Local subcell monolithic DG/FV scheme for nonlinear shallow water equations with source terms on unstructured grids

Sacha Cardonna, Fabien Marche & François Vilar

Institute of Mathematics Alexander Grothendieck, University of Montpellier

Congrès Interdisciplinaire sur les Modèles Avancés de Vagues
Aussois, France – May 2025

IMAG
INSTITUT MONTPELLIERAIN
ALEXANDER GROTHENDIECK

L'AVENIR DEPUIS 1220
UNIVERSITÉ DE MONTPELLIER

Overview of this talk

Local subcell monolithic DG/FV scheme for **nonlinear shallow water equations** with **source terms** on unstructured grids

Some keywords.

- **Local subcell monolithic DG/FV scheme**: combines DG accuracy with FV robustness for stabilization;
- **Nonlinear shallow water equations**: describe the water waves under the hydrostatic assumption;
- **Source terms**: account for geometry and physical effects (e.g., topography, friction)

Table of contents

Table of contents

Table of contents

Shallow water asymptotics

Nonlinear shallow water (NSW) equations

$$\partial_t \mathbf{v} + \nabla_{\boldsymbol{x}} \cdot \mathbb{F}(\mathbf{v}, b) = \mathbf{B}(\mathbf{v}, \nabla_{\boldsymbol{x}} b)$$

$$\Leftrightarrow \begin{cases} \partial_t \eta + \nabla_{\boldsymbol{x}} \cdot \mathbf{q} = 0, \\ \partial_t q + \nabla_{\boldsymbol{x}} \cdot \left(\mathbf{u} \otimes \mathbf{q} + \frac{g\eta}{2}(\eta - 2b)\mathbb{I}_2 \right) = -g\eta \nabla_{\boldsymbol{x}} b \end{cases}$$

- $b : \mathbb{R}^2 \to \mathbb{R}$ is the **topography** parametrization;
- $\mathbf{v} : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathcal{H}^+$ is the vector gathering **total elevation** η and **discharge** $(q_x, q_y)^T$, with $\mathcal{H}^+ = \{(\eta, q_x, q_y) \in \mathbb{R}^3 \mid H := \eta - b \geq 0\}$;
- $\mathbb{F} : \mathcal{H}^+ \times \mathbb{R} \to \mathcal{M}_{2\times 3}(\mathbb{R})$ is the nonlinear **flux** tensor;
- $\mathbf{B} : \mathcal{H}^+ \times \mathbb{R} \to \mathbb{R}^3$ is the **source term** depending on topography.

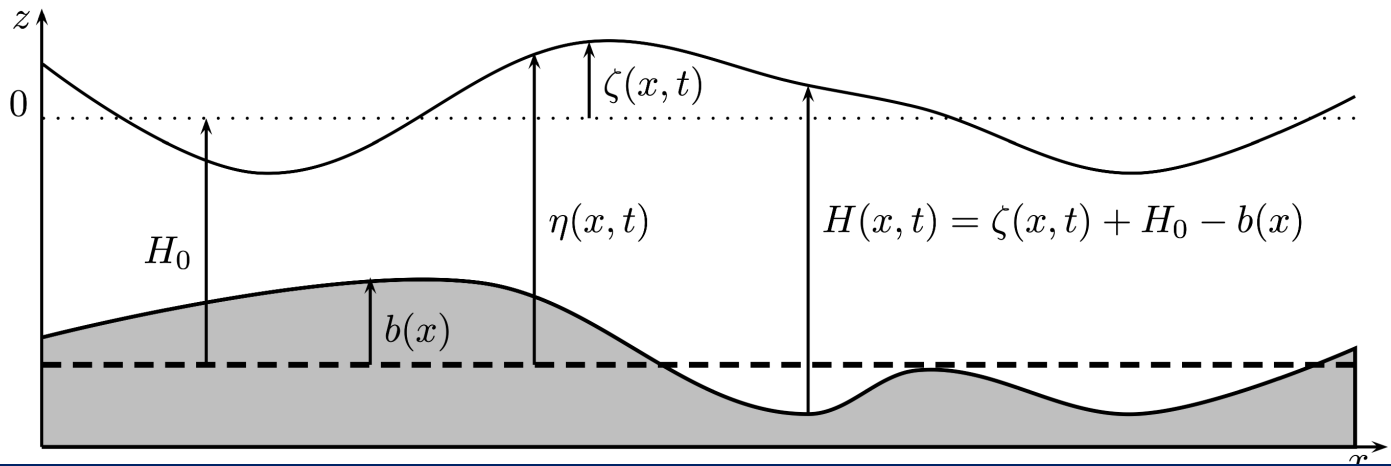


Table of contents

An overview of Finite Volume schemes

Multidimensional conservation law

$$\partial_t \mathbf{U}(\boldsymbol{x},t) + \nabla \cdot \mathbb{F}(\mathbf{U}(\boldsymbol{x},t)) = 0, \quad \mathbf{U} \in \mathbb{R}^m, \quad \boldsymbol{x} \in \Omega, \quad \omega_c \subset \Omega$$

- $\overline{\mathbf{U}}_{\omega_c}(t) = \dfrac{1}{|\omega_c|} \displaystyle\int_{\omega_c} \mathbf{U}(\boldsymbol{x},t)\, d\boldsymbol{x}$
- $\overline{\mathbf{U}}_{\omega_c}(t_{n+1}) = \overline{\mathbf{U}}_{\omega_c}(t_n) - \dfrac{1}{|\omega_c|} \displaystyle\int_{t_n}^{t_{n+1}} \int_{\partial\omega_c} \mathbb{F}(\mathbf{U}(\boldsymbol{x},t)) \cdot \boldsymbol{n}_{\partial\omega_c}\, dS\, dt$

Finite Volume discretization and scheme

- Domain partition: $\Omega = \bigcup_c \omega_c$, with each ω_c a control volume
- $\mathcal{V}_c$: set of neighbors sharing an edge with ω_c
- ℓ_{cv}: length of the interface $\omega_c \cap \omega_v$
- Piecewise constant solution: $\mathbf{U}_c^{n+1} = \mathbf{U}_c^n - \dfrac{\Delta t^n}{|\omega_c|} \displaystyle\sum_{v \in \mathcal{V}_c} \ell_{cv}\, \mathbb{F}_{cv}^*$

  where $\mathbb{F}_{cv}^*$ is a numerical approximation of the flux across the interface.

Finite Volume schemes: pros and cons

Advantages ✔

- Natural conservation across interfaces
- Applicable on general (unstructured) meshes
- Easy to implement for complex geometries
- Robust even on nonlinear problems

Limitations ✖

- Low-order accuracy without reconstruction
- Extension to high-order schemes leads to large stencils
- Limited flexibility for *hp*-adaptivity

An overview of Discontinuous Galerkin methods

Weak formulation

- Partition of the domain: $\mathscr{T}_h := \{\omega_1, \ldots, \omega_{n_{\text{el}}}\}, \quad \overline{\Omega} = \bigcup_{\omega \in \mathscr{T}_h} \overline{\omega}$
- $$\int_{\omega_c} \partial_t \mathbf{U}(\boldsymbol{x}, t)\, \psi(\boldsymbol{x})\, d\boldsymbol{x} - \int_{\omega_c} \mathbb{F}(\mathbf{U}, b) \cdot \nabla_{\boldsymbol{x}} \psi(\boldsymbol{x})\, d\boldsymbol{x} + \int_{\partial\omega_c} \mathbb{F}(\mathbf{U}, b) \cdot \boldsymbol{n}_{\partial\omega_c}\, \psi(s)\, dS = 0, \qquad \forall \psi \in \mathcal{C}_0^1(\omega_c)$$

Discontinuous Galerkin discretization

- Piecewise polynomial solution, discontinuous across interfaces:

$$\mathbf{U}_h^c(\boldsymbol{x}, t) = \sum_{m=1}^{\dim \mathbb{P}^k} \mathbf{U}_m^c(t) \phi_m^c(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \omega_c, \quad \forall t \in [0, t_{\max}],$$

  where the $\mathbf{U}_m^c(t)$ are the local DOFs and $\phi_m^c(\boldsymbol{x})$ are the basis functions
- As in FV framework, numerical flux $\mathbb{F}^*$ replaces $\mathbb{F}(\mathbf{U}) \cdot \boldsymbol{n}_{\partial\omega_c}$ on $\partial\omega_c$

Discontinuous Galerkin methods: pros and cons

Advantages ✔

- High-order accuracy with compact stencils
- Natural conservation across interfaces
- Suited for *hp*-adaptivity
- Well-suited for easy parallel computing
- Flexible for any meshes (unstructured, polytopal, etc.)

Limitations ✖

- More involving to implement than FV methods
- Non-physical oscillations when approaching strong gradients or discontinuities (like every scheme of order ≥ 2)
- Lack of nonlinear stability

Table of contents

Ideal setup for the NSW system

An ideal numerical scheme for the Nonlinear Shallow Water (NSW) equations should be:

- **High-order accurate** to capture smooth solutions and small-scale features;
- **Shock-capturing** to handle discontinuities and strong nonlinearities;
- **Positivity-preserving** to ensure non-negative water height and physical admissibility (i.e. stays in $\mathcal{H}^+$);
- **Well-balanced** to exactly preserve lake at rest steady states;
- **Adaptable to source terms** such as bottom topography and friction effects;
- **Well-suited for unstructured meshes** to deal with complex geometries and realistic domains.

Table of contents

Table of contents

DG formulation through residuals

DG formulation for all $\psi_p^c \in \mathbb{P}^k(\omega_c)$

$$\sum_{m=1}^{N_k} \frac{d\mathbf{v}_m^c}{dt} \int_{\omega_c} \psi_m^c \psi_p^c \, d\boldsymbol{x} - \int_{\omega_c} \mathbb{F} \cdot \nabla_{\boldsymbol{x}} \psi_p^c \, d\boldsymbol{x} + \int_{\partial\omega_c} \mathbb{F}^* \cdot \boldsymbol{n} \, \psi_p^c \, dS = \int_{\omega_c} \mathbf{B} \psi_p^c \, d\boldsymbol{x}$$

Residual DG formulation for any basis function $\psi_m^c \in \mathbb{P}^k(\omega_c)$

$$\boxed{\mathbb{M}_c \frac{d\mathbf{V}_c}{dt} = \Phi_c + \mathbf{S}_c}$$

- $(\mathbf{V}_c)_m = \mathbf{v}_m^c(t)$ **solution moments**
- $(\mathbb{M}_c)_{mp} = \int_{\omega_c} \psi_m^c(\boldsymbol{x}) \, \psi_p^c(\boldsymbol{x}) \, d\boldsymbol{x}$ **local mass matrix**
- $(\Phi_c)_m = \int_{\partial\omega_c} \mathbb{F}^* \cdot \boldsymbol{n} \, \psi_p^c \, dS - \int_{\omega_c} \mathbb{F}(\mathbf{v}_h^c, b_h^c) \cdot \nabla_{\boldsymbol{x}} \psi_p^c \, d\boldsymbol{x}$ **DG residuals**
- $(\mathbf{S}_c)_m = \int_{\omega_c} \mathbf{B}(\mathbf{v}_h^c, \nabla_{\boldsymbol{x}} b_h^c) \, \psi_p^c \, d\boldsymbol{x}$ **source term**

Stabilization principle

- **Classical stabilization:** apply limiters/a posteriori correction on the full cell

  ↪ risks **discarding** a mostly accurate solution due to a **local failure**

- **Subcell approach:** partition each cell into finer subcells to reduce the correction scale

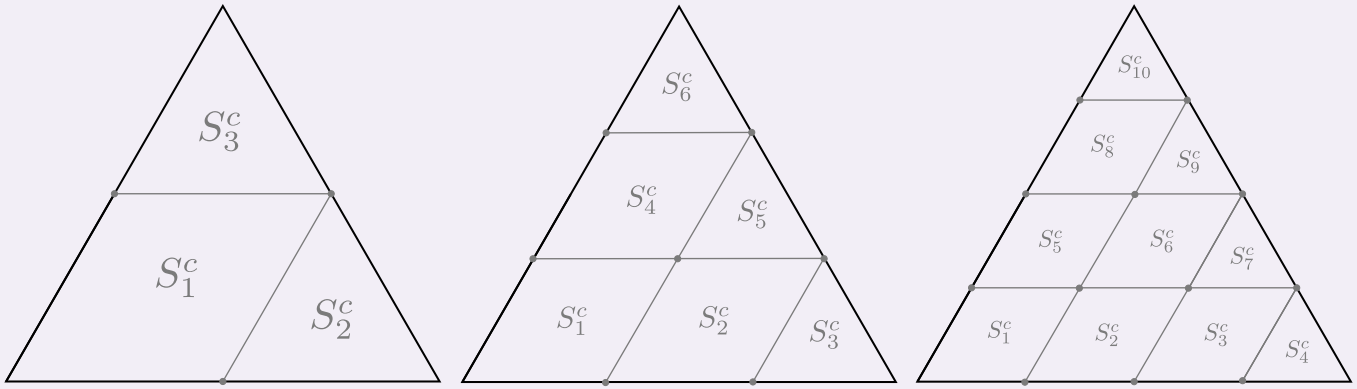
  ↪ enabling a **surgical correction**, meaning only fix what's necessary, preserving as much of the high-order DG content as possible

Theory needed – Reformulation of DG as a subcell FV-like scheme

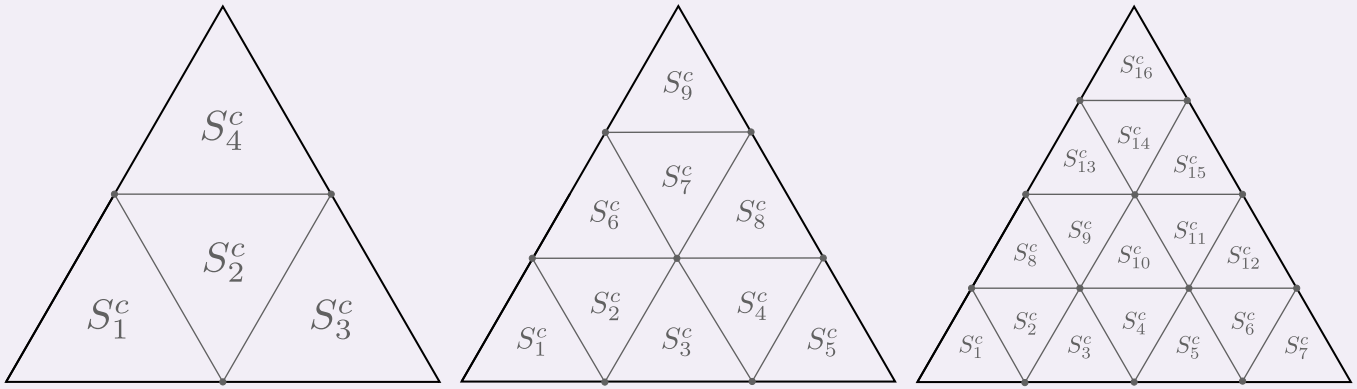
Table of contents

Mesh subdivision

Cell subdivision into $N_s \geq N_k$ subcells



Cell ω_c subdivided into $N_s = N_k$ subcells for $\mathbb{P}^1$ (left), $\mathbb{P}^2$ (center) and $\mathbb{P}^3$ (right) cases



Cell ω_c subdivided into $N_s \geq N_k$ subcells for $\mathbb{P}^1$ (left), $\mathbb{P}^2$ (center) and $\mathbb{P}^3$ (right) cases

A classical mesh ...

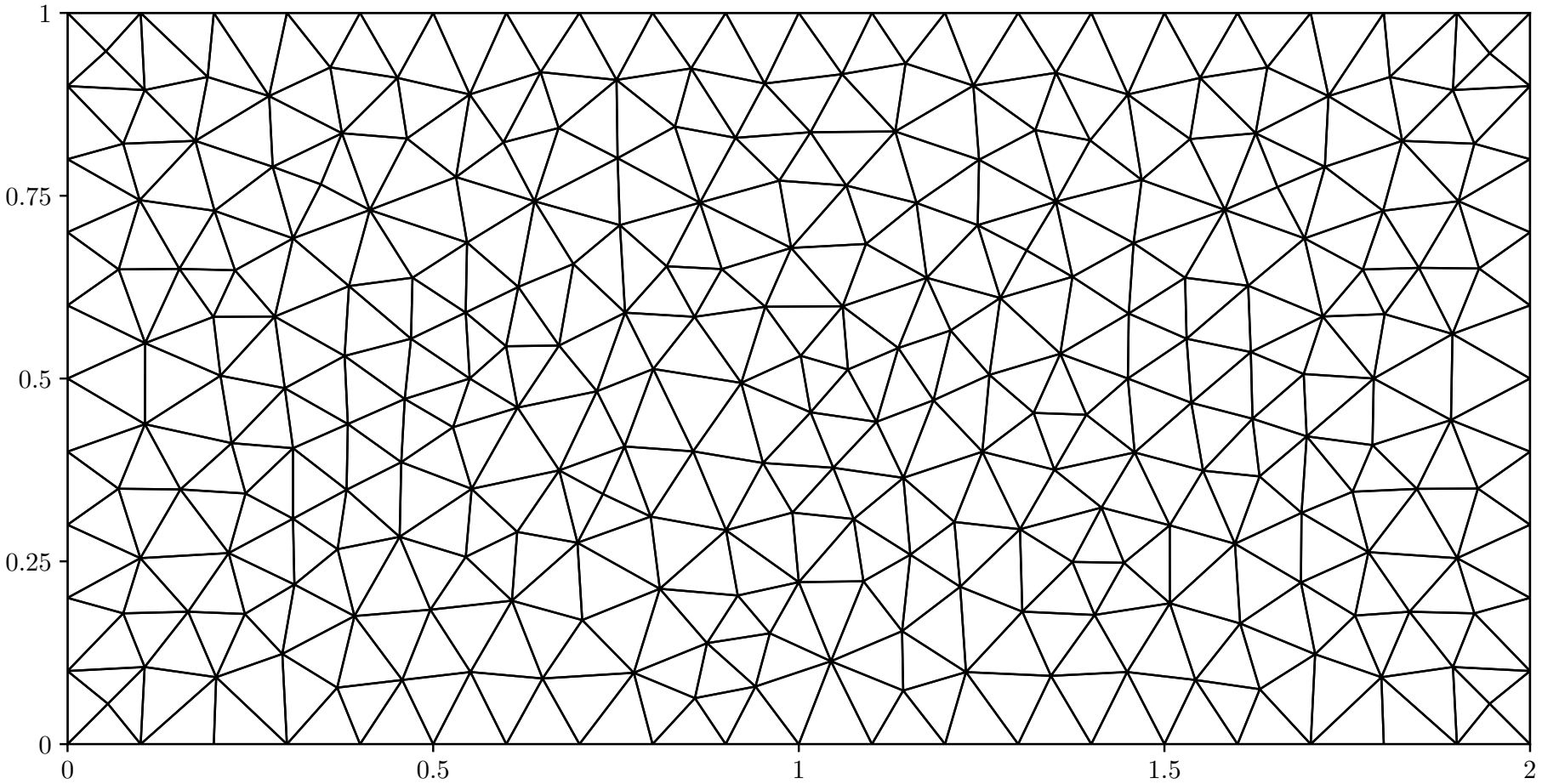


Figure: Unstructured simplicial mesh with $n_{\text{el}} = 350$ cells.

... and its subdivision

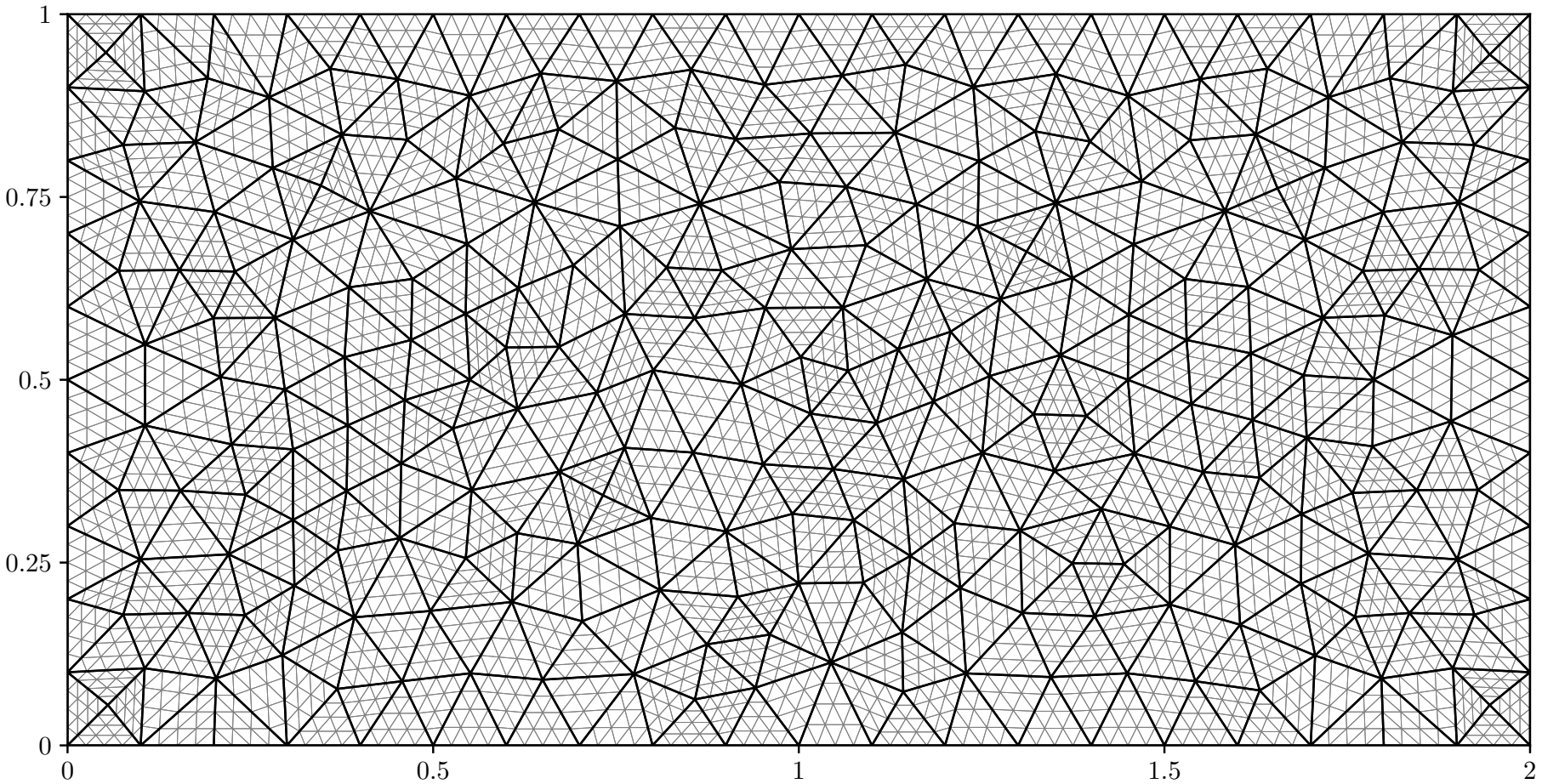


Figure: Unstructured simplicial mesh $\mathbb{P}^3$ subdivision onto triangles with $n_{\text{el}} = 350$ cells.

Subdivision and submean values

Some notations

- For any element $\omega_c \in \mathscr{T}_h$, we define a sub-partition:

$$\mathscr{T}_{\omega_c} := \{S_1^c, \ldots, S_{N_s}^c\}, \quad \overline{\omega}_c = \bigcup_{m=1}^{N_s} \overline{S}_m^c$$

- Γ_{mp}^c: interface between S_m^c and its neighbor S_p^v
- n_f^m: number of faces of subcell S_m^c
- $\mathscr{F}_{S_m^c}$: set of all faces of S_m^c
- n_f^c: total number of subcell faces inside element ω_c
- $\mathcal{V}_m^c$: set of face-neighboring subcells of S_m^c (with $|\mathcal{V}_m^c| = n_f^m$)
- $\breve{\mathcal{V}}_m^c$: subset of $\mathcal{V}_m^c$ containing only neighbors within the same element ω_c

Subneighbors

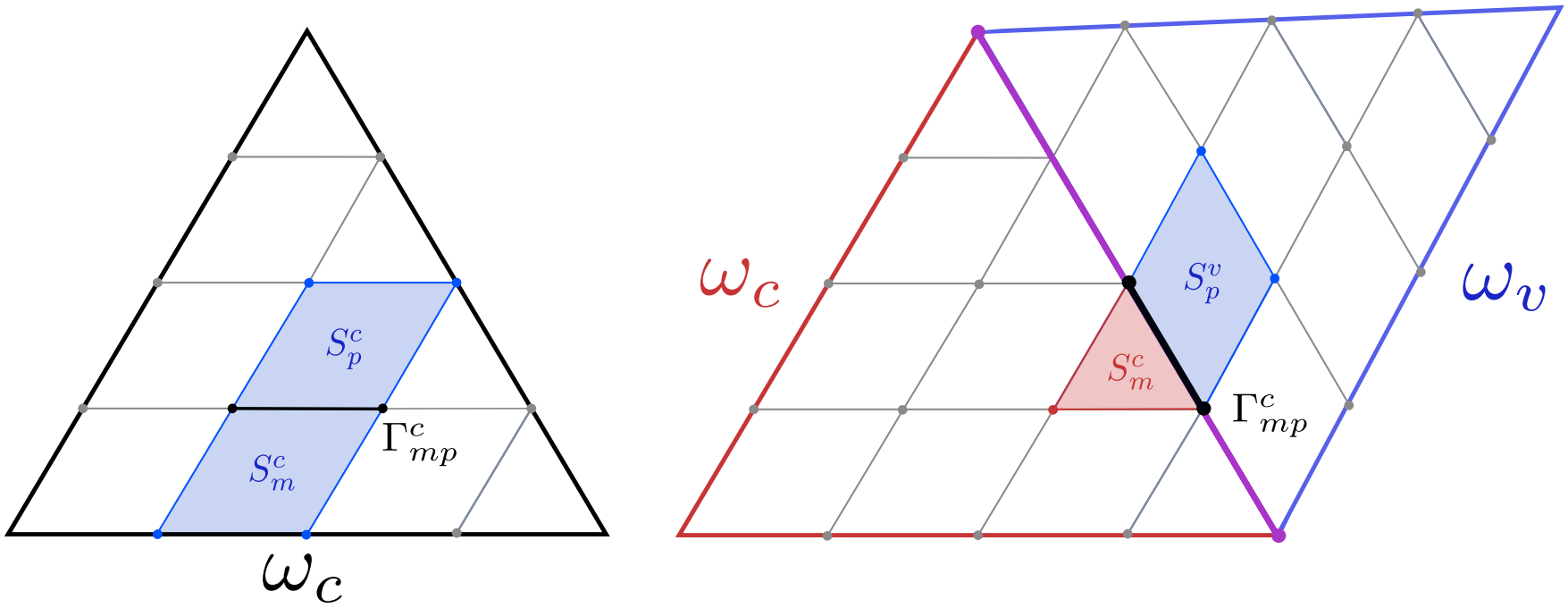


Figure: Two cases: subneighbor S_p inside cell ω_c (left), and subneighbor S_p inside neighbor cell ω_v (right).

Submean values and polynomial moments (1)

Mean value of a function over a subcell $S_m^c \subset \omega_c$

For any $f \in L^2(\omega_c)$, the subcell mean value is $\overline{f}_m^c := \dfrac{1}{|S_m^c|} \displaystyle\int_{S_m^c} f(\boldsymbol{x})\, d\boldsymbol{x}$.

Submean values and projection matrix

- $(\overline{\mathbf{V}}_c)_m = \overline{\mathbf{v}}_m^c(t)$ **submean values**
- $(\mathbb{P}_c)_{mp} = \dfrac{1}{|S_m^c|} \displaystyle\int_{S_m^c} \psi_p^c(\boldsymbol{x})\, d\boldsymbol{x}$ **projection matrix**

$$\overline{\mathbf{v}}_m^c(t) = \frac{1}{|S_m^c|} \sum_{q=1}^{N_k} \mathbf{v}_q^c(t) \int_{S_m^c} \psi_q^c(\boldsymbol{x})\, d\boldsymbol{x} \implies \boxed{\overline{\mathbf{V}}_c = \mathbb{P}_c \mathbf{V}_c}$$

⚠ $\mathbb{P}_c^t \mathbb{P}_c$ has to be **non-singular**, so we use the least-square procedure:

$$\boxed{\mathbf{V}_c = \left(\mathbb{P}_c^t \mathbb{P}_c\right)^{-1} \mathbb{P}_c^t \overline{\mathbf{V}}_c}$$

If $N_s = N_k$, then $\overline{\mathbf{V}}_c = \mathbb{P}_c \mathbf{V}_c \Leftrightarrow \mathbf{V}_c = \mathbb{P}_c^{-1} \mathbf{V}_c$.

Submean values and polynomial moments (2)

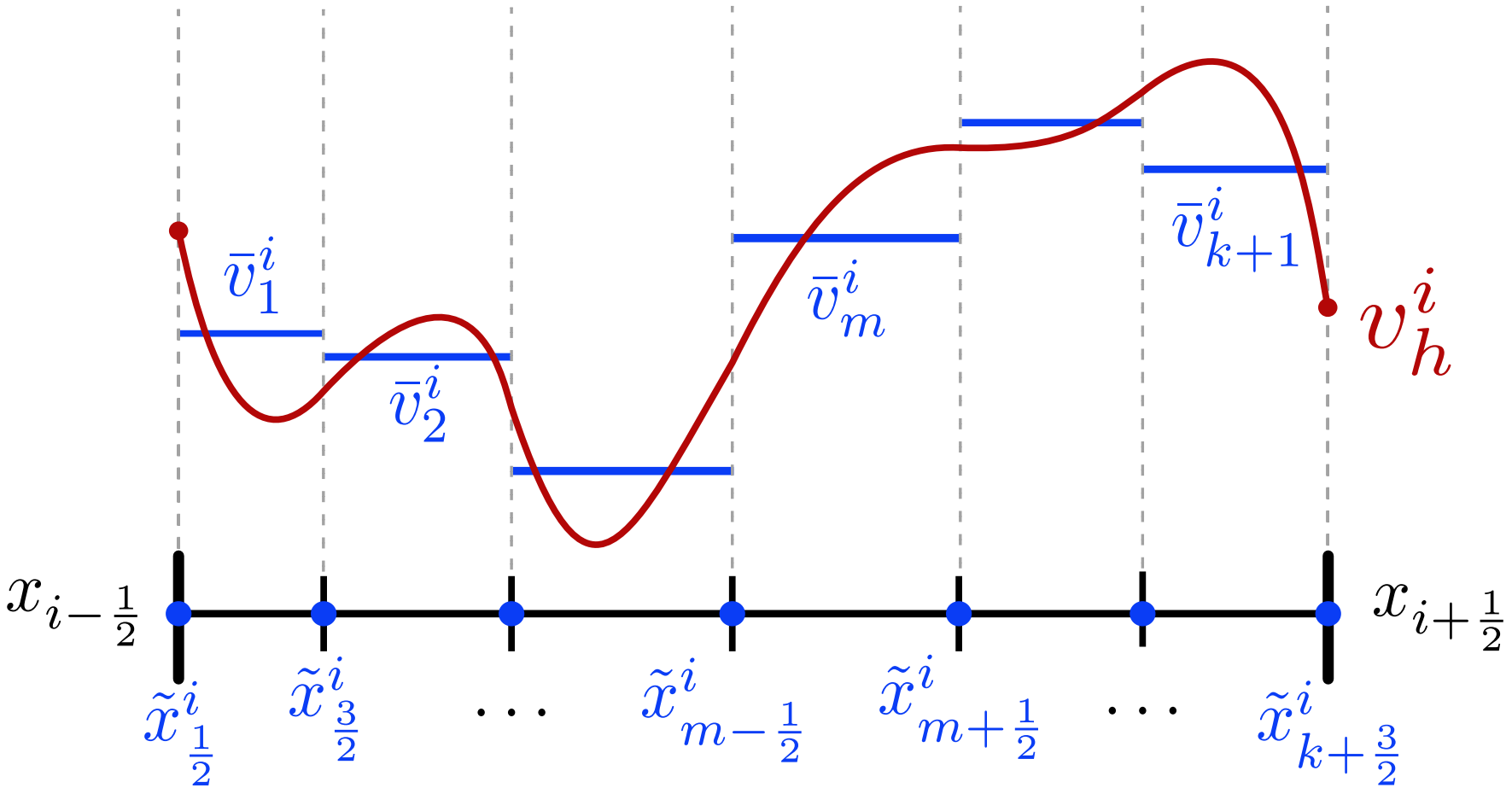


Figure: Piecewise polynomial function v_h^i and associated sub-mean-values (1D case).

Table of contents

Reconstructed DG fluxes (1)

Submean values vector derivative

Since $\mathbb{M}_c \dfrac{d\mathbf{V}_c}{dt} = \Phi_c + \mathbf{S}_c$ and $\overline{\mathbf{V}}_c = \mathbb{P}_c \mathbf{V}_c \implies \boxed{\dfrac{d\overline{\mathbf{V}}_c}{dt} = \mathbb{P}_c \mathbb{M}_c^{-1}(\Phi_c + \mathbf{S}_c)}$

Flux reconstruction to get a FV-like scheme

Let us consider the DG reconstructed flux $\widehat{\mathbb{F}}_n$ such that

$$\frac{d\overline{\mathbf{v}}_m^c}{dt} = -\frac{1}{|S_m^c|}\int_{\partial S_m^c} \widehat{\mathbb{F}}_n(\boldsymbol{x})\, d\boldsymbol{x} + (\mathbb{P}_c\mathbb{M}_c^{-1}\mathbf{S}_c)_m \qquad \text{(FV-like scheme)}$$

$$= -\frac{1}{|S_m^c|}\sum_{S_p^v \in \mathcal{V}_m^c}\int_{\Gamma_{mp}^c} \widehat{\mathbb{F}}_n(\boldsymbol{x})\, d\boldsymbol{x} + (\mathbb{P}_c\mathbb{M}_c^{-1}\mathbf{S}_c)_m \qquad \left(\partial S_m^c = \cup_{S_p^v \in \mathcal{V}_m^c}\Gamma_{mp}^c\right)$$

$$= -\frac{1}{|S_m^c|}\left(\sum_{S_p^v \in \check{\mathcal{V}}_m^c}\int_{\Gamma_{mp}^c} \widehat{\mathbb{F}}_n(\boldsymbol{x})\, d\boldsymbol{x} + \int_{\partial\omega_c \cap \partial S_m^c} \mathbb{F}_n^*\, d\boldsymbol{x}\right) + (\mathbb{P}_c\mathbb{M}_c^{-1}\mathbf{S}_c)_m$$

under the hypothesis that $\widehat{\mathbb{F}}_{n|\partial\omega} = \mathbb{F}^*$ for all $\omega \in \mathscr{T}_h$.

Reconstructed DG fluxes (2)

Interface reconstructed flux

We define $\widehat{\mathbb{F}}_{mp}$ at interface Γ^c_{mp} as: $$\int_{\Gamma^c_{mp}} \widehat{\mathbb{F}}_n(\boldsymbol{x})\, d\boldsymbol{x} = \varepsilon^c_{mp}\widehat{\mathbb{F}}_{mp},$$

where subface orientation is carried through ε^c_{mp}, such that $\varepsilon^c_{pm} = -\varepsilon^c_{mp}$.

Reconstructed flux system

$$\boxed{-\mathbb{A}_c\widehat{\mathbb{F}}_c = \mathbb{D}_c\frac{d\overline{\mathbf{V}}_c}{dt} + \partial\mathbb{F}_c}$$

- $(\widehat{\mathbb{F}}_c)_{mp} = \ell_{mp}\widehat{\mathbb{F}}_{mp}$ — **interior subfaces fluxes**
- $(\mathbb{A}_c)_{mp} = \varepsilon^c_{mp}$ — **adjacency matrix**
- $(\mathbb{D}_c)_m = |S^c_m|$ — **subvolume matrix**
- $(\partial\mathbb{F}_c)_m = \displaystyle\int_{\partial\omega_c\cap\partial S^c_m} \mathbb{F}^*_n\, d\boldsymbol{x}$ — **cell boundary contribution**

⚠ Since $\ker\mathbb{A}_c \neq \{\mathbf{0}\}$, we use a *Graph Laplacian technique*

Reconstructed DG fluxes (3)

Residual definition of reconstructed fluxes

$$\boxed{\widehat{\mathbb{F}}_c = -\mathbb{A}_c^t \mathcal{L}_c^{-1} \left(\mathbb{D}_c \mathbb{P}_c \mathbb{M}_c^{-1} \Phi_c + \partial \mathbb{F}_c \right)}$$

where $\mathcal{L}_c^{-1}$ is the gen. inverse of $\mathbb{L}_c := \mathbb{A}_c \mathbb{A}_c^t$ on the orthogonal of its kernel:

$$\mathcal{L}_c^{-1} = (\mathbb{L}_c + \lambda \Pi)^{-1} - \frac{1}{\lambda} \Pi, \qquad \Pi = \frac{1}{N_s} (1 \otimes 1) \in \mathcal{M}_{N_k}, \qquad \forall \lambda \neq 0$$

R. Abgrall, *Some Remarks about Conservation for Residual Distribution Schemes*. Methods Appl. Math., 18:327-351, 2018.

Few remarks

- **Source term** is excluded in the definition since only flux-dependent integrals are considered in reconstruction;
- **Implementation**: only Φ_c and boundary terms $\partial \mathbb{F}_c$ depend on time, but all the other terms are precomputable;
- **Alternative expression**: using spanning set of subresolution functions $\phi_m^c = p_{\omega_c}^k(\mathbb{1}_m^c)$, where $p_{\omega_c}^k$ is the L^2-projector on cell ω_c.

DG schemes ≡ Subcell FV schemes

Theorem (equivalence of DG and subcell FV schemes)

The NSW-DG residual scheme $\frac{d\mathbf{V}_c}{dt} = \mathbb{M}_c^{-1}(\Phi_c + \mathbf{S}_c)$ can be recast into N_s FV-like subcell schemes as

$$\boxed{\frac{d\overline{\mathbf{V}}_c}{dt} = -\mathbb{D}_c^{-1}\left(\mathbb{A}_c\widehat{\mathbb{F}}_c + \partial\mathbb{F}_c\right) + \overline{\mathbf{S}}_c}$$

where $\overline{\mathbf{S}}_c := \mathbb{P}_c\mathbb{M}_c^{-1}\mathbf{S}_c$ contains the submean values of source term projection, i.e.

$$\overline{\mathbf{B}}_m^c := \frac{1}{|S_m^c|}\int_{S_m^c} p_{\omega_c}^k\left(\mathbf{B}(\mathbf{v}_h, \nabla_{\boldsymbol{x}} b_h)\right)\, d\boldsymbol{x}.$$

DG equivalent semi-discrete scheme on every subcell $S_m^c \subset \omega_c$

$$\frac{d\overline{\mathbf{v}}_m^c}{dt} = -\frac{1}{|S_m^c|}\sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp}\widehat{\mathbb{F}}_{mp} + \overline{\mathbf{B}}_m^c, \qquad \forall m \in [\![1, N_s]\!]$$

Table of contents

Table of contents

Combining DG and FV frameworks (1)

Finite Volume scheme — 👍 robustness 👎 1^{st} order accuracy

+

Discontinuous Galerkin scheme — 👍 k^{th} order accuracy 👎 robustness

⇓

Monolithic DG-FV subcell scheme — 👍 k^{th} order accuracy & robustness

Combining DG and FV frameworks (2)

Our numerical solution should satisfy the following properties:

- **Accuracy**: high-order precision can be required
  ↪ natural in DG schemes; requires mesh refinement in FV schemes
- **Physical admissibility**: in NSW context, the solution should stay in $\mathcal{H}^+$
  ↪ automatic in FV schemes; requires dedicated techniques in DG schemes
- **Stability / No spurious oscillations**: satisfy a discrete maximum principle
  ↪ guaranteed in FV schemes; not ensured by DG schemes (limiters needed)

Idea – blending DG reconstructed fluxes and FV fluxes at subcell scale

Combining DG and FV frameworks (3)

Blended fluxes and blending coefficient

For every face $\Gamma_{mp}^c \in \mathscr{F}_{S_m^c}$, the high-order DG reconstructed flux $\widehat{\mathbb{F}}_{mp}$ and a first-order FV flux $\mathbb{F}_{mp}^{*,\mathrm{FV}}$ are assembled in a convex way:

$$\boxed{\widetilde{\mathbb{F}}_{mp} = \mathbb{F}_{mp}^{*,\mathrm{FV}} + \theta_{mp}\left(\widehat{\mathbb{F}}_{mp} - \mathbb{F}_{mp}^{*,\mathrm{FV}}\right) = \mathbb{F}_{mp}^{*,\mathrm{FV}} + \theta_{mp}\Delta\mathbb{F}_{mp}}$$

⚠ The **blending coefficient** $\theta_{mp} \in [0,1]$ is:

- computed *a priori* on each Γ_{mp}^c, at each time step (or RK stage);
- uniquely defined *i.e.* $\theta_{mp} = \theta_{pm}$, for all $S_p^v \in \mathcal{V}_m^c$.

Monolithic DG-FV subcell scheme with forward Euler time integration

$$\overline{\mathbf{v}}_m^{c,n+1} = \overline{\mathbf{v}}_m^{c,n} - \frac{\Delta t^n}{|S_m^c|}\sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp}\widetilde{\mathbb{F}}_{mp} + \Delta t^n \overline{\mathbf{B}}_m^{c,n}, \qquad \forall m \in [\![1, N_s]\!]$$

Table of contents

Source term treatment

Flowchart of the discretization

Bridging **polynomial degrees of freedom** and **subcell-averaged values**

1. **Subcell averages**: compute $\overline{b}_m^c$ and $\overline{\eta}_m^c$ on each subcell, then reconstruct b_h and η_h via projection matrix $\mathbb{P}_c$;
2. **Projection**: evaluate $\mathbf{B}(\mathbf{v}_h, \nabla_{\boldsymbol{x}} b_h)$ at quadrature nodes, then apply an L^2 projection onto $\mathbb{P}^k$;
3. **Integration**: compute the mean value of the projected source over each subcell:

$$\overline{\mathbf{B}}_m^c := \frac{1}{|S_m^c|} \int_{S_m^c} \mathbf{B}_h \, d\boldsymbol{x}$$

Implementation remark

Formally corresponds to multiplying the DG source integral by $\mathbb{P}_c \mathbb{M}_c^{-1}$:

$$\overline{\mathbf{B}}_m^c = \mathbb{P}_c \mathbb{M}_c^{-1} \left(\int_{\omega_c} \mathbf{B}_h \varphi_h \, d\boldsymbol{x} \right)$$

Generalization to algebraic/geometric source terms

Topography and (nonlinear) friction effects

$$\boxed{\mathbf{S}(\mathbf{v}, b) := \mathbf{B}(\mathbf{v}, \nabla_{\boldsymbol{x}} b) + \mathbf{Fr}(\mathbf{v}, b)}$$

- $\mathbf{B}(\mathbf{v}, \nabla_{\boldsymbol{x}} b) = (0, -g\eta\nabla_{\boldsymbol{x}} b)^t$ **Topography source term**
- $\mathbf{Fr}(\mathbf{v}, b) = \begin{cases} \left(0, -k_f^2\, \mathbf{q}\right)^t, \; k_f > 0 & \textbf{Linear friction law} \\ \left(0, -n_f^2\, \dfrac{\mathbf{q}\,\|\mathbf{q}\|}{(\eta - b)^\gamma}\right)^t, \; n_f, \gamma > 0 & \textbf{Manning friction law} \end{cases}$

❓ Handled the same way as previously → **easily generalizable**

Applications to Serre–Green–Naghdi (SGN) equations

Reformulation: Elliptic problem + NSW with dispersive source term

1. Elliptic problem solved *independently*, using a finite element method;
2. Resulting dispersive source term discretized within the NSW framework.

Table of contents

Reformulation as a Godunov-like scheme

Solution at t^{n+1} as a convex combination of quantities defined at t^n

$$\overline{\mathbf{v}}_m^{c,n+1} = \overline{\mathbf{v}}_m^{c,n} - \frac{\Delta t^n}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp} \widetilde{\mathbb{F}}_{mp} + \Delta t^n \overline{\mathbf{B}}_m^{c,n}$$

$$+ \frac{\Delta t^n}{|S_m^c|} \mathbb{F}\left(\overline{\mathbf{v}}_m^{c,n}, \overline{b}_m^c\right) \cdot \sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp} \boldsymbol{n}_{mp} \pm \frac{\sigma \Delta t^n}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp} \overline{\mathbf{v}}_m^{c,n}$$

$$= \left(1 - \frac{\sigma \Delta t^n}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp}\right) \overline{\mathbf{v}}_m^{c,n} + \frac{\sigma \Delta t^n}{|S_m^c|} \sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp} \widetilde{\mathbf{v}}_{mp}^{*,-} + \Delta t^n \overline{\mathbf{B}}_m^{c,n}$$

- $\widetilde{\mathbf{v}}_{mp}^{*,-}$ are the interior **blended Riemann intermediate states**

$$\widetilde{\mathbf{v}}_{mp}^{*,-} := \overline{\mathbf{v}}_m^{c,n} - \frac{\widetilde{\mathbb{F}}_{mp} - \mathbb{F}\left(\overline{\mathbf{v}}_m^{c,n}, \overline{b}_m^c\right) \cdot \boldsymbol{n}_{mp}}{\sigma} = \mathbf{v}_{mp}^{*,-} - \theta_{mp} \left(\frac{\widehat{\mathbb{F}}_{mp} - \mathbb{F}_{mp}^{*,\mathrm{FV}}}{\sigma}\right);$$

- $\mathbf{v}_{mp}^{*,-}$ are the 1^{st}**-order FV Riemann intermediate states**.

Analytical formula to ensure water height positivity

Relying on 1^{st}-order FV Riemann intermediate states

Proof of the natural **preservation of water-height positivity** for 1^{st}-order elevation Riemann FV states $\eta_{mp}^{*,\pm}$

$\hookrightarrow$ Allows us to rely on the **robustness of FV framework** to ensure the properties we want

Physical admissibility detector

$$\boxed{\theta_{mp}^{\mathcal{H}^+} := \min\left(\theta_{mp}^{\mathcal{H}^+,-}, \theta_{mp}^{\mathcal{H}^+,+}\right)}$$

- $\theta_{mp}^{\mathcal{H}^+,-} := \dfrac{\sigma\left(\eta_{mp}^{*,-} - \overline{b}_m^c\right)}{\Delta\mathbb{F}_{mp}}$ if $\Delta\mathbb{F}_{mp} > 0$, $\theta_{mp}^{\mathcal{H}^+,-} = 1$ else;
- $\theta_{mp}^{\mathcal{H}^+,+} := \dfrac{\sigma\left(\overline{b}_p^v - \eta_{mp}^{*,+}\right)}{\Delta\mathbb{F}_{pm}}$ if $\Delta\mathbb{F}_{pm} < 0$, $\theta_{mp}^{\mathcal{H}^+,+} = 1$ else.

Analytical formulas to prevent spurious oscillations

Mimicking a local maximum principle

$$\alpha_m^c := \min_{S_p^v \in \mathcal{N}(S_m^c)} \left(\overline{\eta}_p^{v,n}, \eta_{mp}^{*,-}\right) \leq \overline{\eta}_m^{c,n+1} \leq \max_{S_p^v \in \mathcal{N}(S_m^c)} \left(\overline{\eta}_p^{v,n}, \eta_{mp}^{*,-}\right) =: \beta_m^c$$

where $\mathcal{P}_m^c$ is the set of vertices $\boldsymbol{x}_p$ of subcell S_m^c and

$$\mathcal{N}(S_m^c) := \bigcup_{\boldsymbol{x}_p \in \mathcal{P}_m^c} \{S_q \mid \boldsymbol{x}_p \in S_q\}$$

Subcell numerical admissibility detector

$$\theta_{mp}^{\text{SubNAD}} := \min\left(1, \left|\frac{\sigma}{\Delta\mathbb{F}_{mp}}\right| \begin{cases} \min\left(\beta_p^v - \eta_{mp}^{*,+},\ \eta_{mp}^{*,-} - \alpha_m^c\right) & \text{if } \Delta\mathbb{F}_{mp} > 0 \\ \min\left(\beta_m^c - \eta_{mp}^{*,-},\ \eta_{mp}^{*,+} - \alpha_p^v\right) & \text{if } \Delta\mathbb{F}_{mp} < 0 \end{cases}\right)$$

⚠ For NSW, no local maximum principle for the conserved variable!

↪ needs to be **relaxed** in the presence of **smooth extremas**

Table of contents

Preservation of steady-states (1)

Why does it matter ?

- **Preserves lake at rest steady states exactly**, avoiding spurious motions;
- **Reduces numerical errors** near equilibrium, especially when small perturbations are present;
- **Essential for wet/dry interfaces**, where small oscillations can destabilize the scheme.

Well-balancing (WB) property

Providing that the integrals of discrete formulation are exactly computed, we have the following result:

$$\forall n \in \mathbb{N}, \quad \forall \eta^e \in \mathbb{R}, \quad (\eta_h^n = \eta^e \text{ and } \mathbf{q}_h^n = \mathbf{0}) \Longrightarrow (\eta_h^{n+1} = \eta^e \text{ and } \mathbf{q}_h^{n+1} = \mathbf{0})$$

Preservation of steady-states (2)

Sketch of proof

Objective: showing that numerical fluxes are cancelling the source term *i.e.*

$$\frac{1}{|S_m^c|}\sum_{S_p^v \in \mathcal{V}_m^c} \ell_{mp}\widetilde{\mathbb{F}}_{mp} = \overline{\mathbf{B}}_m^{c,n} \qquad \text{s.t.} \qquad \overline{\mathbf{v}}_m^{c,n+1} = \overline{\mathbf{v}}_m^{c,n}.$$

- Exact integration required → natural with high-order quadrature;
- Under well-balanced assumptions:
$$\nabla_{\boldsymbol{x}} \cdot \mathbb{F}(\mathbf{v}_c, b_c) = \mathbf{B}(\mathbf{v}_c, \nabla_{\boldsymbol{x}} b_c), \quad \forall \omega_c \in \mathscr{T}_h;$$
- Fluxes $\widehat{\mathbb{F}}_{mp}$ and $\mathbb{F}_{mp}^{*,\mathrm{FV}}$ match the continuous flux $\mathbb{F}_h^c \cdot \boldsymbol{n}_{mp}$ under equilibrium;
- $\widetilde{\mathbb{F}}_{mp}$ is built as a convex combination of these well-balanced fluxes $\hookrightarrow$ preserves equilibrium as well !

Table of contents

Test 1 – Order of accuracy assessment

Steady vortex with $\mathcal{C}^\infty$ topography

- **Domain:** $\Omega = [-5, 5]^2$ **Degree:** $k = 1, 2, 3$ **Mesh:** $n_{\text{el}} = 200 \rightarrow 12800$
- **Goal:** convergence of the scheme on a smooth solution with a consistent discretization of the topography source term

k	1		2		3	
h	$E^\eta_{L^2}$	$q^\eta_{L^2}$	$E^\eta_{L^2}$	$q^\eta_{L^2}$	$E^\eta_{L^2}$	$q^\eta_{L^2}$
1	9.445E-2	2.35	1.529E-2	2.91	4.580E-3	4.19
$\frac{1}{2}$	1.854E-2	2.16	2.039E-3	3.03	2.505E-4	4.10
$\frac{1}{4}$	4.158E-3	2.07	2.491E-4	2.97	1.465E-5	4.00
$\frac{1}{8}$	9.923E-4	–	3.187E-5	–	9.165E-7	–

Figure: L^2-errors between numerical and analytical solutions and convergence rates for η at time $t = 0.1$ sec.

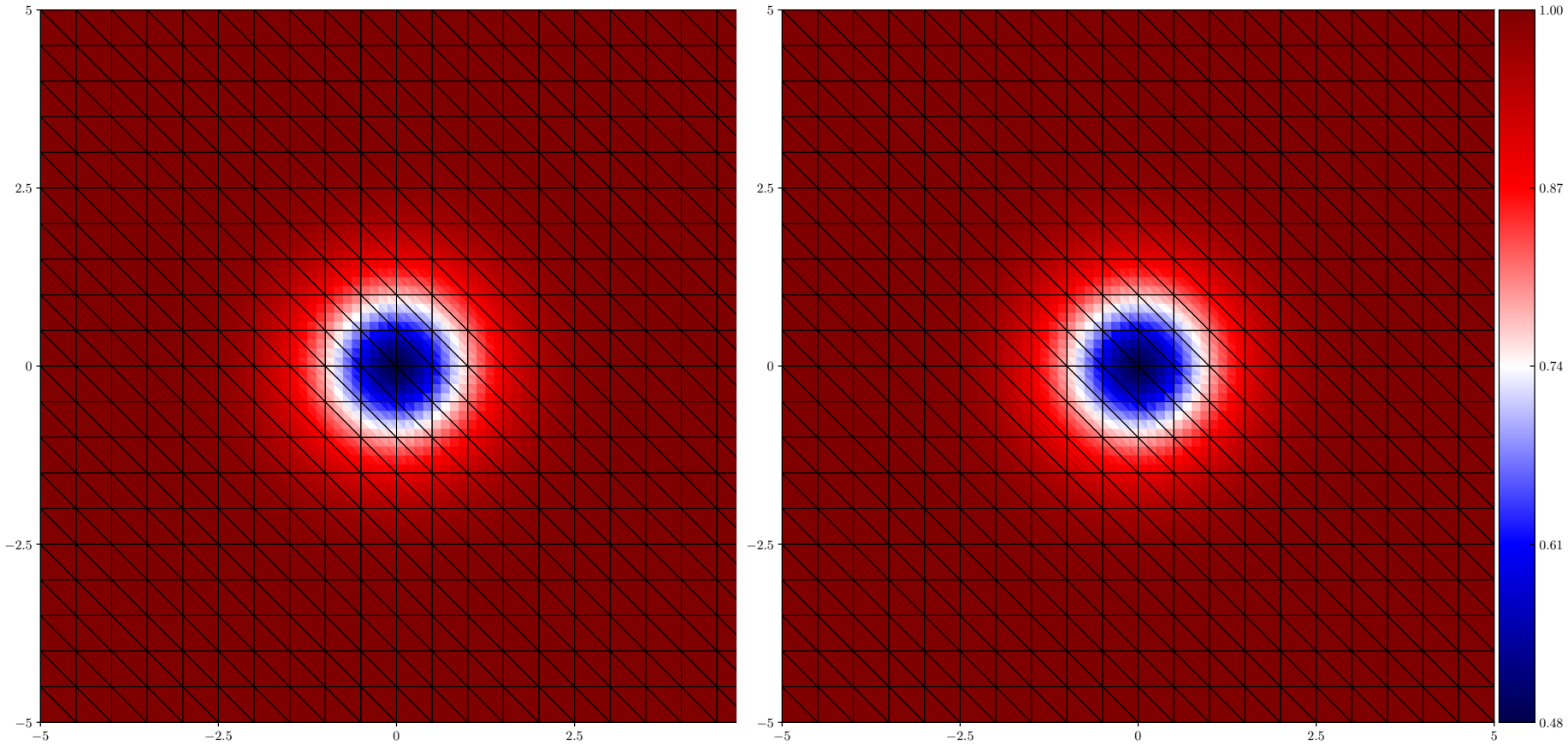


Figure: Steady vortex − Exact (left) and $\mathbb{P}^3$ numerical (right) height at final time $t = 0.1$ sec on 800 cells.

Test 2 – Well-balancing assessment

Well-balancing with dry area

- **Domain:** $\Omega = [0, 2] \times [0, 1]$ **Degree:** $k = 4$ **Mesh:** $n_{\text{el}} = 2064$
- **Goal:** no stability issue, preservation of water-height positivity

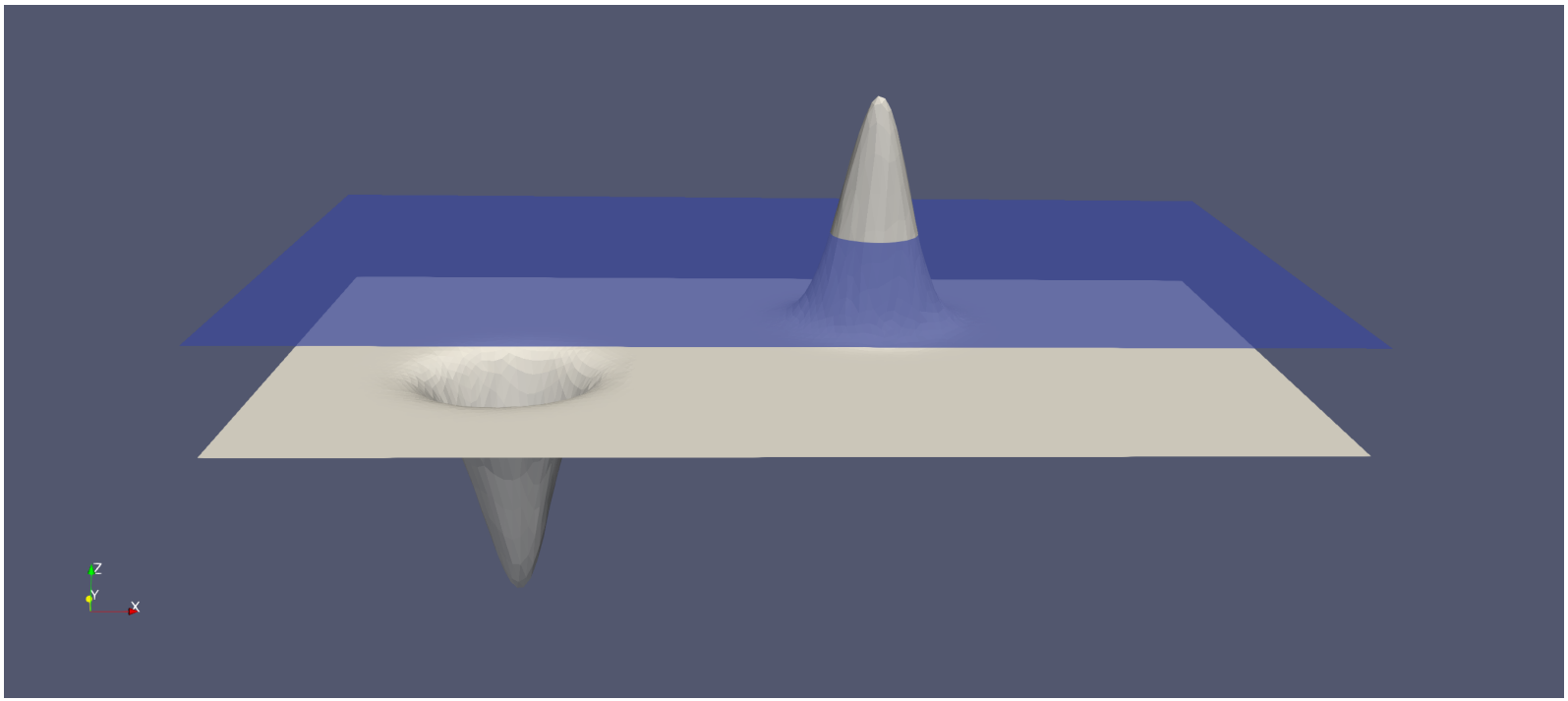


Figure: $\mathbb{P}^4$ initial solution.

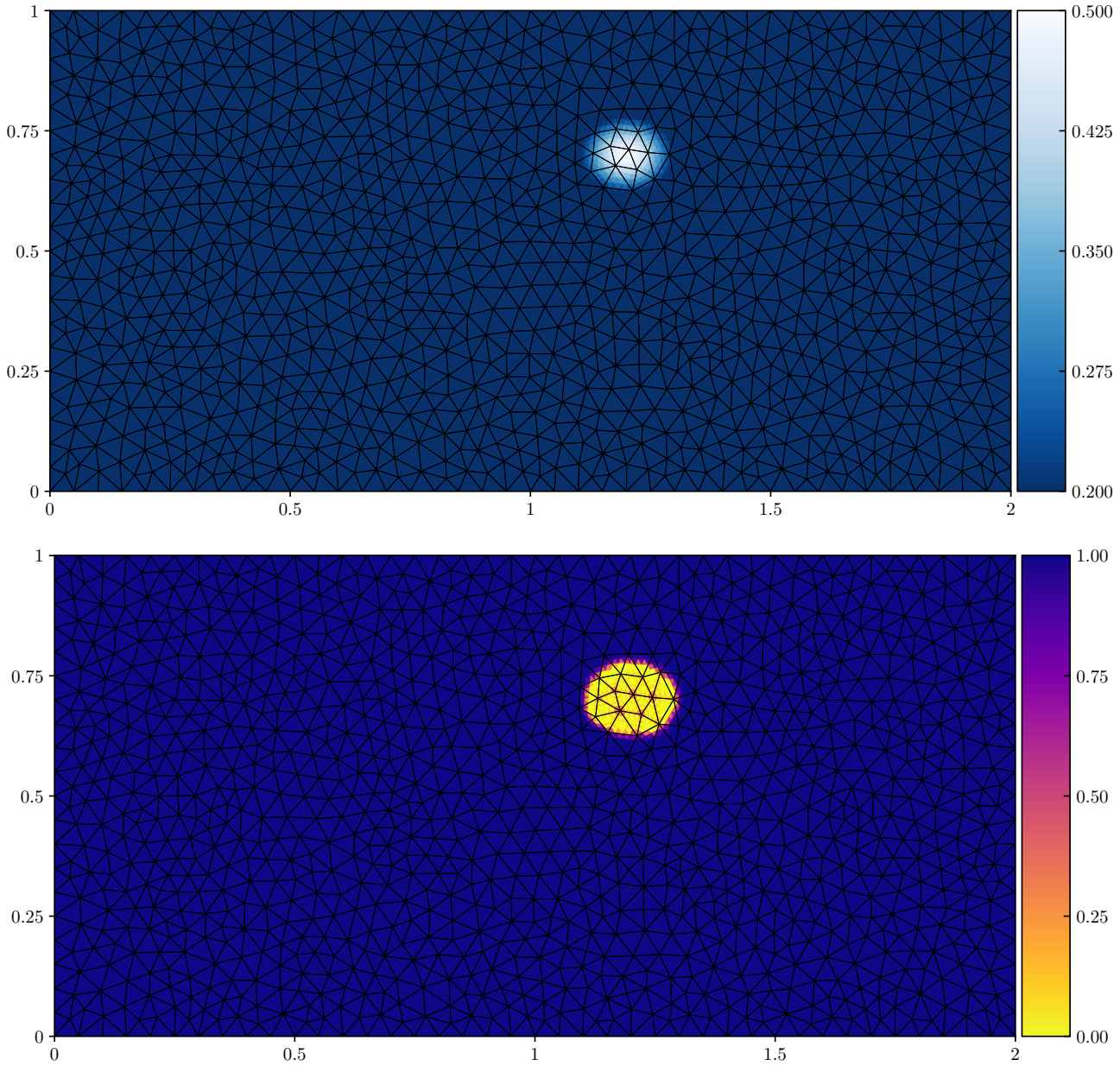


Figure: At $t = 20$ sec, $\mathbb{P}^4$ elevation (top) and map of blending coefficient means per subcell (bottom).

Test 3 – Dam-break problems (1)

Dam-break on a wet bed

- **Domain:** $\Omega = [0, 1000] \times [0, 200]$ **Degree:** $k = 4$ **Mesh:** $n_{\text{el}} = 350$
- **Goal:** handling shock waves and rarefaction fronts

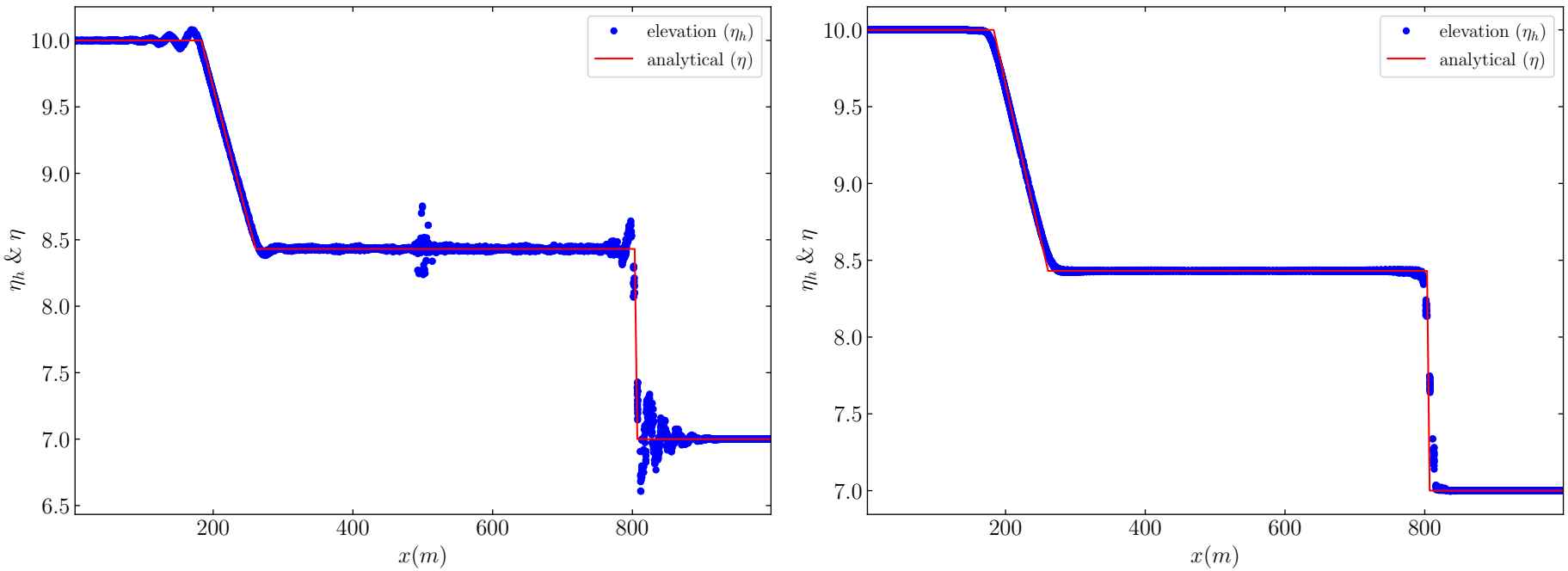


Figure: At $t = 32$ sec, $\mathbb{P}^4$ pure DG elevation (left) and monolithic DG/FV subcells elevation (right).

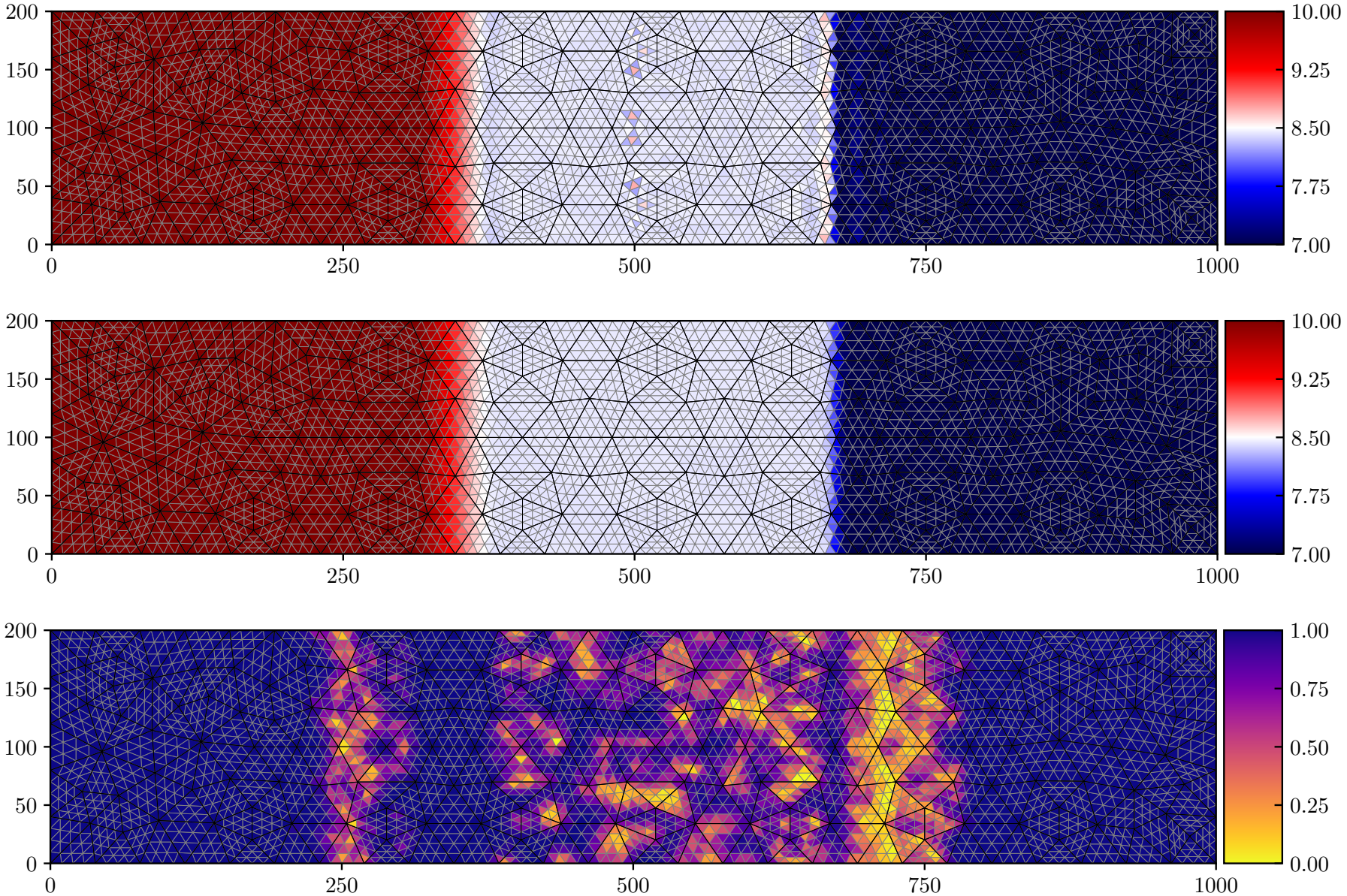


Figure: At $t = 18$ sec, $\mathbb{P}^4$ unlimited DG elevation (top), monolithic DG/FV subcells elevation (center) and map of blending coefficient means per subcell (bottom).

Test 3 – Dam-break problems (2)

Dam-break on a dry bed with friction

- **Domain:** $\Omega = [0, 1000] \times [0, 200]$ **Degree:** $k = 3$ **Mesh:** $n_{\text{el}} = 350$
- **Goal:** treating wet/dry interfaces, supplemented with friction

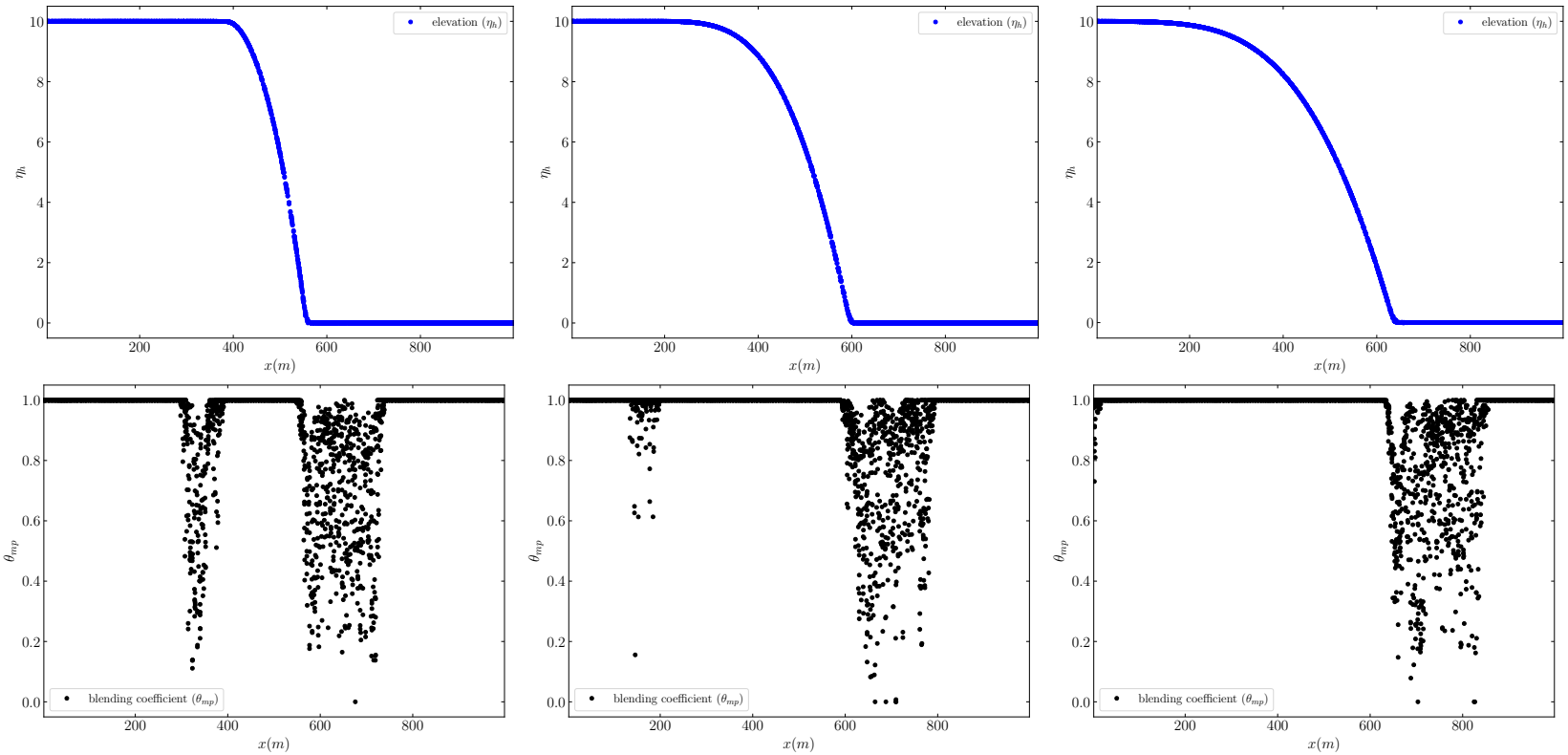


Figure: Snapshots of $\mathbb{P}^3$ free surface elevation and blending density profiles for $t \in [10, 60]$ sec for $k_f = 0.5$.

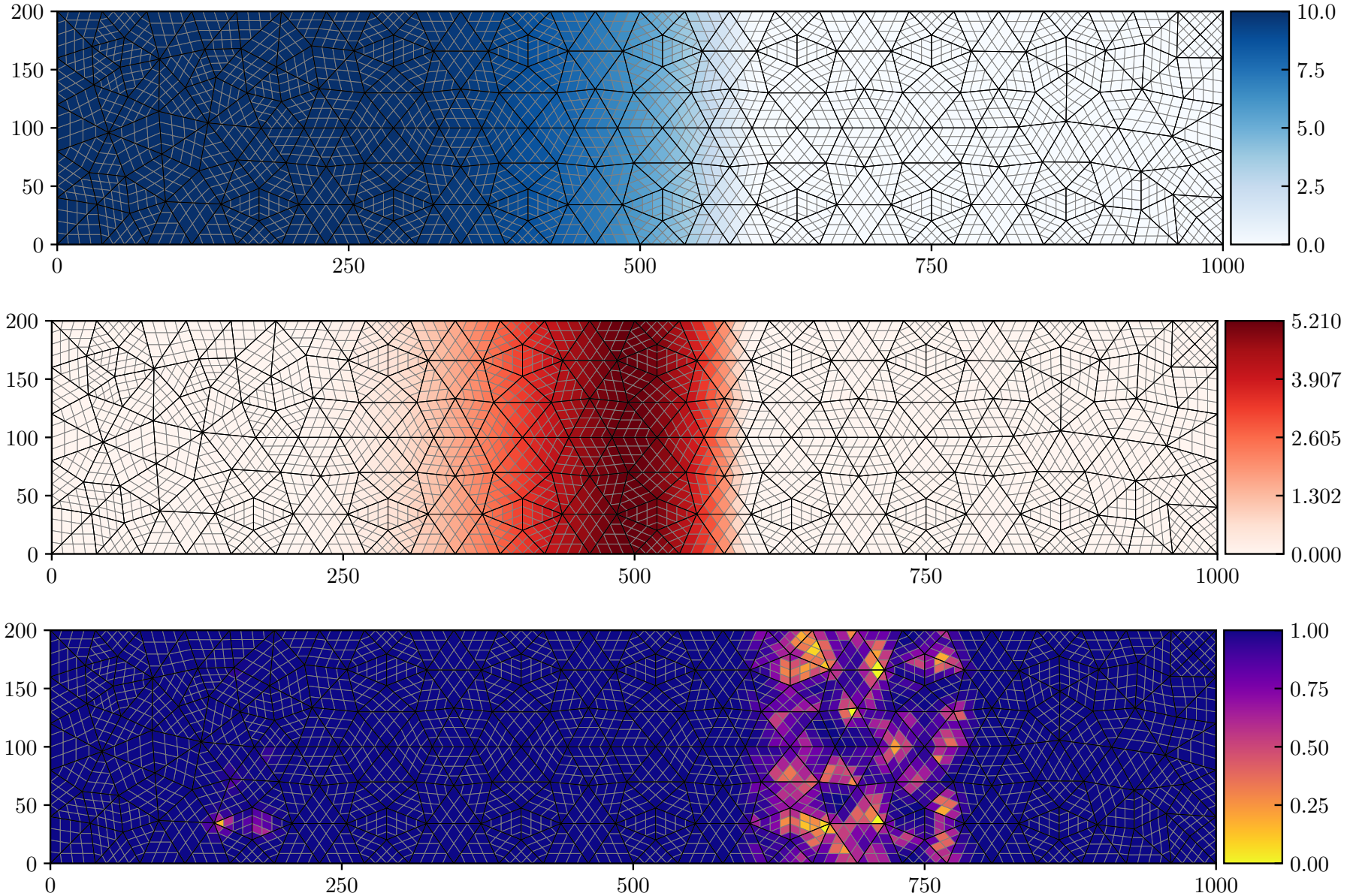


Figure: At $t = 30$ sec, $\mathbb{P}^3$ elevation (top), discharge norm (center) and map of blending coefficient means per subcell (bottom).

Test 4 – Rock-wave interactions

Single wave collapsing on a Gaussian rock

- **Domain:** $\Omega = [5, 25] \times [0, 30]$ **Degree:** $k = 6$ **Mesh:** $n_{\text{el}} = 584$
- **Goal:** assessing robustness and correct shock-capturing in challenging case

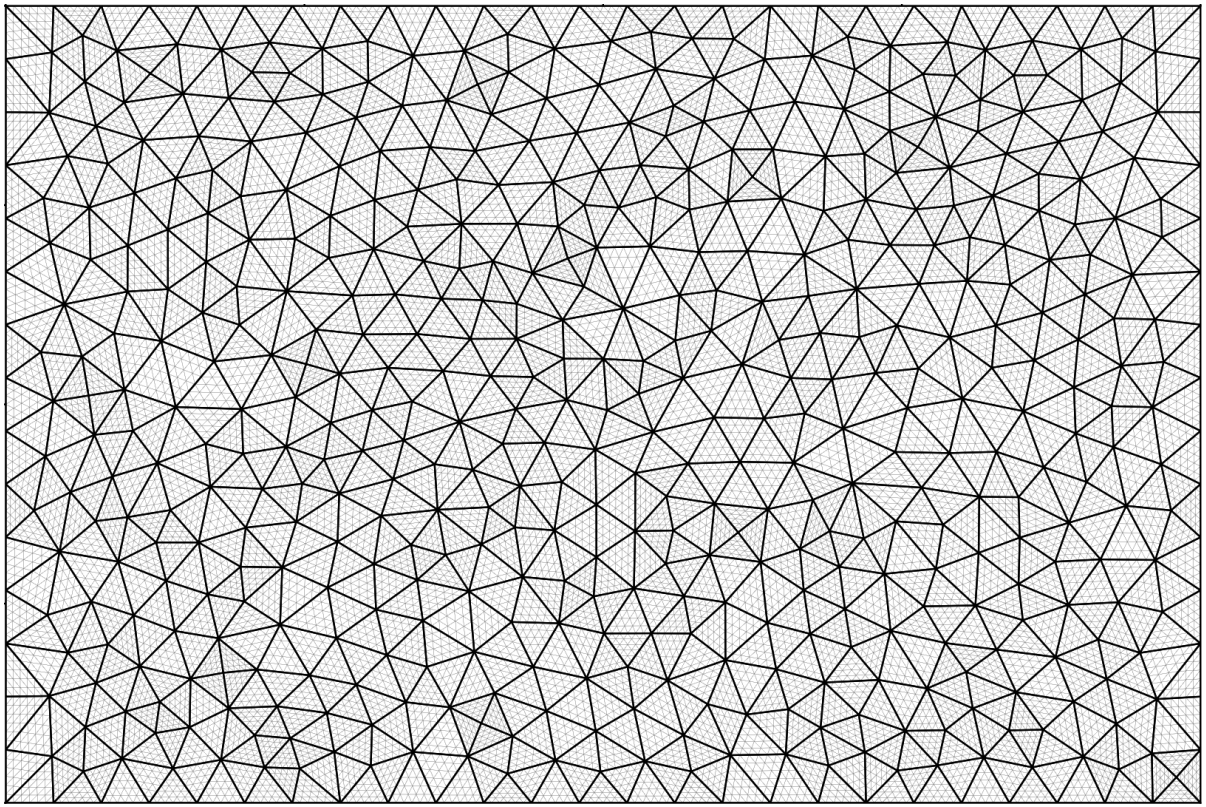


Figure: Unstructured simplicial mesh $\mathbb{P}^6$ subdivision onto triangles with $n_{\text{el}} = 584$ cells.

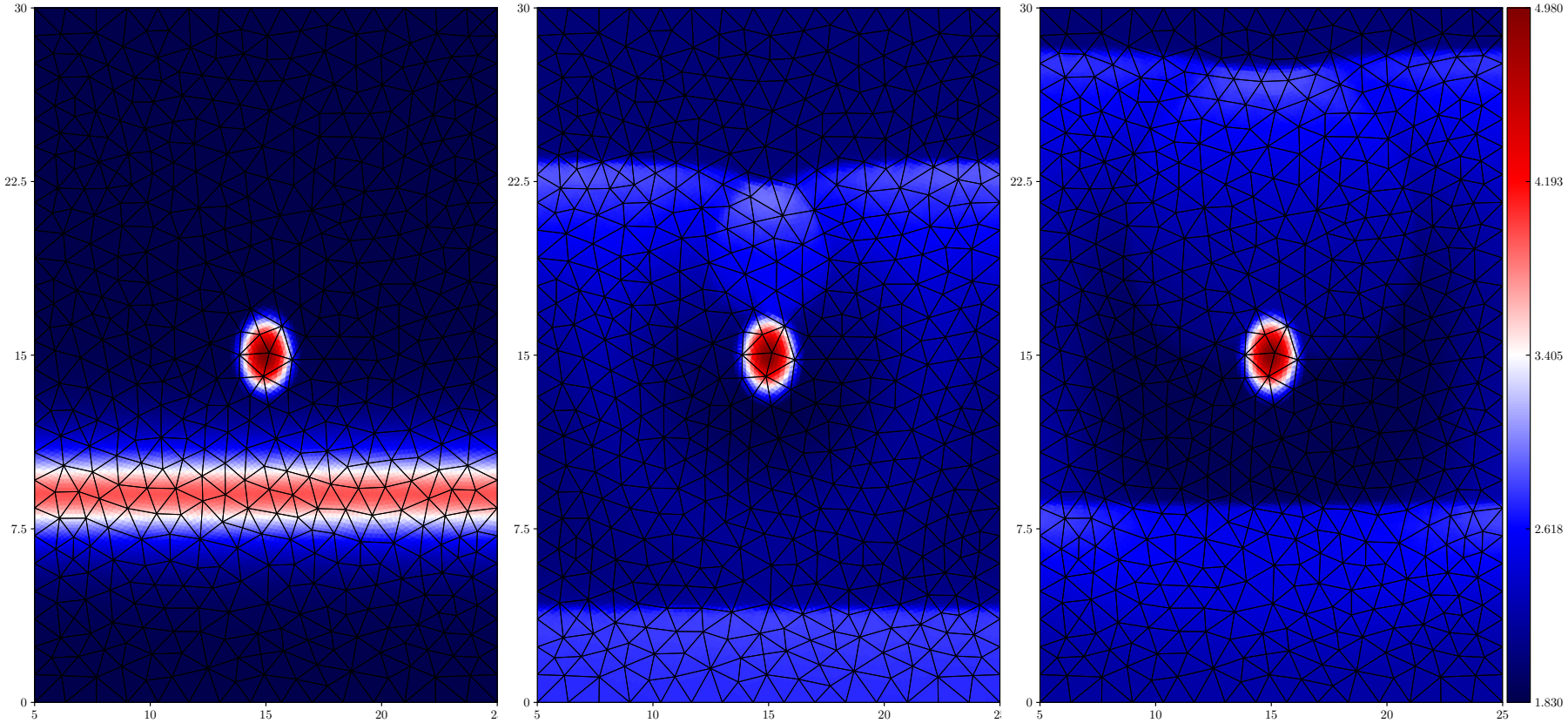


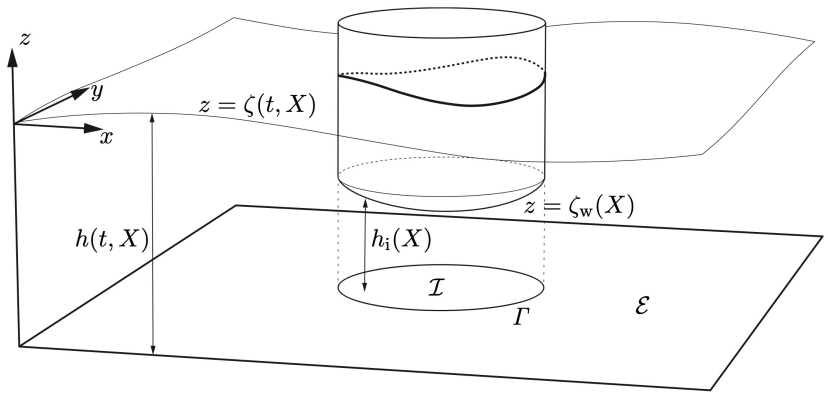
Figure: Snapshots of $\mathbb{P}^6$ elevation at several times (and link to simulation).

Table of contents

Ph.D. objectives

We want an ideal scheme to solve the Nonlinear Shallow Water (NSW) equations, such that we can then study:

wave-structure interactions



From the theory...

to its potential applications...

Ongoing and upcoming work

What has been done...

S.C., A. Haidar, F. Marche & F. Vilar, *Monolithic DG-FV subcell schemes for nonlinear hyperbolic system with source terms. Applications to shallow water asymptotics*. In preparation. 2025.

S.C., F. Marche & F. Vilar, *Local monolithic DG-FV subcell scheme for 2D NSW on unstructured grids*. In preparation. 2025.

... and what remains!

- Designing a mixed **HHO/DG-FV subcells** method for **wave-structure interactions**;
- Adaptation of the method to **moving** or **deforming** meshes via an **ALE framework**;
- Extension to **Green-Naghdi equations** in 2D case.

Thank you for your attention!

Figure: *The Great Wave of Kanagawa*, Hokusai, 1830.

E-mail: *sacha.cardonna@umontpellier.fr*
Website: *sachacardonna.github.io*

Remark about quadrature on subcells

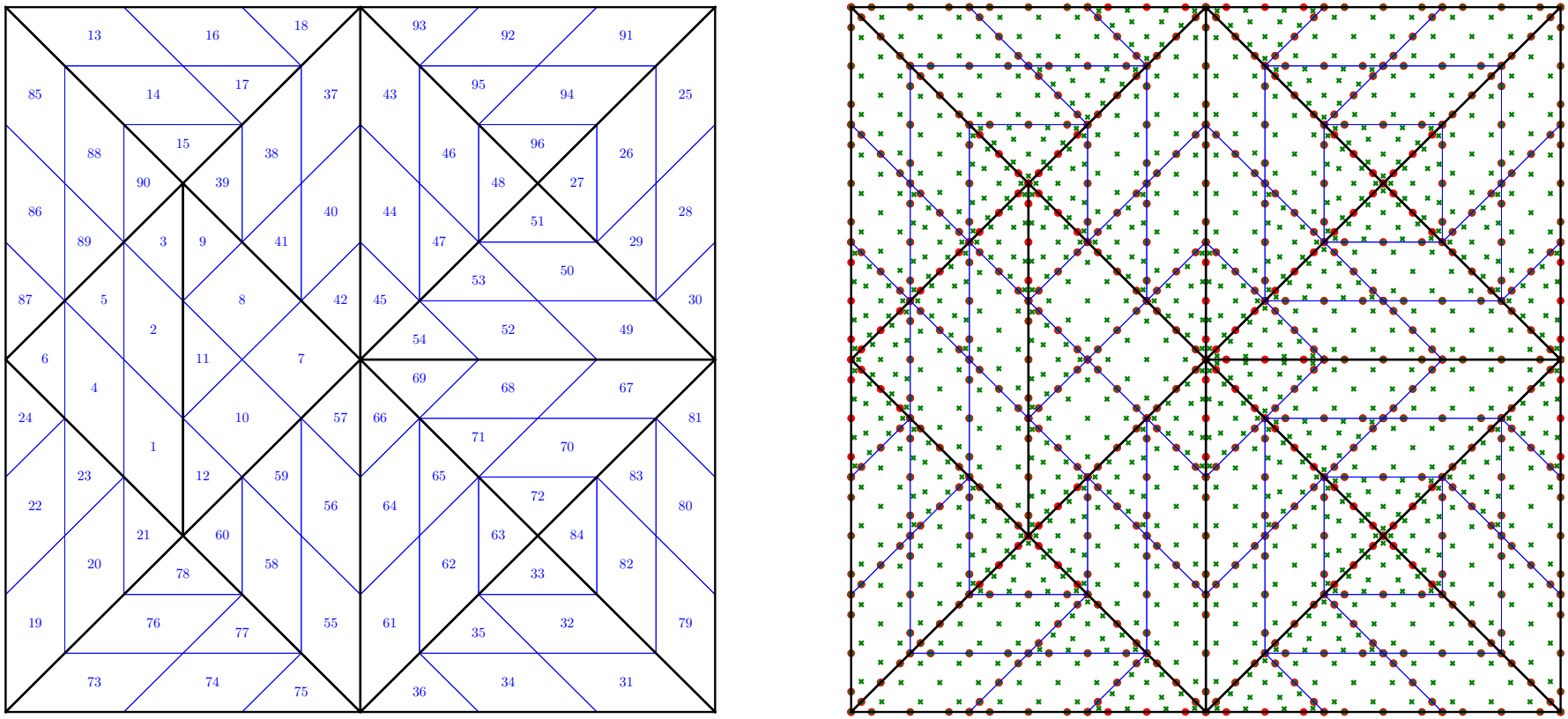


Figure: Subdivision of a coarse mesh into subcells with their global numbering (left), alongside the quadrature points for subcell interiors and faces (right).

Remark about initialization

Initialization strategy

Initialization is performed via **subcell averages** followed by projection using $\mathbb{P}_c$, instead of L^2 projection or interpolation as usually done in DG schemes

$\hookrightarrow$ this guarantees $\mathbf{v}_h \in \mathcal{H}^+$ at $t = 0$, and enforces $\eta_h = b_h$ in dry zones

⚠ Since b_h is discontinuous across cells, **hydrostatic reconstruction** is applied to both DG and subcell FV fluxes.

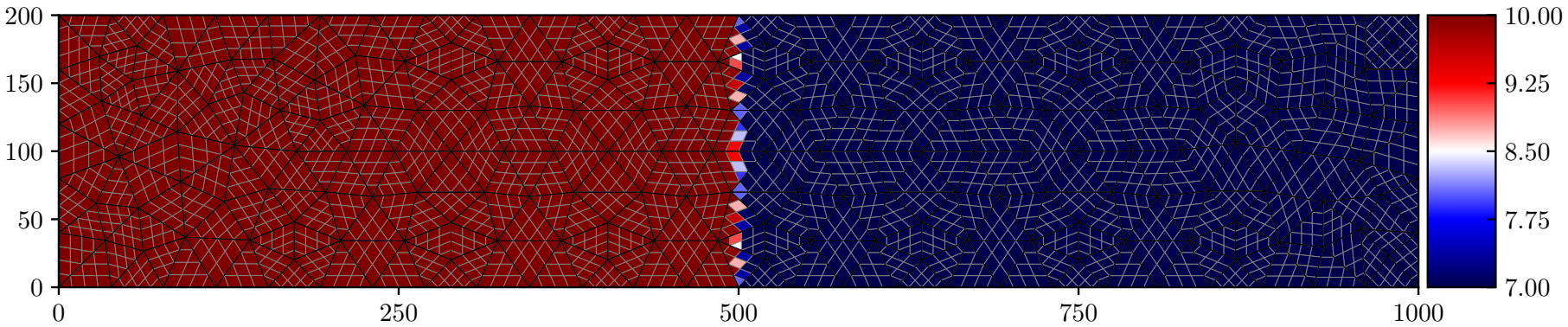


Figure: $\mathbb{P}^3$ dam-break problem initialization.

Remark about hydrostatic reconstruction

Assuring both WB and positivity in numerical fluxes

❓ Hydrostatic reconstruction framework used on both DG and subcell FV fluxes
$\hookrightarrow$ ensures **positivity** of the water height, even for discontinuous topography

At each interface $\Gamma_{cv(k)}$ (resp. subinterface $\Gamma_{mp(k)}$), reconstructed values are defined:

- Topography rec.: $\widetilde{b}_k = \max(b_k^-, b_k^+), \quad \check{b}_k = \widetilde{b}_k - \max(0, \widetilde{b}_k - \eta_k^-)$
- Water height/elevation rec.: $\check{H}_k^\pm = \max(0, \eta_k^\pm - \widetilde{b}_k), \quad \check{\eta}_k^\pm = \check{H}_k^\pm + \widetilde{b}_k$
- Modified states: $\check{\mathbf{v}}_k^\pm = \left(\check{\eta}_k^\pm, \frac{\check{H}_k^\pm}{H_k^\pm} \mathbf{q}_k^\pm \right)^t$

These are then used in a Lax-Friedrichs-type flux $\mathbb{F}^*$, completed by a correction term $\check{\mathbb{F}}_{cv(k)}$ to ensure well-balancing:

$$\mathbb{F}^*_{cv(k)} = \mathbb{F}^*(\check{\mathbf{v}}_k^-, \check{\mathbf{v}}_k^+, \check{b}_k, \check{b}_k, \boldsymbol{n}_{cv(k)}) + \check{\mathbb{F}}_{cv(k)}$$

Remark about source term treatment

Alternative discretization of the source term

$$\boxed{\overline{\mathbf{B}}_m^c = \overline{\mathbf{B}}_m^{c,\mathrm{FV}} + \theta_m^c \left(\overline{\mathbf{B}}_m^{c,\mathrm{DG}} - \overline{\mathbf{B}}_m^{c,\mathrm{FV}} \right)}$$

- $\theta_m^c = \dfrac{1}{\#\mathcal{V}_m^c} \displaystyle\sum_{S_p^v \in \mathcal{V}_m^c} \theta_{mp}$ **subcell global blending**
- $\overline{\mathbf{B}}_m^{c,\mathrm{DG}} = \dfrac{1}{|S_m^c|} \displaystyle\int_{S_m^c} \mathbf{B}_h \, d\boldsymbol{x}$ **DG source term**
- $\overline{\mathbf{B}}_m^{c,\mathrm{FV}} = \dfrac{1}{|S_m^c|} \displaystyle\int_{S_m^c} \mathbf{B}(\overline{\mathbf{v}}_m^c, \nabla_{\boldsymbol{x}} b_h^c) \, d\boldsymbol{x}$ **FV source term**

💬 No significant difference in results → we keep $\overline{\mathbf{B}}_m^c = \overline{\mathbf{B}}_m^{c,\mathrm{DG}}$

Remark about blending smoothening

Why smoothening blending coefficient?

A **sharp switch** between low and high-order fluxes (i.e., $\theta_{mp} = 0$ vs. $\theta_{mp} = 1$) may cause **local oscillations**
$\hookrightarrow$ blending smoothers designed to **mitigate abrupt transitions**

- **Mean-value smoother** (default in experiments):

$$\theta_m^c = \frac{1}{\#\mathcal{V}_m^c} \sum_{S_p^v \in \mathcal{V}_m^c} \theta_{mp}, \quad \widetilde{\theta}_{mp} = \min\left(\theta_{mp}, \frac{1}{\#\mathcal{V}_{mp}} \sum_{S_q^v \in \mathcal{V}_{mp}} \theta_q^v\right)$$

  $\hookrightarrow$ Less diffusive, smoother transitions

- **Minimum-value smoother**:

$$\theta_m^c = \min_{S_p^v \in \mathcal{V}_m^c} \theta_{mp}, \quad \widetilde{\theta}_{mp} = \min\left(\theta_{mp}, \min_{S_q^v \in \mathcal{V}_{mp}} \theta_q^v\right)$$

  $\hookrightarrow$ Stronger damping near discontinuities