*

Then, extending the functions $\Pi^{\mu}_{\mathcal{D}}p$ and S_μ by 0 outside M_μ,

$$\begin{aligned}&\|\mathrm{T}_{\boldsymbol{\xi}_m} S_m(\Pi^m_{\mathcal{D}}p) - S_m(\Pi^m_{\mathcal{D}}p)\|^2_{L^2((0,T)\times\mathbb{R}^d)} + \sum_{i\in I}\Big(\|\mathrm{T}_{\boldsymbol{\xi}^i_f} S_f(\Pi^f_{\mathcal{D}}p)\\ &\quad - S_f(\Pi^f_{\mathcal{D}}p)\|^2_{L^2((0,T)\times\mathcal{P}_i)}\\ &\quad + \sum_{\mathfrak{a}=\mathfrak{a}^{\pm}(i)} \|\mathrm{T}_{\boldsymbol{\xi}^i_f} S_{\mathfrak{a}}(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}p) - S_{\mathfrak{a}}(\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}}p)\|^2_{L^2((0,T)\times\mathcal{P}_i)}\Big)\\ &\quad \lesssim \mathcal{T}_{\mathcal{D}_S}(\boldsymbol{\xi})\sum_{\alpha=1}^{2}\|u^\alpha\|^2_{\mathcal{D}} + |\boldsymbol{\xi}|,\end{aligned}$$

where we recall that $\mathrm{T}_{\boldsymbol{\zeta}} f(\boldsymbol{x}) = f(\boldsymbol{x}+\boldsymbol{\zeta})$, and $\mathcal{T}_{\mathcal{D}_S}$ is given in Definition 3.8.

Proof Let us focus on the matrix Ω, and remember that, as a function of $\mathbf{x}$, S_m is piecewise constant on a polytopal partition $(\Omega_j)_{j\in J_m}$. Write

$$\begin{aligned}\mathrm{T}_{\boldsymbol{\xi}_m} S_m(\Pi^m_{\mathcal{D}}p) - S_m(\Pi^m_{\mathcal{D}}p) &= S_m(\mathbf{x}+\boldsymbol{\xi}_m, \Pi^m_{\mathcal{D}}p(\mathbf{x}+\boldsymbol{\xi}_m,t)) - S_m(\mathbf{x}+\boldsymbol{\xi}_m, \Pi^m_{\mathcal{D}}p(\mathbf{x},t))\\ &\quad + S_m(\mathbf{x}+\boldsymbol{\xi}_m, \Pi^m_{\mathcal{D}}p(\mathbf{x},t)) - S_m(\mathbf{x}, \Pi^m_{\mathcal{D}}p(\mathbf{x},t)).\end{aligned} \tag{27}$$

Let $\Omega_{\boldsymbol{\xi}_m} = \bigcup_j\{\mathbf{x}\in\Omega_j \mid \mathbf{x}+\boldsymbol{\xi}_m \notin \Omega_j\} \cup \{\mathbf{x}\in\mathbb{R}^d\setminus\Omega \mid \mathbf{x}+\boldsymbol{\xi}_m\in\Omega\}$ be the set of points $\mathbf{x}$ that do not belong to the same element Ω_j as their translate $\mathbf{x}+\boldsymbol{\xi}_m$. By assumption on S_m,

$$\sup_{q\in\mathbb{R}} |S_m(\mathbf{x}+\boldsymbol{\xi}_m,q) - S_m(\mathbf{x},q)| \le \begin{cases} 0 & \text{on } \mathbb{R}^d\setminus\Omega_{\boldsymbol{\xi}_m},\\ 1 & \text{on } \Omega_{\boldsymbol{\xi}_m}.\end{cases}$$

Moreover, since each Ω_j is polytopal, $|\Omega_{\boldsymbol{\xi}_m}| \lesssim |\boldsymbol{\xi}_m|$. Hence,

$$\int_0^T\int_{\mathbb{R}^d}\sup_{q\in\mathbb{R}}|S_m(\mathbf{x}+\boldsymbol{\xi}_m,q)-S_m(\mathbf{x},q)|^2\mathrm{d}\mathbf{x}\mathrm{d}t \lesssim |\boldsymbol{\xi}_m|. \tag{28}$$

On the other hand, by definition of $\mathcal{T}_{D_S}$ and the Lipschitz continuity of S_m,

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^d}|S_m(\mathbf{x}+\boldsymbol{\xi}_m,\Pi^m_{\mathcal{D}}p(\mathbf{x}+\boldsymbol{\xi}_m,t)) - S_m(\mathbf{x}+\boldsymbol{\xi}_m,\Pi^m_{\mathcal{D}}p(\mathbf{x},t))|^2\mathrm{d}\mathbf{x}\mathrm{d}t\\ &\lesssim \int_0^T\int_{\mathbb{R}^d}|\Pi^m_{\mathcal{D}}p(\mathbf{x}+\boldsymbol{\xi}_m,t)-\Pi^m_{\mathcal{D}}p(\mathbf{x},t)|^2\mathrm{d}\mathbf{x}\mathrm{d}t \lesssim \|p\|^2_{\mathcal{D}}\mathcal{T}_{D_S}(\boldsymbol{\xi}).\end{aligned} \tag{29}$$

Plugging (28) and (29) into (27) and reasoning similarly for S_f and $S_{\mathfrak{a}}$ concludes the proof. □

Remark 4.10 This proof is the only place where the assumption that each M^j_μ is polytopal is used; this is to ensure that $|\Omega_{\boldsymbol{\xi}_m}| \lesssim |\boldsymbol{\xi}_m|$ (and likewise for fracture and interfacial terms). Obviously, this asssumption on the sets M_μ^j could be relaxed (e.g., into "each M_μ^j has a Lipschitz-continuous boundary"), but assuming that these sets are polytopal is not restrictive for practical applications.

4.2 Initial convergences

We can now state our initial convergence theorem for sequences of solutions to gradient schemes. This theorem does not yet identify the weak limits of such sequences.

Theorem 4.11 (Averaged-in-time convergence of approximate solutions) *Let $(\mathcal{D}^l)_{l\in\mathbb{N}}$ be a coercive, consistent, limit-conforming and compact sequence of space-time gradient discretisations, with piecewise constant reconstructions. Let $(u^{\alpha,l})_{\alpha=1,2\,,l\in\mathbb{N}}$ be such that $(u^{\alpha,l})_{\alpha=1,2} \in [(X_l^0)^{N_l+1}]^2$ is a solution of (5) with $\mathcal{D}=\mathcal{D}^l$. Then, there exists $(\overline{u}^\alpha)_{\alpha=1,2} = (\overline{u}_m^\alpha, \overline{u}_f^\alpha)_{\alpha=1,2} \in [L^2(0,T;V_m^0)\times L^2(0,T;V_f^0)]^2$ such that, up to a subsequence as $l\to\infty$, the convergences (11) and (12) hold.*

Proof Combining Lemmata 4.4 and A.2 immediately gives (11) up to a subsequence. By assumption, $0 \le S_\mu, S_\mathfrak{a} \le 1$ and therefore, by Lemmata 4.8 and 4.9 and the Kolmogorov compactness theorem, there exists a subsequence of $(\Pi_{\mathcal{D}^l}^\mu S_\mu(p^l))_l$ that strongly converges in $L^2((0,T)\times M_\mu)$ and a subsequence of $(\mathbb{T}_{\mathcal{D}^l}^\mathfrak{a} S_\mathfrak{a}(p^l))_l$ that strongly converges in $L^2((0,T)\times\Gamma_\mathfrak{a})$. Also, by assumption, S_μ, $S_\mathfrak{a}$ are non-decreasing functions, which allows us to identify the limits in (12) by applying Corollary A.3. □

Let C_Ω^∞ be the subspace of functions in $C_b^\infty(\Omega\setminus\overline{\Gamma})$ vanishing on a neighbourhood of the boundary $\partial\Omega$. Define also $C_\Gamma^\infty = \gamma_\Gamma(C_0^\infty(\Omega))$ as the image of $C_0^\infty(\Omega)$ through the trace operator $\gamma_\Gamma : H_0^1(\Omega)\to L^2(\Gamma)$.

The following lemma and theorem add a uniform-in-time weak L^2 convergence property to the convergences established in Theorem 4.11.

Lemma 4.12 (Uniform-in-time, weak-in-space translate estimates) *Under the assumptions of Sect. 2.2, let $\mathcal{D}$ be a gradient discretisation with piecewise constant reconstructions $\Pi_\mathcal{D}^\mu$, $\mathbb{T}_\mathcal{D}^\mathfrak{a}$. Let $(u^\alpha)_{\alpha=1,2} \in [(X^0)^{N+1}]^2$ be a solution of the gradient scheme (5), and $p = u^1 - u^2$. Then, for all $\varphi = (\varphi_m,\varphi_f) \in C_\Omega^\infty \times C_\Gamma^\infty$ and all $s,t\in[0,T]$,*

$$\begin{aligned}
&\Bigg|\sum_{\mu\in\{m,f\}} \left\langle d_\mu\phi_\mu \Pi_\mathcal{D}^\mu S_\mu(p)(s) - d_\mu\phi_\mu\Pi_\mathcal{D}^\mu S_\mu(p)(t), \varphi_\mu\right\rangle_{L^2(M_\mu)}\\
&\quad + \sum_{\mathfrak{a}\in\chi} \left\langle \eta\mathbb{T}_\mathcal{D}^\mathfrak{a} S_\mathfrak{a}(p)(s) - \eta\mathbb{T}_\mathcal{D}^\mathfrak{a} S_\mathfrak{a}(p)(t), \gamma_\mathfrak{a}\varphi_m\right\rangle_{L^2(\Gamma_\mathfrak{a})}\Bigg|\\
&\quad \lesssim \mathcal{S}_{\mathcal{D}_S}(\varphi) + (\mathcal{S}_{\mathcal{D}_S}(\varphi) + C_\varphi)\left(1+\sum_{\alpha=1}^2 \|u^\alpha\|_\mathcal{D}^2\right)^{\frac12}\left[|s-t|^{\frac12} + (\Delta t)^{\frac12}\right]. \qquad (30)
\end{aligned}$$

where C_φ only depends on φ, d_f is the width of the fractures, and $d_m = 1$.

Proof Let us introduce an interpolant $\mathrm{P}_{\mathcal{D}_S} : C^\infty_\Omega \times C^\infty_\Gamma \to X^0$ such that, for all $\varphi \in C^\infty_\Omega \times C^\infty_\Gamma$, $s_{\mathcal{D}_S}(\mathrm{P}_{\mathcal{D}_S}\varphi, \varphi) = \mathcal{S}_{\mathcal{D}_S}(\varphi)$. As in the proof of Lemma 4.8, let $n(r) = \min\{k = 1, \ldots, N \mid t_k \geq r\}$ for all $r \in [0, T]$. Denote by L the left-hand side of (30) and introduce $\Pi^\mu_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi$ in the first sum and $\mathbb{T}^\mathfrak{a}_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi$ in the second sum to write

$$L \leq \sum_{\mu\in\{m,f\}} \left(\left| \left\langle d_\mu\phi_\mu \Pi^\mu_{\mathcal{D}} S_\mu(p)(s) - d_\mu\phi_\mu \Pi^\mu_{\mathcal{D}} S_\mu(p)(t), \varphi_\mu - \Pi^\mu_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(M_\mu)} \right| \right) \tag{31}$$

$$+ \sum_{\mathfrak{a}\in\chi} \left(\left| \left\langle \eta \mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(s) - \eta \mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(t), \gamma_\mathfrak{a}\varphi_m - \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(\Gamma_\mathfrak{a})} \right| \right)$$
$$+ \left| \sum_{\mu\in\{m,f\}} \left\langle d_\mu\phi_\mu \left[\Pi^\mu_{\mathcal{D}} S_\mu(p)(s) - \Pi^\mu_{\mathcal{D}} S_\mu(p)(t) \right], \Pi^\mu_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(M_\mu)} \right.$$
$$\left. + \sum_{\mathfrak{a}\in\chi} \left\langle \eta \left[\mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(s) - \mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(t) \right], \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(\Gamma_\mathfrak{a})} \right| \tag{32}$$

$$\lesssim \mathcal{S}_{\mathcal{D}_S}(\varphi) + \left| \sum_{\mu\in\{m,f\}} \left\langle d_\mu\phi_\mu \left[\Pi^\mu_{\mathcal{D}} S_\mu(p)(s) - \Pi^\mu_{\mathcal{D}} S_\mu(p)(t) \right], \Pi^\mu_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(M_\mu)} \right.$$
$$\left. + \sum_{\mathfrak{a}\in\chi} \left\langle \eta \left[\mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(s) - \mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(t) \right], \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(\Gamma_\mathfrak{a})} \right|. \tag{33}$$

Here, the terms (31) and (32) have been estimated by using $0 \leq S_\mu, S_\mathfrak{a} \leq 1$ and the definition of $\mathrm{P}_{\mathcal{D}_S}\varphi$. Let L_1 be the second addend in (33). Assuming that $t < s$, and hence $n(t) \leq n(s)$, write $\Pi^\mu_{\mathcal{D}} S_\mu(p)(s) - \Pi^\mu_{\mathcal{D}} S_\mu(p)(t)$ and $\mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(s) - \mathbb{T}^\mathfrak{a}_{\mathcal{D}} S_\mathfrak{a}(p)(t)$ as the sum of their jumps, and recall the definition (22) of $|\cdot|_{\mathcal{D}_S,*}$ to obtain

$$L_1 \leq \left| \sum_{k=n(t)}^{n(s)-1} \Delta t_{k+\frac{1}{2}} \left(\sum_{\mu\in\{m,f\}} \left\langle d_\mu\phi_\mu \Pi^\mu_{\mathcal{D}} \delta_t S_\mu(p)(t_{k+1}), \Pi^\mu_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(M_\mu)} \right. \right.$$
$$\left. \left. + \sum_{\mathfrak{a}\in\chi} \left\langle \eta \mathbb{T}^\mathfrak{a}_{\mathcal{D}} \delta_t S_\mathfrak{a}(p)(t_{k+1}), \mathbb{T}^\mathfrak{a}_{\mathcal{D}_S}\mathrm{P}_{\mathcal{D}_S}\varphi \right\rangle_{L^2(M_\mu)} \right) \right|$$
$$\leq \sum_{k=n(t)}^{n(s)-1} \Delta t_{k+\frac{1}{2}} \left| \left[\delta_t S_m(p)(t_{k+1}), \delta_t S_f(p)(t_{k+1}), (\delta_t S_\mathfrak{a}(p)(t_{k+1}))_{\mathfrak{a}\in\chi} \right] \right|_{\mathcal{D}_S,*} \|\mathrm{P}_{\mathcal{D}_S}\varphi\|_{\mathcal{D}_S}$$
$$\leq \|\mathrm{P}_{\mathcal{D}_S}\varphi\|_{\mathcal{D}_S} \int_0^T \mathbb{1}_{[t_{n(t)}, t_{n(s)}]}(r) \left| \left[\delta_t S_m(p)(r), \delta_t S_f(p)(r), (\delta_t S_\mathfrak{a}(p)(r))_{\mathfrak{a}\in\chi} \right] \right|_{\mathcal{D}_S,*} \mathrm{d}r.$$

Use now Lemmata 4.6 and the Cauchy–Schwarz inequality to infer

$$L_1 \lesssim \|\mathrm{P}_{\mathcal{D}_S}\varphi\|_{\mathcal{D}_S}\left(1+\sum_{\alpha=1}^{2}\|u^\alpha\|_{\mathcal{D}}^2\right)^{\frac{1}{2}}\left[(s-t)^{\frac{1}{2}}+(\Delta t)^{\frac{1}{2}}\right]. \tag{34}$$

By the triangle inequality,

$$\begin{aligned}\|\mathrm{P}_{\mathcal{D}_S}\varphi\|_{\mathcal{D}_S} &\le \mathcal{S}_{\mathcal{D}_S}(\varphi)+\|\nabla\varphi_m\|_{L^2(\Omega)^d}+\|\nabla_\tau\varphi_f\|_{L^2(\Gamma)^{d-1}}\\ &\quad+\sum_{\mathfrak{a}\in\chi}\|[\![\varphi]\!]_{\mathfrak{a}}\|_{L^2(\Gamma_{\mathfrak{a}})}=\mathcal{S}_{\mathcal{D}_S}(\varphi)+C_\varphi.\end{aligned}$$

Plugging this into (34) and the resulting inequality into (33) concludes the proof. □

Theorem 4.13 (Uniform-in-time, weak-in-space convergence) *Under the assumptions of Theorem 4.11, for all $\mu\in\{m,f\}$ and $\mathfrak{a}\in\chi$, $S_\mu(\overline{p}_\mu):[0,T]\to L^2(M_\mu)$ and $S_{\mathfrak{a}}(\gamma_{\mathfrak{a}}\overline{p}_m):[0,T]\to L^2(\Gamma_{\mathfrak{a}})$ are continuous for the weak topologies of $L^2(M_\mu)$ and $L^2(\Gamma_{\mathfrak{a}})$, respectively, and*

$$\begin{aligned}&\Pi^\mu_{\mathcal{D}^l}S_\mu(p^l)\longrightarrow S_\mu(\overline{p}_\mu)\text{ uniformly in }[0,T]\text{, weakly in }L^2(M_\mu),\\ &\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l}S_{\mathfrak{a}}(p^l)\longrightarrow S_{\mathfrak{a}}(\gamma_{\mathfrak{a}}\overline{p}_m)\text{ uniformly in }[0,T]\text{, weakly in }L^2(\Gamma_{\mathfrak{a}}),\end{aligned} \tag{35}$$

where the definition of the uniform-in-time weak L^2 convergence is recalled in "Appendix A.1".

Proof The proof hinges on the discontinuous Arzelà-Ascoli theorem (Theorem A.1 in the appendix). Consider first the matrix saturation. The space $\mathcal{R}_m=\left\{d_m\phi_m\varphi_m\mid\varphi_m\in C_0^\infty(\Omega\setminus\overline{\Gamma})\right\}$ is dense in $L^2(\Omega)$. Apply (30) to $\varphi=(\varphi_m,0)$. Since $\varphi_f=\gamma_{\mathfrak{a}}\varphi_m=0$, only the term involving S_m remains in the left-hand side. The resulting estimate and the property $0\le S_m\le 1$ show that the sequence of functions $(t\mapsto\Pi^m_{\mathcal{D}^l}S_m(p^l)(t))_{l\in\mathbb{N}}$ satisfies the assumptions of Theorem A.1 with $\mathcal{R}=\mathcal{R}_m$. Hence, $(\Pi^m_{\mathcal{D}^l}S_m(p^l))_{l\in\mathbb{N}}$ has a subsequence that converges uniformly on $[0,T]$ weakly in $L^2(\Omega)$. Given (12), the weak limit of this sequence must be $S_m(\overline{p}_m)$.

A similar reasoning, based on the space $\mathcal{R}_f=\left\{d_f\phi_f\varphi_f\mid\varphi_f\in C_\Gamma^\infty\right\}$—which is dense in $L^2(\Gamma)$—and using $\varphi=(0,\varphi_f)$ in (30), gives the uniform-in-time weak $L^2(\Gamma)$ convergence of $\Pi^f_{\mathcal{D}^l}S_f(p^l)$ towards $S_f(\overline{p}_f)$.

Let us now turn to the convergence of the trace saturations. Take $\varphi_m\in C_\Omega^\infty$ such that the support of $\gamma_{\mathfrak{a}}\varphi_m$ is non empty for exactly one $\mathfrak{a}\in\chi$. Considering $\varphi=(\varphi_m,0)$ in (30) leads to

$$\left|\left\langle\eta\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l}S_{\mathfrak{a}}(p^l)(s)-\eta\mathbb{T}^{\mathfrak{a}}_{\mathcal{D}^l}S_{\mathfrak{a}}(p^l)(t),\gamma_{\mathfrak{a}}\varphi_m\right\rangle_{L^2(\Gamma_{\mathfrak{a}})}\right|$$

$$\lesssim \mathcal{S}_{\mathcal{D}_S}(\varphi) + (\mathcal{S}_{\mathcal{D}_S}(\varphi) + C_\varphi)\left(1 + \sum_{\alpha=1}^{2} \|u^\alpha\|_{\mathcal{D}}^2\right)^{\frac{1}{2}} \left[|s-t|^{\frac{1}{2}} + (\Delta t)^{\frac{1}{2}}\right]$$
$$+ \left|\left\langle d_m\phi_m \Pi_{\mathcal{D}^l}^m S_m(p^l)(s) - d_m\phi_m \Pi_{\mathcal{D}^l}^m S_m(p^l)(t), \varphi_m\right\rangle_{L^2(\Omega)}\right|. \quad (36)$$

Since it was established that $(d_m\phi_m \Pi_{\mathcal{D}^l}^m S_m(p^l))_{l\in\mathbb{N}}$ converges uniformly-in-time weakly in $L^2(\Omega)$, the sequence $(\langle d_m\phi_m \Pi_{\mathcal{D}^l}^m S_m(p^l), \varphi_m\rangle_{L^2(\Omega)})_{l\in\mathbb{N}}$ is equi-continuous and the last term in (36) therefore tends to 0 uniformly in l as $s-t \to 0$. Hence, (36) enables the usage of Theorem A.1, by noticing that $\{\eta\gamma_{\mathfrak{a}}\varphi_m \mid \varphi_m \in C_\Omega^\infty,\ \mathrm{supp}(\gamma_{\mathfrak{b}}\varphi_m) = \emptyset \text{ for all} \mathfrak{b} \in \chi \text{ with } \mathfrak{b} \neq \mathfrak{a}\}$ is dense in $L^2(\Gamma_{\mathfrak{a}})$, and gives the uniform-in-time weak $L^2(\Gamma_{\mathfrak{a}})$ convergence of $\mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} S_{\mathfrak{a}}(p^l)$. □

4.3 Proof of Theorem 4.1

The proof of the main convergence theorem can now be given.

First step: passing to the limit in the gradient scheme.

Let us introduce the family of functions $(\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha})_{\mathfrak{a}\in\chi}^{\alpha=1,2}$:

$$\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}(t,\mathbf{x},\beta) = \left[T_f[kS]_{\mathfrak{a}}^\alpha(\mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} p^l)\beta^+ - T_f[kS]_f^\alpha(\Pi_{\mathcal{D}^l}^f p^l)\beta^-\right](t,\mathbf{x}), \quad \text{for all } \beta \in L^2(\Gamma_{\mathfrak{a}}),$$

and their continuous counterparts $(\mathrm{F}^{\mathfrak{a},\alpha})_{\mathfrak{a}\in\chi}^{\alpha=1,2}$:

$$\mathrm{F}^{\mathfrak{a},\alpha}(t,\mathbf{x},\beta) = \left[T_f[kS]_{\mathfrak{a}}^\alpha(\gamma_{\mathfrak{a}}\overline{p}_m)\beta^+ - T_f[kS]_f^\alpha(\overline{p}_f)\beta^-\right](t,\mathbf{x}), \quad \text{for all} \beta \in L^2(\Gamma_{\mathfrak{a}}).$$

The following properties are easy to check. Firstly, since T_f, $[kS]_{\mathfrak{a}}^\alpha$ and $[kS]_f$ are positive and $s \mapsto s^+$ and $s \mapsto -s^-$ are non-decreasing,

$$\left[\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}(t,\mathbf{x},\beta) - \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}(t,\mathbf{x},\gamma)\right]\left[\beta(t,\mathbf{x}) - \gamma(t,\mathbf{x})\right] \geq 0, \quad \text{for all} \beta,\gamma \in L^2(\Gamma_{\mathfrak{a}}). \quad (37)$$

Secondly, by the convergences (12), for $(\beta_l)_{l\in\mathbb{N}} \subset L^2(\Gamma_{\mathfrak{a}})$ and $\beta \in L^2(\Gamma_{\mathfrak{a}})$,

$$\beta_l \longrightarrow \beta \ \text{in} L^2((0,T)\times\Gamma_{\mathfrak{a}}) \quad \Longrightarrow \quad \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}(\beta_l) \longrightarrow \mathrm{F}^{\mathfrak{a},\alpha}(\beta) \ \text{in } L^2((0,T)\times\Gamma_{\mathfrak{a}}). \quad (38)$$

Thirdly, by Lemma 4.4, the sequences $(\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^l]\!]_{\mathfrak{a},\mathcal{D}^l}))_{l\in\mathbb{N}}$ ($\mathfrak{a} \in \chi$, $\alpha = 1,2$) are bounded in $L^2((0,T)\times\Gamma_{\mathfrak{a}})$ and there exists thus $\rho_{\mathfrak{a}}^\alpha \in L^2((0,T)\times\Gamma_{\mathfrak{a}})$ such that, up to a subsequence,

$$\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l}) \rightharpoonup \rho_{\mathfrak{a}}^\alpha \quad \text{weakly in } L^2((0,T)\times\Gamma_{\mathfrak{a}}). \quad (39)$$

Consider $\varphi^\alpha = (\varphi_m^\alpha, \varphi_f^\alpha) = \sum_{k=1}^b \theta^{\alpha,k} \otimes \psi^{\alpha,k}$, where $(\psi^{\alpha,k})_{k\in\mathbb{N}} = (\psi_m^{\alpha,k}, \psi_f^{\alpha,k})_{k=1,\ldots,b} \in C_\Omega^\infty \times C_\Gamma^\infty$ and $(\theta^{\alpha,k})_{k=1,\ldots,b} \in C_0^\infty([0,T))$. Take $(v_n^{\alpha,l})_{n=0,\ldots,N^l} =$

$(\mathrm{P}_{\mathcal{D}_S^l}\varphi^\alpha(t_n^l))_{n=0,\ldots,N^l} \in (X_l^0)^{N^l+1}$ as "test function" in (5). Here, $\mathrm{P}_{\mathcal{D}_S^l}$ is defined as in the proof of Lemma 4.12. Apply the discrete integration-by-parts of [21, Section D.1.7] on the accumulation terms in (5), let $l \to \infty$ and use standard convergence arguments [19,21] based on Theorem 4.11 to see that

$$\begin{aligned}
&\sum_{\alpha=1}^{2}\Big\{\sum_{\mu\in\{m,f\}}\Big(-\int_0^T\int_{M_\mu}\phi_\mu S_\mu^\alpha(\overline{p}_\mu)\partial_t\varphi_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t\\
&\qquad+\int_0^T\int_{M_\mu}[kS]_\mu^\alpha(\overline{p}_\mu)\ \Lambda_\mu\nabla\overline{u}_\mu^\alpha\cdot\nabla\varphi_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t\\
&\qquad-\int_{M_\mu}\phi_\mu S_\mu^\alpha(\overline{p}_{\mu,0})\varphi_\mu^\alpha(0,\cdot)\mathrm{d}\tau_\mu\Big)\\
&\qquad+\sum_{\mathfrak{a}\in\chi}\Big(\int_0^T\int_{\Gamma_\mathfrak{a}}\rho_\mathfrak{a}^\alpha[\![\varphi^\alpha]\!]_\mathfrak{a}\mathrm{d}\tau \mathrm{d}t-\int_0^T\int_{\Gamma_\mathfrak{a}}\eta S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_m)\partial_t\gamma_\mathfrak{a}\varphi_m^\alpha \mathrm{d}\tau \mathrm{d}t\\
&\qquad-\int_{\Gamma_\mathfrak{a}}\eta S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_{m,0})\gamma_\mathfrak{a}\varphi_m^\alpha(0,\cdot)\mathrm{d}\tau\Big)\Big\}\\
&\qquad=\sum_{\alpha=1}^{2}\sum_{\mu\in\{m,f\}}\int_0^T\int_{M_\mu}h_\mu^\alpha\varphi_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t.
\end{aligned}\tag{40}$$

Note that Equation (40) also holds for any smooth φ^α, by density of tensorial functions in smooth functions [17, Appendix D]. Recalling the weak formulation (2), proving Theorem 4.1 is now all about showing that

$$\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_\mathfrak{a}}\rho_\mathfrak{a}^\alpha[\![\varphi^\alpha]\!]_\mathfrak{a}\mathrm{d}\tau \mathrm{d}t=\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_\mathfrak{a}}\mathrm{F}^{\mathfrak{a},\alpha}([\![\overline{u}^\alpha]\!]_\mathfrak{a})[\![\varphi^\alpha]\!]_\mathfrak{a}\mathrm{d}\tau \mathrm{d}t. \tag{41}$$

This is achieved by using Minty's trick.
Second step: proof that

$$\limsup_{l\to\infty}\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_\mathfrak{a}}\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l})[\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l}\mathrm{d}\tau \mathrm{d}t\le\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_\mathfrak{a}}\rho_\mathfrak{a}^\alpha[\![\overline{u}^\alpha]\!]_\mathfrak{a}\mathrm{d}\tau \mathrm{d}t. \tag{42}$$

Having in mind to employ the energy inequality (16) with $T_0 = T$, we first establish, for $\mu \in \{m, f\}$ and $\mathfrak{a} \in \chi$, the following convergences as $l \to \infty$:

$$\int_0^T\int_{M_\mu}h_\mu^\alpha\Pi_{\mathcal{D}^l}^\mu u^{\alpha,l}\mathrm{d}\tau_\mu \mathrm{d}t\longrightarrow\int_0^T\int_{M_\mu}h_\mu^\alpha\overline{u}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t\,, \tag{43}$$

$$\int_{M_\mu}\phi_\mu B_\mu(S_\mu(\Pi_{\mathcal{D}_S^l}^\mu p_0^l))\mathrm{d}\tau_\mu\longrightarrow\int_{M_\mu}\phi_\mu B_\mu(S_\mu(\overline{p}_{\mu,0}))\mathrm{d}\tau_\mu\,, \tag{44}$$

$$\int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}^\mathfrak{a}_{\mathcal{D}^l_S} p^l_0))\mathrm{d}\tau \longrightarrow \int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_{m,0}))\mathrm{d}\tau. \tag{45}$$

The convergence (43) is obvious by Theorem 4.11. From the choice (4) of the scheme's initial conditions, together with the consistency of the interpolation operator $\mathrm{I}_\mathcal{D}$, $\Pi^\mu_{\mathcal{D}^l_S} p^l_0 \to \overline{p}_{\mu,0}$ in $L^2(M_\mu)$ and $\mathbb{T}^\mathfrak{a}_{\mathcal{D}^l_S} p^l_0 \to \gamma_\mathfrak{a}\overline{p}_{m,0}$ in $L^2(\Gamma_\mathfrak{a})$, as $l \to \infty$. Then, (15) and [27, Lemma A.1] yield (44) and (45).

We further show that

$$\liminf_{l\to\infty} \int_{M_\mu} \phi_\mu B_\mu(S_\mu(\Pi^\mu_{\mathcal{D}^l_S} p^l_{N^l}))\mathrm{d}\tau_\mu \geq \int_{M_\mu} \phi_\mu B_\mu(S_\mu(\overline{p}_\mu)(T))\mathrm{d}\tau_\mu\,, \tag{46}$$

$$\liminf_{l\to\infty} \int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\mathbb{T}^\mathfrak{a}_{\mathcal{D}^l_S} p^l_{N^l}))\mathrm{d}\tau \geq \int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_m)(T))\mathrm{d}\tau\,, \tag{47}$$

$$\begin{aligned}\liminf_{l\to\infty} &\int_0^T \int_{M_\mu} [kS]^\alpha_\mu(\Pi^\mu_{\mathcal{D}^l} p^l)\Lambda_\mu \nabla^\mu_{\mathcal{D}^l} u^{\alpha,l} \cdot \nabla^\mu_{\mathcal{D}^l} u^{\alpha,l} \mathrm{d}\tau_\mu \mathrm{d}t \\ &\geq \int_0^T \int_{M_\mu} [kS]^\alpha_\mu(\overline{p}_\mu)\Lambda_\mu \nabla\overline{u}^\alpha_\mu \cdot \nabla\overline{u}^\alpha_\mu \mathrm{d}\tau_\mu \mathrm{d}t.\end{aligned} \tag{48}$$

By the uniform-in-time weak L^2 convergences of Theorem 4.13, $S_\mu(\Pi^\mu_{\mathcal{D}^l_S} p^l_{N^l}) \rightharpoonup S_\mu(\overline{p}_\mu)(T)$ in $L^2(M_\mu)$ and $S_\mathfrak{a}(\mathbb{T}^\mathfrak{a}_{\mathcal{D}^l_S} p^{N^l}_n) \rightharpoonup S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_m)(T)$ in $L^2(\Gamma_\mathfrak{a})$, as $l \to \infty$. Note also that, since (by assumption) S_μ and $S_\mathfrak{a}$ are not explicitly space-dependent on each open set of the formerly introduced partitions of M_μ and $\Gamma_\mathfrak{a}$, respectively, so are B_μ and $B_\mathfrak{a}$. On these partitions, B_μ and $B_\mathfrak{a}$ are convex l.s.c. and an easy adaptation of [25, Lemma 4.6] (which essentially states the L^2-weak l.s.c. of strongly l.s.c. convex functions on L^2), to account for the terms ϕ_μ and η, thus shows that (46) and (47) hold. To prove (48), apply the Cauchy-Schwarz inequality to write

$$\begin{aligned}&\int_0^T \int_{M_\mu} [kS]^\alpha_\mu(\Pi^\mu_{\mathcal{D}^l} p^l)\Lambda_\mu \nabla\overline{u}^\alpha_\mu \cdot \nabla^\mu_{\mathcal{D}^l} u^{\alpha,l} \mathrm{d}\tau_\mu \mathrm{d}t \\ &\quad\leq \left(\int_0^T \int_{M_\mu} [kS]^\alpha_\mu(\Pi^\mu_{\mathcal{D}^l} p^l)\Lambda_\mu \nabla\overline{u}^\alpha_\mu \cdot \nabla\overline{u}^\alpha_\mu \mathrm{d}\tau_\mu \mathrm{d}t\right)^{\frac{1}{2}} \\ &\quad\times \left(\int_0^T \int_{M_\mu} [kS]^\alpha_\mu(\Pi^\mu_{\mathcal{D}^l} p^l)\Lambda_\mu \nabla^\mu_{\mathcal{D}^l} u^{\alpha,l} \cdot \nabla^\mu_{\mathcal{D}^l} u^{\alpha,l} \mathrm{d}\tau_\mu \mathrm{d}t\right)^{\frac{1}{2}}\end{aligned}$$

and take the inferior limit as $l \to \infty$, using the strong convergence of $[kS]^\alpha_\mu(\Pi^\mu_{\mathcal{D}^l} p^l)$ and weak convergence of $\nabla^\mu_{\mathcal{D}^l} u^{\alpha,l}$ to pass to the limit in the left-hand side and the first term in the right-hand side.

Let us now come back to the proof of (42). Plugging the convergences (43)–(48) into (16) with $T_0 = T$ yields

$$\begin{aligned}
&\limsup_{l\to\infty} \sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_\mathfrak{a}} \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l})[\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}^l} \mathrm{d}\tau \mathrm{d}t \\
&\quad \le \sum_{\mu,\alpha} \Big(\int_0^T \int_{M_\mu} h_\mu^\alpha \overline{u}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t - \int_0^T \int_{M_\mu} [kS]_\mu^\alpha(\overline{p}_\mu)\ \Lambda_\mu \nabla \overline{u}_\mu^\alpha \cdot \nabla \overline{u}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \Big) \\
&\qquad + \sum_\mu \Big(\int_{M_\mu} \phi_\mu B_\mu(S_\mu(\overline{p}_{\mu,0}))\mathrm{d}\tau_\mu - \int_{M_\mu} \phi_\mu B_\mu(S_\mu(\overline{p}_\mu)(T))\mathrm{d}\tau_\mu \Big) \\
&\qquad + \sum_\mathfrak{a} \Big(\int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_{m,0}))\mathrm{d}\tau - \int_{\Gamma_\mathfrak{a}} \eta B_\mathfrak{a}(S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_m)(T))\mathrm{d}\tau \Big). \qquad (49)
\end{aligned}$$

Recall that $C_0^\infty([0,T)) \otimes [C_\Omega^\infty \times C_\Gamma^\infty]$ is dense in $(L^2((0,T) \times M_\mu))_{\mu\in\{m,f\}}$. Owing to "Appendix A.3", we infer from (40) that $\phi_f \partial_t S_f^\alpha(\overline{p}_f) \in L^2(0,T; V_f^{0\prime})$, that $\phi_m \partial_t S_m^\alpha(\overline{p}_m) + \sum_\mathfrak{a} \gamma_\mathfrak{a}^*(\eta \partial_t S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_m)) \in L^2(0,T; V_m^{0\prime})$ (where $\gamma_\mathfrak{a}^*$ is the adjoint of $\gamma_\mathfrak{a}$), and that, for any $\varphi^\alpha \in V$,

$$\begin{aligned}
&\sum_{\alpha=1}^2 \Big\{ \sum_{\mu\in\{m,f\}} \Big(\int_0^T \langle \phi_\mu \partial_t S_\mu^\alpha(\overline{p}_\mu), \varphi_\mu^\alpha \rangle \mathrm{d}t + \int_0^T \int_{M_\mu} [kS]_\mu^\alpha(\overline{p}_\mu)\ \Lambda_\mu \nabla \overline{u}_\mu^\alpha \cdot \nabla \varphi_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t \Big) \\
&\quad + \sum_{\mathfrak{a}\in\chi} \Big(\int_0^T \int_{\Gamma_\mathfrak{a}} \rho_\mathfrak{a}^\alpha [\![\varphi^\alpha]\!]_\mathfrak{a} \mathrm{d}\tau \mathrm{d}t + \int_0^T \langle \eta \partial_t S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a}\overline{p}_m), \gamma_\mathfrak{a}\varphi_m^\alpha \rangle \mathrm{d}t \Big) \Big\} \\
&\quad = \sum_{\alpha=1}^2 \sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} h_\mu^\alpha \varphi_\mu^\alpha \mathrm{d}\tau_\mu.
\end{aligned}$$

Note that the duality product between $(V_f^0)'$ and V_f^0 is taken respective to the measure $\mathrm{d}\tau_f(\mathbf{x}) = d_f(\mathbf{x})\mathrm{d}\tau(\mathbf{x})$, and remember the abuse of notation (57). Apply this to $\varphi^\alpha = (\overline{u}_m^\alpha, \overline{u}_f^\alpha)$. Recalling that $S_\mu^2 = 1 - S_\mu^1$, we have $\partial_t S_\mu^2(\overline{p}_\mu) = -\partial_t S_\mu^1(\overline{p}_\mu)$ and thus

$$\begin{aligned}
&\sum_{\mu\in\{m,f\}} \int_0^T \langle \phi_\mu \partial_t S_\mu(\overline{p}_\mu), \overline{p}_\mu \rangle \mathrm{d}t + \sum_{\mathfrak{a}\in\chi} \int_0^T \langle \eta \partial_t S_\mathfrak{a}(\gamma_\mathfrak{a}\overline{p}_m), \gamma_\mathfrak{a}\overline{p}_m \rangle \mathrm{d}t \\
&\quad + \sum_{\alpha=1}^2 \Big\{ \sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} [kS]_\mu^\alpha(\overline{p}_\mu)\ \Lambda_\mu \nabla \overline{u}_\mu^\alpha \cdot \nabla \overline{u}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t + \sum_{\mathfrak{a}\in\chi} \int_0^T \int_{\Gamma_\mathfrak{a}} \rho_\mathfrak{a}^\alpha [\![\overline{u}^\alpha]\!]_\mathfrak{a} \mathrm{d}\tau \mathrm{d}t \Big\} \\
&\quad = \sum_{\alpha=1}^2 \sum_{\mu\in\{m,f\}} \int_0^T \int_{M_\mu} h_\mu^\alpha \overline{u}_\mu^\alpha \mathrm{d}\tau_\mu \mathrm{d}t. \qquad (50)
\end{aligned}$$

[19, Lemma 3.6] establishes a temporal integration-by-parts property by using arguments purely based on the time variable, and that can easily be adapted to our context, even considering the "combined" time derivatives $\phi_m \partial_t S_m^\alpha(\overline{p}_m) +$

$\sum_{\mathfrak{a}} \gamma_{\mathfrak{a}}^*(\eta \partial_t S_{\mathfrak{a}}^\alpha(\gamma_{\mathfrak{a}} \overline{p}_m))$ and the heterogeneities of the media treated here—i.e. the presence of ϕ_μ, see assumptions in Sect. 2.2. This adaptation yields

$$\int_0^T \langle \phi_f \partial_t S_f^\alpha(\overline{p}_\mu), \overline{p}_f \rangle_{V_f^{0\prime}, V_f^0} \mathrm{d}t = \int_{M_f} \phi_f B_f(S_f(\overline{p}_f)(T)) \mathrm{d}\tau_f \\ - \int_{M_f} \phi_f B_f(S_\mu(\overline{p}_f)(0)) \mathrm{d}\tau_f$$

and

$$\int_0^T \langle \phi_m \partial_t S_m^\alpha(\overline{p}_m), \overline{p}_m \rangle \mathrm{d}t + \sum_{\mathfrak{a} \in \chi} \int_0^T \langle \eta \partial_t S_{\mathfrak{a}}^\alpha(\gamma_{\mathfrak{a}} \overline{p}_m), \gamma_{\mathfrak{a}} \overline{p}_m \rangle \mathrm{d}t \\ = \int_{M_m} \phi_m B_m(S_m(\overline{p}_m)(T)) \mathrm{d}\mathbf{x} - \int_{M_m} \phi_m B_m(S_m(\overline{p}_m)(0)) \mathrm{d}\mathbf{x} \\ + \sum_{\mathfrak{a} \in \chi} \Big(\int_{\Gamma_{\mathfrak{a}}} \eta B_{\mathfrak{a}}(S_{\mathfrak{a}}(\gamma_{\mathfrak{a}} \overline{p}_m)(T)) \mathrm{d}\tau \\ - \int_{\Gamma_{\mathfrak{a}}} \eta B_{\mathfrak{a}}(S_{\mathfrak{a}}(\gamma_{\mathfrak{a}} \overline{p}_m)(0)) \mathrm{d}\tau \Big).$$

Plugging these relations into (50) and using (49) concludes the proof of (42).

Third step: conclusion.

As in the first step, take $\varphi^\alpha = (\varphi_m^\alpha, \varphi_f^\alpha) = \sum_{k=1}^b \theta^{\alpha,k} \otimes \psi^{\alpha,k}$ and set $(v_n^{\alpha,l})_{n=0,\ldots,N^l} = (\mathrm{P}_{\mathcal{D}_S^l} \varphi^\alpha(t_n^l))_{n=0,\ldots,N^l} \in (X_l^0)^{N^l+1}$. Developing the monotonicity property (37) of $\mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}$, integrating over $(0,T) \times \Gamma_{\mathfrak{a}}$ and summing over $\mathfrak{a}, \alpha$ yields

$$\sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}}) [\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau \mathrm{d}t \\ - \sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![v^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}}) ([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}} - [\![v^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}}) \mathrm{d}\tau \mathrm{d}t \\ - \sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}}) [\![v^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau \mathrm{d}t \geq 0.$$

Use (38), (39) and (11) to pass to the limit in the second and third integral terms:

$$\limsup_{l \to \infty} \sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \mathrm{F}_{\mathcal{D}^l}^{\mathfrak{a},\alpha}([\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}}) [\![u^{\alpha,l}]\!]_{\mathfrak{a},\mathcal{D}} \mathrm{d}\tau \mathrm{d}t \\ \geq \sum_{\mathfrak{a},\alpha} \int_0^T \int_{\Gamma_{\mathfrak{a}}} \mathrm{F}^{\mathfrak{a},\alpha}([\![\varphi^\alpha]\!]_{\mathfrak{a}}) ([\![\overline{u}^\alpha]\!]_{\mathfrak{a}} - [\![\varphi^\alpha]\!]_{\mathfrak{a}}) \mathrm{d}\tau \mathrm{d}t$$

$$+\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_{\mathfrak{a}}}\rho^\alpha_{\mathfrak{a}}[\![\varphi^\alpha]\!]_{\mathfrak{a}}\mathrm{d}\tau\mathrm{d}t.$$

Use (42) and the density of the tensorial function spaces $C_0^\infty([0,T))\otimes[C_\Omega^\infty\times C_\Gamma^\infty]$ in $L^2(0,T;V)$ (cf. [11, proposition 2.3]) to obtain

$$\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_{\mathfrak{a}}}\rho^\alpha_{\mathfrak{a}}([\![\overline{u}^\alpha]\!]_{\mathfrak{a}}-[\![\overline{v}^\alpha]\!]_{\mathfrak{a}})\mathrm{d}\tau\mathrm{d}t\geq\sum_{\mathfrak{a},\alpha}\int_0^T\int_{\Gamma_{\mathfrak{a}}}\mathrm{F}^{\mathfrak{a},\alpha}([\![\overline{v}^\alpha]\!]_{\mathfrak{a}})([\![\overline{u}^\alpha]\!]_{\mathfrak{a}}-[\![\overline{v}^\alpha]\!]_{\mathfrak{a}})\mathrm{d}\tau\mathrm{d}t$$

for all $(\overline{v}^\alpha)_{\alpha=1,2}\in L^2(0,T;V)^2$. The conclusion is now standard in the Minty trick (see e.g. [21, Proof of Theorem 3.34]): for any smooth $(\varphi^\alpha)_{\alpha=1,2}$, choose $\overline{v}^\alpha=\overline{u}^\alpha\pm\epsilon\varphi^\alpha$ and let $\epsilon\to 0$ to derive (41) and conclude the proof. □

5 Two-phase flow test cases

We present in this section a series of test cases for two-phase flow through a fractured 2 dimensional reservoir of geometry as shown in Fig. 3. The domain Ω is of extension $(0,10)$m $\times$ $(0,20)$m and the fracture width d_f is assumed constant equal to 1 cm. We consider isotropic permeability in the matrix and in the fracture. The following geological configuration is considered: the matrix and fracture permeabilities are $\lambda_m=0.1$ Darcy and $\lambda_f=100$ Darcy, respectively; the matrix and fracture porosities are $\phi_m=0.2$ and $\phi_f=0.4$, respectively.

Initially, the reservoir is saturated with water (density $\rho^2=1000\,\mathrm{kg/m^3}$, viscosity $\kappa^2=0.001\,\mathrm{Pa.s}$) and oil (density $\rho^1=700\,\mathrm{kg/m^3}$, viscosity $\kappa^1=0.005\,\mathrm{Pa.s}$) is injected from below. Also, hydrostatic distribution of pressure is assumed. The oil then rises by gravity, thanks to its lower density compared to water. At the lower boundary of the domain, we impose constant capillary pressure of 0.1 bar and water pressure of 3 bar; at the upper boundary, the capillary pressure is constant equal to 0 bar and the water pressure is 1 bar. Elsewhere, homogeneous Neumann conditions are imposed.

We use the VAG scheme to obtain solutions for the DFM. We refer to [11] for a presentation of the scheme as a gradient scheme, and for proofs that, under standard regularity assumptions on the meshes, the corresponding sequences of gradient discretisations are coercive, GD-consistent, limit-conforming and compact. The tests are driven on a triangular mesh extended to a 3D mesh with one layer of prisms (we use a 3D implementation of the VAG scheme). The resulting numbers of cells and degrees of freedom are exhibited in Table 1. The mesh size is of order $10d_f$.

The non-linear system of equations occurring at each time step is solved via a Newton algorithm with relaxation. To solve the linear system obtained at each step of the Newton iteration, we use the sequential version of the SuperLU direct sparse solver [15,16]. The stopping criterion on the L^1 relative residual is $\mathrm{crit}^{\mathrm{rel}}_{\mathrm{Newton}}$. To ensure well defined values for the capillary pressure, after each Newton iteration, we project the (oil) saturation on the interval $[0,1-10^{-14}]$. The time stepping is progressive, i.e. after each iteration, the upcoming time step is deduced by multiplying the previous

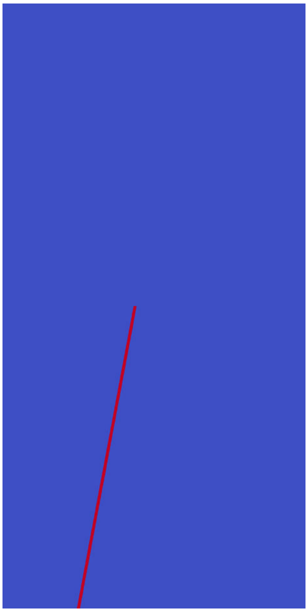

Fig. 3 Geometry of the reservoir under consideration. Fracture in red and matrix domain in blue. $\Omega = (0, 10)\text{m} \times (0, 20)\text{m}$ and $d_f = 0.01\text{m}$

Table 1 **Nb Cells** is the number of cells of the mesh; **Nb DOF** is the number of discrete unknowns; **Nb DOF el.** is the number of discrete unknowns after elimination of cell unknowns without fill-in. Time steps used in the simulations in days (d)

Nb Cells	Nb DOF	Nb DOF el.	$\text{crit}^{\text{rel}}_{\text{Newton}}$	Δt_{max} for $0 \le t \le 1/2$ d	Δt_{max} for $1/2$ d $< t \le 10$ d
5082	10610	5528	$1.E{-}6$	0.01 d	0.19 d

one by 2, while imposing a maximal time step Δt_{max}. If at a given time iteration the Newton algorithm does not converge after 35 iterations, then the actual time step is divided by 4 and the time iteration is repeated. The number of time step failures at the end of a simulation is indicated by $\mathbf{N}_{\text{Chop}}$.

Inside the matrix domain the capillary pressure function is given by Corey's law $p_m = -a_m \log(1-S_m)$ with $a_m = 1$ bar. Inside the fracture network, we suppose $p_f = -a_f \log(1 - S_f)$ with $a_f = 0.02$ bar. The matrix and fracture relative permeabilities of each phase α are given by Corey's laws $k^\alpha_{r,m}(S^\alpha_m) = (S^\alpha_m)^2$ and $k^\alpha_{r,f}(S^\alpha_f) = S^\alpha_f$, and the phase mobilities are defined by $k^\alpha_\mu(S^\alpha_\mu) = \frac{1}{\kappa^\alpha} k^\alpha_{r,m}(S^\alpha_\mu)$, $\mu \in \{m, f\}$ (see Fig. 4). The phase saturations at the interfacial layers are defined by the interpolation

$$S^\alpha_{\mathfrak{a}} = \theta S^\alpha_m + (1-\theta) S^\alpha_f, \tag{51}$$

with parameter $\theta \in [0, 1]$. The mapping $S^\alpha_{\mathfrak{a}} : [0, +\infty) \to [0, 1)$ is a diffeomorphism so the choice

$$[kS]^\alpha_{\mathfrak{a}} = \theta k^\alpha_m(S^\alpha_m) + (1-\theta) k^\alpha_f(S^\alpha_f).$$

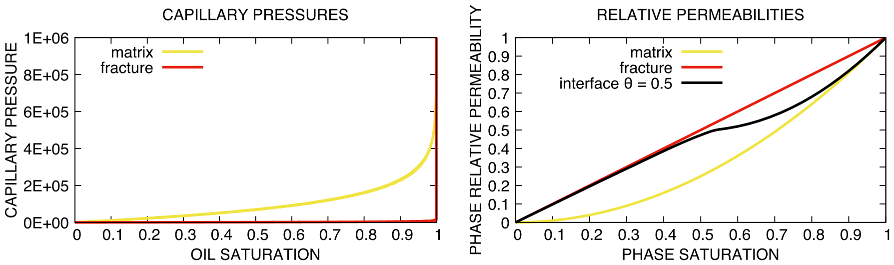

Fig. 4 Curves for capillary pressures and relative permeabilities

is valid, since this function can be written as $k_{\mathfrak{a}}^{\alpha}(S_{\mathfrak{a}}^{\alpha})$ with $k_{\mathfrak{a}}^{\alpha}(\xi) = \theta k_m^{\alpha}(S_m^{\alpha} \circ (S_{\mathfrak{a}}^{\alpha})^{-1}(\xi)) + (1-\theta)k_f^{\alpha}(S_f^{\alpha} \circ (S_{\mathfrak{a}}^{\alpha})^{-1}(\xi))$. Finally, the interfacial porosity $\phi_{\mathfrak{a}}$ is set to 0.2 and

$$d_{\mathfrak{a}} = \frac{d_f}{2}\varepsilon,$$

with parameter $\varepsilon > 0$. The parameter η is then defined by $\eta = \phi_{\mathfrak{a}} d_{\mathfrak{a}}$.

Let us start with some remarks. From the capillary pressure functions (cf. Fig. 4), it is obvious that for given p, the one-sided jump of the oil saturation is negative, i.e.

$$S_m(p) - S_f(p) < 0. \tag{52}$$

To account for the interfacial zone properly, the mobilities have to be adjusted by choosing the model parameter θ depending on the rock type characteristics of the layer. Obviously, $\theta = 0$ refers to a fracture rock type and $\theta = 1$ to a matrix rock type.

On the other hand, with larger η, the volume of the interfacial layers gets augmented and the interfacial accumulation terms play a more important role. The availability of the supplementary volume has a direct impact on the phase front speed inside the fracture during its filling: (51)–(52) show that the volume of oil in the interfacial layers is strictly decreasing as a function of θ, given a distribution of capillary pressures. This indicates that, from the accumulation point of view, the fracture front speed should grow with growing θ, and this effect should be enhanced by a larger η.

Figure 5a indicates that, for a fixed $\theta = 0, 0.5, 1$, the solutions are not sensitive to small variations of ε. Quantitatively, we see that the solution for $\varepsilon = 0.1$ is close to the solution for $\varepsilon = 10^{-6}$. With respect to the computational performance exposed in Table 2, we thus see that choosing $\varepsilon = 0.1$ is a good compromise between accuracy and cost. This point is presented in more detail for the intermediate rock type, i.e. $\theta = 0.5$, in Fig. 6. Figure 5b confirms the aforementioned feature of extended (large ε) interfacial layers to delay the propagation of the oil in the drain. As suggested, this effect is even more important, with decreasing θ. In Fig. 5c, we study the impact of the choice of the interfacial mobility for parameters $\theta = 0, 0.5, 1$ on the solution. Here, the interfacial accumulation is negligible due to an ε close to zero. Let us remark

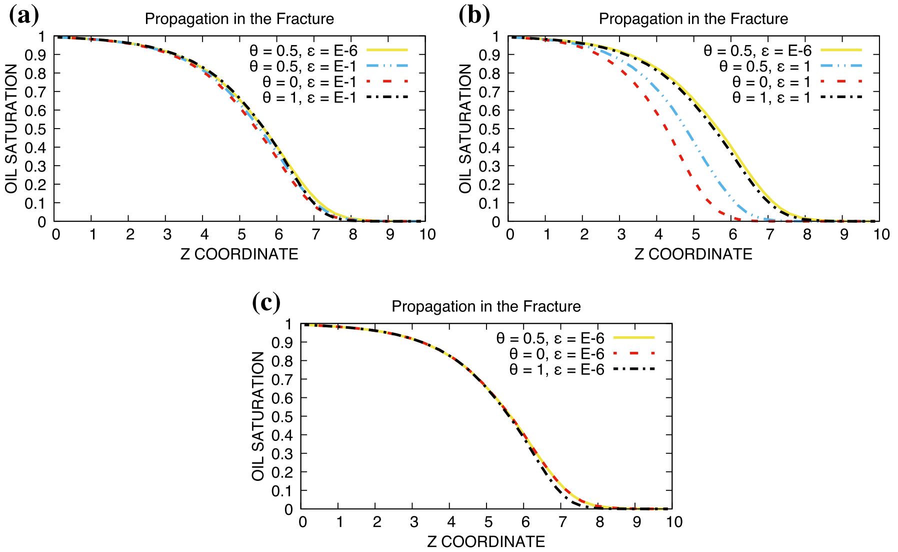

Fig. 5 Fracture oil saturation for time $t = 6$h

Table 2 Computational cost

θ	0				0.5				1			
ε	1	1.E−1	1.E−6	0	1	1.E−1	1.E−6	0	1	1.E−1	1.E−6	0
$\mathbf{N}_{\Delta t}$	125	125	125	–	125	125	125	–	183	284	377	–
$\mathbf{N}_{Newton}$	506	521	547		513	521	546		674	892	1410	
$\mathbf{N}_{\text{Chop}}$	0	0	0		0	0	0		22	61	94	
CPU	147	160	159		151	152	170		402	860	1402	

that in the limit of a vanishing interfacial layer, i.e. $\eta = 0$, we aim at recovering the fracture mobilities for the mass exchange fluxes between the matrix-fracture interface and the fracture. Hence, in this case, the right choice of θ would be 0. We observe that changing the mobilities does not much influence the solution, due to the fact that fluxes are mostly oriented from the fracture towards the interfacial layers. The regions where a difference is observed in the fracture oil front for the different models are those with a small positive oil saturation. There, the relative permeabilities for $\theta = 0$ and $\theta = 0.5$ are very close and the difference to $\theta = 1$ is at its peak; this explains the behaviour of the fracture front for the three models.

Table 2 shows that the computational cost increases with decreasing ε and that, in the case of $\varepsilon = 0$, the Jacobian becomes singular. Furthermore, the efficiency severely deteriorates for $\theta = 1$. In this case, $S'_{\mathfrak{a}}(p)$ is (significantly) smaller during the filling of the fracture (for capillary pressures p below a characteristic $p_1 \in \mathbb{R}^+$), since $S'_m(p) \ll S'_f(p)$. When oil fluxes oriented from the fracture to the interface are present, the Jacobian is thus ill-conditioned.

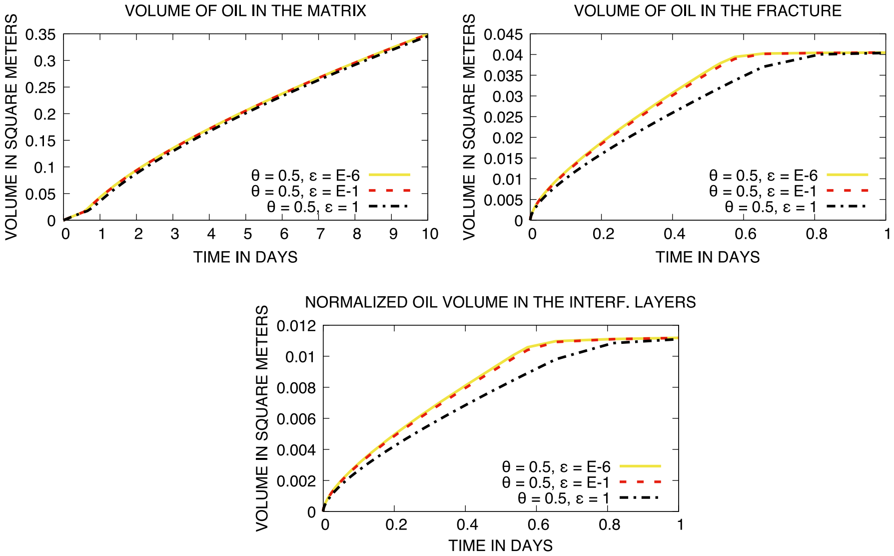

Fig. 6 Volume occupied by oil in the matrix, fracture and oil volume normalised by ε in the interfacial layers, for $\theta = 0.5$, as a function of time

6 Conclusion

We introduced a new discrete fracture matrix model for two phase Darcy flow, permitting pressure discontinuity at the matrix-fracture interfaces. It respects the heterogeneities of the media and between the matrix and the fractures, since it takes into account saturation jumps due to different capillary pressure curves in the respective domains. It also considers damaged layers located at the matrix-fracture interfaces. Another feature of the model are upwind fluxes between these interfacial layers and the fractures. The upwinding is needed for transport dominated flow in normal direction to the fractures. The extension to gravity is straightforward (cf. [12]).

We developed the numerical analysis of the model in the framework of the gradient discretisation method, which contains for example the VAG and HMM schemes. Based on compactness arguments, we showed in Theorem 4.1 the strong L^2 convergence of the saturations and the weak L^2 and H^1 convergences for the pressures to a solution of Model (1). In Theorem 4.13, we established uniform-in-time, weak L^2 in space convergence for the saturations, a result that is extended to uniform-in-time, strong L^2 in space convergence in [26].

Finally, we presented a series of test cases, with the objective to study the impact of the interfacial layer on the solution. The observed behaviour of the solutions for the different situations corresponds to the expectations. It exhibits significant differences, during the filling of the fracture, for large interfacial layers and small differences for small layers. In terms of computational cost, we saw that the presence of a damaged zone at the matrix-fracture interface is needed in order to solve the linear system of the discrete problem, occurring at each time step. We also observed that for a large contrast between the drain's and the interfacial layer's capillary pressures, the simulation becomes expensive. Therefore, we see that, in order to cope with both, fractures acting as drains or as barriers, the possibility to deal with mixed rock types for the damaged zone is essential.

Acknowledgements The authors would like to thank Total S.A. and the Australian Research Council's Discovery Projects funding scheme (Project Number DP170100605) for supporting this work.

A Appendix

A.1 Uniform-in-time weak L^2 convergence

Let A be a subset of $\mathbb{R}^n$, endowed with the standard Lebesgue measure, and $\{\varphi_\ell : \ell \in \mathbb{N}\}$ be a dense countable set in $L^2(A)$. On any bounded ball of $L^2(A)$, the weak topology can be defined by the following distance:

$$\mathrm{dist}(v, w) = \sum_{\ell \in \mathbb{N}} \frac{\min(1, |\langle v - w, \varphi_\ell \rangle_{L^2(A)}|)}{2^\ell}.$$

A sequence $(v_m)_{m\in\mathbb{N}}$ of bounded functions $[0, T] \to L^2(A)$ converges uniformly on $[0, T]$ weakly in $L^2(A)$ to some v if it converges uniformly for the weak topology of $L^2(A)$, meaning that, for all $\phi \in L^2(A)$, $\langle v_m(\cdot), \phi\rangle_{L^2(A)} \to \langle v(\cdot), \phi\rangle_{L^2(A)}$ uniformly on $[0, T]$ as $m \to \infty$.

With this introductory material, the following result is a consequence of [21, Theorem 4.26] or [19, Theorem 6.2] (see also the reasoning at the end of [19, Proof of Theorem 3.1]).

Theorem A.1 (Discontinuous weak L^2 Ascoli–Arzela theorem) *Let $\mathcal{R}$ be a dense subset of $L^2(A)$ and $(v_m)_{m\in\mathbb{N}}$ be a sequence of functions $[0, T] \to L^2(A)$ such that*

- $\sup_{m\in\mathbb{N}} \sup_{t\in[0,T]} \|v_m(t)\|_{L^2(A)} < +\infty$,
- *for all $\varphi \in \mathcal{R}$, there exist $\omega_\varphi : [0, T]^2 \to [0, \infty)$ and $(\delta_m(\varphi))_{m\in\mathbb{N}} \subset [0, \infty)$ satisfying*

$$\omega_\varphi(s, t) \to 0 \text{ as } s - t \to 0\,,\ \delta_m(\varphi) \to 0 \text{ as } m \to \infty\,, \text{ and}$$
$$\forall (s, t) \in [0, T]^2\,,\ \forall m \in \mathbb{N}\,,\ |\langle v_m(s) - v_m(t), \varphi\rangle_{L^2(A)}| \le \delta_m(\varphi) + \omega_\varphi(s, t).$$

Then, there exists a function $v : [0, T] \to L^2(A)$ such that, up to a subsequence as $m \to \infty$, $v_m \to v$ uniformly on $[0, T]$ weakly in $L^2(A)$. Moreover, v is continuous on $[0, T]$ for the weak topology of $L^2(A)$.

A.2 Generic results on gradient discretisations

The following lemma is a classical result in the context of the standard gradient discretisation method, see e.g. [21, Lemma 4.7]. We give a sketch of its proof for gradient discretisations adapted to discrete fracture matrix model.

Lemma A.2 *(Regularity of the limit) Let $(\mathcal{D}^l)_{l\in\mathbb{N}}$ be a coercive and limit-conforming sequence of gradient discretisations, and let $(v^l)_{l\in\mathbb{N}}$ be such that $v^l \in (X_l^0)^{N_l+1}$, where N_l is the number of time steps of $\mathcal{D}^l$. We assume that $(\|v^l\|_{\mathcal{D}^l})_{l\in\mathbb{N}}$ is bounded. Then, there exists $\overline{v} = (\overline{v}_m, \overline{v}_f) \in L^2(0,T;V_m^0) \times L^2(0,T;V_f^0)$ such that, up to a subsequence, the following weak convergences hold:*

$$\begin{cases} \Pi_{\mathcal{D}^l}^{\mu} v^l \rightharpoonup \overline{v}_\mu & \text{in } L^2((0,T)\times M_\mu)\,,\ \text{for } \mu\in\{m,f\},\\ \nabla_{\mathcal{D}^l}^{\mu} v^l \rightharpoonup \nabla\overline{v}_\mu & \text{in } L^2((0,T)\times M_\mu)^d\,,\ \text{for } \mu\in\{m,f\},\\ \mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} v^l \rightharpoonup \gamma_{\mathfrak{a}}\overline{v}_m & \text{in } L^2((0,T)\times\Gamma_{\mathfrak{a}}),\ \text{for all } \mathfrak{a}\in\chi,\\ [\![v^l]\!]_{\mathfrak{a},\mathcal{D}^l} \rightharpoonup [\![\overline{v}]\!]_{\mathfrak{a}} & \text{in } L^2((0,T)\times\Gamma_{\mathfrak{a}}),\ \text{for all } \mathfrak{a}\in\chi. \end{cases} \tag{53}$$

Proof By coercivity and since $(\|v^l\|_{\mathcal{D}^l})_{l\in\mathbb{N}}$ is bounded, all the sequences in (53) are bounded in their respective spaces. Up to a subsequence, we can therefore assume that there exists $\overline{v}_\mu \in L^2((0,T)\times M_\mu)$, $\boldsymbol{\xi}_\mu \in L^2((0,T)\times M_\mu)^d$, $\beta_{\mathfrak{a}} \in L^2((0,T)\times\Gamma_{\mathfrak{a}})$ and $j_{\mathfrak{a}} \in L^2((0,T)\times\Gamma_{\mathfrak{a}})$ such that $\Pi_{\mathcal{D}^l}^{\mu} v^l \rightharpoonup \overline{v}_\mu$, $\nabla_{\mathcal{D}^l}^{\mu} v^l \rightharpoonup \boldsymbol{\xi}_\mu$, $\mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} v^l \rightharpoonup \beta_{\mathfrak{a}}$ and $[\![v^l]\!]_{\mathfrak{a},\mathcal{D}^l} \rightharpoonup j_{\mathfrak{a}}$ weakly in their respective L^2 spaces as $l\to\infty$.

Take $\mathbf{q} \in \mathbf{C}_\Omega^\infty \times \mathbf{C}_\Gamma^\infty$, $\varphi_{\mathfrak{a}} \in C_0^\infty(\Gamma_{\mathfrak{a}})$ and $\rho \in C_0^\infty(0,T)$. For F a function of $\mathbf{x}$, set $(\rho\otimes F)(t,\mathbf{x}) = \rho(t)F(\mathbf{x})$. The definition of $\mathcal{W}_{\mathcal{D}_S^l}$ (see Definition 3.6) yields

$$\begin{aligned} &\Bigg| \int_0^T \int_\Omega \Big(\nabla_{\mathcal{D}^l}^{m} v^l \cdot (\rho\otimes\mathbf{q}_m) + (\Pi_{\mathcal{D}^l}^{m} v^l)\mathrm{div}(\rho\otimes\mathbf{q}_m) \Big) \mathrm{d}\mathbf{x}\mathrm{d}t \\ &\quad + \int_0^T \int_\Gamma \Big(\nabla_{\mathcal{D}^l}^{f} v^l \cdot (\rho\otimes\mathbf{q}_f) + (\Pi_{\mathcal{D}^l}^{f} v^l)\mathrm{div}_\tau(\rho\otimes\mathbf{q}_f) \Big) \mathrm{d}\tau(\mathbf{x})\mathrm{d}t \\ &\quad - \sum_{\mathfrak{a}\in\chi} \int_0^T \int_{\Gamma_{\mathfrak{a}}} (\rho\otimes(\mathbf{q}_m\cdot\mathbf{n}_{\mathfrak{a}}))\mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} v^l \mathrm{d}\tau(\mathbf{x})\mathrm{d}t \\ &\quad + \sum_{\mathfrak{a}\in\chi} \int_0^T \int_{\Gamma_{\mathfrak{a}}} (\rho\otimes\varphi_{\mathfrak{a}}) \Big(\mathbb{T}_{\mathcal{D}^l}^{\mathfrak{a}} v^l - \Pi_{\mathcal{D}^l}^{f} v^l - [\![v^l]\!]_{\mathfrak{a},\mathcal{D}^l} \Big) \mathrm{d}\tau(\mathbf{x})\, dt \Bigg| \\ &\leq \|v^l\|_{\mathcal{D}^l} \|\rho\|_{L^2(0,T)} \mathcal{W}_{\mathcal{D}_S^l}(\mathbf{q},\varphi_{\mathfrak{a}}). \end{aligned}$$

The limit-conformity shows that the right-hand side of this inequality tends to 0. Hence,

$$\begin{aligned} &\int_0^T \int_\Omega \Big(\boldsymbol{\xi}_m \cdot (\rho\otimes\mathbf{q}_m) + \overline{v}_m \mathrm{div}(\rho\otimes\mathbf{q}_m) \Big) \mathrm{d}\mathbf{x}\mathrm{d}t \\ &\quad + \int_0^T \int_\Gamma \Big(\boldsymbol{\xi}_f \cdot (\rho\otimes\mathbf{q}_f) + \overline{v}_f \mathrm{div}_\tau(\rho\otimes\mathbf{q}_f) \Big) \mathrm{d}\tau(\mathbf{x})\mathrm{d}t \\ &\quad - \sum_{\mathfrak{a}\in\chi} \int_0^T \int_{\Gamma_{\mathfrak{a}}} (\rho\otimes(\mathbf{q}_m\cdot\mathbf{n}_{\mathfrak{a}}))\beta_{\mathfrak{a}} \mathrm{d}\tau(\mathbf{x})\mathrm{d}t \end{aligned}$$

$$+\sum_{\mathfrak{a}\in\chi}\int_0^T\int_{\Gamma_\mathfrak{a}}(\rho\otimes\varphi_\mathfrak{a})\Big(\beta_\mathfrak{a}-\overline{v}_f-j_\mathfrak{a}\Big)\mathrm{d}\tau(\mathbf{x})\,dt=0.$$

Applying this to $(\mathbf{q},\varphi_\mathfrak{a})=((\mathbf{q}_m,0),0)$ with $\mathbf{q}_m\in C_0^\infty(\Omega\setminus\overline{\Gamma})^d$, and using the density of tensorial functions $\{\sum_{r=1}^N\rho_r\otimes\mathbf{q}_m\,:\,N\in\mathbb{N},\ \rho_r\in C_0^\infty(0,T),\ \mathbf{q}_m\in C_0^\infty(\Omega\setminus\overline{\Gamma})^d\}$ in $C_0^\infty((0,T)\times(\Omega\setminus\overline{\Gamma}))^d$ (see [17, Appendix D]) shows that $\boldsymbol{\xi}_m=\nabla\overline{v}_m$. With $(\mathbf{q},\varphi_\mathfrak{a})=((0,\mathbf{q}_f),0)$ where $\mathbf{q}_f\in C_0^\infty(\overline{\Gamma}_i)^{d-1}$, we obtain $\boldsymbol{\xi}_f=\nabla\overline{v}_f$. Considering now $(\mathbf{q},\varphi_\mathfrak{a})=((\mathbf{q}_m,0),0)$ with $\mathbf{q}_m\in C_b^\infty(\Omega\setminus\overline{\Gamma})^d$ and applying the divergence theorem gives $\beta_\mathfrak{a}=\gamma_\mathfrak{a}\overline{v}_m$. Finally, taking $(\mathbf{q},\varphi_\mathfrak{a})=((0,0),\varphi_\mathfrak{a})$ with a general $\varphi_\mathfrak{a}\in C_0^\infty(\Gamma_\mathfrak{a})$ yields $j_\mathfrak{a}=\beta_\mathfrak{a}-\overline{v}_f=\gamma_\mathfrak{a}\overline{v}_m-\overline{v}_f=[\![\overline{v}]\!]_\mathfrak{a}$. □

With [29, Lemma 3.6], we can state the following.

Corollary A.3 *Under the assumptions of Lemma A.2, if $g_\mu:\mathbb{R}\to\mathbb{R}$ ($\mu\in\{m,f\}$) and $g_\mathfrak{a}:\mathbb{R}\to\mathbb{R}$ ($\mathfrak{a}\in\chi$) are continuous, non-decreasing functions and if $(\Pi_{\mathcal{D}^l}^\mu g_\mu(v^l))_l$ strongly converges in $L^2((0,T)\times M_\mu)$ and $(\mathbb{T}_{\mathcal{D}^l}^\mathfrak{a}g_\mathfrak{a}(v^l))_l$ strongly converges in $L^2((0,T)\times\Gamma_\mathfrak{a})$, then*

$$\begin{cases}\Pi_{\mathcal{D}^l}^\mu g_\mu(v^l)\to g_\mu(\overline{v}_\mu) & \text{in } L^2((0,T)\times M_\mu),\\ \mathbb{T}_{\mathcal{D}^l}^\mathfrak{a}g_\mathfrak{a}(v^l)\to g_\mathfrak{a}(\gamma_\mathfrak{a}\overline{v}_m) & \text{in } L^2((0,T)\times\Gamma_\mathfrak{a}).\end{cases}$$

A.3 Identification of time derivatives

We discuss here how weak formulations, with derivatives on test functions, enable us to recover some regularity properties on time derivatives of quantities of interest.

Let us start with a classical situation, similar to [19, Remark 1.1]. Let (M,ν) be a measured space and E be a Banach space densely embedded in $L^2(M)$, so that $E\hookrightarrow L^2(M)\hookrightarrow E'$. Assume also that E' is separable. Let $\mathcal{L}:L^2(0,T;E)\to\mathbb{R}$ be a continuous linear form and let $\mathcal{E}\subset C_0^1([0,T);E)$ be such that $\mathcal{E}_0=\{\Phi\in\mathcal{E}\,:\,\Phi(0,\cdot)=0\}$ is dense in $L^2(0,T;E)$. Suppose that $U\in L^2(0,T;E)$ and $U_0\in L^2(M)$ satisfy, for all $\Phi\in\mathcal{E}$,

$$-\int_0^T\int_M U(t,\mathbf{x})\partial_t\Phi(t,\mathbf{x})\mathrm{d}\nu(\mathbf{x})\mathrm{d}t+\int_M U_0(\mathbf{x})\Phi(0,\mathbf{x})\mathrm{d}\nu(\mathbf{x})=\mathcal{L}(\Phi). \qquad (54)$$

This relation shows that

$$\Xi:\Phi\mapsto-\int_0^T\int_M U(t,\mathbf{x})\partial_t\Phi(t,\mathbf{x})\mathrm{d}\nu(\mathbf{x})\mathrm{d}t$$

is linear (equal to $\mathcal{L}$) on $\mathcal{E}_0$, and continuous for the topology of $L^2(0,T;E)$. By density of $\mathcal{E}_0$ in this space, Ξ can be extended into an element of $(L^2(0,T;E))'=L^2(0,T;E')$ (see [17, Theorem 1.4.1]). We denote this element by $\partial_t U$, as it clearly corresponds to the distributional derivative of U [17, Section 2.1.2]. By [17, Section 2.5.2] this shows that $U : [0, T] \to L^2(M)$ is continuous and, using [17, Proposition 2.5.2] to integrate by parts in (54), that $U(0) = U_0$ and

$$\forall \Phi \in \mathcal{E},\ \langle \partial_t U, \Phi \rangle_{L^2(0,T;E'),L^2(0,T;E)} \mathrm{d}t = \int_0^T \langle \partial_t U(t), \Phi(t) \rangle_{E',E} \mathrm{d}t = \mathcal{L}(\Phi). \tag{55}$$

By density of $\mathcal{E}$ in $L^2(0, T; E)$, this relation actually holds for any $\Phi \in L^2(0, T; E)$.

We now consider the setting in the proof of Theorem 4.1 (see Section 4.3). Fixing $M = M_f$, $\mathrm{d}\nu = \mathrm{d}\tau_f$, $E = V_f^0$, $\mathcal{E} = C^1([0, T]; C_\Gamma^\infty)$ and

$$\mathcal{L}(\Phi) = \int_0^T \int_{M_f} h_f^\alpha \Phi \mathrm{d}\tau_f \mathrm{d}t - \int_0^T \int_{M_f} [kS]_f^\alpha(\overline{p}_f)\, \Lambda_f \nabla \overline{u}_f^\alpha \cdot \nabla \Phi \mathrm{d}\tau_f \mathrm{d}t$$
$$+ \sum_{\mathfrak{a} \in \chi} \Big(\int_0^T \int_{\Gamma_\mathfrak{a}} \rho_\mathfrak{a}^\alpha (-\Phi) \mathrm{d}\tau \mathrm{d}t \Big),$$

and using (40) with $\varphi_m^\alpha = 0$ and $\varphi_f^\alpha = \Phi$, $\varphi_f^\beta = 0$, for $\alpha, \beta = 1, 2$ with $\alpha \neq \beta$, this identifies $\partial_t(\phi_f S_f^\alpha(\overline{p}_f)) = \phi_f \partial_t S_f^\alpha(\overline{p}_f)$ as an element of $L^2(0, T; V_f^{0\prime})$.

Let us now deal with a slightly more complicated case, in which the time derivatives of two functions need to be combined to exhibit a certain regularity. With the same M and E as above, take (N, λ) a measured space and $\gamma : E \to L^2(N)$ a continuous linear mapping. Assume that $U \in L^2(0, T; E)$, $V \in L^2(0, T; L^2(N))$, $U_0 \in L^2(M)$ and $V_0 \in L^2(N)$, satisfy, for all $\phi \in \mathcal{E}$,

$$- \int_0^T \int_M U(t, \mathbf{x}) \partial_t \Phi(t, \mathbf{x}) \mathrm{d}\nu(\mathbf{x}) \mathrm{d}t - \int_0^T \int_N V(t, \mathbf{x}) \partial_t \gamma(\Phi(t))(\mathbf{x}) \mathrm{d}\lambda(\mathbf{x}) \mathrm{d}t$$
$$+ \int_M U_0(\mathbf{x}) \Phi(0, \mathbf{x}) \mathrm{d}\nu(\mathbf{x}) + \int_N V_0(\mathbf{x}) \gamma(\Phi(0))(\mathbf{x}) \mathrm{d}\lambda(\mathbf{x}) = \mathcal{L}(\Phi). \tag{56}$$

The same reasoning as above shows that

$$\widetilde{\Xi} : \Phi \mapsto - \int_0^T \int_M U(t, \mathbf{x}) \partial_t \Phi(t, \mathbf{x}) \mathrm{d}\nu(\mathbf{x}) \mathrm{d}t - \int_0^T \int_N V(t, \mathbf{x}) \partial_t \gamma(\Phi(t))(\mathbf{x}) \mathrm{d}\lambda(\mathbf{x}) \mathrm{d}t$$

can be extended into a linear continuous form on $L^2(0, T; E)$. Letting $\gamma^* : L^2(N) \to E'$ be the adjoint of γ (that is, $\langle g, \gamma(\Phi) \rangle_{L^2(N)} = \langle \gamma^* g, \Phi \rangle_{E',E}$ for all $g \in L^2(N)$ and $\Phi \in E$), the form $\widetilde{\Xi}$ is naturally denoted by $\partial_t U + \gamma^* \partial_t V$. Note that, in this sum, the two terms cannot be separated and it cannot, for example, be asserted that $\partial_t U \in L^2(0, T; E')$ and $\gamma^* \partial_t V \in L^2(0, T; E')$. Then, a reasoning similar to the one in [17] shows that $U + \gamma^* V : [0, T] \to L^2(M)$ is continuous with value $U_0 + \gamma^* V_0$ at $t = 0$, and that, for all $\Phi \in L^2(0, T; E)$,

$$\langle \partial_t U + \gamma^* \partial_t V, \Phi \rangle_{L^2(0,T;E'),L^2(0,T;E)} = \mathcal{L}(\Phi).$$

To write more natural equations, in the paper we sometimes make an abuse of notation and separate the two derivatives. We then write

$$\begin{aligned}\langle \partial_t U + \gamma^* \partial_t V, \Phi \rangle_{L^2(0,T;E'),L^2(0,T;E)} &= \int_0^T \langle \partial_t U, \Phi \rangle \mathrm{d}t + \int_0^T \langle \gamma^* \partial_t V, \Phi \rangle \mathrm{d}t \\ &= \int_0^T \langle \partial_t U, \Phi \rangle \mathrm{d}t + \int_0^T \langle \partial_t V, \gamma \Phi \rangle \mathrm{d}t,\end{aligned} \tag{57}$$

where, in the right-hand side, the duality brackets do not have indices, to avoid claiming that $\partial_t U \in L^2(0, T; E')$ or $\gamma^* \partial_t V \in L^2(0, T; E')$, and to remember that these two terms must be understood together.

Used in (40) with $\gamma = \gamma_\mathfrak{a}$ for all $\mathfrak{a} \in \chi$, the above reasoning and notations enable us to identify the (combined) time derivatives of $\phi_m S_m^\alpha(\overline{p}_m)$ and $\sum_\mathfrak{a} \eta S_\mathfrak{a}^\alpha(\gamma_\mathfrak{a} \overline{p}_m)$ as an element of $L^2(0, T; V_m^{0'})$.

References

1. Aghili, J., Brenner, K., Hennicker, J., Masson, R., Trenty, L.: Two-phase discrete fracture matrix models with nonlinear transmission conditions (2018). https://hal.archives-ouvertes.fr/hal-01764432. Accessed 3 May 2018
2. Ahmed, R., Edwards, M., Lamine, S., Huisman, B., Pal, M.: Control-volume distributed multi-point flux approximation coupled with a lower-dimensional fracture model. J. Comput. Phys. **284**, 462–489 (2015)
3. Ahmed, R., Edwards, M.G., Lamine, S., Huisman, B.A., Pal, M.: Three-dimensional control-volume distributed multi-point flux approximation coupled with a lower-dimensional surface fracture model. J. Comput. Phys. **303**, 470–497 (2015)
4. Alboin, C., Jaffré, J., Roberts, J., Serres, C.: Modeling fractures as interfaces for flow and transport in porous media. SIAM J. Sci. Comput. **295**, 13–24 (2002)
5. Angot, P., Boyer, F., Hubert, F.: Asymptotic and numerical modelling of flows in fractured porous media. ESAIM. Math. Model. Numer. Anal. **43**(2), 239–275 (2009)
6. Antonietti, P.F., Formaggia, L., Scotti, A., Verani, M., Verzott, N.: Mimetic finite difference approximation of flows in fractured porous media. ESAIM M2AN **50**, 809–832 (2016)
7. Berrone, S., Pieraccini, S., Scialò, S.: An optimization approach for large scale simulations of discrete fracture network flows. J. Comput. Phys. **256**, 838–853 (2014)
8. Bogdanov, I.I., Mourzenko, V.V., Thovert, J.-F., Adler, P.M.: Two-phase flow through fractured porous media. Phys. Rev. E **68**(2), 026703 (2003)
9. Brenner, K., Groza, M., Guichard, C., Lebeau, G., Masson, R.: Gradient discretization of hybrid-dimensional Darcy flows in fractured porous media. Numer. Math. **134**(3), 569–609 (2016)
10. Brenner, K., Groza, M., Guichard, C., Masson, R.: Vertex approximate gradient scheme for hybrid-dimensional two-phase Darcy flows in fractured porous media. ESAIM. Math. Model. Numer. Anal. **2**(49), 303–330 (2015)
11. Brenner, K., Hennicker, J., Masson, R., Samier, P.: Gradient discretization of hybrid-dimensional Darcy flow in fractured porous media with discontinuous pressures at matrix-fracture interfaces. IMA J. Numer. Anal. **37**(3), 1551–1585 (2017)
12. Brenner, K., Hennicker, J., Masson, R., Samier, P.: Hybrid dimensional modelling of two-phase flow through fractured porous media with enhanced matrix fracture transmission conditions. J. Comput. Phys. **357**, 100–124 (2018)
13. Cheng, H.M., Droniou, J., Le, K.-N.: Convergence analysis of a family of ELLAM schemes for a fully coupled model of miscible displacement in porous media, pp. 1–38 (2017)

14. D'Angelo, C., Scotti, A.: A mixed finite element method for Darcy flow in fractured porous media with non-matching grids. ESAIM. Math. Model. Numer. Anal. **46**(2), 465–489 (2012)
15. Demmel, J., Eisenstat, S., Gilbert, J., Li, X., Liu, J.: A supernodal approach to sparse partial pivoting. SIAM J. Matrix Anal. Appl. **20**(3), 720–750 (1999)
16. Demmel, J., Gilbert, J., Grigori, L., Li, X., Shao, M., Yamazaki, I.: Technical Report LBNL-44289, Lawrence Berkeley National Laboratory, SuperLU Users' Guide, September (1999). http://crd.lbl.gov/~xiaoye/SuperLU
17. Droniou, J.: Intégration et espaces de sobolev à valeurs vectorielles. Polycopiés de l'Ecole Doctorale de Maths-Info de Marseille (2001). https://hal.archives-ouvertes.fr/hal-01382368. Accessed 19 Dec 2016
18. Droniou, J.: Finite volume schemes for diffusion equations: introduction to and review of modern methods. Math. Models Methods Appl. Sci. **24**(8), 1575–1619 (2014)
19. Droniou, J., Eymard, R.: Uniform-in-time convergence of numerical methods for non-linear degenerate parabolic equations. Numer. Math. **132**(4), 721–766 (2016)
20. Droniou, J., Eymard, R., Feron, P.: Gradient Schemes for Stokes problem. IMA J. Numer. Anal. **36**(4), 1636–1669 (2016)
21. Droniou, J., Eymard, R., Gallouët, T., Guichard, C., Herbin, R.: The Gradient Discretisation Method. Mathematics and Applications. Springer, Heidelberg (2018). **(To appear)**
22. Droniou, J., Eymard, R., Gallouët, T., Herbin, R.: A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. Math. Models Methods Appl. Sci. **20**(2), 265–295 (2010)
23. Droniou, J., Eymard, R., Gallouët, T., Herbin, R.: Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations. Math. Models Methods Appl. Sci. **13**(23), 2395–2432 (2013)
24. Droniou, J., Eymard, R., Herbin, R.: Gradient schemes: generic tools for the numerical analysis of diffusion equations. M2AN Math. Model. Numer. Anal. **50**(3), 749–781 (2016). **(Special issue—Polyhedral discretization for PDE)**
25. Droniou, J., Eymard, R., Talbot, K.S.: Convergence in $C([0;T];L^2(\Omega))$ of weak solutions to perturbed doubly degenerate parabolic equations. J. Differ. Equ. **260**(11), 7821–7860 (2016)
26. Droniou, J., Hennicker, J., Masson, R.: Uniform-in-time convergence of numerical schemes for a two-phase discrete fracture model, vol. 199, pp. 275–283 (2017)
27. Droniou, J., Talbot, K.S.: On a miscible displacement model in porous media flow with measure data. SIAM J. Math. Anal. **46**(5), 3158–3175 (2014)
28. Eymard, R., Guichard, C., Herbin, R.: Small-stencil 3D schemes for diffusive flows in porous media. ESAIM. Math. Model. Numer. Anal. **46**(2), 265–290 (2012)
29. Eymard, R., Guichard, C., Herbin, R., Masson, R.: Gradient schemes for two-phase flow in heterogeneous porous media and Richards equation. ZAMM Z. Angew. Math. Mech. **94**(7–8), 560–585 (2014)
30. Faille, I., Fumagalli, A., Jaffré, J., Roberts, J.E.: Model reduction and discretization using hybrid finite volumes of flow in porous media containing faults. Comput. Geosci. **20**, 317–339 (2016)
31. Flauraud, E., Nataf, F., Faille, I., Masson, R.: Domain decomposition for an asymptotic geological fault modeling. C. R. Mec. **331**(12), 849–855 (2003)
32. Fumagalli, A., Scotti, A.: A reduced model for flow and transport in fractured porous media with non-matching grids. In: Cangiani, A., Davidchack, R., Georgoulis, E., Gorban, A., Levesley, J., Tretyakov, M. (eds) Numerical Mathematics and Advanced Applications 2011. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-33134-3_53
33. Haegland, H., Assteerawatt, A., Dahle, H., Eigestad, G., Helmig, R.: Comparison of cell- and vertex-centered discretization methods for flow in a two-dimensional discrete-fracture-matrix system. Adv. Water Resour. **32**, 1740–1755 (2009)
34. Hoteit, H., Firoozabadi, A.: An efficient numerical model for incompressible two-phase flow in fractured media. Adv. Water Resour. **31**(6), 891–905 (2008)
35. Jaffré, J., Mnejja, M., Roberts, J.: A discrete fracture model for two-phase flow with matrix-fracture interaction. Procedia Comput. Sci. **4**, 967–973 (2011)
36. Karimi-Fard, M., Durlofsky, L., Aziz, K.: An efficient discrete-fracture model applicable for general-purpose reservoir simulators. SPE J. **9**(02), 227–236 (2004)
37. Martin, V., Jaffré, J., Roberts, J.E.: Modeling fractures and barriers as interfaces for flow in porous media. SIAM J. Sci. Comput. **26**(5), 1667–1691 (2005)

38. Matthai, S.K., Mezentsev, A.A., Belayneh, M.: Finite element—node-centered finite-volume two-phase-flow experiments with fractured rock represented by unstructured hybrid-element meshes. SPE Reserv. Eval. Eng. **10**(06), 740–756 (2007)
39. Monteagudo, J.E., Firoozabadi, A.: Control-volume model for simulation of water injection in fractured media: incorporating matrix heterogeneity and reservoir wettability effects. SPE J. **12**(03), 355–366 (2007)
40. Reichenberger, V., Jakobs, H., Bastian, P., Helmig, R.: A mixed-dimensional finite volume method for two-phase flow in fractured porous media. Adv. Water Resour. **29**(7), 1020–1036 (2006)
41. Sandve, T., Berre, I., Nordbotten, J.: An efficient multi-point flux approximation method for discrete fracture-matrix simulations. J. Comput. Phys. **231**(9), 3784–3800 (2012)
42. Schwenck, N., Flemisch, B., Helmig, R., Wohlmuth, B.: Dimensionally reduced flow models in fractured porous media: crossings and boundaries. Comput. Geosci. **19**, 1219–1230 (2015)
43. Tunc, X., Faille, I., Gallouët, T., Cacas, M.C., Havé, P.: A model for conductive faults with non-matching grids. Comput. Geosci. **16**(2), 277–296 (2012)
44. Xing, F., Masson, R., Lopez, S.: Parallel vertex approximate gradient discretization of hybrid-dimensional Darcy flow and transport in discrete fracture networks. Comput. Geosci. **21**, 595–617 (2017)
45. Xing, F., Masson, R., Lopez, S.: Parallel numerical modelling of hybrid-dimensional compositional non-isothermal Darcy flows in fractured porous media. J. Comput. Phys. **345**, 637–664 (2018)