



**HAL**  
open science

# Efficient preconditioners for solving dynamical optimal transport via interior point methods

Enrico Facca, Gabriele Todeschi, Andrea Natale, Michele Benzi

► **To cite this version:**

Enrico Facca, Gabriele Todeschi, Andrea Natale, Michele Benzi. Efficient preconditioners for solving dynamical optimal transport via interior point methods. 2022. hal-03766668v1

**HAL Id: hal-03766668**

**<https://inria.hal.science/hal-03766668v1>**

Preprint submitted on 1 Sep 2022 (v1), last revised 22 Jan 2024 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# EFFICIENT PRECONDITIONERS FOR SOLVING DYNAMICAL OPTIMAL TRANSPORT VIA INTERIOR POINT METHODS

---

**Enrico Facca**

Univ. Lille, Inria, CNRS, UMR 8524  
Laboratoire Paul Painlevé  
Lille, France  
enrico.facca@inria.fr

**Gabriele Todeschi**

Univ. Grenoble Alpes, ISTERRE  
F-38058 Grenoble, France  
gabriele.todeschi@univ-grenoble-alpes.fr

**Andrea Natale**

Univ. Lille, Inria, CNRS, UMR 8524  
Laboratoire Paul Painlevé  
Lille, France  
andrea.natalea@inria.fr

**Michele Benzi**

Scuola Normale Superiore  
Piazza dei Cavalieri, 7, 56126  
Pisa, Italy  
michele.benzi@sns.it

## ABSTRACT

In this paper we address the numerical solution of the quadratic optimal transport problem in its dynamical form, the so-called Benamou-Brenier formulation. When solved using interior point methods, the main computational bottleneck is the solution of large saddle point linear systems arising from the associated Newton-Raphson scheme. The main purpose of this paper is to design efficient preconditioners to solve these linear systems via iterative methods. Among the proposed preconditioners, we introduce one based on the partial commutation of the operators that compose the dual Schur complement of these saddle point linear systems, which we refer as  $\mathcal{BB}$ -preconditioner. A series of numerical tests show that the  $\mathcal{BB}$ -preconditioner is the most efficient among those presented, with a CPU-time scaling only slightly more than linearly with respect to the number of unknowns used to discretize the problem.

**Keywords** Optimal transport · Benamou-Brenier formulation · Saddle point problem · Algebraic multigrid methods · Preconditioners

## 1 Introduction

Optimal transport deals with the problem of finding the optimal way to reallocate one nonnegative density into another by minimizing the total cost of displacement in space. In recent years, numerous contributions have been made to the study of this problem, both on the theoretical and computational level. We suggest, for example, the monographs [1, 2, 3, 4] for a detailed presentation of the subject. Due to these advances, optimal transport is nowadays an established tool for many applications including, for example, the analysis of partial differential equations (PDE) [5], physical modeling [2], data science and machine learning [4], economics [6], and inverse problems [7].

When the cost of displacement per unit mass is given by the square of the Euclidean distance, the problem can be recast dynamically, as shown by Benamou and Brenier [8]. Consider a compact and convex domain  $\Omega \subset \mathbb{R}^d$  and two nonnegative densities  $\rho^{\text{in}}$  and  $\rho^{\text{f}}$  in  $L^1(\Omega)$ , with  $\int_{\Omega} \rho^{\text{in}} dx = \int_{\Omega} \rho^{\text{f}} dx$ . In order to transport the former to the latter, we aim to find a time-dependent density  $\rho : [0, 1] \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  and a velocity field  $v : [0, 1] \times \Omega \rightarrow \mathbb{R}^d$  that solve the

following minimization problem:

$$\min_{\rho, v} \int_0^1 \int_{\Omega} \frac{\rho |v|^2}{2} : \begin{cases} \partial_t \rho + \operatorname{div}(\rho v) = 0 & \text{in } [0, 1] \times \Omega \\ \rho v \cdot \hat{n} = 0 & \text{on } [0, 1] \times \partial\Omega \\ \rho(0, \cdot) = \rho^{\text{in}}, \rho(1, \cdot) = \rho^{\text{f}} \end{cases} \quad (1)$$

The total kinetic energy represents the cost of displacement. Thanks to a change of variables  $(\rho, v) \rightarrow (\rho, m = \rho v)$ , (1) can be rewritten as a convex optimization problem. From the optimality conditions of the problem, one can deduce that the optimal velocity field is the gradient of a potential  $\phi : [0, 1] \times \Omega \rightarrow \mathbb{R}$ . The potential and the density itself are given as the solution of the following system of PDEs:

$$-\partial_t \rho - \operatorname{div}(\rho \nabla \phi) = 0, \quad (2a)$$

$$\partial_t \phi + \frac{\|\nabla \phi\|^2}{2} + s = 0, \quad (2b)$$

$$\rho \geq 0, s \geq 0, \rho s = 0, \quad (2c)$$

with boundary conditions  $\rho(0, \cdot) = \rho^{\text{in}}, \rho(1, \cdot) = \rho^{\text{f}}, \rho \nabla \phi \cdot \hat{n} = 0$  on  $[0, 1] \times \partial\Omega$ . The auxiliary variable  $s : [0, 1] \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  is related to the positivity constraint on  $\rho$ .

The Benamou-Brenier formulation offers several advantages. In addition to the total displacement cost, it also directly provides how to continuously reallocate the mass. This also constitutes a natural way of defining interpolations between densities. Furthermore, it draws a clear link between optimal transport and continuum mechanics, and it is naturally suited for Eulerian discretizations. Finally, it can be easily generalized to other problems by penalizing/constraining the evolution  $\rho$  (such as in variational mean-field games and planning problems [9], or unbalanced optimal transport [10]). On the other hand, the numerical solution of (1), or its system of optimality conditions (2), poses significant challenges. Although it is a convex optimization problem, it is nonlinear and, in general, nonsmooth for vanishing densities. Moreover, it is a time-space boundary value problem and there is a positivity constraint on  $\rho$  to take into account. In the original paper [8], a strategy based on centered finite differences and the Alternating Direction Method of Multipliers (ADMM) was used. Since then, more works in this direction have followed, using finite differences or finite elements and more general proximal splitting techniques [11, 12, 13, 14]. These approaches are extremely robust but not particularly efficient, especially when high accuracy is required and nonregular grids are used.

In [15] the problem has been discretized using finite volumes, following a "first discretize, then optimize" approach in order to preserve the variational structure, and solved using an Interior Point (IP) method [16, 17]. The optimization problem is relaxed by adding a logarithmic barrier function scaled by a parameter  $\mu > 0$ , that is,  $-\mu \int_0^1 \int_{\Omega} \log(\rho)$ . This perturbation provides smoothness by enforcing the strict positivity of the density and uniqueness of the solution. The problem can then be effectively solved by solving the relaxed optimality conditions via the Newton method, and the original unperturbed solution is retrieved by repeating this procedure while reducing the relaxation term  $\mu$  to zero. The combination of the finite volume discretization and the use of the IP method allows one to tackle general domain discretizations and solve efficiently the problem up to high accuracy. The authors furthermore considered a two-level discretization of the domain  $\Omega$ , in order to discretize the density  $\rho$  and the potential  $\phi$  separately, which alleviates some checkerboard instabilities that may appear when the same grid is used to discretize both variables. The same issue has been pointed out in [18, 19] when dealing with optimal transport with unitary displacement cost given by the Euclidean distance.

From a computational point of view, the most demanding task in [15] is to solve a sequence of saddle point linear systems in the following form

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ \mathcal{B} & -\mathcal{C} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix},$$

generated by the Newton method. The goal of this paper is to find an efficient preconditioning technique to solve these linear systems using iterative methods. The literature on this problem is relatively scarce. The most closely related work is [20], where the authors study linear algebra approaches involved in the solution of mean-field games systems via the Newton method. Our problem may in fact be seen as the vanishing viscosity limit of a particular mean-field game. However, their linear algebra approaches do not fit into the discretization scheme used in [15]. Moreover, they lose efficiency for vanishing viscosities, that is when the two problems are more similar.

In this paper we consider four preconditioners: the Hermitian skew-Hermitian splitting (HSS) preconditioner [21], a preconditioner based on the approximation of the inverse of the primal Schur complement  $\mathcal{S}_p = \mathcal{A} + \mathcal{B}\mathcal{C}^{-1}\mathcal{B}^T$ , the SIMPLE preconditioner [22], and one preconditioner inspired by the ideas in [23], where we approximate the

inverse of the dual Schur complement  $\mathcal{S}_d = -\mathcal{C} - \mathcal{B}\mathcal{A}^{-1}\mathcal{B}^T$  looking at the differential nature of the operators that compose  $\mathcal{S}_d$  (note that, formally, we cannot write  $\mathcal{A}^{-1}$  since in our problem the matrix  $\mathcal{A}$  is singular). A series of numerical experiments suggest that the latter, named  $\mathcal{B}\mathcal{B}$ -preconditioner, is the most efficient among those presented in this paper.

The rest of the paper is structured as follows. In section 2 we summarize the discretization method proposed in [15]. Then in section 3 we recall the IP approach used to solve the optimization problem and the nonlinear and linear problems associated with it. Finally, in section 4 we present the preconditioners described in this paper, together with some numerical experiments where we solve a specific test case for different time and space refinements. In particular, in section 4.6 we compare the CPU time required by the preconditioners in solving different test cases.

## 2 Discrete setting

In this section, we summarize the discretization considered in [15] that combines the use of finite volumes with staggered temporal grids, so we can write the discrete counterpart of system eq. (2).

### 2.1 Spatial discretization

In [15], the authors considered two different discretizations of the domain  $\Omega$ , which is assumed to be polygonal, both admissible for Two Point Flux Approximation (TPFA) finite volumes, according to [24, Definition 9.1]. At each time  $t \in [0, 1]$ , the variable  $\rho(t)$  is discretized on a Delaunay triangulation, with the further hypothesis that only acute angles appear. The variable  $\phi(t)$  is discretized on a finer grid obtained from the previous one by dividing each triangular cell into three quadrilateral cells, joining the edges midpoints to the triangle's circumcenter [15, Figure 1]. Both meshes consist of two sets  $(\mathcal{T}, \mathcal{E})$ , the set of cells  $c$  and edges  $e$ , respectively. To distinguish the coarser mesh from the finer one, we denote the former by  $(\mathcal{T}', \mathcal{E}')$ , with the same notation for all its elements. Since the discrete model involves two different spatial discretizations, we introduce the injection operator  $\mathbf{J} \in \mathbb{R}^{N_{\mathcal{T}}, N_{\mathcal{T}'}}$

$$\mathbf{J}[i, j] = \begin{cases} 1 & \text{if } c_i \subset c'_j, \\ 0 & \text{else.} \end{cases}$$

Due to the no flux boundary condition, boundary edges are not relevant to the discrete model. We will then consider by convention the sets  $\mathcal{E}$  and  $\mathcal{E}'$  without boundary edges. Let us denote by  $N_{\mathcal{T}'}$  and  $N_{\mathcal{E}'}$  the total number of cells and (internal) edges of the coarser mesh. The total number of cells and edges of the finer mesh are then  $N_{\mathcal{T}} = 3N_{\mathcal{T}'}$  and  $N_{\mathcal{E}} = 2N_{\mathcal{E}'} + 3N_{\mathcal{T}'}$ .

In order to define the discrete differential operators associated to the finite volume discretization, we fix an arbitrary orientation on the set of edges  $\mathcal{E}$  and define the matrix  $\mathbf{E} \in \mathbb{R}^{N_{\mathcal{T}}, N_{\mathcal{E}}}$  given by

$$\mathbf{E}[i, k] = \begin{cases} 1 & \text{if cell } c_i \text{ is the left side of edge } e_k, \\ -1 & \text{if cell } c_i \text{ is the right side of edge } e_k. \end{cases}$$

The discrete gradient and divergence  $\nabla \in \mathbb{R}^{N_{\mathcal{E}}, N_{\mathcal{T}}}$  and  $\text{div} \in \mathbb{R}^{N_{\mathcal{T}}, N_{\mathcal{E}}}$  are given by

$$\begin{aligned} \nabla &= \nabla_{\mathcal{E}, \mathcal{T}} = \text{Diag}(|\mathbf{w}|)^{-1} \mathbf{E}^T, \\ \text{div} &= \text{div}_{\mathcal{T}, \mathcal{E}} = -\nabla^T \text{Diag}(|\mathbf{w}||e|) = -\mathbf{E} \text{Diag}(|e|), \end{aligned}$$

where  $|\mathbf{w}| \in \mathbb{R}^{N_{\mathcal{E}}}$  and  $|e| \in \mathbb{R}^{N_{\mathcal{E}}}$  are the vectors of distances between the circumcenters of adjacent cells and the lengths of the edges, respectively. Moreover, we define the mass matrices as

$$\begin{aligned} \mathbf{M} &:= \text{Diag}(|c|), \quad |c| := (|c_i|)_{i=1}^{N_{\mathcal{T}}}, \\ \mathbf{M}' &:= \text{Diag}(|c'|), \quad |c'| := (|c'_i|)_{i=1}^{N_{\mathcal{T}'}} \end{aligned}$$

where  $|c_i|$  denotes the area of the triangle  $c_i$ .

The TPFA finite volume discretization and the two-level grids require the introduction of two additional operators to match the dimension between the different spaces. The first is a reconstruction operator  $\mathbf{R}_{\mathcal{E}} : \mathbb{R}^{N_{\mathcal{T}'}} \rightarrow \mathbb{R}^{N_{\mathcal{E}}}$ , mapping a positive density  $\rho$  defined on the cells of the coarse grid to a variable defined on the edges of the finer one. In [15] this reconstruction first lifts the density  $\rho$  into the finer space via the operator  $\mathbf{J}$  and then averages the values of adjacent cells. The authors considered two types of averages, a weighted arithmetic mean and a weighted harmonic mean. Using the former, the linear reconstruction operator explicitly writes

$$(\mathbf{R}_{\mathcal{E}}\rho)_e := \lambda_e(\mathbf{J}\rho)_{c_i} + (1 - \lambda_e)(\mathbf{J}\rho)_{c_j}$$

for the two cells  $c_i, c_j$  sharing the edge  $e$ . Moreover, they introduced a further operator  $\mathbf{R}_{\mathcal{T}} : \mathbb{R}^{N_{\mathcal{E}}} \rightarrow \mathbb{R}^{N_{\mathcal{T}'}}$  that maps the variables defined at the edges of the finer grid to the variables defined at the cells of the coarser one. To preserve the variational structure of the discrete problem, it is defined as

$$\mathbf{R}_{\mathcal{T}} := (\mathbf{R}_{\mathcal{E}})^T \text{Diag}(|\mathbf{w}||e|).$$

**Remark 1.** *In this study, we do not consider the harmonic reconstruction introduced in [15] in order to avoid complicating the exposition. From numerical experiments, this nonlinearity does not add any further difficulty from a linear algebra point of view. See also remark 5.*

## 2.2 Temporal discretization

In [15], the temporal discretization is based on staggered temporal grids. The time interval  $[0, 1]$  is divided into  $K + 1$  sub-intervals  $[t^k, t^{k+1}]$  of equal length  $\Delta t = 1/(K + 1)$ , for  $k = 0, \dots, K \geq 1$ , with  $t^0 = 0$  and  $t^{K+1} = 1$ . The variables  $\rho$  and  $s$  are defined at time  $t^k$  for  $k = 1, \dots, K$ , while  $\phi$  is discretized at each instant  $(t^k + t^{k+1})/2$  for  $k = 0, \dots, K$ .

Combining this temporal and spatial discretization, the discrete counterpart of the potential  $\phi$ , the density  $\rho$ , and the slack variable  $s$  in equation eq. (2) are the vectors  $\phi \in \mathbb{R}^n$ ,  $\rho \in \mathbb{R}_{\geq 0}^m$ , and  $\mathbf{s} \in \mathbb{R}_{\geq 0}^m$  with  $n = N_{\mathcal{T}}(K + 1)$  and  $m = N_{\mathcal{T}'}K$  given by

$$\begin{aligned} \phi &= (\phi^1; \dots; \phi^{K+1}), \quad \phi^k \in \mathbb{R}^{N_{\mathcal{T}}}, \\ \rho &= (\rho^1; \dots; \rho^K), \quad \rho^k \in \mathbb{R}_{\geq 0}^{N_{\mathcal{T}'}} , \\ \mathbf{s} &= (\mathbf{s}^1; \dots; \mathbf{s}^K), \quad \mathbf{s}^k \in \mathbb{R}_{\geq 0}^{N_{\mathcal{T}'}} , \end{aligned}$$

where we use the symbol  $;$  to denote the concatenation of vectors. Moreover, we will denote by  $\mathbf{x}^k$  the  $k$ -slice of a concatenated vector  $\mathbf{x}$ , which corresponds to its  $k$ -th time portion. The discrete counterparts of the initial and final densities are the two vectors  $\rho^0$  and  $\rho^{K+1}$  in  $\mathbb{R}^{N_{\mathcal{T}'}}$  given by

$$\rho_i^0 = \frac{1}{|c_{i'}|} \int_{c_{i'}} \rho^{\text{in}} dx, \quad \rho_i^{K+1} = \frac{1}{|c_{i'}|} \int_{c_{i'}} \rho^{\text{f}} dx, \quad i = 1, \dots, N_{\mathcal{T}'}$$

## 2.3 Discrete nonlinear system of equations

Due to the IP strategy, the complementarity constraint between  $\rho$  and  $s$  in eq. (2c) is relaxed by a parameter  $\mu > 0$ . In this way, the two quantities are forced to be strictly positive, making the problem simpler and easier to solve by using an inexact Newton method. The solution of the original problem is recovered using a continuation method with the parameter  $\mu$  going to zero.

At each IP iteration, the optimality conditions for the discrete variational problem are described by the following nonlinear system of equations

$$F_{\phi}(\phi, \rho) := (F_{\phi}^1; \dots; F_{\phi}^{K+1}) = \mathbf{0} \in \mathbb{R}^{N_{\mathcal{T}}(K+1)}, \quad (3a)$$

$$F_{\rho}(\phi, \rho, \mathbf{s}) := (F_{\rho}^1; \dots; F_{\rho}^K) = \mathbf{0} \in \mathbb{R}^{N_{\mathcal{T}'}K}, \quad (3b)$$

$$F_{\mathbf{s}}(\rho, \mathbf{s}) := (F_{\mathbf{s}}^1; \dots; F_{\mathbf{s}}^K) = \mathbf{0} \in \mathbb{R}^{N_{\mathcal{T}'}K}, \quad (3c)$$

where for  $k = 1, \dots, K + 1$  the functions  $F_{\rho}^k$  are given by

$$F_{\phi}^k(\phi, \rho) = -M\mathbf{J} \left( \frac{\rho^k - \rho^{k-1}}{\Delta t} \right) - \text{div} \left( \mathbf{R}_{\mathcal{E}} \left( \frac{\rho^k + \rho^{k-1}}{2} \right) \odot \nabla \phi^k \right), \quad (4)$$

while for  $k = 1, \dots, K$  the functions  $F_{\phi}^k$  and  $F_{\mathbf{s}}^k$  are given by

$$F_{\rho}^k(\phi, \rho, \mathbf{s}) = M' \mathbf{J}^T \left( \frac{\phi^{k+1} - \phi^k}{\Delta t} \right) + \frac{1}{4} \mathbf{R}_{\mathcal{T}} \left( (\nabla \phi^k)^2 + (\nabla \phi^{k+1})^2 \right) + M' \mathbf{s}^k, \quad (5)$$

$$F_{\mathbf{s}}^k(\rho, \mathbf{s}) = \rho^k \odot \mathbf{s}^k - \mu \mathbf{1}. \quad (6)$$

**Remark 2.** *Thanks to the finite volume discretization, the conservative structure of the continuity equation is preserved. Then, due to the no-flux boundary conditions, which are encoded explicitly in the divergence operator, the equation  $F_{\phi}^k(\phi, \rho) = \mathbf{0}$  implies*

$$\mathbf{1}^T F_{\phi}^k = |\mathbf{c}|^T \mathbf{J} (\rho^k - \rho^{k-1}) = |\mathbf{c}'|^T (\rho^k - \rho^{k-1}) = 0 \quad \forall k = 1, \dots, K.$$

*Hence, at each intermediate time step, the discrete mass  $|\mathbf{c}'|^T \rho^k$  is preserved and is equal to the mass of the discrete initial and final densities.*

### 3 Inexact Newton method and linear algebra problems

The nonlinear system of equations eq. (3) is solved using an inexact Newton method. Each Newton iteration requires the solution of a linear system in the form

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T & \\ \mathcal{B} & \text{Diag}(s) & \mathcal{M}' \\ & & \text{Diag}(\rho) \end{pmatrix} \begin{pmatrix} \delta\phi \\ \delta\rho \\ \delta s \end{pmatrix} = \begin{pmatrix} f \\ g \\ h \end{pmatrix} = - \begin{pmatrix} F_\phi \\ F_\rho \\ F_s \end{pmatrix}, \quad (7)$$

where the block matrix in equation eq. (7) is the Jacobian matrix of  $(F_\phi; F_\rho; F_s)$ . We denote by  $\mathcal{M}'$  and  $\mathcal{M}$  the matrices given by

$$\mathcal{M}' = \text{Block Diag}((M')_{k=1}^K), \quad \mathcal{M} = \text{Block Diag}((M)_{k=1}^{K+1}).$$

The matrices  $\mathcal{A}, \mathcal{B}, \mathcal{B}^T$  in eq. (7) are the finite-dimensional version of the following differential operators (up to a multiplication by a mass matrix)

$$\mathcal{A} \approx -\text{div}(\rho \nabla), \quad \mathcal{B} \approx \partial_t + \nabla \phi \cdot \nabla, \quad \mathcal{B}^T \approx -\partial_t - \text{div}(\cdot \nabla \phi).$$

According to the discretization scheme described in section 2, matrix  $\mathcal{A} \in \mathbb{R}^{n,n}$  is a block diagonal matrix with  $K+1$  blocks. Each block is a weighted Laplacian matrix given by

$$A^k = -\text{div} \text{Diag}(\tilde{\rho}^k) \nabla \in \mathbb{R}^{N_\tau, N_\tau}, \quad \tilde{\rho}^k := R_\mathcal{E} \left( \frac{\rho^k + \rho^{k-1}}{2} \right),$$

which we can write equivalently as  $\mathcal{A} = -\text{Div}_x \text{Diag}((\tilde{\rho}^1; \dots; \tilde{\rho}^k; \dots; \tilde{\rho}^{K+1})) \mathcal{D}_x$  where

$$\text{Div}_x = \text{Block Diag}((\text{div})_{k=1}^{K+1}), \quad \mathcal{D}_x = \text{Block Diag}((\nabla)_{k=1}^{K+1}). \quad (8)$$

Matrix  $\mathcal{B}$  is a block bidiagonal matrix in  $\mathbb{R}^{m,n}$  given by

$$\mathcal{B} = \mathcal{M}' \mathcal{D}_t \mathcal{J}^T + \mathcal{H} \mathcal{G} \mathcal{D}_x,$$

where the matrices  $\mathcal{D}_t, \mathcal{J}, \mathcal{H}$  and  $\mathcal{G}$  are

$$\mathcal{D}_t = \frac{1}{\Delta t} \begin{pmatrix} -I & I & & \\ & \ddots & \ddots & \\ & & -I & I \end{pmatrix}, \quad \mathcal{J} = \text{Block Diag}((J)_{k=1}^{K+1}), \quad (9)$$

$$\mathcal{H} = \frac{1}{2} \begin{pmatrix} I & I & & \\ & \ddots & \ddots & \\ & & I & I \end{pmatrix}, \quad \mathcal{G} = \text{Block Diag}((G^k)_{k=1}^{K+1}), \quad (10)$$

with  $(G^k)_{k=1, \dots, K+1}$  given by

$$G^k := R_\mathcal{T} \text{Diag}(\nabla \phi^k).$$

**Remark 3.** The kernel of the block diagonal matrix  $\mathcal{A}$  has dimension  $K+1$ , since each block  $A^k$  has the constant vectors as kernel. On the other hand,  $\text{Ker}(\mathcal{A})$  is included in  $\text{Ker}(\mathcal{H} \mathcal{G} \mathcal{D}_x)$  and therefore

$$\text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{B}) = \text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{M}' \mathcal{D}_t \mathcal{J}^T) = \langle \mathbf{1} \rangle \in \mathbb{R}^n,$$

thus the Jacobian matrix is singular. This singularity can be removed grounding one entry of the solution  $\delta\phi$ . However, this procedure often affects negatively the performance of the preconditioner used. Therefore we avoid using this approach in this paper, since iterative methods work also when dealing with singular matrices [25].

#### 3.1 Reduction to a saddle point linear system

In order to solve the linear system in eq. (7) it is convenient to eliminate the variable  $\delta s$  writing

$$\delta s = (\text{Diag}(\rho))^{-1} (h - \text{Diag}(s) \delta \rho)$$

and reduce it into the following saddle point system

$$\mathcal{J} \begin{pmatrix} \delta\phi \\ \delta\rho \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ \mathcal{B} & -\mathcal{C} \end{pmatrix} \begin{pmatrix} \delta\phi \\ \delta\rho \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \tilde{\mathbf{g}} \end{pmatrix}, \quad (11)$$

where the matrix  $\mathcal{C} \in \mathbb{R}^{m,m}$  and the vector  $\tilde{\mathbf{g}} \in \mathbb{R}^m$  are given by

$$\mathcal{C} := -\mathcal{M}'\text{Diag}(\rho)^{-1}\text{Diag}(s), \quad \tilde{\mathbf{g}} := \mathbf{g} - \mathcal{M}'\text{Diag}(\rho)^{-1}\mathbf{h}.$$

The matrix  $\mathcal{J}$  in eq. (11) has typically higher conditioning number than the Jacobian matrix in eq. (7), in particular when approaching the optimal solution, i.e. as  $s \odot \rho = \mu\mathbf{1} \rightarrow \mathbf{0}$ . However, designing an efficient preconditioner for the fully coupled system in eq. (7) can be harder than for the standard saddle point linear system in eq. (11), for which different preconditioning approaches are described in [26].

We adopt an Inexact Newton approach, which means that we seek for a solution  $(\delta\phi, \delta\rho)$  such that

$$\|\mathcal{J}(\delta\phi; \delta\rho) - (\mathbf{f}; \tilde{\mathbf{g}})\| \leq \varepsilon_{\text{out}}\|(\mathbf{f}; \mathbf{g}; \mathbf{h})\|.$$

The linear system residual is scaled by the right-hand side of the original system to avoid oversolving issues we faced using  $\|(\mathbf{f}; \tilde{\mathbf{g}})\|$ . In our experiments we use a fixed tolerance  $\varepsilon_{\text{out}} = 1e - 5$ . This value may be quite small compared to the literature for an inexact Newton approach but we want to avoid to undermine the effectiveness of the non linear solver and focus on the linear algebra only.

**Remark 4.** From remark 2, we know the solution  $\rho$  to equation eq. (3) must satisfy  $|\mathbf{c}'|^T \rho^k = |\mathbf{c}'|^T \rho^0$  for all  $k = 1, \dots, K$ . If this condition is satisfied for the starting point of the Newton scheme, then the increment  $\delta\rho$  solution to equations (7) satisfies

$$|\mathbf{c}'|^T \delta\rho^k = 0 \quad \forall k = 1, \dots, K,$$

at each Newton iteration. To see this, sum both sides of the first block of equations of system (11), for each time step  $k$ , and use the no-flux boundary conditions in the divergence operator. Due to our inexact Newton steps, this condition is not verified exactly, but only up to the tolerance chosen. We can nevertheless impose it by renormalizing  $\rho$  after each Newton update.

**Remark 5.** When the harmonic mean is used to reconstruct the density, the block (2, 2) in linear system in eq. (7) is no longer zero and, when the system is reduced, the matrix  $\mathcal{C}$  is no longer diagonal. However, all the preconditioning techniques we present in the following sections can again be applied replacing  $\mathcal{C}$  with its diagonal, without affecting significantly the performances.

## 4 Preconditioning approaches and numerical results

In this section we present four preconditioners to solve the linear system in eq. (11) and a series of numerical experiments used to measure their performance. All of these preconditioners involve the usage of inexact inner solvers, therefore, they may not be linear and must be applied as right preconditioners within a FGMRES cycle [27]. In all cases, the inner solver we adopted is the AGgregation-based Algebraic MultiGrid (AGMG) method described by [28].

We adopt three measures to evaluate the preconditioner performance for each Newton cycle associated to each IP iteration.

1. **Outer/Lin.sys.:** the average number of outer FGMRES iterations, calculated as the number of outer iterations divided by the number of linear systems solved for each Newton cycle.
2. **CPU/Lin.sys.:** the average CPU time (measured in seconds) required for the solution of the linear systems, computed as the total CPU time required to solve all linear systems within a Newton cycle divided by the number of linear systems solved (all experiments were conducted single core, on a machine equipped with a Intel® Xeon® 2.8 GHz CPU and 128 GiB of RAM memory).
3. **Inner/Outer:** the average number of inner iterations per preconditioner application. This number is computed by dividing the cumulative number of AGMG iterations required to solve the linear systems involved in the preconditioner application (for each preconditioner, we will specify which linear systems we will refer to) by the total number of outer iterations for each Newton cycle.

### 4.1 Test cases

The IP algorithm and the linear solver strategies described in the next sections are tested on three numerical experiments taken from [15, Section 5.2], all set in the domain  $\Omega = [0, 1] \times [0, 1]$ . These are

1. Gaussian densities  $\rho^{\text{in}}$  and  $\rho^{\text{f}}$  with the same variance.
2. Translation of a compactly supported smooth sinusoidal density  $\rho^{\text{in}}$ .
3. Compression of a compactly supported smooth sinusoidal density  $\rho^{\text{in}}$ .

These test cases are of increasing difficulty for a numerical solver. In the first one, the variance of the Gaussian densities is chosen sufficiently high in order to ensure the presence of mass everywhere in the domain, which consequently ensures smoothness in the optimum. Notice that the solution is not a translation, which is the case for the second test, where the solution is compactly supported and the problem is not smooth. Finally, in the third test, the optimal transport is the compression of an initial density onto a smaller support. The latter situation has been shown experimentally to exhibit severe instabilities in the discrete solution, and this, in turn, has a negative effect on the discrete solver.

Each problem is discretized using four different meshes, taken from [29], having  $N_{\mathcal{T}'} = 56, 224, 896, 3584$  cells, where each mesh is a refinement of the other. These meshes will be denoted by  $\mathcal{T}_{h_0}, \mathcal{T}_{h_0/2}, \mathcal{T}_{h_0/4}, \mathcal{T}_{h_0/8}$ , where  $h_0$  is the typical mesh size of the coarsest mesh. The corresponding number of subcells is  $N_{\mathcal{T}} = 168, 672, 2688, 10752$ . For each grid, we consider four time steps,  $\Delta t = 1/(K+1)$  with  $K = 8, 16, 32, 64$ . Notice that  $h_0 = 1/8$ , so that the diagonal pairing of the space and time steps provides a uniform discretization of the time-space domain  $[0, 1] \times \Omega$ .

Convergence is achieved when the relaxation parameter  $\mu$  reaches the value  $1e-6$ , which corresponds to ten iterations of IP, with the choice of parameters described in [15]. The nonlinear system eq. (3) is solved with the same tolerance in order to have a robust implementation for the IP solver. Each IP step requires between 3 and 7 inexact Newton iterations. These numbers are practically insensitive to the mesh and the time step adopted. The total number of Newton steps (which corresponds to the total number of linear systems to be solved) ranges between 40 and 50. The preconditioning approaches used to solve these linear systems are described in the next sections.

For each preconditioner, we will present a table reporting the metrics mentioned above while the relaxation parameter  $\mu$  is reduced for all combination of spatial and temporal discretizations. We will show these detailed results only for the second test case, since it captures the main challenges involved in the linear algebra problems. The numerical results for the other two cases will be summarized in section 4.6, where we will compare the performance of the preconditioners only in terms of total CPU time.

The code is written in MATLAB<sup>®</sup> and the source is available at this repository, where it is possible to reproduce the experiments presented in this paper. The AGMG solver is interfaced via MEX.

## 4.2 HSS Preconditioner

The first proposed approach consists in the Hermitian skew-Hermitian splitting (HSS) preconditioner described in [21], building on the scheme proposed in [30]. It consists in rewriting the linear system in eq. (11) in the form

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ -\mathcal{B} & \mathcal{C} \end{pmatrix} \begin{pmatrix} \delta\phi \\ \delta\rho \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ -\tilde{\mathbf{g}} \end{pmatrix}. \quad (12)$$

Then, fixing a relaxation parameter  $\alpha > 0$ , the HSS preconditioner for the linear system in eq. (12) is given by

$$\mathbf{P}_{HSS}^{-1} := (\mathcal{H}_\alpha)^{-1} (\mathcal{K}_\alpha)^{-1} \quad (13)$$

where

$$\mathcal{H}_\alpha = \begin{pmatrix} \mathcal{A} + \alpha \mathbf{I}_n & \\ & \mathcal{C} + \alpha \mathbf{I}_m \end{pmatrix} \quad \text{and} \quad \mathcal{K}_\alpha = \begin{pmatrix} \alpha \mathbf{I}_n & \mathcal{B}^T \\ -\mathcal{B} & \alpha \mathbf{I}_m \end{pmatrix}. \quad (14)$$

Note that in eq. (13) the order of the factor is reversed with respect to the original HSS preconditioner in [21], since this swapped preconditioner gave better results.

The application of the preconditioner  $\mathbf{P}_{HSS}^{-1}$  requires the (approximate) inversion of the matrices  $\mathcal{H}_\alpha$  and  $\mathcal{K}_\alpha$  in eq. (14). The first requires the solution of  $K+1$  weighted and shifted Laplacian systems in the form of

$$\left( \mathbf{A}^k + \alpha \mathbf{I}_{N_{\mathcal{T}}} \right) \mathbf{x}^k = \mathbf{f}^k \in \mathbb{R}^{N_{\mathcal{T}}}, \quad (15)$$

and the inversion of a diagonal matrix. The second inversion can be done in two ways, forming the primal or the dual Schur complement. However, the second approach, that involves the solution of a linear system in the following form

$$\left( \alpha \mathbf{I}_m + \frac{1}{\alpha} \mathcal{B} \mathcal{B}^T \right) \mathbf{y} = \mathbf{c} \in \mathbb{R}^m, \quad (16)$$



turned out to be more efficient. Linear systems in eqs. (15) and (16) are solved using the AGMG solver with precision  $0 < \varepsilon_{\text{in}} \leq \varepsilon_{\text{out}} = 1e - 5$ . After a few numerical tests we found that the parameters  $\alpha = 0.5$  and  $\varepsilon_{\text{in}} = 1e - 1$  give the best performance for all the choices of time step and mesh size. Moreover, as suggested in [21], we apply this preconditioner after scaling the linear system in eq. (12) by its diagonal, leading to a significant reduction in the number of iterations.

In table 1 we summarize the results obtained using the HSS preconditioner for different combinations of time and space discretizations. For each of these configurations, we report the quantities measuring the preconditioner performance defined in items 1 to 3 for each Newton cycle as the relaxation parameter  $\mu$  decreases. The number **Inner/Outer** in table 1 refers to the AGMG iterations to solve the linear system in eq. (16), since the number of AGMG iterations to solve those in eq. (15) is practically constant, ranging between 1 and 2. For a fixed grid, as the relaxation parameter  $\mu$  decreases, we observed a reduction of **Outer/Lin.sys.** and a small increase of **Inner/Outer**, the latter ranging between 1 and 5.7. These two effects yield to a practically constant CPU time required to solve each linear system. Unfortunately, the total number of outer FGMRES iterations approximately doubles when we consider a refined grid.

These results show that the HSS preconditioner is rather robust, but its application to problems with grids with a large number of cells may become too costly. In fact, the CPU time approximately scales linearly with respect to the time steps number but quadratically with respect to the cells number.

### 4.3 Preconditioner based on the primal Schur complement

In this section we consider a preconditioner for the linear system in eq. (11) based on the primal Schur complement  $\mathcal{S} = \mathcal{A} + \mathcal{B}^T \mathcal{C}^{-1} \mathcal{B}$ , similarly to the approach used in [31] for the solution of the OTP on graphs with cost equal to the shortest path distance. It is given by

$$P^{-1} = \begin{pmatrix} I & \\ \mathcal{C}^{-1} \mathcal{B} & I \end{pmatrix} \begin{pmatrix} \mathcal{S}^{-1} & \\ & -\mathcal{C}^{-1} \end{pmatrix} \begin{pmatrix} I & \mathcal{B}^T \mathcal{C}^{-1} \\ & I \end{pmatrix}. \quad (17)$$

The application of the preconditioner in eq. (17) requires the inversion of  $\mathcal{C}$ , which is diagonal, and the (approximate) solution of the linear system

$$\mathcal{S} \mathbf{x} = \mathbf{b}, \quad (18)$$

with  $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ . The matrix  $\mathcal{S}$  is a block tridiagonal matrix that can be written as

$$\mathcal{S} = \underbrace{-\mathcal{D}iv_x \text{Diag}(\tilde{\rho}) \mathcal{D}_x}_{\mathcal{A}} \underbrace{-\mathcal{D}iv_x \mathcal{G} \mathcal{C}^{-1} \mathcal{G} \mathcal{D}_x}_{=:\mathcal{S}_{xx}} + \underbrace{\mathcal{D}_t^T \mathcal{C}^{-1} \mathcal{D}_t}_{=:\mathcal{S}_{tt}} + \underbrace{\mathcal{D}_t^T \mathcal{C}^{-1} \mathcal{G}^T \mathcal{D}_x + (\mathcal{D}_t^T \mathcal{C}^{-1} \mathcal{G}^T \mathcal{D}_x)^T}_{=:\mathcal{S}_{tx} + \mathcal{S}_{tx}^T}$$

and it is the sum of a  $\rho$ -weighted spatial Laplacian matrix  $\mathcal{A}$ , an anisotropic spatial Laplacian  $\mathcal{S}_{xx} \approx -\text{div}(\rho/s \nabla \phi \otimes \nabla \phi \nabla)$ , a weighted temporal Laplacian with Neumann boundary condition  $\mathcal{S}_{tt}$ , and a time-space operator  $\mathcal{S}_{tx} + \mathcal{S}_{tx}^T$ . This structure poses significant challenges in the solution of the linear system in eq. (18). We again adopt the AGMG solver to solve it, with relative accuracy  $\varepsilon_{\text{in}} = 1e - 1 < \varepsilon_{\text{out}} = 1e - 5$ .

In table 2 we summarize the results obtained using the preconditioner presented in this section for different combinations of time and space discretizations. We use the metrics items 1 to 3 used for the HSS preconditioner, where in this case (**Inner/Outer**) refers to the inner iterations required to solve the linear system in eq. (18). We also include the percentage of CPU time used by the AGMG algorithm to build the sequence of coarse matrices used by the multigrid solver, before Flexible GMRES is used.

The number of average outer iterations **Outer/Lin.sys.** remains practically constant using different grids and time steps. However, the average number of inner iterations is strongly influenced by the relationship between the mesh and the time step used ( $\Delta t = 1/K$  in our experiments). The phenomenon is particularly evident for the third mesh,  $\mathcal{T}_{h_0/8}$ . For  $K = 16$ , during the last IP iterations, the average number of inner iterations increases progressively, reaching  $\approx 134$  at the last IP iteration. For  $K = 32$ , this number remains between 2.1 and 2.8. The same phenomenon occurs using the finest grid  $\mathcal{T}_{h_0/8}$ , passing from  $K = 32$  to  $K = 64$ . We attribute this to an improper construction of the coarse matrix sequence used by the multigrid solver due to the poor scaling of the temporal and spatial components of  $\hat{\mathcal{S}}$ . The AGMG is based on an aggregation-based coarsening approach, hence if these components are unbalanced the coarse matrices may not reflect the infinite-dimensional operators discretized by our problem, compromising the performance of the multigrid solver. In fact, taking the time-space couplings ( $h = h_0/4, \Delta t = 1/32$ ) and ( $h = h_0/8, \Delta t = 1/64$ ), the average number of inner iterations remains between two and three.

$K$	8	16	32	64	8	16	32	64	8	16	32	64
$\mu$	Outer/Lin.sys.				CPU/Lin.sys.				Inner/Outer			
1	31	35	34	34	0.8	1.8	3.4	6.2	1.0	1.5	2.0	2.8
2e-1	25	26	26	27	0.6	1.6	3.0	5.7	1.3	2.0	2.6	3.7
4e-2	21	22	22	23	0.6	1.5	2.8	5.4	1.9	2.3	3.0	3.8
8e-3	19	19	19	20	0.7	1.4	2.5	5.2	2.0	2.3	3.1	4.1
2e-3	19	20	21	20	0.7	1.4	2.8	5.2	2.0	2.4	3.3	4.3
3e-4	19	20	21	22	0.8	1.4	2.8	5.2	2.0	2.6	3.5	4.6
6e-5	20	22	22	22	0.8	1.5	2.8	5.4	2.1	2.7	3.7	4.5
1e-5	22	23	23	23	0.8	1.5	2.8	5.4	2.0	2.7	3.7	4.3
3e-6	22	23	23	24	0.8	1.5	2.8	5.5	2.0	2.8	3.5	4.1
5e-7	22	23	23	24	0.8	1.5	2.8	5.6	2.1	2.6	3.4	4.3
	22	24	24	24	0.7	1.5	2.8	5.5	1.7	2.3	3.1	3.9
1	63	71	72	75	1.0	2.0	4.0	8.2	1.0	1.0	1.5	2.0
2e-1	49	51	53	55	1.1	1.9	3.9	7.9	1.0	1.4	2.1	2.8
4e-2	38	39	39	40	1.0	1.7	3.6	7.2	1.5	2.1	2.7	3.6
8e-3	33	32	33	34	1.0	1.8	3.6	7.0	2.0	2.4	3.2	4.3
2e-3	32	30	30	30	0.9	1.8	3.5	6.5	2.0	2.3	3.2	4.4
3e-4	31	29	28	30	1.0	1.8	3.5	6.8	2.0	2.3	3.5	4.6
6e-5	30	28	30	32	1.0	1.8	3.4	7.0	2.0	2.4	3.6	5.0
1e-5	30	29	30	34	1.0	1.8	3.5	7.0	2.0	2.9	4.1	5.4
3e-6	30	30	30	32	1.0	1.8	3.5	6.8	2.0	3.0	4.1	5.5
5e-7	29	32	33	34	1.0	1.9	3.5	7.0	2.1	2.9	4.0	5.3
	38	39	39	41	1.0	1.8	3.6	7.2	1.6	2.0	2.8	3.8
1	192	190	201	218	4.6	8.4	17.8	44.0	1.0	1.0	1.0	1.5
2e-1	118	122	128	133	3.6	7.0	15.7	34.3	1.0	1.0	1.4	2.2
4e-2	88	89	91	92	3.2	6.4	13.6	32.0	1.6	1.6	2.4	3.1
8e-3	70	72	72	72	3.0	6.0	13.2	30.0	2.0	2.0	2.8	3.8
2e-3	67	66	63	61	3.2	6.2	13.2	27.5	2.1	2.1	3.0	4.5
3e-4	64	65	62	58	3.5	6.8	13.2	27.5	2.2	2.2	3.0	4.8
6e-5	63	64	62	56	3.5	6.8	13.8	27.5	2.3	2.2	3.1	5.0
1e-5	62	63	63	58	3.5	7.0	14.0	27.5	2.3	2.2	3.5	5.2
3e-6	58	60	62	57	3.3	6.8	13.8	27.5	2.5	2.3	4.0	5.6
5e-7	58	59	62	57	3.7	6.8	13.5	25.0	2.6	2.7	4.2	5.9
	90	90	92	92	3.5	6.8	14.3	30.9	1.6	1.6	2.3	3.2
1	641	620	656	707	46.0	86.0	200.0	433.3	1.0	1.0	1.0	1.0
2e-1	340	361	379	395	28.6	60.0	135.7	357.1	1.0	1.0	1.0	1.4
4e-2	230	236	245	248	26.0	51.7	111.7	283.3	1.6	1.7	1.7	2.6
8e-3	169	175	177	177	22.0	44.0	98.0	260.0	2.1	2.1	2.3	2.9
2e-3	164	165	150	139	37.5	55.0	105.0	217.5	2.8	2.8	2.9	3.6
3e-4	158	161	151	139	45.0	57.5	115.0	232.5	3.0	3.0	3.0	4.0
6e-5	147	162	156	147	47.5	60.0	122.5	247.5	3.1	3.0	3.0	4.0
1e-5	142	162	159	145	46.7	62.5	127.5	240.0	3.2	3.0	3.0	4.2
3e-6	137	159	159	142	46.7	63.3	130.0	245.0	3.2	3.0	3.0	4.2
5e-7	129	153	161	140	43.3	60.0	130.0	240.0	3.2	3.0	3.0	4.3
	248	254	256	264	37.2	59.8	128.3	286.3	1.8	1.8	1.8	2.2

Table 1: Numerical results using HSS preconditioner for the linear system in eq. (12) within a FGMRES cycle. Each subtable reports the results obtained for a given mesh, from the coarsest  $\mathcal{T}_{h_0}$  (top) to the finest  $\mathcal{T}_{h_0/8}$  (bottom). Each subtable contains 3 blocks of 4 columns. Each block reports, from left to right, the averaged number of outer FGMRES iterations, the average CPU time per linear system, and the averaged number of inner iterations for solving the linear system in eq. (16) (see items 1 to 3). Within each block, each column reports the related metric for one choice of time step  $\Delta t = 1/K$  with  $K = 8, 16, 32, 64$  as the relaxation parameter (reported in the leftmost column) is reduced. A final row summarizes the same quantities averaged on the whole simulation. We highlighted in gray the spatial-temporal combination providing the uniform discretization.

$K$	8	16	32	64	8	16	32	64	8	16	32	64	8	16	32	64
$\mu$	Outer/Lin.sys.				Inner/Outer				CPU/Lin.sys.				% Preprocess			
1	5	5	5	2	1.7	2.0	2.2	1.5	0.1	0.1	0.2	0.4	59	66	62	80
2e-1	5	5	5	1	2.2	2.0	1.9	1.0	0.1	0.1	0.3	0.5	66	68	69	81
4e-2	5	5	4	1	2.0	1.6	1.6	1.0	0.1	0.2	0.3	0.4	71	71	71	82
8e-3	5	5	4	1	1.5	1.6	1.5	1.0	0.1	0.2	0.3	0.5	75	74	72	84
2e-3	4	4	4	1	1.5	1.6	1.9	1.0	0.1	0.2	0.3	0.5	78	76	71	83
3e-4	4	4	4	2	1.5	1.4	1.7	1.2	0.1	0.2	0.3	0.4	78	77	72	82
6e-5	4	4	3	3	1.4	1.5	1.7	1.1	0.1	0.2	0.3	0.5	78	75	72	78
1e-5	4	3	4	3	1.5	1.5	1.5	1.2	0.1	0.2	0.3	0.5	76	77	73	78
3e-6	3	3	3	2	1.3	1.5	1.3	1.0	0.1	0.2	0.3	0.5	81	78	74	80
5e-7	3	3	3	2	1.3	1.4	1.3	1.3	0.1	0.2	0.3	0.5	81	80	78	76
	4.3	4.2	3.9	1.7	1.7	1.7	1.7	1.1	0.1	0.2	0.3	0.5	74	74	71	80
1	6	4	5	5	2.1	2.2	2.0	2.7	0.1	0.2	0.4	1.3	54	51	37	48
2e-1	5	4	5	4	2.6	2.5	2.4	2.6	0.2	0.3	0.7	3.6	52	47	54	72
4e-2	6	6	6	5	2.5	2.0	2.2	1.8	0.2	0.4	1.2	4.2	55	55	58	81
8e-3	8	7	5	5	2.2	1.5	2.1	1.6	0.3	0.6	1.4	4.2	54	62	67	81
2e-3	6	5	5	4	3.9	1.7	1.8	1.8	0.4	0.6	1.5	4.5	51	71	71	78
3e-4	5	4	4	4	6.1	1.7	1.9	1.8	0.5	0.6	1.6	4.0	52	72	72	81
6e-5	4	4	4	4	6.9	1.7	1.9	1.8	0.4	0.6	1.6	4.2	51	72	76	81
1e-5	4	4	4	4	6.8	1.6	1.9	1.8	0.5	0.7	1.6	4.5	53	77	79	83
3e-6	4	4	4	4	7.1	1.7	1.7	1.8	0.5	0.6	1.7	4.2	50	76	76	82
5e-7	4	4	4	4	5.7	1.6	1.5	1.7	0.4	0.6	1.6	4.5	58	76	80	83
	5.1	4.6	4.6	4.3	4.0	1.9	2.0	2.0	0.3	0.5	1.3	3.9	53	67	70	79
1	5	5	5	5	2.8	2.6	2.1	2.2	0.4	0.8	1.8	3.2	38	35	36	29
2e-1	5	5	5	5	2.9	2.7	2.8	3.0	0.5	1.0	2.3	8.9	35	34	34	56
4e-2	6	6	7	6	3.1	3.0	2.3	2.4	0.7	1.9	4.0	19.0	35	41	47	74
8e-3	6	7	7	6	2.9	3.4	2.3	2.2	1.0	2.6	5.8	30.0	46	47	62	80
2e-3	6	8	7	6	2.4	5.3	2.1	1.9	1.1	3.8	6.8	35.0	56	39	70	86
3e-4	6	7	6	5	3.1	11.2	2.2	2.0	1.4	6.0	7.5	37.5	52	28	73	87
6e-5	6	6	5	4	6.4	22.2	2.5	2.1	1.8	8.5	7.8	37.5	41	21	74	93
1e-5	5	5	4	4	11.4	62.1	2.4	2.0	2.5	20.8	8.8	42.5	29	9	77	94
3e-6	5	5	4	4	16.5	93.5	2.2	1.9	3.2	27.5	8.2	40.0	24	7	79	88
5e-7	5	6	4	4	37.9	121.2	2.1	2.0	6.3	42.5	9.8	50.0	13	5	82	90
	5.4	5.7	5.3	4.9	7.0	27.1	2.3	2.2	1.6	10.2	5.8	28.1	31	12	69	85
1	4	4	5	5	3.0	3.0	2.8	2.3	1.5	3.0	6.4	13.2	34	32	31	30
2e-1	5	5	5	5	3.3	3.3	3.1	2.9	2.0	4.1	9.3	20.0	31	29	26	36
4e-2	6	6	6	6	4.5	3.9	4.0	2.7	3.6	6.8	20.0	83.3	24	32	38	76
8e-3	7	6	7	7	4.7	4.1	5.3	2.7	5.0	10.0	40.0	220.0	33	44	60	88
2e-3	7	7	8	8	3.9	3.2	7.7	2.3	5.5	12.5	55.0	230.0	43	56	55	87
3e-4	6	7	8	7	3.8	3.6	15.4	2.3	5.8	14.8	85.0	425.0	48	61	38	94
6e-5	6	6	7	6	3.7	7.2	53.4	2.5	5.8	20.5	192.5	300.0	52	50	21	92
1e-5	5	5	6	4	4.6	12.2	125.4	2.7	6.3	25.7	375.0	260.0	58	42	12	92
3e-6	5	5	7	4	4.3	30.1	151.8	2.7	6.7	40.0	475.0	300.0	55	28	10	92
5e-7	4	5	11	4	3.4	81.4	169.2	2.5	6.0	90.0	925.0	300.0	67	14	6	92
	5.4	5.6	6.7	5.5	3.9	10.4	55.7	2.6	4.4	18.0	188.2	194.3	45	32	13	89

Table 2: Numerical results using the preconditioner based on the primal Schur complement for linear system eq. (11) within a Flexible GMRES cycle. Each subtable reports the results obtain for a given mesh, from the coarsest  $\mathcal{T}_{h_0}$  (top) to the finest  $\mathcal{T}_{h_0/8}$  (bottom). Each subtable contains four blocks with four columns. The first three blocks report, from left to right, averaged number of outer iterations for solving eq. (11), averaged number of inner iterations for solving eq. (20), averaged CPU time per linear system (see the items 1 to 3 in section 4). The last block reports the percentage of this time spent during the preprocess phase. Each column reports these metrics using  $\Delta t = 1/K$  with  $K = 8, 16, 32, 64$ . Each column entry reports these metrics as the relaxation parameter  $\mu$  (reported in the leftmost column) is reduced. A final row summarizes the same quantities averaged on the whole simulation. We highlighted in gray the spatial-temporal coupling providing the uniform discretization.

Nevertheless, even in best-case scenarios, we experimented that the preprocess phase is the most demanding in terms of CPU time. Typically, more than 50% of the time is spent building the sequence of coarse matrices, reaching peaks of 90% during the last IP steps using the finest discretization. We attribute this additional preprocess time to the fact that the matrix  $\mathcal{S}$  becomes highly ill-conditioned when the relaxation parameter  $\mu$  reaches  $1e-6$ . In fact, matrix  $\mathcal{C}$  contains the term  $\text{Diag}(s)(\text{Diag}(\rho))^{-1}$  where the term  $s \odot \rho = \mu \mathbf{1} \rightarrow 0$ . As a consequence,  $\mathcal{C}^{-1}$  contains values that vary by several orders of magnitude. Similar considerations can be found in [32]. Unfortunately, it was not possible to access the hierarchy of matrices built by the AGMG solver to validate this hypothesis. Building an ad hoc geometric multigrid preconditioner could, in principle, cope with this issue. However, its design goes beyond the purpose of this paper.

In order to alleviate these problems, we tried to neglect some components of  $\mathcal{S}$ . We tried to remove the extra-diagonal or lower-diagonal part of  $\mathcal{S}$ , or the time-space operator  $\mathcal{S}_{tx} + \mathcal{S}_{tx}^T$ . However, none of these approaches worked. All numerical experiments suggest that none of these term can be neglected without altering the spectral property of the matrix, and consequently affecting the effectiveness of the preconditioner.

#### 4.4 The SIMPLE preconditioner

In this section, we recall the classical SIMPLE preconditioner described in [22, 33] and based on the approximation of the inverse of the dual Schur complement  $\mathcal{S} = -(\mathcal{C} + \mathcal{B}\mathcal{A}^{-1}\mathcal{B}^T)$  (we recall that, formally, we cannot write  $\mathcal{A}^{-1}$  since the matrix  $\mathcal{A}$  is singular). The SIMPLE preconditioner is given by

$$P^{-1} = \begin{pmatrix} I & -\hat{\mathcal{A}}^{-1}\mathcal{B}^T \\ & I \end{pmatrix} \begin{pmatrix} \hat{\mathcal{A}}^{-1} & \\ & \hat{\mathcal{S}}^{-1} \end{pmatrix} \begin{pmatrix} I & \\ -\mathcal{B}\hat{\mathcal{A}}^{-1} & I \end{pmatrix} \quad (19)$$

where

$$\hat{\mathcal{A}} = \text{Diag}(\mathcal{A}), \quad \hat{\mathcal{S}} = -(\mathcal{C} + \mathcal{B}^T \hat{\mathcal{A}}^{-1} \mathcal{B}).$$

The main computational cost of applying this preconditioner is to solve linear systems in the form

$$\hat{\mathcal{S}}\mathbf{y} = \mathbf{c}. \quad (20)$$

The matrix  $\hat{\mathcal{S}}$  can be formed explicitly and is block tridiagonal. We use the AGMG solver with tolerance  $\varepsilon_{\text{in}} = 1e-1$  to solve the linear system in eq. (20).

The results using the SIMPLE preconditioner are summarized in table 3. Despite its simplicity, the proposed preconditioner turned out to be rather efficient and robust and the IP algorithm always reaches convergence. The average number of inner iterations **Inner/Outer** remains practically constant, independent of the time step and the mesh size used or the IP step considered. The preprocess time required by the AGMG solver is limited, approximately accounting for 1% of the CPU time spent in the linear system solution. The number of outer iterations per IP step remains practically constant when we change the time step  $\Delta t = 1/K$ . Surprisingly, the preconditioner becomes more efficient as the IP relaxation term  $\mu$  is lowered to zero.

Unfortunately, the number of outer iterations more than doubles at each mesh refinement. This phenomenon can also be explained by looking at the structure of the preconditioned matrix

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ \mathcal{B} & -\mathcal{C} \end{pmatrix} P^{-1} = \begin{pmatrix} I + (\mathcal{A}\hat{\mathcal{A}}^{-1} - I)(I + \mathcal{B}^T \hat{\mathcal{S}}_{\varepsilon_{\text{in}}}^{-1} \mathcal{B}\hat{\mathcal{A}}^{-1}) & (\mathcal{A}\hat{\mathcal{A}}^{-1} - I)\mathcal{B}^T \hat{\mathcal{S}}_{\varepsilon_{\text{in}}}^{-1} \\ (I - \hat{\mathcal{S}}_{\varepsilon_{\text{in}}}^{-1})\mathcal{B}\hat{\mathcal{A}}^{-1} & \hat{\mathcal{S}}_{\varepsilon_{\text{in}}}^{-1} \end{pmatrix}$$

and studying its condition number. The preconditioned matrix is approximately a block upper triangular matrix, whose bottom-right block becomes close to the identity as  $\varepsilon_{\text{in}} \rightarrow 0$ , with  $m$  eigenvalues close to one. The upper left block is the identity plus a perturbation where the term  $(\mathcal{A}\hat{\mathcal{A}}^{-1} - I)$  appears. This term grows rapidly as  $h \rightarrow 0$  as shown in [34].

#### 4.5 The $\mathcal{B}\mathcal{B}$ -preconditioner

We present now the fourth preconditioner we considered in this work. Our idea is to define a preconditioner similar to the one given in eq. (19) but using a better approximation of the pseudo-inverse of  $\mathcal{A}$ . Since this matrix is block diagonal and each block  $\mathcal{A}^k$  is a weighted Laplacian, we can approximate the action of its pseudo-inverse by solving separately  $K + 1$  linear systems in the form

$$\mathcal{A}^k \mathbf{x}^k = \mathbf{b}^k \quad (21)$$

using multigrid solvers. However, two issues arise.

$K$	8	16	32	64	8	16	32	64	8	16	32	64
$\mu$	Outer/Lin.sys.				CPU/Lin.sys.				Inner/Outer			
1	45	51	51	53	0.2	0.1	0.3	0.6	1.0	1.0	1.7	2.1
2e-1	35	36	36	37	0.1	0.2	0.4	0.5	1.1	1.5	2.2	2.9
4e-2	32	32	32	34	0.1	0.2	0.3	0.6	1.3	2.1	2.6	3.0
8e-3	30	30	29	29	0.1	0.2	0.3	0.6	1.6	2.1	2.9	3.5
2e-3	30	29	28	28	0.1	0.2	0.3	0.5	1.7	2.6	3.5	4.3
3e-4	29	28	29	32	0.1	0.2	0.2	0.6	1.8	2.7	3.4	3.0
6e-5	28	29	28	38	0.1	0.2	0.2	0.7	1.8	2.5	3.3	2.0
1e-5	27	29	33	40	0.1	0.2	0.3	0.7	1.8	2.2	2.2	1.9
3e-6	26	35	38	36	0.1	0.3	0.4	0.7	1.8	1.3	1.4	1.5
5e-7	28	33	35	37	0.1	0.3	0.4	0.7	1.3	1.1	1.3	1.6
	32	34	35	37	0.1	0.2	0.3	0.6	1.4	1.8	2.3	2.4
1	99	108	109	106	0.4	0.7	1.3	3.4	1.0	1.0	1.0	1.8
2e-1	71	72	73	73	0.4	0.8	1.4	3.3	1.0	1.1	1.6	2.4
4e-2	65	67	68	67	0.5	0.8	1.6	3.4	1.2	1.4	2.2	3.3
8e-3	63	60	61	62	0.7	0.8	1.8	3.4	1.3	1.8	2.5	3.3
2e-3	64	59	57	52	0.7	0.9	1.9	3.0	1.6	1.9	2.5	3.4
3e-4	64	61	50	54	0.6	1.0	1.7	3.0	1.6	1.8	2.5	3.9
6e-5	66	53	52	47	0.6	0.8	1.5	2.4	1.6	2.1	2.9	4.2
1e-5	59	58	53	59	0.5	0.9	1.6	3.2	1.4	1.9	3.1	2.9
3e-6	59	57	68	70	0.6	0.9	2.3	4.0	1.4	1.8	1.9	2.0
5e-7	56	63	81	65	0.5	1.2	2.8	3.8	1.4	1.2	1.3	2.0
	68	67	68	66	0.5	0.9	1.7	3.3	1.3	1.5	2.0	2.7
1	260	251	231	231	3.0	5.0	8.5	18.3	1.0	1.0	1.0	1.0
2e-1	134	136	146	148	2.0	3.7	8.3	18.6	1.0	1.0	1.1	1.5
4e-2	114	120	122	124	2.2	4.0	8.4	19.8	1.2	1.1	1.5	2.3
8e-3	94	97	115	115	2.0	4.0	9.2	20.0	1.4	1.3	1.8	2.7
2e-3	125	95	102	105	3.2	4.5	9.0	22.0	1.8	1.7	2.0	2.9
3e-4	134	109	110	107	3.8	5.2	10.5	22.8	1.8	1.8	2.0	2.8
6e-5	123	116	113	106	3.8	6.0	11.2	22.8	1.8	1.7	1.9	2.8
1e-5	123	122	128	103	3.5	6.5	12.2	21.0	1.7	1.6	2.1	3.2
3e-6	128	110	132	192	3.7	5.8	12.2	42.5	1.7	1.6	2.0	1.5
5e-7	136	121	147	221	4.0	6.5	17.0	52.5	1.5	1.6	1.3	1.2
	138	130	139	148	3.0	5.0	10.3	25.0	1.4	1.3	1.5	1.9
1	599	480	499	484	21.7	33.3	66.7	157.1	1.0	1.0	1.0	1.0
2e-1	291	292	302	327	13.4	27.1	52.9	142.9	1.0	1.0	1.0	1.0
4e-2	225	218	218	218	13.6	26.7	48.3	123.3	1.2	1.4	1.1	1.6
8e-3	200	177	188	188	14.2	24.0	50.0	118.0	1.5	1.5	1.3	1.8
2e-3	240	202	170	185	30.0	30.0	55.0	135.0	1.9	1.8	1.8	2.2
3e-4	263	237	199	185	45.0	40.0	72.5	155.0	1.9	2.1	1.9	2.1
6e-5	326	282	230	214	70.0	52.5	92.5	190.0	1.9	2.1	1.9	2.1
1e-5	329	349	264	224	76.7	70.0	110.0	200.0	1.8	2.0	1.8	2.1
3e-6	337	372	270	231	83.3	76.7	115.0	207.5	1.8	1.9	1.7	1.9
5e-7	300	323	238	251	80.0	66.7	100.0	247.5	1.9	1.8	1.7	1.5
	318	292	268	266	37.8	40.0	72.7	163.4	1.4	1.5	1.4	1.5

Table 3: Numerical results using the SIMPLE preconditioner for linear system in eq. (11) within a Flexible GMRES cycle. Each subtable reports the results obtain for a given mesh, from the coarsest  $\mathcal{T}_{h_0}$  (top) to the finest  $\mathcal{T}_{h_0/8}$  (bottom). Each subtable contains 3 blocks of 4 columns. Each block reports one metric used to measure the preconditioner performance that are, from left to right, the averaged number of outer FGMRES iterations, the average CPU-time per linear system, and the averaged number of inner iterations for solving linear system in eq. (20) (see items 1 to 3). Within all blocks, each column reports the related metric for one choice of time step  $\Delta t = 1/K$  with  $K = 8, 16, 32, 64$  as the relaxation parameter (reported in the leftmost column). A final row summarizes the same quantities averaged on the whole simulation. We highlighted in gray the spatial-temporal combination providing the uniform discretization.

The first issue is that each block  $\mathcal{A}^k$  is singular with kernel  $\langle \mathbf{1} \in \mathbb{R}^{N\tau} \rangle$ , so we first need to ensure the well-posedness of linear systems in eq. (21). This means that we must guarantee

$$\mathbf{b}^k \in \text{Im}(\mathcal{A}^k) = \text{Ker}(\mathcal{A}^k)^\perp = \{\mathbf{1} \in \mathbb{R}^{N\tau}\}^\perp, \quad k = 1, \dots, K+1. \quad (22)$$

It may not be possible to guarantee condition (22) within a FMGRES cycle while using inexact solvers to approximate the pseudo-inverse of  $\mathcal{A}$  and the dual Schur complement. Moreover, even if the system is compatible, we do not know a priori how to select particular solutions  $\mathbf{x}^k$ , because this selection should fix the increments  $\delta\phi^k$  according to the linearized Hamilton-Jacobi equation in eq. (3b). Neither grounding the linear systems (21) is a feasible solution, since we do not know at which values we should ground the increments  $\delta\phi^k$ .

The second issue is the design of a better approximation of the inverse of the matrix  $\mathcal{S} = -(\mathcal{C} + \mathcal{B}\mathcal{A}^{-1}\mathcal{B}^T)$ . Firstly,  $\mathcal{S}$  is not well defined since  $\mathcal{A}$  is singular and  $\text{Im}(\mathcal{B}^T)$  is not orthogonal to  $\text{Ker}(\mathcal{A})$ . Moreover, even overcoming this first issue, the matrix  $\mathcal{S}$  becomes dense, thus too costly to form and invert.

In the next subsections we will show how to cope with both issues starting from a reformulation of the continuous nonlinear system in eq. (2).

#### 4.5.1 Splitting formulation

In order to cope with the first issue, we split the solution  $\phi$  of eq. (2) into two components, a new potential  $\bar{\phi} : \Omega \times [0, 1] \rightarrow \mathbb{R}$  satisfying the constraint  $\int_\Omega \bar{\phi}(t) dx = 0$  for all  $t \in [0, 1]$  and a correction term:

$$\phi(t, x) = \bar{\phi}(t, x) + \int_0^t \lambda(s) ds \quad (23)$$

where  $\lambda : [0, 1] \rightarrow \mathbb{R}$  depends only on the time variable. In the new unknowns  $(\bar{\phi}, \rho, s, \lambda)$  the system of PDEs (2) becomes

$$\begin{aligned} -\partial_t \rho - \text{div}_x(\rho \nabla_x \bar{\phi}) &= 0, \\ \partial_t \bar{\phi} + \frac{\|\nabla_x \bar{\phi}\|^2}{2} + s + \lambda &= 0, \\ \rho \geq 0, s \geq 0, \rho s &= 0, \\ \int_\Omega \bar{\phi} dx &= 0. \end{aligned} \quad (24)$$

Following the same steps described in section 2, the (relaxed) discrete counterpart of equation in eq. (24) is a nonlinear system of equations with unknowns  $(\bar{\phi}, \rho, s, \lambda) \in (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^K)$  given by

$$\begin{aligned} F_\phi(\bar{\phi}, \rho) &= (F_\phi^1; \dots; F_\phi^{K+1}) = \mathbf{0} \in \mathbb{R}^n, \\ \bar{F}_\rho(\bar{\phi}, \rho, s, \lambda) &= (\bar{F}_\rho^1; \dots; \bar{F}_\rho^K) = \mathbf{0} \in \mathbb{R}^m, \\ F_s(\rho, s) &= (F_s^1; \dots; F_s^K) = \mathbf{0} \in \mathbb{R}^m, \\ F_\lambda(\bar{\phi}) &= (F_\lambda^1; \dots; F_\lambda^K) = \mathbf{0} \in \mathbb{R}^{K+1}, \end{aligned}$$

where  $F_\phi^k$  and  $F_s^k$  are defined in eqs. (4) and (6), while  $\bar{F}_\rho^k$  and  $F_\lambda^k$  are given by

$$\begin{aligned} \bar{F}_\rho^k(\rho, \bar{\phi}, s, \lambda) &:= F_\rho^k(\rho, \bar{\phi}, s) + \mathcal{M}' \lambda^k, & k = 1, \dots, K, \\ \bar{F}_\lambda^k(\bar{\phi}^k) &:= |\mathbf{c}|^T \bar{\phi}^k, & k = 1, \dots, K+1, \end{aligned}$$

with  $F_\rho^k$  as given in eq. (5). Note that there are  $n + 2m + K + 1$  equations and  $n + 2m + K$  unknowns. The additional equation fixes the global constant for the potential  $\bar{\phi}$ . The linear system arising from the Newton method becomes

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T & & & \\ \mathcal{B} & & \mathcal{M}' & & \\ & \text{Diag}(s) & \text{Diag}(\rho) & & \\ \mathcal{E}\mathcal{M} & & & \mathcal{M}'\mathcal{E}'^T & \end{pmatrix} \begin{pmatrix} \delta\bar{\phi} \\ \delta\rho \\ \delta s \\ \delta\lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \bar{\mathbf{g}} \\ \mathbf{h} \\ \mathbf{i} \end{pmatrix} = - \begin{pmatrix} F_\phi \\ \bar{F}_\rho \\ F_s \\ F_\lambda \end{pmatrix}, \quad (25)$$

where matrices  $\mathcal{E} \in \mathbb{R}^{K+1, N_\tau}$  and  $\mathcal{E}' \in \mathbb{R}^{K, N_{\tau'}}$  are given by

$$\mathcal{E} = \begin{pmatrix} \mathbf{1}_{N_\tau}^T & & \\ & \ddots & \\ & & \mathbf{1}_{N_\tau}^T \end{pmatrix}, \quad \mathcal{E}' = \begin{pmatrix} \mathbf{1}_{N_{\tau'}}^T & & \\ & \ddots & \\ & & \mathbf{1}_{N_{\tau'}}^T \end{pmatrix}.$$

Once the vectors  $(\delta\bar{\phi}; \delta\lambda)$  in eq. (26) are found,  $\delta\phi$  can be retrieved in the original system eq. (11) using the formula

$$\delta\phi^k = \delta\bar{\phi}^k + \sum_{i=1}^k \Delta t^i \delta\lambda^i,$$

which is the finite-dimensional counterpart of eq. (23).

#### 4.5.2 New saddle point linear system

After eliminating the unknown  $s$  as in section 3.1, the linear system in eq. (25) becomes

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T & \\ \mathcal{B} & -\mathcal{C} & \mathcal{M}'\mathcal{E}'^T \\ \mathcal{E}\mathcal{M} & & \end{pmatrix} \begin{pmatrix} \delta\bar{\phi} \\ \delta\rho \\ \delta\lambda \end{pmatrix} = \begin{pmatrix} f \\ \tilde{g} \\ i \end{pmatrix}. \quad (26)$$

The variable  $\delta\lambda$  can be eliminated as well by multiplying by  $\mathcal{E}'$  the second block of equations in eq. (25), yielding

$$\mathcal{E}'(\mathcal{B}\delta\bar{\phi} - \mathcal{C}\delta\rho) + \underbrace{\mathcal{E}'\mathcal{M}'\mathcal{E}'^T}_{|\Omega|I_K} \delta\lambda = \mathcal{E}'\tilde{g}.$$

Plugging this expression in eq. (26), the linear system in the unknowns  $(\delta\bar{\phi}; \delta\rho)$  becomes

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \\ \mathcal{P}\mathcal{B} & -\mathcal{P}\mathcal{C} \\ \mathcal{E}\mathcal{M} & \end{pmatrix} \begin{pmatrix} \delta\bar{\phi} \\ \delta\rho \end{pmatrix} = \begin{pmatrix} f \\ \mathcal{P}\tilde{g} \\ i \end{pmatrix}. \quad (27)$$

where the matrix  $\mathcal{P}$ , which is given by

$$\mathcal{P} := I - \frac{1}{|\Omega|} \mathcal{M}'\mathcal{E}'^T \mathcal{E}',$$

is a projector (indeed  $\mathcal{P}^2 = \mathcal{P}$ ). Moreover, its transpose  $\mathcal{P}^T$  is the discretization of the zero mean-projector, which maps a function  $g : [0, 1] \times \Omega \rightarrow \mathbb{R}$  to  $g - \int_\Omega g / |\Omega|$ .

The linear system in (27) is not overdetermined since  $\text{Ker}(\mathcal{A}) \subset \text{Ker}(\mathcal{P}\mathcal{B})$  and the rows of  $\mathcal{E}\mathcal{M}$  are  $K + 1$  vectors linear independent from the rows of  $\mathcal{A}$  or  $\mathcal{P}\mathcal{B}$ . Recalling remark 4,  $\delta\rho \in \text{Ker}(\mathcal{P})$  or equivalently  $\delta\rho \in \text{Im}(\mathcal{P}^T)$  and, since  $\mathcal{P}^T$  is also a projector, we can write  $\delta\rho = \mathcal{P}^T \delta\rho$ . Thus, we can rewrite system (27) as

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^T \mathcal{P}^T \\ \mathcal{P}\mathcal{B} & -\mathcal{P}\mathcal{C}\mathcal{P}^T \end{pmatrix} \begin{pmatrix} \delta\bar{\phi} \\ \delta\rho \end{pmatrix} = \begin{pmatrix} f \\ \mathcal{P}\tilde{g} \end{pmatrix}, \quad (28)$$

where we neglect the constraints  $\mathcal{E}\mathcal{M}\delta\bar{\phi} = i$ . For any solution  $(\delta\bar{\phi}; \delta\rho)$ , the solution  $(\delta\bar{\phi}; \delta\rho)$  of eq. (27) can be recovered by simply correcting the mean of  $\delta\bar{\phi}$ .

The linear system in eq. (28) does not suffer from the first issue mentioned above associated to the singularity of matrix  $\mathcal{A}$ . In fact, note that  $\text{Im}(\mathcal{B}^T \mathcal{P}^T) \perp \text{Ker}(\mathcal{A})$ , since  $\text{Ker}(\mathcal{A}) \subset \text{Ker}(\mathcal{P}\mathcal{B})$ . This ensures that the dual Schur complement  $\mathcal{S} = -\mathcal{P}(\mathcal{C} + \mathcal{B}\mathcal{A}^+\mathcal{B}^T)\mathcal{P}^T$  (where  $\mathcal{A}^+$  denotes the Moore-Penrose inverse of  $\mathcal{A}$ ) is well defined. Moreover the following (ideal) block triangular preconditioner

$$\mathcal{P}_t = \begin{pmatrix} \mathcal{A} & \mathcal{B}^T \mathcal{P}^T \\ & -\mathcal{S} \end{pmatrix} \quad (29)$$

is well defined as well. In fact, given a vector  $(f, g)$  with  $f \in \text{Ker}(\mathcal{A})^\perp$  and  $g \in \text{Im}(\mathcal{P})$ , the application of this preconditioner requires to compute the vector  $(x; y)$  solving the following two linear systems

$$-\mathcal{S}y = g, \quad (30)$$

$$\mathcal{A}x = f - \mathcal{B}^T \mathcal{P}^T y. \quad (31)$$

Thanks to the presence of the projector  $\mathcal{P}^T$  and since  $\text{Im}(\mathcal{B}^T \mathcal{P}^T) \perp \text{Ker}(\mathcal{A})$ , the linear system in eq. (31) is well defined. We recall that matrix  $\mathcal{A}$  is block diagonal, thus in our implementation we solve this system solving with respect to each block with the AGMG solver with tolerance  $\varepsilon_{\text{in}} = 5e - 2$ .

Now, we are left with second issue mentioned above, how to approximate the Schur complement, which is addressed in the next paragraphs.

### 4.5.3 Approximating the inverse of the dual Schur complement

Our first attempt was to approximate the dual Schur complement  $\mathcal{S}$  (and then its inverse) with  $\mathcal{P}(\mathcal{C} + \mathcal{B}(\text{Diag}(\mathcal{A}))^{-1} \mathcal{B}^T) \mathcal{P}^T$ , but it led to poor performance. We then tried to use the Least Square Commutator approach described in [23]. Unfortunately, it is not possible to apply this approach directly due to the presence of the matrix  $\mathcal{C} = \mathcal{M}' \text{Diag}(s) \text{Diag}(\rho)^{-1}$  which has no clear ‘‘differential’’ meaning. We tried to apply the Algebraically stabilized Least Square Commutator proposed in [35, Section 4.2], but this leads to poor results, for different combinations of the relaxation parameters that appear in such approach. We attribute this phenomenon to the fact that the matrix  $\mathcal{C}$  is not a stabilization operator for the saddle point system, as assumed in [35].

In order to devise a better approximation of the Schur complement, here we follow the idea in [36, 23] of looking at the infinite-dimensional differential operators associated to the matrices that compose the dual Schur complement and try to deduce a possible approximation of its inverse. The following proposition shows a differential identity that goes in this direction.

**Proposition 1.** *Let  $\rho : [0, 1] \times \Omega \rightarrow \mathbb{R}_{>0}$ ,  $\phi : [0, 1] \times \Omega \rightarrow \mathbb{R}$ , smooth enough. Denoting by  $-\Delta_\rho = -\text{div}(\rho \nabla)$ , the following operator identity holds:*

$$\begin{aligned} (\partial_t + \text{div}(\cdot \nabla \phi))(-\Delta_\rho) &= -\Delta_\rho(\partial_t + \nabla \phi \cdot \nabla) - \text{div}((\partial_t \rho + \text{div}(\rho \nabla \phi)) \nabla) \\ &\quad + 2 \text{div}(\rho \nabla^2 \phi \nabla) \end{aligned} \quad (32)$$

*Proof.* Consider any function  $g_0 : [0, 1] \times \Omega \rightarrow \mathbb{R}$ , such that  $\int_\Omega g_0(t, \cdot) = 0$  for all  $t \in [0, 1]$  and let us define  $u : [0, 1] \times \Omega \rightarrow \mathbb{R}$  given by

$$-\text{div}(\rho \nabla u) = g_0 \quad (33)$$

with zero Neumann boundary conditions. Multiplying both sides of equation eq. (33) by  $\partial_i \phi$  (where  $\partial_i = \partial_{x_i}$ ) and taking the divergence, we obtain

$$-\sum_{i,j} \partial_i (\partial_i \phi \partial_j (\rho \partial_j u)) = \text{div}(g_0 \nabla \phi).$$

Using the identity  $\partial_i \phi \partial_j (\rho \partial_j u) = \partial_j (\partial_i \phi \rho \partial_j u) - \partial_j (\partial_i \phi) \rho \partial_j u$ , we get

$$\begin{aligned} \text{div}(g_0 \nabla \phi) &= -\sum_{i,j} \partial_i (\partial_j (\partial_i \phi \rho \partial_j u) - \partial_j (\partial_i \phi) \rho \partial_j u) \\ &= -\sum_{j,i} \partial_j (\partial_{i,j} u \rho \partial_i \phi) - \sum_{j,i} \partial_j (\partial_i (\rho \partial_i \phi) \partial_j u) + \text{div}(\rho \nabla^2 \phi \nabla u) \\ &= -\sum_j \partial_j \left( \sum_i \partial_{i,j} u \rho \partial_i \phi \right) - \sum_j \partial_j \left( \sum_i \partial_i (\rho \partial_i \phi) \partial_j u \right) \\ &\quad + \text{div}(\rho \nabla^2 \phi \nabla u) \\ &= -\text{div}(\rho \nabla^2 u \nabla \phi) - \text{div}(\text{div}(\rho \nabla \phi) \nabla u) + \text{div}(\rho \nabla^2 \phi \nabla u) \end{aligned}$$

Using the identity  $\nabla(\nabla \phi \cdot \nabla u) = \nabla^2 \phi \nabla u + \nabla^2 u \nabla \phi$ , we obtain

$$\text{div}(g_0 \nabla \phi) = -\text{div}(\rho \nabla(\nabla \phi \cdot \nabla u)) - \text{div}(\text{div}(\rho \nabla \phi) \nabla u) + 2 \text{div}(\rho \nabla^2 \phi \nabla u). \quad (34)$$

Now, differentiating with respect to time the expression in eq. (33), we obtain

$$-\text{div}(\partial_t \rho \nabla u + \rho \nabla \partial_t u) = \partial_t g_0. \quad (35)$$

Combining eq. (34), eq. (35) we obtain

$$\begin{aligned} \partial_t g_0 + \text{div}(g_0 \nabla \phi) &= -\text{div}(\rho \nabla(\partial_t u + \nabla \phi \cdot \nabla u)) - \text{div}((\partial_t \rho + \text{div}(\rho \nabla \phi)) \nabla u) \\ &\quad + 2 \text{div}(\rho \nabla^2 \phi \nabla u). \end{aligned}$$

Since  $g_0 = -\Delta_\rho u$  by eq. (33), we proved eq. (32).  $\square$



Our approximation of the Schur complement is based on neglecting the terms  $-\operatorname{div}((\partial_t \rho + \operatorname{div}(\rho \nabla \phi)) \nabla + 2 \operatorname{div}(\rho \nabla^2 \phi \nabla))$  on the right-hand side of eq. (32). The first term may be assumed to be small since it contains the continuity equation, which is the first nonlinear equation in the system eq. (2a). This term is small at the initial and final Newton steps within each IP iteration, and experimental observations showed it remains small during the intermediate Newton iterations. The second term can be neglected under certain assumptions. If the optimal transport is a translation this term is null. When the transported measures  $\rho^{\text{in}}$  and  $\rho^{\text{f}}$  are Gaussian [37], or more generally log-concave probability densities [38, 39], the largest eigenvalue of  $\nabla^2 \phi$  solving eq. (2) is bounded. However, note that in general this term may be non null, as for example in Test-case 3. Supposing that both terms can be neglected, we get the following approximate identity:

$$(\partial_t + \operatorname{div}(\cdot \nabla \phi))(-\Delta_\rho) \approx -\Delta_\rho(\partial_t + \nabla \phi \cdot \nabla).$$

In our finite-dimensional problem, this expression translates into the approximate matrix identity

$$-\mathcal{B}^T \mathcal{M}'^{-1} \tilde{\mathcal{A}} \approx \mathcal{A} \mathcal{M}^{-1} \tilde{\mathcal{B}}, \quad (36)$$

with  $\tilde{\mathcal{A}} \in \mathbb{R}^{m,m}$  and  $\tilde{\mathcal{B}} \in \mathbb{R}^{n,m}$  discretize the operators  $-\Delta_\rho$  and  $\partial_t + \nabla \phi \cdot \nabla$ , respectively, similarly to the matrices  $\mathcal{A}$  and  $\mathcal{B}$  but with different dimensions. The scaling factors  $\mathcal{M}'^{-1}$  and  $\mathcal{M}$  in eq. (36) are introduced to match the scaling dimensions of the matrices  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{B}^T$ . The most natural definition of  $\tilde{\mathcal{A}}$  and  $\tilde{\mathcal{B}}$  is the following:

$$\begin{aligned} \tilde{\mathcal{A}} &= \text{Block Diag} \left( \left( \tilde{\mathcal{A}}^k \right)_{k=1}^K \right), \quad \tilde{\mathcal{A}}^k := \operatorname{div}_{\mathcal{T}', \mathcal{E}'} \operatorname{Diag} \left( R_{\mathcal{E}'} [\rho^k] \right) \nabla_{\mathcal{E}', \mathcal{T}'}, \\ \tilde{\mathcal{B}} &:= \mathcal{J} \mathcal{D}_t^T \mathcal{M}' + \mathcal{G} \tilde{\mathcal{D}}_x \mathcal{J}^T \mathcal{H}^T, \end{aligned}$$

where the matrices  $\mathcal{D}_t$ ,  $\mathcal{J}$ ,  $\mathcal{H}$ , and  $\mathcal{G}$  are defined in eqs. (9) and (10), and

$$\tilde{\mathcal{D}}_x = \text{Block Diag} \left( (\nabla)_{k=1}^K \right)$$

having one block less than  $\mathcal{D}_x$ , which is defined in eq. (8).

Using the approximate identity in eq. (36), the linear system in eq. (30) becomes

$$-g = \mathcal{S}y = \mathcal{P} \left( \mathcal{C} + \mathcal{B} \mathcal{A}^+ \mathcal{B}^T \right) \mathcal{P}^T y \approx \mathcal{P} \left( \mathcal{C} - \mathcal{B} \mathcal{M}^{-1} \tilde{\mathcal{B}} \tilde{\mathcal{A}}^+ \mathcal{M}' \right) \mathcal{P}^T y$$

Since  $\operatorname{Ker}(\mathcal{P} \mathcal{M}') = \operatorname{Im}(\mathcal{E}^T)$  are discrete functions constant in space, we have that  $\operatorname{Im}(\mathcal{M}' \mathcal{P}^T) \subset \operatorname{Im}(\tilde{\mathcal{A}})$  and the last expression is equal to

$$\mathcal{P} \left( \mathcal{C} \mathcal{M}'^{-1} \tilde{\mathcal{A}} \tilde{\mathcal{A}}^+ \mathcal{M}' - \mathcal{B} \mathcal{M}^{-1} \tilde{\mathcal{B}} \tilde{\mathcal{A}}^+ \mathcal{M}' \right) \mathcal{P}^T y = -g.$$

Hence, we only have to solve the linear system

$$\mathcal{P} \left( \mathcal{C} \mathcal{M}'^{-1} \tilde{\mathcal{A}} - \mathcal{B} \mathcal{M}^{-1} \tilde{\mathcal{B}} \right) \tilde{\mathcal{A}}^+ \mathcal{M}' \mathcal{P}^T y = -g,$$

which is compatible since  $g \in \operatorname{Im}(\mathcal{P})$ . Moreover, since the solution  $\delta \rho$  in eq. (28) belongs to the image of  $\mathcal{P}^T$ , we may look for a solution  $y \in \operatorname{Im}(\mathcal{P}^T)$ . By setting

$$z = \tilde{\mathcal{A}}^+ \mathcal{M}' \mathcal{P}^T y = \tilde{\mathcal{A}}^+ \mathcal{M}' y$$

we can compute  $y$  by first solving the linear system

$$(\mathcal{C} \mathcal{M}'^{-1} \tilde{\mathcal{A}} - \mathcal{B} \mathcal{M}^{-1} \tilde{\mathcal{B}}) z = -g, \quad (37)$$

and then computing  $y = \mathcal{M}'^{-1} \tilde{\mathcal{A}} z$ .

We will refer to the resulting preconditioner as “ $\mathcal{B}\mathcal{B}$ -preconditioner”, due to presence of this composed operator in the approximate factorization of the Schur complement. This choice was done in analogy with the “ $\mathcal{B}\mathcal{F}\mathcal{B}t$ -preconditioner” introduced in [36], whose ideas inspired the preconditioner described in this section.

#### 4.5.4 Numerical results

In table 4 we summarized the results for the preconditioner described in this section. The number of averaged inner iterations **Inner/Outer** refers to the solution of the linear system eq. (37), which is solved using the AGMG solver with inner tolerance  $\varepsilon_{\text{in}} = 1e - 1$ . Unfortunately, the number of outer iterations tends to increase as  $\mu \rightarrow 0$  and, in some cases with  $\mu \leq 1e - 4$ , to exceed the 400 limit we fixed for the FGMRES solver. However, for  $\mu \approx 1e - 5$ , this preconditioner is the only one scaling well with respect to the temporal and spatial discretization. The number of averaged outer iterations increases only slightly when the mesh is refined or the time steps is halved. The number **Inner/Outer** remains bounded between 1.7 and 5.7 iterations per preconditioner applications. All these results produce a CPU time that scales well with respect to the number of degrees of freedom used to discretize the problem.

$K$	8	16	32	64	8	16	32	64	8	16	32	64
$\mu$	Outer/Lin.sys.				CPU/Lin.sys.				Inner/Outer			
1	6	8	10	11	0.9	1.1	2.2	4.2	2.0	2.1	2.5	3.1
2e-1	7	8	9	10	0.5	1.0	1.8	3.6	2.2	2.6	3.2	4.0
4e-2	9	11	12	14	0.5	1.0	1.9	3.8	2.2	2.7	3.4	4.3
8e-3	13	14	19	27	0.6	1.0	2.1	4.2	2.1	2.8	3.4	4.2
2e-3	19	22	33	55	0.6	1.1	2.4	5.8	2.3	3.2	3.9	4.2
3e-4	25	33	57	120	0.7	1.2	3.0	8.8	4.0	5.1	5.6	5.4
6e-5	36	48	92	†	0.8	1.4	3.8	†	6.2	8.3	9.6	†
1e-5	60	76	148	†	0.9	1.9	5.8	†	10.0	13.5	16.5	†
3e-6	98	176	†	†	1.2	3.8	†	†	16.5	26.9	†	†
5e-7	†	†	†	†	†	†	†	†	†	†	†	†
	28	41	43	35	0.7	1.4	2.8	4.8	9.5	16.1	10.2	4.7
1	6	8	9	10	0.6	1.0	2.0	4.2	2.6	2.2	2.3	2.9
2e-1	8	8	9	9	0.6	1.1	2.1	4.3	2.2	2.6	3.2	4.0
4e-2	11	12	13	14	0.6	1.2	2.4	4.6	2.1	2.7	3.4	4.1
8e-3	14	17	18	21	0.7	1.2	2.6	5.2	2.1	2.7	3.5	4.2
2e-3	18	22	25	35	0.7	1.4	2.8	6.2	2.1	2.7	3.4	3.9
3e-4	25	34	45	68	0.8	1.6	3.5	9.0	2.2	2.5	3.1	4.0
6e-5	37	52	82	†	0.9	2.1	5.4	†	2.8	4.0	5.1	†
1e-5	60	91	154	†	1.2	3.0	9.0	†	4.9	5.7	7.0	†
3e-6	100	150	250	†	1.8	5.5	17.2	†	9.1	13.1	14.0	†
5e-7	174	†	†	†	3.5	†	†	†	16.3	†	†	†
	41	40	60	23	1.1	1.9	4.8	5.4	9.4	7.4	8.6	3.9
1	6	7	8	9	0.7	1.4	2.8	6.2	3.1	2.9	2.4	2.6
2e-1	8	8	9	9	0.8	1.6	3.0	6.3	2.8	2.7	2.9	3.4
4e-2	12	13	15	16	0.9	1.8	3.8	8.0	2.7	2.4	3.0	3.9
8e-3	16	18	22	22	1.0	2.0	4.4	9.4	2.5	2.4	3.0	4.1
2e-3	21	23	27	28	1.2	2.2	5.0	10.8	2.7	2.4	3.2	4.4
3e-4	27	29	37	42	1.3	2.5	6.0	14.0	2.6	2.5	3.1	4.0
6e-5	38	39	55	73	1.6	3.0	8.0	21.5	2.2	2.4	2.9	4.4
1e-5	53	58	80	121	2.0	4.0	11.2	42.5	2.2	2.6	4.0	8.5
3e-6	73	87	120	355	2.7	6.2	20.2	147.5	3.4	4.3	7.8	15.3
5e-7	96	124	180	†	4.0	11.5	47.5	†	6.5	10.0	19.2	†
	29	37	50	66	1.4	3.4	10.2	26.4	3.5	5.1	9.3	10.6
1	5	6	7	8	1.4	3.0	6.0	13.3	4.1	3.5	3.0	2.5
2e-1	8	8	9	10	1.6	3.3	7.1	15.7	3.2	2.9	2.9	2.9
4e-2	14	14	17	20	2.2	4.4	10.0	25.0	3.0	2.9	2.5	3.1
8e-3	20	20	23	31	2.8	5.4	12.8	34.0	3.2	2.9	2.5	3.2
2e-3	26	26	31	41	3.5	6.5	15.5	42.5	3.1	2.8	2.6	3.4
3e-4	39	34	38	50	5.0	8.0	18.0	50.0	2.6	2.8	2.5	3.4
6e-5	57	47	53	62	7.2	10.0	23.5	60.0	2.4	2.5	2.3	3.2
1e-5	87	73	75	115	10.3	14.3	32.5	102.0	2.3	2.2	2.4	3.0
3e-6	132	111	124	190	13.7	21.0	55.0	197.5	2.8	2.3	3.0	4.8
5e-7	188	154	223	†	21.0	32.7	122.5	†	3.7	3.5	5.5	†
	46	40	53	52	5.6	9.0	27.1	53.8	3.0	2.8	3.7	3.7

Table 4: Numerical results using the  $\mathcal{BB}$ -preconditioner defined in eq. (29) for the linear system eq. (28) within a Flexible GMRES cycle. Each sub-table reports the results obtained for a given mesh, from the coarsest  $\mathcal{T}_{h_0}$  (top) to the finest  $\mathcal{T}_{h_0/8}$  (bottom). Each sub-table contains 3 blocks of 4 columns. Each block reports one metric used to measure the preconditioner performance that are, from left to right, the averaged number of outer FGMRES iterations, the average CPU time per linear system, and the averaged number of inner iterations for solving the linear system eq. (37) (see items 1 to 3). Within each block, each column reports the related metric for one choice of time step  $\Delta t = 1/K$  with  $K = 8, 16, 32, 64$  as the relaxation parameter (reported in the leftmost column). The † symbol denotes those IP iterations where the solver failed. A final row summarizes the same quantities averaged on the whole simulation (in case of failure, until the last successful IP iteration). We highlighted in gray the spatial-temporal combination providing the uniform discretization.

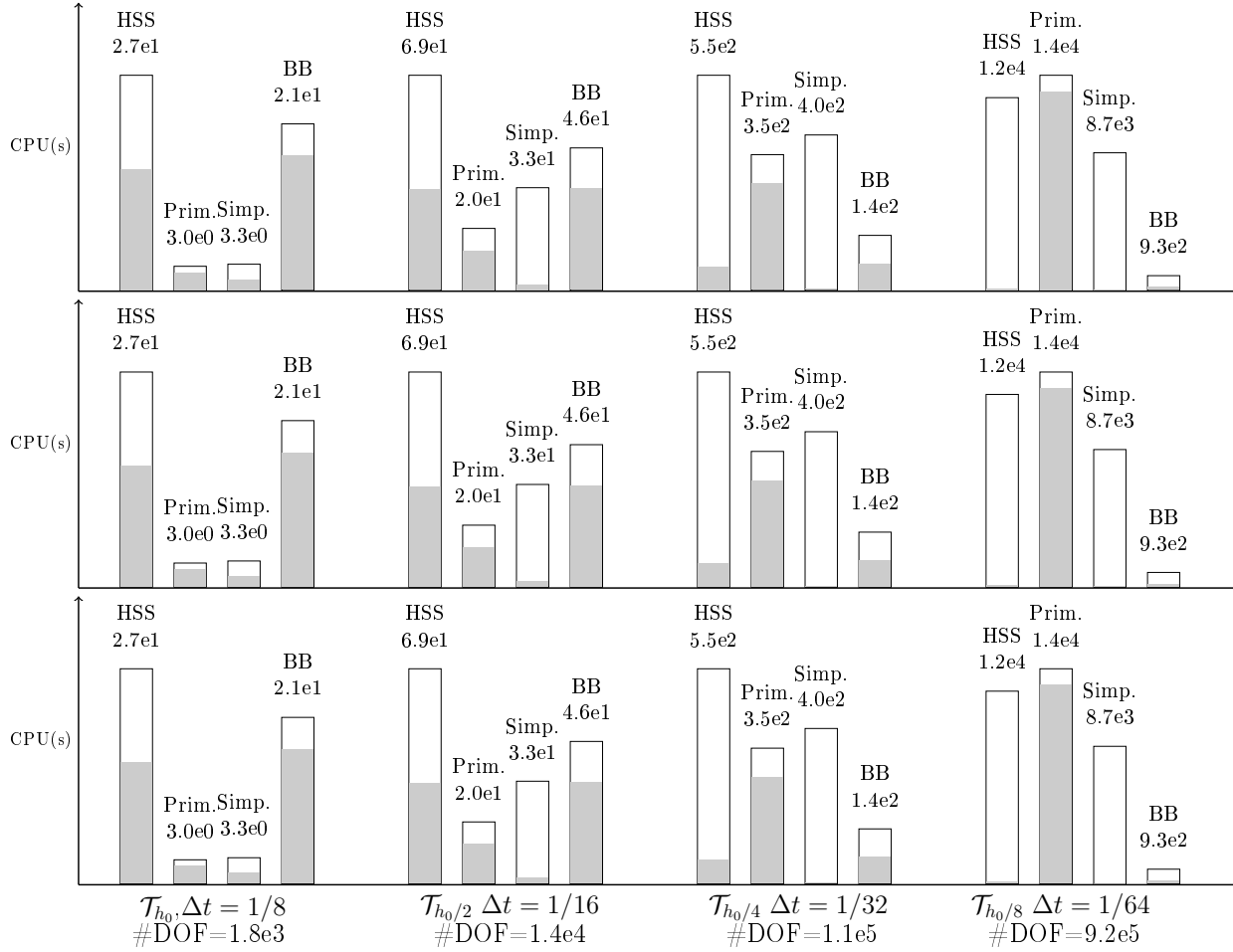


Figure 1: Comparison of CPU time spent in solving linear algebra problems to reach an IP relaxation  $\mu \approx 1e-5$  using the preconditioners approaches presented in this paper. They are denoted by HSS (the Hermitian skew-Hermitian preconditioner in section 4.2), Prim. (the preconditioner based on the primal Schur complement in section 4.3), Simp. (the SIMPLE preconditioner in section 4.4), and the **BB**-preconditioner. The results refers to Test-cases 1,2, and 3 (from top to bottom panel) and using different time-space discretization (left to right). The columns height is normalized by the maximum among the four preconditioners. The gray portion of the column denotes the part of the preprocess time of each preconditioner.

#### 4.6 Summary of numerical results

In this section, we summarize the numerical results to compare the pros and cons of the preconditioning approaches proposed in this paper. In fig. 1 we compare the total CPU time (y-axis) required to achieve IP relaxation  $\mu \approx 1e-5$  (first eight IP iterations) with respect to the total number of degrees of freedom used, while halving the time step and the mesh size used. This summarizes the data in tables 1 to 4 with gray background color. We included in this comparison the results obtain for the test cases 1 and 3. From fig. 1 we can observe that the preconditioner described in section 4.5 is the most efficient among those presented in this paper when larger problems are considered.

The HSS and the SIMPLE preconditioners are rather robust with respect to different time steps and relaxation parameters  $\mu$ . However, they become inefficient when large grids are used. In fact, the CPU time approximately scales linearly with respect to the number of time steps and the number of interior point steps, but quadratically with respect to the number of cells. The SIMPLE preconditioner is the only one that becomes faster in the last IP iterations. Thus, it must be the only viable option when accurate solutions of the OT problem are required, possibly combined with more efficient approaches for the initial IP steps.

The behavior of the preconditioner based on the primal Schur complement described in section 4.3 is strongly influenced by the inner solver used to solve the block linear system 18. In our case, we used the AGMG solver whose

performance is strongly influenced by the balance of the spatial and temporal component of the primal Schur complement. This approach can result in an efficient and scalable approach if we could drastically reduce the time required to assemble the sequence of coarse matrices used by the multigrid solver.

The **BB**-preconditioner described in section 4.5 is the only one able to tackle efficiently large scale problems, since the number of inner iterations required per each Krylov step remains rather constant, while the number of outer iterations increases only slightly using smaller time-step and finer grids. This holds as long as the IP tolerance is approximately  $1e - 5$ . Using smaller values, the number of inner and outer iterations increases rapidly, leading in some cases to the solver's failure.

## 5 Conclusions

We presented different preconditioners for solving via iterative methods the saddle point linear systems arising from the solution of the Benamou-Brenier formulation via IP methods. The most efficient approach turned out to be a triangular preconditioner where the inverse of dual Schur complement is approximated exploiting a partial commutation of its components, named **BB**-preconditioner. While this approach loses efficiency in the latest IP steps, it is the only one scaling well with respect to time and space discretization size. In a series of numerical experiments we showed that, if the IP method is stopped when the relaxation parameter  $\mu$  is approximately  $1e - 5$ , we can solve our problem with a good scaling of the CPU time with respect to the number of degrees of freedom. This value of  $\mu$  is quite reasonable, and consistent with discretization error. Moreover, the **BB**-preconditioner is well suited for parallelization, since its application requires the solution of a sequence of decoupled weighted Laplacians and the solution of the large but sparse linear system (37).

Let us also remark that if further density-dependent energies are included in eq. (1), the structure of the problem does not change and the **BB**-preconditioner can be applied without major modifications.

## References

- [1] Cédric Villani. *Topics in Optimal Transportation*. Graduate studies in mathematics. American Mathematical Society, 2003.
- [2] Filippo Santambrogio. Optimal Transport for Applied Mathematicians. *Birkäuser, NY*, 55(58-63), 2015.
- [3] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on Optimal Transport*. Springer, 2021.
- [4] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.
- [5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in Metric Spaces and in the Space of Probability Measures*. Springer Science & Business Media, 2005.
- [6] Alfred Galichon. Optimal transport methods in economics. In *Optimal Transport Methods in Economics*. Princeton University Press, 2016.
- [7] Ludovic Métivier, Romain Brossier, Félix Kpadonou, Jérémie Messud, and Arnaud Pladys. A review of the use of optimal transport distances for high resolution seismic imaging based on the full waveform. *arXiv preprint arXiv:2204.08514*, 2022.
- [8] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.
- [9] Yves Achdou, Fabio Camilli, and Italo Capuzzo-Dolcetta. Mean field games: numerical methods for the planning problem. *SIAM Journal on Control and Optimization*, 50(1):77–109, 2012.
- [10] Aymeric Baradat and Hugo Lavenant. Regularized unbalanced optimal transport as entropy minimization with respect to branching brownian motion. *arXiv preprint arXiv:2111.01666*, 2021.
- [11] Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. Optimal transport with proximal splitting. *SIAM Journal on Imaging Sciences*, 7(1):212–238, 2014.
- [12] Hugo Lavenant, Sebastian Claiici, Edward Chien, and Justin Solomon. Dynamical optimal transport on discrete surfaces. *ACM Transactions on Graphics (TOG)*, 37(6):1–16, 2018.
- [13] Jose A Carrillo, Katy Craig, Li Wang, and Chaozhen Wei. Primal dual methods for wasserstein gradient flows. *Foundations of Computational Mathematics*, 22(2):389–443, 2022.
- [14] Andrea Natale and Gabriele Todeschi. A mixed finite element discretization of dynamical optimal transport. *Journal of Scientific Computing*, 91(2):1–26, 2022.

- [15] Andrea Natale and Gabriele Todeschi. Computation of optimal transport with finite volumes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 55(5):1847–1871, 2021.
- [16] Stephen J Wright. *Primal-Dual Interior-Point Methods*. SIAM, 1997.
- [17] Stephen Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, 2004.
- [18] Enrico Facca, Franco Cardin, and Mario Putti. Towards a stationary Monge-Kantorovich dynamics: The Physarum Polycephalum experience. *SIAM Journal on Applied Mathematics*, 78(2):651–676, 2018.
- [19] Enrico Facca, Sara Daneri, Franco Cardin, and Mario Putti. Numerical solution of Monge–Kantorovich equations via a dynamic formulation. *Journal of Scientific Computing*, 82(3):1–26, 2020.
- [20] Yves Achdou and Victor Perez. Iterative strategies for solving linearized discrete mean field games systems. *Networks and Heterogeneous Media*, 7(2):197–217, 2012.
- [21] Michele Benzi and Gene H. Golub. A preconditioner for generalized saddle point problems. *SIAM Journal on Matrix Analysis and Applications*, 26(1):20–41, 2004.
- [22] S.V Patankar and D.B Spalding. A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. *International Journal of Heat and Mass Transfer*, 15(10):1787–1806, 1972.
- [23] Howard Elman, Victoria E. Howle, John Shadid, Robert Shuttleworth, and Ray Tuminaro. Block preconditioners based on approximate commutators. *SIAM Journal on Scientific Computing*, 27(5):1651–1668, 2006.
- [24] Robert Eymard, Thierry Gallouët, and Raphaële Herbin. Finite volume methods. *Handbook of numerical analysis*, 7:713–1018, 2000.
- [25] Henk A Van der Vorst. *Iterative Krylov Methods for Large Linear Systems*. Number 13. Cambridge University Press, 2003.
- [26] Michele Benzi, Gene H. Golub, and Jörg Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.
- [27] Youcef Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM Journal on Scientific & Statistical Computing*, 14(2):461–469, 1993.
- [28] Yvan Notay. An aggregation-based algebraic multigrid method. *Electronic Transactions on Numerical Analysis*, 37:123–146, 2010.
- [29] Fvcav, benchmark, 2008.
- [30] Zhong-Zhi Bai, Gene H. Golub, and Michael K. Ng. Hermitian and skew-hermitian splitting methods for non-hermitian positive definite linear systems. *SIAM Journal on Matrix Analysis and Applications*, 24(3):603–626, 2003.
- [31] Enrico Facca and Michele Benzi. Fast iterative solution of the optimal transport problem on graphs. *SIAM Journal on Scientific Computing*, 43(3):A2295–A2319, 2021.
- [32] Michele Benzi, Eldad Haber, and Lauren Taralli. Multilevel algorithms for large-scale interior point methods. *SIAM J. Sci. Comput.*, 31(6):4152–4175, 2009.
- [33] Suhas V. Patankar. *Numerical Heat Transfer and Fluid Flow*. Hemisphere Pub. Corp. ; McGraw-Hill Washington : New York, 1980.
- [34] Lennard Kamenski, Weizhang Huang, and Hongguo Xu. Conditioning of finite element equations with arbitrary anisotropic meshes. *Mathematics of computation*, 83(289):2187–2211, 2014.
- [35] Howard Elman, Victoria E. Howle, John Shadid, David Silvester, and Ray Tuminaro. Least squares preconditioners for stabilized discretizations of the Navier-Stokes equations. *SIAM Journal on Scientific Computing*, 30(1):290–311, dec 2007.
- [36] Howard C Elman. Preconditioning strategies for models of incompressible flow. *Journal of Scientific Computing*, 25(1):347–366, 2005.
- [37] Stefán Ingi Valdimarsson. On the hessian of the optimal transport potential. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, Ser. 5, 6(3):441–456, 2007.
- [38] Luis A. Caffarelli. Monotonicity properties of optimal transportation and the FKG and related inequalities. *Communications in Mathematical Physics*, 214(3):547–563, 2000.
- [39] Maria Colombo, Alessio Figalli, and Yash Jhaveri. Lipschitz changes of variables between perturbations of log-concave measures. *Annali Scuola Normale Superiore - Classe di scienze*, pages 1491–1519, 2017.