UiO : University of Oslo

Karl Erik Holter

Robust preconditioning of multiphysics problems and interstitial fluid flow

Thesis submitted for the degree of Philosophiae Doctor

Department of Informatics
Faculty of Mathematics and Natural Sciences

University of Oslo
Simula Research Laboratory

2021

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2394

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Acknowledgments

"Whenever there is a difficult question that you cannot answer, there is also a simpler question that you cannot answer." —Achi Brandt

First of all, I would like to thank my supervisors Kent-Andre Mardal, Unn Haukvik and Anders Dale, who have all been forthcoming with both guidance and care as appropriate. I am particularly grateful to Kent for his ability to explain mathematics in language even a lapsed algebraic geometer could understand, the book recommendations, the occasional (but generally resonant) piece of life advice, the frequent laughs, and for showing me a level of insight, intuition and enthusiasm about mathematics that I'll forever aspire to. I would also like to thank Miroslav Kuchta, who has always been willing to share of his time and extensive knowledge of matters both mathematical and computational.

As a PhD student employed as part of the SUURPh programme, I am thankful to the Norwegian Ministry of Research and Education for funding the programme, and to Marie Rognes and Gabriel Balaban for encouraging me to apply. I would also like to thank everyone involved in the programme, including in particular Andy Edwards, Rachel Thomas, Elin Backe Christophersen and Kimberly McCabe for considerable and impressively varied support.

During the SUURPh programme, I also had the opportunity to visit the University of California, San Diego, and I would like to thank Anna Devor, Martin Thunemann and Kim Weldy for making my stay at the Devor Lab possible, pleasant and educational.

My time as a Ph.D. student was also significantly improved by my fellow class of SUURPh students Alessio Buccino, Aslak Bergersen, Eleonora Piersanti, Jonas van den Brink, Karoline Jæger, Liubov Nikitushkina, Solveig Næss, Tristan Stöber and Viviane Timmermann, and I would like to thank them for being there to share in the ups and downs involved in a Ph.D. Getting to meet you all was definitely among the ups.

Finally, I would like to express my deepest gratitude and appreciation to my family and to Marthe Fallang, the person who came with me and with whom I want to go. Their limitless love and belief in me has been a great comfort throughout the years in general and during the time spent working on this thesis in particular.

Karl Erik Holter
Oslo, April 2021

List of Papers

Paper I

Holter, et al.

Holter, Kehlet, Devor, Sejnowski, Dale, Omholt, Ottersen, Nagelhus, Mardal and Pettersen. 'Interstitial solute transport in 3D reconstructed neuropil occurs by diffusion rather than bulk flow'.

In: *Proceedings of the National Academy of Sciences* **114.37** (2017), pp. 9894–9899. DOI: 10.1073/pnas.1706942114.

Paper II

Holter, Kuchta, and Mardal. 'Sub-voxel perfusion modeling in terms of coupled 3d-1d problem'.

arXiv: 1803.04896. In: *ENUMATH 2017 Proceedings.*

Paper III

Holter, Kuchta, and Mardal. 'Robust preconditioning of monolithically coupled multiphysics problems'.

arXiv: 2001.05527 *In submission.*

Paper IV

Holter, Kuchta, and Mardal. 'Robust preconditioning for coupled Stokes-Darcy problems with the Darcy problem in primal form'.

arXiv: 2001.05529 In: *Computers & Mathematics with Applications 2020.* DOI: 10.1016/j.camwa.2020.08.021

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, we motivate the problems considered in this thesis. I have made a best effort to keep the chapter brief and accessible to readers without extensive background knowledge, but this has required some details to be elided. It is my hope that the reader will be capable of filling them in on their own as required.

We start out by recalling how the finite element method can reduce the problem of approximating the solution to a linear partial differential equation to that of solving a system of linear equations. Next, we explain why the resulting systems of linear equations can be difficult to solve despite being finite dimensional. We therefore introduce the idea of preconditioning as a way of making the system of linear equations tractable. We then recall the framework of operator preconditioning, which connects the problem of obtaining a preconditioner for the system of linear equations with the problem of finding the "correct" function spaces for the continuous PDE.

Having established the framework in which we will be working, we recall the celebrated criteria of Babuska and Brezzi for stability of weak problems in Section 1.3.1. We use this in Section 1.4 to prove Lemma 1.4.1, which shows that using Lagrange multipliers to enforce boundary conditions yields a stable problem if and only if enforcing boundary conditions strongly yields a stable problem. This way of enforcing boundary conditions requires the introduction of trace operators, which can only be properly done using fractional Sobolev spaces. These are therefore introduced in Section 1.4.1.

We are then in a position to introduce the idea of multiphysics problems in Section 1.5, which are of immense interest in applications and the main subject of this thesis. A natural way of modeling multiphysics problems is as compositions of multiple "single physics" subproblems coupled via Lagrange multipliers, which will be seen to result in a saddle point problem. We state and prove simple criteria which can be used to prove such problems well-posed under appropriate conditions. Finally, in Section 1.6 we discuss methods suitable for solving coupled multiphysics problems.

1.1 Preliminaries

We will use $L^2(\Omega)$ to denote the Sobolev (and Hilbert) space of functions for which the L^2-norm $\|f\|^2_{L^2(\Omega)} := \int_\Omega f^2$ is finite. Similarly, we define the space $H^1(\Omega)$ with the norm $\|f\|^2_{H^1(\Omega)} := \|f\|^2_{L^2(\Omega)} + \|\nabla f\|^2_{L^2(\Omega)}$, and $H(\mathrm{div}, \Omega)$ by $\|\mathbf{f}\|^2_{H(\mathrm{div},\Omega)} = \|\mathbf{f}\|^2_{L^2(\Omega)} + \|\nabla \cdot \mathbf{f}\|^2_{L^2(\Omega)}$. For a Hilbert space U, we let U^* denote its dual space, defined as the vector space of all *continuous* linear functionals $f : U \to \mathbb{R}$, and define on it the operator norm $\|f\|_{U^*} := \sup_{u \in U} \frac{f(u)}{\|u\|_U}$. By the Riesz representation theorem, U and U^* are isometric, and we let $R_U^{-1} : U \to U^*$

denote the map $u \to (u, -)_U$. We emphasize that this isomorphism is "weakly non-canonical" in the sense that it depends on the inner product of U.

As the choice of inner product will be central to several arguments, we will always write inner products as $(u, v)_U$, indicating the space with a subscript . For $u \in U, v \in U^*$, we will let $\langle v, u\rangle_U \to v(u)$ denote the duality pairing $U^* \times U \to \mathbb{R}$ evaluating a linear functional at a point in its domain, as distinguished from the inner product $(u_1, u_2)_U$, where $u_1, u_2 \in U$. In cases where it is clear from context, we will omit the subscript from the duality pairing.

1.2 Preconditioning for the finite element method

1.2.1 A brief overview of the finite element method

The finite element method is perhaps the most widely used technique for numerically solving partial differential equations. In this section, we outline very briefly how and why it works, and what challenges it leaves unsolved for practitioners. For concreteness' sake, suppose Ω is a given bounded domain, V some given space of functions on Ω and we want to find $u \in V$ so that

$$-\Delta u(x) = f(x) \text{ for all } x \in \Omega \tag{1.1}$$

Here, the unknown u and the known right hand side f are both functions $\Omega \to \mathbb{R}$, and the equation is assumed to hold on all points of Ω. Accordingly, Equation (1.1) is called the *strong form* of the PDE. The first step in applying the finite element method is to rewrite the problem in *weak form* by multiplying by an arbitrary test function v and integrating over Ω. Doing this, we see that any solution u of the strong form must also have the property that

$$\int_\Omega -v\Delta u = \int_\Omega fv \text{ for all test functions } v \in V \tag{1.2}$$

For reasons which will later become clear, we integrate by parts and rewrite as

$$\int_\Omega \nabla u \nabla v = \int_{\partial\Omega} v\frac{\partial u}{\partial n} + \int_\Omega fv \text{ for all test functions } v \in V \tag{1.3}$$

What makes this form useful for applications is that we no longer require the u to be as smooth for this to be well-defined, and it now suffices that $u, v \in H^1(\Omega)$. We will use it to obtain a system of linear equations whose solution yields an approximation of the exact solution. To do this, we choose some triangulation $\Omega_h \subset \Omega$, and replace the space V by the finite-dimensional subspace V_h of piecewise linear functions[1] with break points only at the edges of the triangulations, as shown in Figure 1.1.

Such a function is clearly determined by its values at the vertices of the triangles, and we see that a basis for V_h consists of the functions ϕ_i which are 1

[1] We could also choose different spaces V_h. For example, we could use piecewise quadratic functions. The choice of exactly what functions we are considering on each triangle is our choice of finite element.

at vertex i and 0 at all other vertices. To obtain the requisite system of linear equations, then, denote the (unknown) values of the (approximate) solution u_h at the vertices $v_1, \ldots, v_N$ by $u_1, \ldots, u_N$. Then we have that

$$u_h = u_1\phi_1 + \ldots + u_N\phi_N$$

Substituting this into Equation (1.3), we see that by choosing $v = \phi_i$ for $i = 1, \ldots, N$, we obtain N different equations involving the unknowns u_i. Because ϕ_i, f are known functions, all other quantities involved can be computed, and by linearity, the resulting system is linear, and can be written in the form $\mathbf{Au} = \mathbf{f}$ where now $\mathbf{A} \in \mathbb{R}^{N\times N}$, and $\mathbf{u}, \mathbf{f} \in \mathbb{R}^N$. Now, if we can solve this system for $\mathbf{u}$, we have that $u_h = u_1\phi_1 + \ldots + u_N\phi_N$ is an approximate solution of the original problem, completing the reduction of the (approximate) original problem to a linear system of equations.

1.2.2 The computational cost of the finite element method

We have so far not shown that the approximation u_h, for which we derived a linear system in the preceding section, is in fact close to the true solution u. It might be expected that it is closer the more triangles we use, and that in the limit $N \to \infty$, $u_h \to u$. In this case, this is true, and the error $\|u - u_h\|_{H^1(\Omega)}$ is of order $\mathcal{O}(h)$, where h is the length scale of the triangles involved. For now, we will take this for granted, and return to the question of what happens when we replace a continuous problem by a discrete one later.

Suppose, then, that we are using the finite element method to solve a PDE. We have an error threshold of ϵ, and per the reasoning above, we must therefore choose the triangles to be of size $\mathcal{O}(\epsilon)$. If Ω is a 2D domain as in the figure, this means that we require the number of vertices N to be $\mathcal{O}(\epsilon^{-2})$. From this, we see that the size of the systems we want to solve grows quadratically in the inverse of the approximation error we want. This means that the systems quickly become very large, so the computational cost of application of the finite element method quickly comes to depend on the algorithm chosen to solve the resulting linear system.

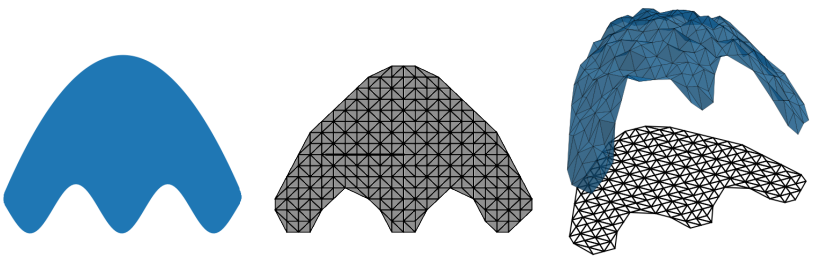

Figure 1.1: A domain Ω on the left, a discretization Ω_h in the middle, and a piecewise linear function u_h on the right.

Before proceeding, we observe that although the matrix $\mathbf{A}$ described above is nominally a $N \times N$ matrix, most of its entries will be zero. The reason for this is that if two vertices i and j are far away from each other, we see from Equation (1.3) that $\mathbf{A}_{i,j}$ will be zero, as only vertices which share some triangle "interact". Assuming that the number of triangles each vertex is in is roughly constant, this means that the number of nonzero elements Z of our matrix will be $\mathcal{O}(N)$.

Methods for the solution of the general linear problem $\mathbf{Ax} = \mathbf{b}$ can, broadly speaking, be divided into direct methods and iterative methods. Direct methods solve the system in "one shot". The traditional pen-and-paper method of Gaussian elimination taught in schools is an example of a direct method, as are techniques based on factorizing $\mathbf{A}$ like the LU composition. The performance of such methods is outside the scope of this thesis, but commonly the sparse problem matrices used for the finite element method on 3D domains result in $\mathcal{O}(N^2)$ time complexity and superlinear memory complexity [88]. See, however, [2] for specifics on the algorithm used in Figure 1.2, and [33], [3] for discussions of recent advances.

Iterative methods operate by starting with some initial guess $\mathbf{x}_0$, and repeatedly computing better guesses $\mathbf{x}_{n+1}$ until a sufficiently good guess is obtained. An example[2] is the Richardson iteration, where

$$\mathbf{x}_{n+1} = \mathbf{x}_n + r\left(\mathbf{b} - \mathbf{A}\mathbf{x}_n\right) \tag{1.4}$$

for an appropriate real number r chosen to ensure the sequence converges to the true solution. We see from Equation (1.4) that the computational cost of a single iteration is dominated by the cost of computiting $\mathbf{Ax}$. As A is sparse, this can be seen to be $\mathcal{O}(N)$, meaning that the total cost becomes $\mathcal{O}(TN)$, where T is the required number of iterations to reach an acceptable error.

1.2.2.1 Convergence properties of the Richardson iteration

In order to determine the total computational cost of our method, we are therefore forced to consider how rapidly the sequence $\mathbf{x}_n$ converges. Suppose x is the exact solution, so that $\mathbf{Ax} = \mathbf{b}$. Then

$$\mathbf{x}_{n+1} - \mathbf{x} = \mathbf{x}_n - \mathbf{x} + r(\mathbf{b} - \mathbf{A}\mathbf{x}_n) = (\mathbf{I} - r\mathbf{A})(\mathbf{x}_n - \mathbf{x})$$

From this, we see that the error $\mathbf{e}_n := \mathbf{x}_n - \mathbf{x} = (\mathbf{I} - r\mathbf{A})^n \mathbf{e}_0$. Thus the convergence properties of the Richardson iteration depend on the matrix $\mathbf{M}_r := \mathbf{I} - r\mathbf{A}$, and it can be shown that the convergence rate will be $\|\lambda_r\|$, where λ_r is the eigenvalue of $\mathbf{M}_r$ of largest magnitude. Hence we must choose r so that $\|\lambda_r\| < 1$ to ensure the method converges, and preferably so that it is as small as possible.

Supposing that $\mathbf{A}$ is symmetric and positive definite, it has real eigenvalues $\lambda_1 \leq \ldots \leq \lambda_N$, and we see that the eigenvalues of $\mathbf{M}_r$ are $1 - r\lambda_1 \geq \ldots \geq 1 - r\lambda_N$.

[2] The choice of the Richardson iteration here is for ease of exposition, not because it is the most commonly used iterative method.

Hence the convergence rate is optimal[3] if

$$1 - r\lambda_1 = -(1 - r\lambda_N) \Rightarrow r = \frac{2}{\lambda_1 + \lambda_N},$$

in which case the convergence rate is

$$1 - r\lambda_1 = 1 - \frac{2}{1 + \kappa(\mathbf{A})}$$

where we have introduced the *condition number* $\kappa(\mathbf{A}) := \frac{\lambda_N}{\lambda_1}$. This means that if the condition number is large, the convergence rate will be just barely below 1, and thus very slow. Hence while if the condition number is close to 1, the convergence rate will be close to 0 and thus very fast.

We are now in a position to estimate what happens to the required number of Richardson iterations as N increases. As we have established that $\epsilon_n = \mathcal{O}\left(\left(1 - \frac{2}{1+\kappa(\mathbf{A})}\right)^n\right)$, using the approximation $\ln(1+x) \approx x$ on $x = \frac{2}{1+\kappa(\mathbf{A})}$ we see that

$$\epsilon_n = \mathcal{O}\left(\exp\left(\frac{-2n}{1 + \kappa(\mathbf{A})}\right)\right).$$

Recalling that $\mathbf{A}$ depends on N, we see that the number of iterations T required to reach some acceptable error threshold will have the same growth rate as $\kappa(\mathbf{A})$ as N increases. In particular, if $\kappa(\mathbf{A})$ is bounded, so is the required number of iterations.

1.2.2.2 Condition numbers of finite element matrices

As the final step of our estimate, it remains to estimate $\kappa(\mathbf{A})$. For the matrix $\mathbf{A}$ defined earlier as the coeffient matrix of Equation (1.3) discretized by piecewise linear elements into N triangles, the condition number will have growth rate $\mathcal{O}(N)$. We shall not prove this formally, but give a heuristic argument as to why it is true: note that the solutions of the continuous eigenproblem $-\Delta u = \lambda u$ on the real plane are $\lambda_{n,m} = n^2 + m^2, u_{n,m} = \sin(nx + my)$. As $\sin(nx + my)$ has oscillations with a length scale $\mathcal{O}\left(1/\max\{n, m\}\right)$, not all of these will be representable in our discretized space where the triangles have sides of size $h = \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. We would therefore expect the largest eigenvalue of our discretized operator $\mathcal{A}$ to be $\mathcal{O}\left(\frac{1}{h^2}\right) = \mathcal{O}\left(N\right)$, and the smallest eigenvalue to be constant, yielding the expected condition number.

This means that the required number of iterations is $T = \mathcal{O}(\kappa(\mathbf{A})) = \mathcal{O}(N)$. Thus the total cost of using Richardson iteration to solve our system is $\mathcal{O}(TN) = \mathcal{O}(N^2)$. As we require $N = \mathcal{O}(\epsilon^{-2})$ to obtain an H^1 error of $\mathcal{O}(\epsilon)$, this means that our algorithm has a total running time of $\mathcal{O}(\epsilon^{-4})$, and that halving the error will require 16 times as much computation. This quickly becomes prohibitive - assuming for ease of comparison that one computational step is a nanosecond, a

[3]The condition is saying that the eigenvalues of $\mathbf{M}_r$ are clustered about 0 in such a way that the largest and smallest are equidistant to 0.

reasonable error tolerance of $\epsilon = 10^{-4}$ means that ϵ^{-4} steps is required, which is more than 100 days. On the other hand, if we had been able to keep T constant, we would be looking at (a constant times) ϵ^2 instead, which is about a second. In three dimensions, we would have $N = \mathcal{O}(\epsilon^{-3})$ instead, making the difference even more pronounced.

We remark that although the exact convergence rate obtained here is only valid for the Richardson iteration for a symmetric positive definite matrix in particular, techniques obtaining convergence criteria on iterative methods by looking at the spectrum of eigenvalues of the operator can be extended to more general cases, and overall the conclusion that iterative methods converge faster when $\kappa(\mathbf{A})$ is close to 1 holds.[4] However, a more general tool for deriving bounds on iteration counts of iterative methods is the *field of values* or *numerical range* of an operator $A : V \to V^*$, defined as the range of $\frac{\langle Ax,x\rangle}{(x,x)_V}$ over nonzero $x \in V$. It evidently contains the eigenvalues of $R_V^{-1}A$, meaning it contains the spectrum, and per the Toeplitz-Haussdorff theorem (see e.g. [60]) it is also convex, and hence contains the convex hull of the (complex) spectrum. We refer the reader to [14] for details of the properties of the field of values, and to [15, 41] for examples of how it can be used to analyze preconditioners for PDEs.

1.2.3 Operator preconditioning of finite element matrices

We are thus led to pose the following question: Could we replace the system $\mathbf{Ax} = \mathbf{b}$ by the *preconditioned* system $\mathbf{PAx} = \mathbf{Pb}$ for some nonsingular matrix $\mathbf{P}$ so that $\mathbf{PA}$ has a smaller condition number than $\mathbf{A}$? Evidently the solution is the same, but if we could choose the *preconditioner* $\mathbf{P}$ so that $\mathbf{PA}$ has a bounded condition number, per the reasoning of Section 1.2.2.1 we would have an $\mathcal{O}(N)$ algorithm as opposed to an $\mathcal{O}(N^2)$ one. An obvious choice here is, of course, $\mathbf{P} = \mathbf{A}^{-1}$. This makes $\kappa(\mathbf{PA}) = 1$, which would be good. However, because actually computing $\mathbf{P}$ in this case means inverting $\mathbf{A}$, it is useless for our purposes, as using it would be equivalent to just solving our original system. For $\mathbf{P}$ to be a suitable preconditioner, then, we therefore also require that $\mathbf{Px}$ be efficiently computable, as we need to compute it at every iteration of our iterative method.

As we have established that a good preconditioner $\mathbf{P}$ is one for which the eigenvalues of $\mathbf{PA}$ are bounded above and below, a good preconditioner must be spectrally equivalent to $\mathbf{A}^{-1}$. It is, then, perhaps not surprising that one way to obtain good preconditioners is through the framework of *operator preconditioning*, on which we shall spend the remainder of this section. Our exposition will follow that of [71] and [86] closely, and the interested reader is referred there for a more thorough treatment.

[4] In the case of a symmetric (resp. symmetric positive definite) matrix, using the MINRES (resp. CG) iterative methods instead will yield faster convergence. Letting $\kappa := \kappa(\mathbf{A})$, we then obtain convergence bounds of $2\left(\frac{\kappa-1}{\kappa+1}\right)^{\frac{N}{2}}$ (resp. $2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{N}$) instead, see e.g. [82]. The conclusion that a bounded condition number implies a bounded number of iterations thus remains valid also for MINRES and CG.

We start out by examining the apparent contradiction[5] between the fact that the discrete problem we ended up with had a poor condition number, and the fact that the continuous problem we started with, $-\Delta u = f$, enjoys several guarantees of well-posedness. We state without proof the following theorem, and later return to ways in which such theorems can be established:

Theorem 1.2.1 (Well-posedness of the Poisson problem)**.** *Let $a(u,v) = \int_\Omega \nabla u \cdot \nabla v$, and $L(v) = \int_\Omega fv$. Then the problem of finding $u \in H_0^1(\Omega)$ so that*

$$a(u,v) = L(v) \text{ for any } v \in H_0^1(\Omega)$$

is well-posed in the sense that for any f, there exists a unique solution $u \in H_0^1(\Omega)$, and there exists a constant C depending only on Ω for which we have the bound

$$\|u\|_{H^1(\Omega)} \leq C\|f\|_{H^{-1}(\Omega)}$$

Denoting by $A : H_0^1(\Omega) \to \left(H_0^1(\Omega)\right)^*$ the map sending any $u \in H_0^1(\Omega)$ to the linear functional $v \to a(u,v)$, we observe that this theorem implies that the operator norm of $\|A^{-1}\|$ is bounded. As an alternative characterization of the condition number which extends to the continuous case is as $\|A\|\|A^{-1}\|$, the apparent contradiction is that this theorem implies that the continuous condition number is bounded, while in the discrete case, we saw that this was not so.

One resolution might be that replacing the space $V = H_0^1(\Omega)$ by some discrete subspace V_h (as we did by considering only piecewise linear functions) takes away the well-posedness properties, but in this case, this is incorrect - the above theorem remains valid when V is replaced by V_h, with the constant C independent of the discretization scale h.

The resolution to the apparent contradiction is that the operator norm $\|A\|$ of our map $V \to V^*$, defined as $\sup_{v \in V} \frac{\|Av\|_{V^*}}{\|v\|_V}$ depends on the norm we choose for V. In the continuous case, in which we have a bounded condition number, the norm chosen is the H^1 norm, and we are treating $\mathbf{A}$ as a map $V \to V^*$. In order for the definitions of $\kappa(\mathbf{A})$ in terms of eigenvalues and in terms of operator norms to be equivalent, the norms used for the operator norm must in fact be the standard Euclidean ℓ^2 norm, and indeed, if we compute the constant C of the above theorem in this norm, it will be seen to depend on h, and we will end up with the same dependence of the condition number on h.

An instructive way of thinking about what is going on is that by defining the condition number in terms of the eigenvalues λ for which $\mathbf{Ax} = \lambda\mathbf{x}$, we are comparing a vector $\mathbf{Ax}$ in the codomain of $\mathbf{A}$ with a vector $\lambda\mathbf{x}$ in the domain of $\mathbf{A}$. Hence we are making an implicit identification between the domain V and the codomain V^*, which in this case lead us to an implicit choice of the ℓ^2 norm.

If we want to choose other norms, we must make this identification explicit. Suppose we have some norm $\|\cdot\|_V$ in mind, and let R be the matrix for which $\|\mathbf{u}\|_V^2 = \mathbf{u}^\mathbf{T}\mathbf{Ru}$. Here, R can be thought of as a map $V \to V^*$, with inverse being the Riesz map of the space V, and if we define the condition number of $\mathbf{A}$

[5] Of course, as will become clear, there is no literal contradiction here.

not in terms of the standard eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{x}$, but the generalized eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{Rx}$, there is no implicit identification being made on our behalf, as we are comparing two vectors in $\mathbf{V}^*$. This results in a definition of the condition number which is, in fact, equivalent to the one involving the operator norm of $A : V \to V^*$ and its inverse.

Looking at the generalized eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{Rx}$, this can be recast as a standard eigenvalue problem $\mathbf{R}^{-1}\mathbf{Ax} = \lambda\mathbf{x}$, meaning that the V-condition number of A is in fact equal to the "regular" condition number of $\mathbf{R}^{-1}\mathbf{A}$. From our discussion of preconditioners, we see that this is precisely equivalent to R^{-1} being a good preconditioner for A.

Summarizing, we have argued that the problem of finding a good preconditioner R^{-1} for the (continuous) problem is solved by finding a space V on which the problem $\mathbf{Au} = \mathbf{f}$ is well-posed in the sense of Theorem 1.2.1, and choosing R^{-1} to be the matrix of the Riesz map of V. Additionally, we need some effective procedure for computing $\mathbf{R}^{-1}\mathbf{x}$ for any vector $\mathbf{x}$, and we require the stability estimate to remain valid with a constant independent of h when we replace V by our discretization V_h. Neither of these two caveats are mere technicalities, but the beauty of this connection is that it relates the search for a good preconditioner for the discrete problem to the choice of "stable" spaces for the continuous problem.

Here, we use "choice" because the operators corresponding to the weak forms some PDEs may be viewed as stable isomorphisms $V \to V^*$ for several different spaces V. All such V can then be used to obtain an operator preconditioner, but evidently some may be easier to implement or compute than others. We also remark that although the focus of this thesis is preconditioners for matrices arising from PDE problems, other matrices also exist. For a survey of preconditioning techniques for a "general matrix", i.e. methods not making any assumption about the problem giving rise to the matrix as we do here, see [108].

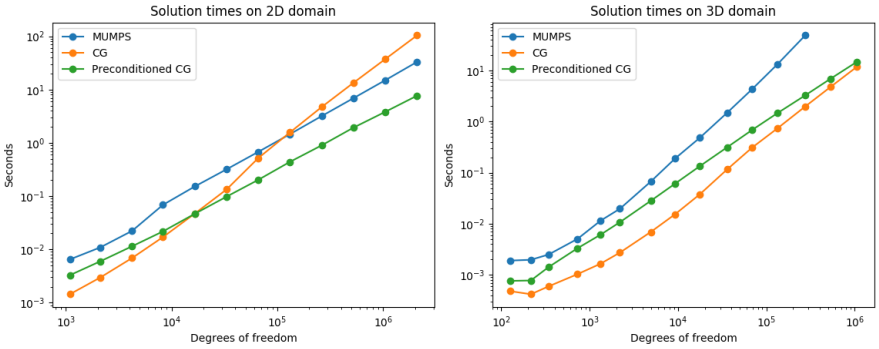

Figure 1.2: Comparison of solution times for solving a Poisson problem with N degrees of freedom for different solution methods.

Example 1.2.2 (Comparison of solution methods for the Poisson problem)**.** In this example, we perform a simple comparison of direct and iterative methods

for the solution of the Poisson problem

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \partial\Omega \end{aligned}$$

with Ω being the unit square and $f(x, y) = (x^3 + 2y^2 + 1)\sin(x + y^2)$. Using the FEniCS software suite [83], we discretize this by a $M \times M$ uniform grid and P1 elements, and compare three solvers: a direct (MUMPS) solver, a conjugate gradient (CG) method with no preconditioner, and a CG solver preconditioned by an AMG preconditioner, with the iterative solvers having a fixed convergence criterion of a relative or absolute error decrease of 10^{-10}. The CPU time required for successively larger grid sizes is shown in Figure 1.2.

From the plot, we see that the unpreconditioned iterative method seem to have complexity $\mathcal{O}(N^2)$, while the preconditioned iterative method is $\mathcal{O}(N)$. The direct method appears to have performance close to $\mathcal{O}(N)$. On the right, a similar experiment is shown for a 3D domain. Here, the iterative methods can be seen to have similar performance, but the direct method appears closer to $\mathcal{O}(N^2)$. In both cases, memory requirements for the direct method became the limiting factor prohibiting a more extensive comparison.

1.3 Stability of weak problems

Having established a rough equivalence between a good preconditioner and a choice of a norm in which the weak problem is stable, it is only natural to turn to the question of when a particular weak problem is stable. We shall also see that the tools for answering this question will let us address the issue elided earlier of when a problem which is stable on the full, "continuous" space V becomes unstable on a discrete subspace $V_h \subset V$.

We shall occasionally use the terms *stable* or *well-posed*. By a *stable* or *well-posed problem*, we mean a weak form $a(u, v) : U \times V \to \mathbb{R}$ problem such that for any $L \in V^*$, there exists a unique solution $u \in U$ so that $a(u, v) = L(v)$ for all $v \in V$, and so that for some constant C we have the estimate $\|x\|_U \leq C\|L\|_{V^*}$.

Before proceeding, we briefly comment on the importance of distinguishing between U and U^*. Per the Riesz representation theorem, the two are isomorphic (indeed, isometric) via the map $u \in U$ and $(u, -)_U \in U^*$, but care must be taken to not confuse the two. As we argued in the previous section, doing so is, implicitly, identifying the two according to the Riesz map defined by some inner product, which may not be the one we want to use. We will therefore strive to be precise and explicit about which Riesz maps we are using, at the cost of being somewhat verbose.

1.3.1 General stability results

Suppose, then, that we have some linear operator A between Hilbert spaces acting on U, and we are interested in finding an inverse. This corresponds to, for

any f in its codomain, finding an u in U for which $Au = f$. Letting V denote the *dual* of its codomain, $A : U \to V^*$ induces a bilinear form $a : U \times V \to \mathbb{R}$ defined by $a(u, v) := \langle Au, v\rangle$, and f a linear functional $L(v) : V \to \mathbb{R}$ defined by $L(v) = \langle f, v\rangle$. By the Riesz representation theorem, the problem of finding $u \in U$ so that $Au = b$ is then equivalent to the problem of finding $u \in U$ so that $a(u, v) = L(v)$ for all $v \in V$.

The following theorem is a very minor rephrasing of the seminal result of [5]:

Theorem 1.3.1. *Let U, V be two Hilbert spaces, and $a(u, v) : U \times V \to \mathbb{R}$ a continuous bilinear form, with the constant β so that $\|a(u, v)\| \leq \beta\|u\|_U\|v\|_V$. Suppose there exists a constant $\alpha > 0$ so that b satisfies the following inf-sup conditions:*

- $\inf\limits_{u\in U} \sup\limits_{v\in V} \frac{a(u,v)}{\|u\|_U\|v\|_V} \geq \alpha$
- $\inf\limits_{v\in V} \sup\limits_{u\in U} \frac{a(u,v)}{\|v\|_V\|u\|_U} \geq \alpha$

Then for any $f \in V^*$, there exists a unique $u \in U$ so that for any $v \in V$, $a(u, v) = (f, v)$ for all $v \in V$. Moreover, we have the estimate

$$\frac{1}{\beta}\|f\|_{V^*} \leq \|u\|_U \leq \frac{1}{\alpha}\|f\|_{V^*}$$

Conversely, suppose that a is a bilinear form so that for any $f \in V^*$, the weak problem $a(u, v) = (f, v)$ has a unique solution u satisfying the above estimate. Then a is continuous with boundedness constant β, and the above two inf-sup conditions are satisfied.

Remark 1.3.2. Observe that when $U = V$ and a is symmetric, the two inf-sup conditions are obviously equivalent. However, in the general case, the values of the two inf-sups above need not be equal. For a simple counterexample, let $U := V \times V$, and define $a((w_1, w_2), v)$ by $(w_1 - w_2, v)_V$. Then the second condition holds with $\alpha = 1$, while choosing $u = (w, w)$ shows that the first condition does not hold for any $\alpha > 0$. In this case, the weak problem can always be solved, with solutions $u = (R_V(f) + w, w)$ for any $w \in V$, but the solution is not unique.

However, in the case that the weak problem is uniquely solvable with the given bound (which is implied by the two conditions holding with possibly different $\alpha_1, \alpha_2 > 0$), the values of the two inf-sups can be shown to be equal. To see this, define $A : U \to V^*$ by $\langle Au, v\rangle = a(u, v)$. Then solvability implies that there exists an inverse map $A^{-1} : V^* \to U$. Hence

$$\inf_{u\in U}\sup_{v\in V} \frac{\langle Au, v\rangle}{\|u\|_U\|v\|_V} = \inf_{w\in V^*}\sup_{v\in V} \frac{\langle w, v\rangle}{\|A^{-1}w\|_U\|v\|_V} = \inf_{w\in V} \frac{\|w\|^2_{V^*}}{\|A^{-1}w\|} = \frac{1}{\|A^{-1}\|}$$

Next, consider the adjoint $(A^*)^{-1} = (A^{-1})^*$. Per the same argument as above, the value of the other inf-sup condition is $\frac{1}{\|(A^*)^{-1}\|}$, and as adjoints have the same operator norm, we are done.

Note that there is no conflict between this and the case considered above where we have solvability but not uniqueness, as in this case we do not have a two-sided inverse A^{-1}, only a right inverse E. The above argument still shows that the first inf-sup condition holds with constant at least $\frac{1}{\|E\|}$, but when we take the adjoint, E^* is a *left* inverse to A^*, not a right inverse.

The above result is a generalization of the following, where $V = U$:

Theorem 1.3.3 (Lax-Milgram)**.** *Let $a : U \times U \to \mathbb{R}$ be a bilinear form satisfying the following two conditions:*

- $a(u, v) \leq \beta\|u\|\|v\|$ *(a is bounded)*
- $a(u, u) \geq \alpha\|u\|^2$ *(a is coercive)*

Then for any $f \in U^*$, there exists a unique $u \in U$ for which $a(u, v) = (f, v)$ for any $v \in U$, and we have

$$\frac{1}{\beta}\|f\|_{U^*} \leq \|u\|_U \leq \frac{1}{\alpha}\|f\|_{U^*}$$

Note, however, that due to the stronger hypotheses assumed in Lax-Milgram, the converse is not true, with saddle point problems being a counterexample we shall return to later. In the special case where a is symmetric, there is a particularly straightforward proof of the above: by the two conditions, we can define a new inner product on V by $(u, v)_a := a(u, v)$, giving us a new Hilbert space structure V_a on V. Then the existence of a solution is simply the Riesz representation theorem. In the general case, the following very nice proof is adapted from [71]:

Proof. Let $A : U \to U^*$ be the linear operator $u \to a(u, -)$, and $R_U : U^* \to U$ be the Riesz map. The idea of the proof is to consider the mapping $T_c : u \to u - cR(Au - f)$, which is seen to be exactly the Richardson iteration preconditioned by a Riesz map considered earlier. The hypotheses then let us show that this is a contraction for some $c > 0$ and show that Banach's fixed point theorem applies to guarantee the existence of a fixed point, which will then necessarily be a solution to the problem.

We emphasize that the Riesz map is crucial to the argument, as without it, the residual and the unknown will live in different spaces, and T_c cannot be defined. Writing $w := u_1 - u_2$, we see that for $c > 0$,

$$\|T_c u_1 - T_c u_2\|_U^2 = \|w\|_U^2 + c^2\|R_U Aw\|_U^2 - 2c(w, R_U Aw)_U \leq (1 - 2c\alpha + c^2\beta)\|w\|_U^2$$

where the last inequality uses $\|R_U Aw\|_U = \|Aw\|_{U^*} \leq \beta\|w\|_U$ and

$$-2c(w, R_U Aw)_U = -2c\langle w, Aw\rangle = -2c\, a(w, w) \leq -2c\alpha\|w\|^2$$

As $\alpha, \beta > 0$, the constant is minimized by $c = \frac{\alpha}{\beta}$ for which we get $\|T_c\| \leq \left(1 - \frac{\alpha^2}{\beta^2}\right) < 1$, so T_c is a contraction. Hence by the Banach fixed point theorem and the fact that U is a Hilbert space, it has a unique fixed point, so we are done. ∎

Next, we apply the same idea to give a (to our knowledge novel) proof of Theorem 1.3.1. In this case, attempting to define the map T_c from the Richardson iteration for the problem $R_V A u = R_V f$ will not work, as $u \in U$, but $R_V(Au - f) \in V$. However, by applying the adjoint A^* and preconditioning by R_U, we get back to U again and the method works.

$$\begin{array}{ccc} U & \xrightarrow{A} & V^* \\ {\scriptstyle R_U}\uparrow & & \downarrow{\scriptstyle R_V} \\ U^* & \xleftarrow[A^*]{} & V \end{array}$$

In effect, this corresponds to (ignoring Riesz maps for legibility) solving the normal equations $A^*Au = A^*f$ instead of $Au = f$. The inf-sup conditions will be seen to ensure that these are equivalent.

Proof. Let $f \in V^*$ be arbitrary. As before, let $A : U \to V^*$ be the operator so that $a(u, v) = \langle Au, v\rangle_V$, and $A^* : V \to U^*$ its adjoint, meaning $a(u, v) = \langle u, A^*v\rangle_U$. Define $A' : V^* \to U$ by $A' = R_U A^* R_V$. As R_U, R_V are isometries, $\|A'\| = \|A^*\|$.

Finding a solution to the weak problem is then equivalent to finding an f for which $Au = f$. Observe that the boundedness condition implies $\|A\|, \|A^*\| \leq \beta$, and that the inf-sup conditions imply that $\|Au\|_{V^*} \geq \alpha\|u\|_U, \|A^*v\|_U \geq \alpha\|v\|_{V^*}$. In particular, A^* is injective, so A' is too. For $c \in \mathbb{R}$, define $T_c : U \to U$ by

$$T_c u = u - c(A'Au - A'f)$$

Defining $w := u_1 - u_2$, we see that if $c > 0$,

$$\|T_c u_1 - T_c u_2\|_U^2 = \|w\|_U^2 2c(A'Aw, w)_U + c^2\|A'Aw\|_U^2 \leq \|w\|_U^2 \left(1 - 2c\,\alpha^2 + c^2\beta^4\right)$$

Here, we used that

$$\begin{aligned}(A'Aw, w)_U &= (R_U A^* R_V Aw, Aw)_U = \langle A^* R_V Aw, w\rangle_U = \langle R_V Aw, Aw\rangle_{V^*} \\ &= (Aw, Aw)_{V^*} = \|Aw\|_{V^*}^2 \geq \alpha^2\|w\|_U^2.\end{aligned}$$

The constant is minimized by choosing $c = \frac{\alpha^2}{\beta^4}$, for which we get

$$\|T_c w\|_U^2 \leq \|w\|_U^2 \left(1 - \frac{\alpha^4}{\beta^4}\right).$$

This constant is evidently less than one, so T_c is a contraction. Per the Banach fixed point theorem, it must therefore have a unique fixed point u_0, for which $A'Au_0 = A'f$, meaning that $A'(Au_0 - f) = 0$. As A' is injective, this means that $Au_0 = f$, showing the existence of a solution. The bounds on the solution follows from the fact that $\alpha\|u\|_U \leq \|Au\|_{V^*} \leq \beta\|u\|_U$, so we are done.

The converse implication is straightforward, with the inf-sup conditions established by the argumentation just before Theorem 1.3.3, and the boundedness following from the fact that for any u, defining $f = Au$ we have that u is the unique solution of $Au = f$, per assumption satisfying $\frac{1}{\beta}\|f\|_{V^*} \leq \|u\|_U$. ■

We emphasize that in the course of the proof, both inf-sup conditions were used, with the one where the supremum is over V used to establish that T_c is a contraction, and the one where the supremum is over U is used to establish injectivity of A^*.

Finally, we present the celebrated stability result of Brezzi for (symmetric) saddle point problems, i.e. problems whose operator takes the form $\begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix}$, which can be written in the weak form $a_{sp}\left((u,p),(v,q)\right) = a(u,v)+b(v,q)+b(u,p)$ for some bilinear forms $a : U \times U \to \mathbb{R}$, $b : U \times Q \to \mathbb{R}$

Theorem 1.3.4 (Brezzi, [29])**.** *Let U, Q be Hilbert spaces, and $a(u,v), b(v,q)$ be bilinear forms, and a is symmetric. Define the bilinear form*

$$a_{sp}\left((u,p),(v,q)\right) = a(u,v) + b(v,q) + b(u,p)$$

and let $B : U \to Q^*$ be the operator so that $b(u,p) = \langle Bu, p\rangle$. Let $K = \ker B$. Suppose a, b satisfy the following:

- $a(u,v) \leq \beta_a \|u\|_U \|v\|_U$ *for all* $u, v \in U$
- $b(u,p) \leq \beta_b \|u\|_U \|p\|_Q$ *for all* $p, u \in U$
- $\inf\limits_{u\in K} \sup\limits_{v\in K} a(u,v) \geq \alpha_a$
- $\inf\limits_{p\in Q} \sup\limits_{u\in U} \frac{b(u,p)}{\|u\|_U \|p\|_Q} \geq \alpha_b$

Then there exists a constant C depending only on $\beta_a, \beta_b, \alpha_a, \alpha_b$ so that the problem of finding (u,p) so that $a_{sp}((u,p),(v,q)) = \langle (f,g),(v,q)\rangle$ is uniquely solvable for all $(v,q) \in V \times Q$, with the solution (u,p) satisfying $\|(u,p)\|_{V\times Q} \leq C\|(f,g)\|_{V^*\times Q^*}$.

Remark 1.3.5. Something very similar to a converse to the above theorem is also true: if the problem with a_{sp} is uniquely solvable with a bound $\frac{1}{\alpha}$, then both inf-sup conditions hold with $\alpha_i = \alpha$. The reason this is not quite a converse because the constant C above need not be $\frac{1}{\alpha}$. For an explicit bound on this constant, see [113].

Interestingly, although Theorem 1.3.4 can be used to characterize stable discretizations for a saddle point problems, the resulting characterization is not sufficient to ensure stability of the associated eigenvalue problem, and additional conditions may be required, as established in [20].

1.3.2 Restricting stability results to a discrete subspace

As observed in e.g. [71], Lax-Milgram "plays nice" with conforming discretizations $U_h \subset U$ in the following sense: given a bilinear form b, let $A : U \to U^*$ be the operator $u \to b(u,-)$. From $\kappa(A) = \|A\|\|A^{-1}\|$, we see that if b satisfies the hypotheses of Lax-Milgram, we have $\kappa(A) \leq \frac{\beta}{\alpha}$. Now, if our discretized operator

$b_h : U_h \times U_h \to \mathbb{R}$ is simple restriction[6] of b to U_h, both the properties in the hypothesis of Lax-Milgram remains true over U_h, meaning that $\kappa(A_h) \leq \kappa(A)$. Hence, in a certain sense, discretizing a problem to which Lax-Milgram applies will never make it harder.

However, the hypotheses of Theorem 1.3.1 are not inheritable by discrete subspaces. While the coercivity and boundedness constants of Theorem 1.3.3 can be characterized by expressions of the form $\inf_{u\in U} f(u)$ or $\sup_{u,v\in U} g(u,v)$, the inf-sup conditions in the hypotheses of Theorem 1.3.1 take the form $\inf\limits_{u\in U}\sup\limits_{v\in V} h(u,v)$. Hence (when $U_h \subset U$, $V_h \subset V$), although $\inf\limits_{U} \leq \inf\limits_{U_h}$ and $\sup\limits_{U} \geq \sup\limits_{U_h}$, there is no corresponding inequality relating $\inf\limits_{U}\sup\limits_{V}$ and $\inf\limits_{U_h}\sup\limits_{V_h}$.

This means that for a general well-posed continuous problem, we cannot say from Theorem 1.3.1 alone whether discretizing by a conforming discretization will make it easier or harder, and it may (and empirically does) depend on the form of the problem and the discretization. Because Theorem 1.3.1 is an equivalence, this shows that we cannot find "one discretization to rule them all", but are inevitably forced to consider the problem when discretizing to ensure we do not ruin well-posedness properties.

Even in the case where we have chosen a good discretization so that the discrete problem is actually well-posed, however, it is not obvious that its solution must be close to the continuous solution. Fortunately, up to a constant relating to the well-posedness of the problem, the continuous and the discrete solutions are as close as they could be. This is guaranteed by Céa's lemma when the weak form is coercive, but we give here a slightly more general result from [113].

Theorem 1.3.6. *Suppose $U_h \subset U, V_h \subset V$ are discrete spaces so that the weak problem given by $a : U \times V \to \mathbb{R}$ is well-posed both in the discrete and the continuous case, in the sense that there exists a constant α_i so that for any $f_i \in V_i^*$, there exists a unique $u_i \in U_i$ so that $Au_i = f_i$, with $\|u_i\|_{U_i} \leq \frac{1}{\alpha_i}\|f_i\|_{U_i^*}$ both for $V_i = U, U_h$.*

Then, for any f, if we let $u \in U$ be the solution of the continuous problem and $u_h \in U_h$ be the solution of the discrete problem (where f is simply restricted to V_h), we have

$$\|u - u_h\|_U \leq \frac{\|a\|}{\alpha_h} \inf_{v_h \in V_h} \|u - v_h\|_U$$

To prove this, we require the fact that if P is linear, idempotent ($P^2 = P$), with P neither being zero nor the identity (i.e., when P is a nontrivial projection), we have $\|I - P\| = \|P\|$. A discussion of this identity with several proofs is found in [102].

Proof (From [113]). Let $P : U \to U_h$ be the map $u \to u_h$, i.e. sending any u to the u_h such that $a(u_h, v_h) = a(u, v_h)$ for all $v_h \in V_h$. Evidently, it is linear and

[6] Observe that this is *not* the case if, for example, a mesh-dependent stabilization term is added to the problem, as is the case for SUPG methods for the convection-diffusion problem.

idempotent.[7] First, observe that $u - u_h = (I-P)u = (I-P)(u-v_h)$ for any v_h. Hence, $\|u-u_h\|_U = \|(I-P)(u-v_h)\|_U \leq \|I-P\|_U \|u-v_h\|_U = \|P\|\|u-v_h\|_U$. We therefore only need to prove $\|P\| \leq \frac{\|a\|}{\alpha_h}$, which follows from

$$\|Pu\|_U = \|u_h\|_U \leq \frac{1}{\alpha_h}\|f\|_{U_h^*} \leq \frac{1}{\alpha_h}\|f\|_{U^*} \leq \frac{\|a\|}{\alpha_h}\|u\|_U$$

■

1.3.3 Parameter-independent stability and weighted spaces

Observe that if α, β are the constants of Theorem 1.3.1, we have $\kappa(A) = \|A\|\|A^{-1}\| \leq \frac{\beta}{\alpha}$. Recalling the discussion of Section 1.2, we see that we may efficiently invert an operator satisfying the hypotheses of Theorem 1.3.1 using the Riesz map. The number of iterations required is roughly speaking linear in the condition number, meaning that we also require that β is not too big and α is not too small.

Suppose we are solving a problem which involves some material parameter μ, which may appear in the problem operator A. The constants α, β above can then turn out to depend on μ, meaning the condition number does too. It can, then, happen that the condition number becomes large for some values of μ, meaning that our iterative method for solving it will be inefficient for those values of μ. Although the problem is stable in the sense that the hypotheses of Section 1.2 are satisfied, this situation is clearly unsatisfactory. If we avoid it, meaning that we can choose our space so that the constants α, β are independent of μ, we shall say that the resulting preconditioner is *parameter robust* to variations in μ.

A natural way to obtain a parameter robust preconditioner is simply to incorporate the parameter into the definition of the problem space. In the framework of operator preconditioning, the preconditioner will depend on the problem space, meaning that it can incorporate the dependence on the problem parameters.

For a Hilbert space V, and a positive number c, we define the *weighted space* cV to be the space consisting of the elements of V, with norm $\|u\|_{cV} = c\|u\|_V$. Evidently, the norm of cV is equivalent to the norm of V, but the equivalence constants will involve c, making it a natural way to "factor out" dependencies on problem parameters.

Additionally, for Hilbert spaces U, V contained in some common ambient space, we also define the space $U \cap V$ and $U + V$ with norms

$$\|u\|^2_{U\cap V} := \|u\|^2_U + \|u\|^2_V, \qquad \|u\|^2_{U+V} := \inf_{\substack{u_1\in U, u_2 \in V \\ u = u_1+u_2}} \|u_1\|^2_{U_1} + \|u_2\|^2_{U_2}.$$

Then $(cV)^* = \frac{1}{c}V^*$, and $(U\cap V)^* = U^* + V^*$. With these definitions, it is often possible to define spaces in which the material parameters do not appear in the stability constants, as the example in the next section will show.

[7] If $U = V$ and $a(\cdot,\cdot)$ can be thought of as an inner product, P is just projection from $U \to U_h$ in the a-inner product, but the result also holds in the general case.

We remark that the consideration of very large or very small parameters is not purely of theoretical interest. In applications such as biomechanics, problems coupling behavior at differing length scales can involve parameters spanning many orders of magnitude. As shown in Paper III, coupled problems involving the dm-scale macro-circulation to the μm-scale microcirculation can involve parameter variation of order 10^{10} arising from the large difference in length scales of the coupled problems. In order to solve such problems, preconditioners which are robust across a large span of the parameter space are therefore of paramount practical importance. Parameter dependence can also become a practical necessity for problem formulations involving e.g. discretization or stabilization parameters. See [19] for some examples of such cases arising from virtual element method formulations.

1.3.4 Example: Stability of the primal and mixed Poisson problem

To make this concrete, we will apply our stability results to the primal and the mixed form of the Poisson problem $-\mu\Delta p = f$ on a compact, smooth domain Ω. For simplicity, we suppose we have the homogeneous Dirichlet condition $p = 0$ on all of the boundary, which corresponds to considering the space $H_0^1(\Omega)$ instead of $H^1(\Omega)$. For readability, we omit Ω from our notation where appropriate.

Write $a_p(p,q) = \int_\Omega \mu\nabla p \cdot \nabla q,\ L(q) = \int_\Omega fq$. Then the weak form of our problem is to find $p \in H_0^1(\Omega)$ so that $a_p(p,q) = L(q)$ for all $q \in H_0^1(\Omega)$. From the similarity of the weak form to the H^1 inner product, it seems plausible that the problem will be stable in the H^1-norm. We will show that it is stable on the space $\sqrt{\mu}H^1$.

Indeed, recalling the Poincarè inequality $\|u\|_{L^2}^2 \leq C\|\nabla u\|_{L^2}^2$, where C is a constant depending only on the domain, we see that

$$a_p(p,p) = \mu\|\nabla p\|_{L^2}^2 = \|\nabla p\|_{\sqrt{\mu}L^2}^2 \geq D\|p\|_{\sqrt{\mu}H^1}^2$$

where $D = \frac{1}{2}\min\{1, \frac{1}{C}\}$ only depends on Ω. This establishes coercivity with constant $\alpha = D$. From the Cauchy-Schwarz-inequality, we have

$$a_p(p,q) \leq \mu\|p\|_{L^2}\|q\|_{L^2} \leq \mu\|p\|_{H^1}\|q\|_{H^1} = \|p\|_{\sqrt{\mu}H^1}\|q\|_{\sqrt{\mu}H^1},$$

establishing boundedness with constant $\beta = 1$. Hence the Lax-Milgram theorem shows that the problem is stable on $\sqrt{\mu}H^1(\Omega)$, and letting A be the operator associated to a_p, we have $\kappa(A) \leq \frac{\beta}{\alpha} = \frac{1}{D}$, which depends only on the domain. Thus, using the Riesz map of $\sqrt{\mu}H^1$ as a preconditioner, we get an efficient iterative method.

We remark that if we had instead used the space H^1, ignoring μ, α and β would both include a factor μ, meaning that our bound on $\kappa(A)$ remains the same even with the unweighted space. To see that this is not always so, consider instead the mixed version of the problem, obtained by writing $\mathbf{u} = -\nabla p$. This takes the form $\mu\nabla\cdot\mathbf{u} = f$ with the additional constraint that $\mathbf{u} = \nabla p$, with the

weak formulation given by

$$a_m((\mathbf{u},p),(\mathbf{v},q)) = \int_\Omega \mu\mathbf{u}\cdot\mathbf{v} + p\nabla\cdot\mathbf{v} + q\nabla\cdot\mathbf{u},\ L((\mathbf{v},q)) = \int_\Omega fq.$$

This is a symmetric saddle-point problem to which Theorem 1.3.4 can be applied[8]. We will show that a_m satisfies the hypotheses of the criterion on the space $\sqrt{\mu}H(\mathrm{div}) \times \frac{1}{\sqrt{\mu}}L^2$, showing that the problems are well-posed in the corresponding norms. As will be clear from the proof, this choice of spaces and weighting can be motivated by looking at what operators should be bounded, and choosing the spaces of their arguments so that they are.

Write $a(\mathbf{u},\mathbf{v}) = \int_\Omega \mu\mathbf{u}\cdot\mathbf{v}$, $b(\mathbf{u},\mathbf{p}) = \int_\Omega p\nabla\cdot\mathbf{u}$. First, we see that by the Cauchy-Schwarz inequality,

$$a(\mathbf{u},\mathbf{v}) = (\mathbf{u},\mathbf{v})_{\sqrt{\mu}L^2} \leq \|\mathbf{u}\|_{\sqrt{\mu}L^2}\|\mathbf{v}\|_{\sqrt{\mu}L^2} \leq \|\mathbf{u}\|_{\sqrt{\mu}H(\mathrm{div})}\|\mathbf{v}\|_{\sqrt{\mu}H(\mathrm{div})}$$

Hence a is bounded with constant $\beta_a = 1$. Similarly,

$$\begin{aligned} b(\mathbf{u},p) &= (\nabla\cdot\mathbf{u},p)_{L^2} \leq \|\nabla\cdot\mathbf{u}\|_{L^2}\|p\|_{L^2} \\ &= \|\nabla\cdot\mathbf{u}\|_{\sqrt{\mu}L^2}\|p\|_{\frac{1}{\sqrt{\mu}}L^2} \leq \|\mathbf{u}\|_{\sqrt{\mu}H(\mathrm{div})}\|p\|_{\frac{1}{\sqrt{\mu}}L^2} \end{aligned}$$

showing that b is bounded with constant $\beta_b = 1$. Next, we show that a is coercive on the kernel of b. If $\mathbf{u}$ is in the kernel of b, then $\nabla\cdot\mathbf{u} = 0$, hence $\|u\|_{H(\mathrm{div})} = \|u\|_{L^2}$. Therefore

$$a(\mathbf{u},\mathbf{u}) = \|u\|^2_{\sqrt{\mu}L^2} = \|u\|^2_{\sqrt{\mu}H(\mathrm{div})},$$

so a is coercive with $\alpha_a = 1$.

Finally, we need to establish the inf-sup condition, which in some sense is the hard part - while the other conditions followed naturally from our choice of space, the inf-sup condition does not. To do this, we need a right inverse to the divergence operator.

Lemma 1.3.7. *Suppose Ω is a smooth, 2D compact domain. For any $p \in L^2(\Omega)$, there exists a $\mathbf{u} \in H(div,\Omega)$ so that $\nabla\cdot\mathbf{u} = p$, and $\|\mathbf{u}\|_{H(div,\Omega)} \leq C\|p\|_{L^2(\Omega)}$ for some constant C depending only on Ω.*

Such a $\mathbf{u}$ can e.g. be obtained by solving the Poisson problem $-\Delta\phi = p$, and setting $\mathbf{u} = -\nabla\phi$. Provided we have a bound $\|\phi\|_{H^1(\Omega)} \leq C\|p\|_{L^2(\Omega)}$, this then yields that $\|u\|_{H(\mathrm{div},\Omega)} \leq (1+C)\|p\|_{L^2(\Omega)}$. We remark that the existence of a right inverse is complicated slightly in the case where we have Neumann conditions on p instead of Dirichlet conditions, as they then translate to Dirichlet conditions on u. In this case, we must modify the solution u defined above to ensure it is in $H^1_0(\Omega)$. Another source of complication is when the boundary is not smooth. For a thorough treatment of these difficulties and a proof, see e.g. [48].

[8] Observe that irrespective of our choice of space, Theorem 1.3.3 does not apply to a_m - because $a_m((\mathbf{0},p),(\mathbf{0},p)) = 0$, for all p, a_m will never be coercive.

The above lemma lets us define a linear map $E : L^2(\Omega) \to H(\mathrm{div}, \Omega)$ sending any $q \in L^2(\Omega)$ to the function $\mathbf{u}$ defined by the above lemma. This is, essentially, all we need to prove the inf-sup condition. To properly handle the parameter weighting and to demonstrate a general approach we shall use later, we complete the proof in a slightly cumbersome way. Given any $q \in \frac{1}{\sqrt{\mu}} L^2(\Omega)$, let $\mathbf{u}_q := E R^{-1}_{\frac{1}{\sqrt{\mu}} L^2} q$, where $R^{-1}_{\frac{1}{\sqrt{\mu}} L^2} : \frac{1}{\sqrt{\mu}} L^2(\Omega) \to \sqrt{\mu} L^2(\Omega)$ is the inverse Riesz map. As the Riesz map is an isometry, we have that $\|R^{-1}_{\frac{1}{\mu} L^2} q\|_{\sqrt{\mu} L^2} = \|q\|_{\frac{1}{\sqrt{\mu}} L^2}$. Thus $\|E R^{-1} q\|_{\sqrt{\mu} H(\mathrm{div})} \leq C \|R^{-1}_{\frac{1}{\sqrt{\mu}} L^2} q\|_{\sqrt{\mu} L^2} = C \|q\|_{\frac{1}{\sqrt{\mu}} L^2}$. Further, note from the definition of the Riesz map that $(\nabla \cdot \mathbf{u}_q, q)_{L^2} = \left(R^{-1}_{\frac{1}{\mu} L^2} q, q \right)_{L^2} = \|q\|^2_{\frac{1}{\mu} L^2}$. Hence, we establish the inf-sup condition with constant $\alpha_b = \frac{1}{C}$

$$\sup_{\mathbf{u}} \frac{(\nabla \cdot \mathbf{u}, q)_{L^2}}{\|\mathbf{u}\|_{H(\mathrm{div})}} \geq \frac{(\nabla \cdot \mathbf{u_q}, q)_{L^2}}{\|\mathbf{u}_q\|_{H(\mathrm{div})}} = \frac{\|q\|^2_{L^2}}{\|\mathbf{u}_q\|_{H(\mathrm{div})}} \geq \frac{1}{C} \|q\|_{L^2}$$

From the lemma, C is here some constant depending only on Ω. As $\beta_a, \beta_b, \alpha_a$ were all 1, we see that all our constants are independent of μ. Hence the bound on the condition number of the problem operator, being expressible in terms of the constants, must also be independent of μ, so we are done. As we have shown that the problem is well-posed on the space $V \times Q := \sqrt{\mu} H\,(\mathrm{div}) \times \frac{1}{\sqrt{\mu}} L^2$, the appropriate operator preconditioner is

$$\begin{pmatrix} \mu(\nabla \cdot \mathbf{u}, \nabla \cdot \mathbf{v}) & \\ & \frac{1}{\mu}(p, q) \end{pmatrix}^{-1}$$

To emphasize the role of the norms in our argument, let $A : V \to V^*$ where $V = \sqrt{\mu} H^1_0$ be the problem operator for the primal Poisson problem mapping $p \to a_m(p, -)$. Then $Ap = -\nabla p$. In the above, we proved that A has a continuous inverse as a map $V \to V^*$, by proving that it is bounded and coercive. The choice of space V is core to this property and to our argument. To see this, note that the operator $-\nabla$, understood as a (partial) map $L^2 \to L^2$, will in general be unbounded, as its spectrum is unbounded. This does not contradict what we have showed, as H^1 is a stronger norm on the domain of A than L^2, and $(H^1)^*$ is a weaker norm on the codomain than L^2, meaning that A can be unbounded as an operator $A : L^2 \to L^2$ despite being bounded as an operator $H^1 \to (H^1)^*$. This distinction highlights the importance of being clear about the domain and codomain of the operators we are working with.

1.4 Enforcing boundary conditions with trace operators

Suppose we want to enforce the boundary condition $u = f$ on (some subset of) $\partial\Omega$ for a general PDE written in the form $Au = f$. A common way to do this is to choose an arbitrary function F on Ω which equals f on $\partial\Omega$, introduce the new unknown $u_0 := u - F$, and instead solve the problem $Au_0 + AF$ with the condition $u_0 = 0$ on Ω. This homogeneous Dirichlet condition can then be

enforced by encoding it into the function space used, as the set of functions satisfying it form a vector space, unlike in the inhomogeneous space. Accordingly, we can define, say, the space $H_0^1(\Omega)$ of the functions in $H^1(\Omega)$ which are 0 on the boundary. In the discrete case, this is often easy, as restriction to this space simply corresponds to dropping the degrees of freedom on the boundary.

Although the above is very practical from an implementational point of view, we shall later be interested in a way of imposing boundary conditions which is easier to couple with other problems. This becomes important when, for example, we want to use knowledge that a problem is well-posed on its own to prove well-posedness of the coupled problem, where the normal boundary conditions are replaced by coupling conditions to another problem. Accordingly, we will in this section develop tools for enforcing the constraint $u = f$ on $\partial\Omega$ directly in the formulation of the weak problem, meaning they do not need to be encoded in the function space of u. Another method which also avoids needing to modify the function space of u is the widely-used Nitsche's method, which accomplishes this through the introduction of a mesh-dependent term in the variational form. As we will not use Nitsche's method in this thesis, we refer the reader to e.g. [62] for an introduction of the method and an overview of some of its many applications.

A proper description of the trace operators we will need requires defining the fractional Sobolev space $H^{0.5}(\partial\Omega)$ of one half times differentiable(!) functions. Because this is somewhat technical, we will first motivate trace operators by explaining how we plan to use them. Let us suppose we can define a bounded, surjective operator $T : H_0^1(\Omega) \to X$ which restricts u to its boundary value $u|_{\partial\Omega}$. (We will later explain why this is not obvious.)

Suppose also that some weak problem $a(u, v) = \langle f, v\rangle$ is well-posed for $u, v \in H_0^1(\Omega)$. Consider now the saddle point problem $a_T\left((u, \lambda), (v, \mu)\right) = \langle f, v\rangle + \langle g, \mu\rangle$, where

$$a_T\left((u, \lambda), (v, \mu)\right) = a(u, v) + \langle Tu, \mu\rangle + \langle \lambda, Tv\rangle.$$

Here the function space has been enlarged from $H_0^1(\Omega)$ to $H^1(\Omega)$, and an additional unknown λ has been added. In the case where a is symmetric, the problem of finding u so that $a(u, v) = \langle f, v\rangle$ for all $v \in H_0^1(\Omega)$ is equivalent to minimizing $J(u) := \langle Au, u\rangle - \langle f, u\rangle$ over $u \in H_0^1(\Omega)$, and the problem defined by a_T can be seen as minimizing the same functional over all of $H^1(\Omega)$, with the constraint $Tu = 0$ enforced by the Lagrange multiplier λ.

In this respect, we may think of the problems as being the same, differing mainly in whether we we want to enforce the constraint Tu strongly or weakly. The following lemma justifies this, and is a straightforward application of Brezzi theory. For ease of notation we do it in full generality.

Lemma 1.4.1. *Suppose $T : U \to X^*$ is a surjective, bounded operator with kernel K, and that $A : U \to U^*$ is a bounded linear operator which restricts to $A_0 : K \to K^*$. Define the block operator $A_T := \begin{pmatrix} A & T^* \\ T & \end{pmatrix} : U \times X \to U^* \times X^*$. Then A_T has a bounded inverse if and only if A_0 does.*

Proof. Define $a(u,v) := \langle Au, v\rangle_U$ and $a_T\left((u,\lambda),(v,\mu)\right) := a(u,v) + \langle Tu,\mu\rangle_X + \langle Tv,\lambda\rangle_X$. By Theorem 1.3.4, A admits a bounded inverse if and only if a satisfies two inf-sup conditions on $K \times K$ and T satisfies the inf-sup condition $\inf_{\mu\in X}\sup_{u\in V}\frac{\langle Tu,\mu\rangle}{\|u\|_U\|\mu\|_X} \geq \alpha$. Because T is surjective and bounded, it admits a bounded right inverse E, meaning that setting $u = ER_X^{-1}\mu$ we get

$$\inf_{\mu\in X}\sup_{u\in V}\frac{\langle Tu,\mu\rangle}{\|u\|_U\|\mu\|_X} \geq \inf_{\mu\in X}\frac{\langle TER_X^{-1}\mu,\mu\rangle}{\|ER_X^{-1}\mu\|_U\|\mu\|_X} \geq \inf_{\mu\in X}\frac{\|\mu\|_X^2}{\|E\|\|\mu\|_X^2} = \frac{1}{\|E\|}$$

Hence T satisfies the requisite inf-sup condition with constant $\frac{1}{\|E\|}$. This means that A_T admits a right inverse if and only if a satisfies the two inf-sup conditions on K, which by Theorem 1.3.1 happens if and only if A_0 admits a bounded inverse, so we are done. ■

This lemma acts as a guarantee that we are free to enforce a boundary condition of $u = 0$ on the boundary as we like - if the problem is well-posed when we enforce it strongly by only considering test functions $u \in H_0^1(\Omega) = \ker T$, it will also be well-posed if we enforce them with the Lagrange multiplier λ. This is not much of a surprise, but means that when we couple multiphysics problems together by use of problems where the boundary condition is enforced with a multiplier, we cannot end up losing well-posedness.

1.4.1 Trace operators and fractional Sobolev spaces

It is easy to miss that the condition $u = f$ on $\partial\Omega$ is not necessarily well-defined in the context of weak problems whose solutions lie in Sobolev spaces. Indeed, the element of such spaces are not "normal" functions $f : \Omega \to \mathbb{R}$, but "generalized" functions, or rather equivalence classes of normal functions defined by an appropriate norm. For example, $L^2(\Omega)$ consists of all functions on Ω whose L^2-norm is finite, where functions with zero L^2 norm are identified. (Otherwise, the resulting norm would only be a seminorm.). As $\|f\|_{L^2}^2 = \int_\Omega f^2$, we see that modifying f at a single point (or, indeed, at a set of measure zero) will not change the norm, and accordingly, evaluation at $\partial\Omega$ cannot be well-defined for a general $u \in L^2(\Omega)$.

Because the H^1 norm would be changed by modification at a set of measure zero, it seems plausible that requiring more regularity will let us restrict our functions to a set of measure zero, and indeed we shall later see that H^1 functions are regular enough that a trace operator can be defined, yielding the following theorem:

Theorem 1.4.2 (Trace theorem [59, p. 1.5.1])**.** *For a domain Ω with Lipschitz continuous boundary, there exists a surjective, bounded operator $T : H^1(\Omega) \to H^{0.5}(\partial\Omega)$ which is simple restriction on the subspace of smooth functions. Accordingly, there exists a bounded right inverse E.*

Evidently, a proper formulation of the above requires a precise definition of the fractional Sobolev space $H^{0.5}(\partial\Omega)$. Before doing so in the next section, however, we remark that a common "introductory" version of the trace theorem given in e.g. [50] has $L^2(\partial\Omega)$ as the target space, sidestepping the need for introducing fractional Sobolev spaces. While conceptually simpler, this will not be suitable for our purposes. The reason for this is that although T remains continuous (as we are weakening the norm), it is no longer surjective, as we are enlarging the space. Accordingly, T will no longer have a right inverse, and the proof of Lemma 1.4.1 will not go through. In order to use our desired approach to model multiphysics problems, we require a space which accurately characterizes the image of H^1 under the trace operator, making fractional Sobolev spaces a necessity if we want our multiphysics problems to be well-posed.

We remark that although the condition that Ω has a smooth boundary is necessary, similar results hold in cases where this condition is relaxed. See, for example, [59, Section 1.5.2] for a similar result in the case of a plane polygon, which is not smooth at the corners.

1.4.1.1 H^s via Fourier series

The space H^s for $0 < s < 1$ can be defined in multiple ways, and the interested reader is referred to [45] for a general overview, and to [40] for an in-depth treatment of the two definitions we give in this section. In this section, we shall motivate a definition in terms of the Fourier transform, sacrificing some generality in the interest of intuition.

To begin with, recall the definition of the Fourier transform of an absolutely square-integrable function $f(x) : \mathbb{R} \to \mathbb{C}$ as $\hat{f}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ix\xi}\,\mathrm{dx}$ By integration by parts and integrability, it then follows that $\hat{f}'(\xi) = i\xi\hat{f}(\xi)$, or more generally that $\widehat{f^{(n)}}(\xi) = (i\xi)^n\hat{f}(\xi)$. This means that up to a constant, differentiation in the real domain becomes multiplication by the Fourier variable in the Fourier domain. This suggests that we may perform "half a differentiation" by multiplying by $(i\xi)^{\frac{1}{2}}$ in the Fourier domain.

If we intuitively want a H^s norm to be of the form $\|f\|_{H^s}^2 = \|f\|_{L^2}^2 + \|\frac{d^s f}{dx^s}\|_{L^2}^2$, similarly to the H^1 norm, this suggests that we define a norm H^s by

$$\|f\|_{H^s}^2 = \int_{-\infty}^{\infty}(1 + |\xi|^{2s})|\hat{f}(\xi)|^2\,\mathrm{d}\xi = \|\hat{f}\|_{L^2}^2 + \|(i\xi)^s\hat{f}\|_{L^2}^2$$

Although we here assumed that f was a function of one variable, a similar definition may be carried through in the Sobolev space setting when f is a distribution on $\mathbb{R}^n$. See [45, Section 3.1, Proposition 3.3] or [40, Section 4.1] for a rigorous treatment.

1.4.1.2 H^s via eigenvalues of the Laplacian

In this section, we shall motivate a second definition for fractional Sobolev spaces made by considering eigenvalues of the Laplacian. We refer the reader to

[40, Lemma 4.11] for a result establishing equivalence[9] between this definition and the one given in the previous section. First, observe that for any $u \in H^1(\Omega)$, $(u, -)_{L^2}$ is a linear functional on $H^1_0(\Omega)$. Accordingly, we can define a linear operator[10] $S : H^1_0(\Omega) \to H^1_0(\Omega)$ so that $(Su, v)_{H^1} = (u, v)_{L^2}$, where now $(a, b)_{H^1} = (\nabla a, \nabla b)_{L^2}$. Choose[11] an orthonormal basis of H^1 of eigenvectors $e_1, \ldots$ of S with corresponding positive eigenvalues $\lambda_1, \ldots$. Now, for an $s \in [-1, 1]$ and any $u \in H^1$, write $u = \sum_i c_i e_i$, and define

$$\|u\|_s^2 := \sum_i \lambda_i^{-s} c_i^2$$

Evidently, for $s = 0$ we get the norm on L^2, and for $s = 1$ we get the norm on H^1_0. Accordingly, we may think of the space defined by the norm resulting from numbers $0 < s < 1$ as being more restrictive than L^2, but not quite as restrictive as H^1_0. We remark that one advantage of this definition is that it is reasonably straightforward to implement, as it is defined in terms of eigenvalues of the operators we are interested in for operator preconditioning.

1.5 Multiphysics problems

In this section, we introduce the idea of a multiphysics problem, which is a problem involving multiple kinds of physics. We give the coupled Darcy-Stokes problem as example. Having done so, we move on to describe a general approach to a formulation of the coupled multiphysics problem, and give results proving its stability conditional on stability of the subproblems. We conclude with a review of other possible solution approaches to multiphysics problems.

1.5.1 The Darcy and Stokes problems

For simplicity, we consider an incompressible fluid which is therefore of constant density. Suppose we want to model fluid transport in soil or some other porous solid. This can be done by the *Darcy equations*, which assert that the amount of fluid moving through a patch of soil is proportional to the pressure difference across it. Writing $\mathbf{u}$ for the (vectorial) fluid velocity at a point and p for the pressure, they read as follows:

$$\frac{\mu}{\kappa}\mathbf{u} - \nabla p = 0 \tag{1.5}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{1.6}$$

[9] We emphasize that this should be understood as equivalence of norms, not equality of norms, which we in general do not have. See [40, Example 4.15] for an example demonstrating this.

[10] Heuristically, from the identity $(-\Delta u, v)_{L^2} = (\nabla u, \nabla v)_{L^2} + \text{boundary term}$, we see that S can be thought of as being similar to $(-\Delta)^{-1}$. Note, however, that the identity is only valid for sufficiently smooth functions.

[11] Such a basis exists because S is self-adjoint and compact, and hence diagonalizable.

Here, the first equation enforces conservation of momentum and decribes the relationship between pressure differentials and the resulting water movement, while the second equation enforces conservation of mass. The material parameters involved are the *viscosity* μ of the fluid, which describes how thick it is, and the *permeability* κ of the soil, which can be thought of as measuring how much of an obstacle it presents to the fluid. Plastic is very impermeable, which is why we can make buckets out of it, while paper is more permeable. We see that the thinner the fluid and the more permeable the solid, the easier it is for fluid to move through the soil.

Consider now how the fluid will move if we remove it from the porous medium and place it in a bucket. In the bucket, there is no soil the fluid has to move through, and we might imagine that this would be described by letting $\kappa \to \infty$. However, Darcy's law is no longer valid in this regime, as it would make no sense for the fluid to reach an infinite velocity if lightly tapped. Under this regime, the celebrated *Navier-Stokes equations* govern the motion of the fluid:

$$\rho \frac{\partial \mathbf{u}}{\partial t} = -\rho(\mathbf{u} \cdot \nabla)\mathbf{u} - \nabla p + \mu \Delta \mathbf{u} \tag{1.7}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{1.8}$$

Again, the first equation enforces conservation of momentum, and the second equation enforces conservation of mass. Provided the fluid velocity is not too great,[12] we may disregard the inertial term and be left with the Stokes equations, which are linear:

$$-\mu \Delta \mathbf{u} - \nabla p = 0 \tag{1.9}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{1.10}$$

A way to think about the differing regimes is that in both regimes, the fluid resists being moved. In the the porous flow regime governed by the Darcy equation, the primary component of this resistance is the fact that in order to move, the fluid needs to go through the medium, which is fixed, so when the fluid is pushed by a force, the medium pushes back. Under the free flow regime, there is no medium, and the viscosity of the fluid itself becomes more significant.

We remark that both problems are incomplete as they stand, in the sense that they do not determine a unique solution and hence are not well posed. To obtain uniqueness, we must supplement the equations by appropriate boundary conditions.

We also remark that the Darcy problem is essentially just a vectorial version of the mixed Poisson problem discussed in Section 1.3.4. Accordingly, the appropriate function space for the Darcy problem in the sense of operator preconditioners is $(u, p) \in \sqrt{\frac{\mu}{\kappa}} H(\operatorname{div}, \Omega) \times \sqrt{\frac{\kappa}{\mu}} L^2(\Omega)$, meaning the operator

[12] More precisely, that the Reynolds number is not too high.

preconditioner becomes

$$\begin{pmatrix} -\frac{\kappa}{\mu}(I - \nabla\nabla\cdot)^{-1} & \\ & \frac{\mu}{\kappa}I^{-1} \end{pmatrix}.$$

If one tries to apply Theorem 1.3.4 in the same manner as we did in Section 1.3.4, it will be seen that the appropriate function space for the Stokes problem is $(u, p) \in \sqrt{\mu}H^1(\Omega) \times \frac{1}{\sqrt{\mu}}L^2(\Omega)$. Hence, the appropriate operator preconditioner is

$$\begin{pmatrix} -\frac{1}{\mu}(\Delta)^{-1} & \\ & \mu I^{-1} \end{pmatrix}$$

1.5.2 Brinkman problem

The Brinkman problem is, effectively, an interpolation problem between the Stokes and Darcy problems.

$$-\nu\Delta\mathbf{u} + \frac{\mu}{K}\mathbf{u} - \nabla p = 0 \tag{1.11}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{1.12}$$

Here, ν is the effective viscosity, μ is the dynamic viscosity, and K is the permeability. Writing this in block form, we see that it looks like

$$\begin{pmatrix} aI - b\Delta & -\nabla \\ \nabla\cdot & \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ p \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ g \end{pmatrix}$$

where $a = \nu, b = \frac{\mu}{K}$ can, perhaps, be thought of as weighting the Stokes and Darcy operators - for $b = 0$ we get a Stokes problem, and for $a = 0$ we get a Darcy problem. The Brinkman problem models flow which is neither entirely free nor entirely porous. It can also be used as an alternative multiphysics model to the coupled Darcy-Stokes problem we shall define in the next section by having these material parameters vary in the domain. For example, we can let $a = 1, b = 0$ in the Stokes region, and $a = 0, b = 1$ in the Darcy region, or to avoid a discontinuity have a, b vary smoothly, although the latter comes at the cost of somewhat obscuring exactly what kind of physics we are prescribing near the interface. However, finding a suitable finite element which is robust in all parameters becomes harder, as we effectively require it to work both for the Darcy and the Stokes subproblems. Examples of such elements are given in [69, 85].

1.5.3 The coupled Darcy-Stokes problem

Although the problems described in the previous section model porous and free flow well individually, examples abound of physical systems in which the water transport is neither entirely porous nor entirely free. For example, in a porous rock with open cracks, the movement of water in the cracks will be characterized

by the Stokes problem, and in the rock, it will be characterized by the Darcy problem. Hence, in order to model the cracked rock, we need to solve the coupled problem consisting of both the Darcy and the Stokes problem alongside coupling conditions describing the interaction between the two.

Denoting the domain modeled by the Darcy problem by Ω_p, the domain modeled by the Stokes domain by Ω_f and their common interface by Γ, the coupled problem reads as follows:

$$\frac{\mu}{K}\mathbf{u}_p - \nabla p_p = 0 \qquad \text{in } \Omega_p, \tag{1.13a}$$

$$\nabla \cdot \mathbf{u}_p = 0 \qquad \text{in } \Omega_p, \tag{1.13b}$$

$$-\mu\Delta\mathbf{u}_f - \nabla p_f = 0 \qquad \text{in } \Omega_f, \tag{1.13c}$$

$$\nabla \cdot \mathbf{u}_f = 0 \qquad \text{in } \Omega_f, \tag{1.13d}$$

$$\mathbf{u}_p \cdot \mathbf{n}_\Gamma - \mathbf{u}_f \cdot \mathbf{n}_\Gamma = 0 \qquad \text{on } \Gamma, \tag{1.13e}$$

$$\mu\frac{\partial \mathbf{u}_f}{\partial \mathbf{n}_\Gamma} \cdot \mathbf{n}_\Gamma + p_f = p_p \qquad \text{on } \Gamma, \tag{1.13f}$$

$$-\mu\frac{\partial \mathbf{u}_f}{\partial \mathbf{n}_\Gamma} \cdot \boldsymbol{\tau} - D\mathbf{u}_f \cdot \boldsymbol{\tau} = 0 \qquad \text{on } \Gamma. \tag{1.13g}$$

Observe here that Equations (1.13a) to (1.13d) are just the Darcy and Stokes problems on their respective domains. Equations (1.13e) to (1.13g) are the coupling conditions, with Equation (1.13e) enforcing conservation of mass and Equation (1.13f) enforcing balance of stress. As a third interface condition is required to have a well-posed problem, we also include Equation (1.13g) the Beaver-Joseph-Saffman condition. It was originally based on experimental data, though has since been theoretically justified as a transition condition between porous and free flow regimes. See [90] for a theoretical justification and a review of its history.

We also remark that, again, the resulting system must be completed by boundary conditions on the rest of the boundary. However, observe that on the interface, we do not enforce a boundary condition, reflecting the fact that we do not prescribe, but want to model, what happens there. A priori, it is not obvious that resulting system is well-posed, and we shall treat the question of whether it is in Section 1.5.4.

1.5.4 Coupling multiphysics problems with Lagrange multipliers

In this section, we develop an abstract formulation of the coupled multiphysics problem and give a stability result for it in Theorem 1.5.3. The reader is encouraged to keep in mind the Stokes and Darcy problems coupled via the normal trace as a motivating example. Suppose we have two individual problems ($i = 1, 2$) described by the operators $A_i : U_i \to U_i^*$ acting on function spaces U_i, each problem modeling a kind of physical behavior, and that we want to pose the coupled problem jointly modeling the two kinds of physical behavior. Suppose also that we have operators $T_i : U_i \to X_i^*$ so that the coupling conditions of

interest between the two problems take the form $T_1 u_i + T_2 u_2 = 0$, read as an equality in $(X_1 \cap X_2)^*$. Recalling that we may enforce the condition $T_i u_i = g_i$ weakly by solving the problem

$$\begin{pmatrix} A_i & T_i^* \\ T_i & \end{pmatrix} \begin{pmatrix} u_i \\ p_i \end{pmatrix} = \begin{pmatrix} f_i \\ g_i \end{pmatrix} \text{ on } U_i \times X_i$$

we see that the condition $T_1 u_1 + T_2 u_2 = 0$ may be enforced by solving

$$\begin{pmatrix} A_1 & & T_1^* \\ & A_2 & T_2^* \\ T_1 & T_2 & \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ p \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ 0 \end{pmatrix} \text{ on } U_1 \times U_2 \times (X_1 \cap X_2) \tag{1.14}$$

A natural question to ask, then, is what hypotheses are required on the subproblems A_i for the coupled problem to be stable in the sense of Section 1.3. Evidently, we must have stability of each subproblem, but as will be seen in the following example, this is insufficient.

Example 1.5.1. To see this, choose any two vector spaces M, N, and let $U = M \times N$ with projection operators π_M, π_N. Let now $A : U \to U^*$ be defined by $\langle Au, v\rangle = (\pi_M u, \pi_M v)_M$, and $B : U \to N^*$ be defined by $\langle Bu, n\rangle = (\pi_N u, n)_N$. Then by Brezzi theory the problem

$$\begin{pmatrix} A & B^* \\ B & \end{pmatrix} \begin{pmatrix} u \\ n \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \text{ on } U \times N$$

is trivially stable. However, if we couple together two copies of this problem in the manner of Equation (1.14), we get the problem

$$\begin{pmatrix} A & & B^* \\ & A & B^* \\ B & B & \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ g \end{pmatrix} \text{ on } U \times U \times N,$$

which is not well posed. Indeed, the coupling condition now only imply that $\pi_N u_1 + \pi_N u_2 = 0$, meaning that we may add (resp. subtract) any $n \in N$ to u_1 (resp. u_2). Because $A_i = \pi_M$ cannot see the effect of this at all, the problem operator cannot in general be injective, and the solution is only determined in the case where $f_1, f_2 \in U^*$ agree on N, and even then the solution is only determined up to a constant in N.

In this case, we see that the information added by the projections was essential to the solvability of the problem, as can be seen by considering the problem $A_i u_i = f_i$. Although it is solvable if considered as a problem in M^* alone, it is not solvable on all of U^*.

Although the above example is quite artificial, consider also the following similar-looking, yet more natural example:

Example 1.5.2. Suppose that Ω is a disk, split into two parts Ω_1, Ω_2 along a diameter Γ. Suppose we were to write up the Poisson problem on Ω_i as an

operator A_i with right hand side f_i. Then, letting T be the trace to Γ, it seems plausible that the problem of finding a pair u_1, u_2 so that $A_i u_i = f_i$ for $i = 1, 2$ and $T_1 u_1 = T_1 u_2$ should be similar (if, as we shall see, not equivalent) to just solving the Poisson problem on the full domain, which we have every reason to think should be well posed.

In this case, we do not have the same problem. For normal definitions of A_i, the constraint $T_i u_i = g_i$ is not, in fact, required to solve the problem, and can be substituted by e.g. a Neumann condition on Γ instead. Accordingly, the problem $A_i u_i = f_i$ is well-posed not only on the kernel of T_i, but on all of U_i, meaning that we do not run into the issue described above, and the coupled problem will be well-posed.

We remark also that the splitting procedure described above will not, in general, actually result in something equivalent to the "whole" Poisson problem. To see this, observe that in the "split" formulation, we may choose f_1, f_2 so that the solution u_1, u_2 and test functions, although H^1 on either subdomain, need not glue together to a function in $H^1(\Omega_1 \cup \Omega_2)$.

Note also that in the case where we only have Neumann conditions on the boundary of Ω, the coupled problem will only be solvable up to a constant, while each subproblem will be fully solvable, seeing as we there have the condition $T_i u_i = g_i$.

The above two examples show that coupling subproblems may break the inf-sup properties of the diagonal block of the resulting coupled saddle-point problem. However, the following lemma, proven in Paper 3 for the case of a saddle-point problem, shows that the coupled problem satisfies all the other assumptions of Theorem 1.3.4.

Theorem 1.5.3. *Suppose that each of the operators* $P_i : \begin{pmatrix} A_i & T_i^* \\ T_i & \end{pmatrix} : U_i \times X_i \to U_i^* \times X_i^*$ *admit bounded inverses* P_i^{-1}*. and satisfy the strengthened*[13] *inf-sup condition* $\inf_{u \in U_i} \sup_{v \in U_i} \langle A_i u_i, v_i \rangle \geq \alpha_i, \inf_{v \in U_i} \sup_{b \in U_i} \langle A_i u_i, v_i \rangle \geq \alpha_i,$*. Then the map*

$$\begin{pmatrix} A_1 & & T_1^* \\ & A_2 & T_2^* \\ T_1 & T_2 & \end{pmatrix} : U_1 \times U_2 \times (X_1 \cap X_2) \to U_1^* \times U_2^* \times (X_1 \cap X_2)^*$$

has a bounded inverse, with constant depending only on $\|P_i\|, \|P_i^{-1}\|, \alpha_i$*.*

Proof. By Theorem 1.3.4, it suffices to show the diagonal block $\begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}$ satisfies both inf-sup conditions. From the assumption that we have strengthened

[13] Note that this condition is stronger than the assumptions of Theorem 1.3.4, which only assumes this condition for u, v in $\ker T_i$. The condition may be interpreted as A_i being invertible as an operator $U_i \to U_i^*$, while invertibility of P_i only requires it to be invertible on $\ker T_i$.

inf-sup conditions, these hold with constant $\min_i \alpha_i$. It therefore remains to prove the inf-sup condition

$$\inf_{X_1\cap X_2}\sup_{U_1\times U_2}\frac{(T_1u_1,w)+(T_2u_2,w)}{\|w\|_{X_1\cap X_2}\|(u_1,u_2)\|_{U_1\times U_2}}\geq\alpha$$

Given any $w\in X_1\cap X_2$, let $\begin{pmatrix}u_i^w\\ \lambda_i^w\end{pmatrix}=P_i^{-1}\begin{pmatrix}0\\ R_{X_i}^{-1}w\end{pmatrix}$ where $R_{X_i}^{-1}:X_i\to X_i^*$ is the inverse Riesz map. Then by construction $T_iu_i=R_{X_i}^{-1}w$. Because P_i^{-1} is bounded, we have that $\|u_i^w\|_{U_i}\leq\|P_i^{-1}\|\|R_{X_i}^{-1}w\|_{X_i^*}=\|P_i^{-1}\|\|w\|_{X_i}$.

Hence, we have that

$$\begin{aligned}\sup_{U_1\times U_2}\frac{(T_1u_1,w)+(T_2u_2,w)}{\|w\|_{X_1\cap X_2}\|(u_1,u_2)\|_{U_1\times U_2}}&\geq\frac{(T_1u_1^w,w)+(T_2u_2^w,w)}{\|w\|_{X_1\cap X_2}\|(u_1^w,u_2^w)\|_{U_1\times U_2}}\\&=\frac{(R_{X_1}^{-1}w,w)+(R_{X_w}^{-1}w,w)}{\|w\|_{X_1\cap X_2}\|(u_1^w,u_2^w)\|_{U_1\times U_2}}\\&=\frac{\|w\|_{X_1}^2+\|w\|_{X_2}^2}{\|w\|_{X_1\cap X_2}\|(u_1^w,u_2^w)\|_{U_1\times U_2}}\\&=\frac{\|w\|_{X_1\cap X_2}^2}{\|w\|_{X_1\cap X_2}\|(u_1^w,u_2^w)\|_{U_1\times U_2}}\\&=\frac{\|w\|_{X_1\cap X_2}}{\|(u_1^w,u_2^w)\|_{U_1\times U_2}}\end{aligned}$$

Now, because

$$\begin{aligned}\|(u_1^w,u_2^w)\|_{U_1\times U_2}^2=\|u_1^w\|_{U_1}^2+\|u_2^w\|_{U_2}^2&\leq\max_i\|P_i^{-1}\|^2\left(\|R_{X_1}^{-1}w\|_{X_1^*}^2+\|R_{X_2}^{-1}w\|_{X_2^*}^2\right)\\&=\max_i\|P_i^{-1}\|^2\left(\|w\|_{X_1}^2+\|w\|_{X_2^*}^2\right)=\max_i\|P_i^{-1}\|^2\|w\|_{X_1\cap X_2}^2\end{aligned}$$

where we used that the Riesz map is an isometry, we have established the inf-sup condition with constant $\frac{1}{\max_i\|P_i^{-1}\|}$. Finally, stability of P_i implies boundedness of all the blocks by Theorem 1.3.4, meaning that A_i,T_i,T_i^* are bounded with constants dependent only on $\|P_i\|,\|P_i^{-1}\|$. Hence we have established all the hypotheses of Theorem 1.3.4, meaning that $\begin{pmatrix}A_1 & & T_1^*\\ & A_2 & T_2^*\\ T_1 & T_2 & \end{pmatrix}$ has a bounded inverse as desired. ■

We remark that the above theorem may be applied also when U_i are discrete spaces. In this case, an advantage of the formulation is that it does not require any consistency conditions between the discretization chosen for U_1,U_2. Using a term we will define in Section 1.6, this makes the formulation suitable for a monolithic non-unified method for solving the multiphysics problem.

Example 1.5.4. We describe here how Theorem 1.5.3 may be used to prove stability of the coupled Darcy-Stokes problem, conditional on stability of their subproblems and Dirichlet conditions being applied on some part of each subproblem boundary. Because a more thorough analysis is given in Paper III, we shall not explicitly describe the function spaces or norms the problems are well-posed in.

We recall first the weak formulation of the Darcy and Stokes problems. If written in block form, they will look as follows:

$$\begin{pmatrix} -\frac{\mu}{K} & -\nabla \\ \nabla\cdot & \end{pmatrix}\begin{pmatrix}\mathbf{u}_p \\ p_p\end{pmatrix} = \begin{pmatrix}\mathbf{f}_1 \\ f_2\end{pmatrix} \text{ and } \begin{pmatrix} -\mu\Delta & -\nabla \\ \nabla\cdot & \end{pmatrix}\begin{pmatrix}\mathbf{u}_f \\ p_f\end{pmatrix} = \begin{pmatrix}\mathbf{g}_1 \\ g_2\end{pmatrix}$$

Suppose we want to couple them with the interface condition $\mathbf{u}_f \cdot \mathbf{n}_\Gamma = \mathbf{u}_p \cdot \mathbf{n}_\Gamma$ on Γ[14]. Recalling the normal trace operator T_n, this condition is just $T_n\mathbf{u}_f - T_n\mathbf{u}_p$. If we enforce these boundary conditions with the Lagrange multiplier λ, the Darcy and Stokes problems are equivalent to the following Darcy and Stokes subproblems with Lagrange multiplier:

$$\begin{pmatrix} -\frac{\mu}{K} & -\nabla & (T_n)^* \\ \nabla\cdot & & \\ T_n & & \end{pmatrix}\begin{pmatrix}\mathbf{u}_p \\ p_p \\ \lambda\end{pmatrix} = \begin{pmatrix}\mathbf{f}_1 \\ f_2 \\ f_3\end{pmatrix} \text{ and } \begin{pmatrix} -\mu\Delta & -\nabla & (T_n)^* \\ \nabla\cdot & & \\ T_n & & \end{pmatrix}\begin{pmatrix}\mathbf{u}_f \\ p_f \\ \lambda\end{pmatrix} = \begin{pmatrix}\mathbf{g}_1 \\ g_2 \\ g_3\end{pmatrix}$$

Writing

$$A_1 = \begin{pmatrix} -\frac{\mu}{K} & -\nabla \\ \nabla\cdot & \end{pmatrix}, \quad A_2 = \begin{pmatrix} -\mu\Delta & -\nabla \\ \nabla\cdot & \end{pmatrix}, \quad T_1 = -T_2 = T_n$$

we see[15] that these problems are in a form to which Theorem 1.5.3 applies. What it shows is that the coupled problem

$$\begin{pmatrix} -\frac{\mu}{K} & & -\nabla & & (T_n)^* \\ & -\mu\Delta & & -\nabla & (-T_n)^* \\ \nabla\cdot & & & & \\ & \nabla\cdot & & & \\ T_n & -T_n & & & \end{pmatrix}\begin{pmatrix}\mathbf{u}_p \\ \mathbf{u}_f \\ p_p \\ p_f \\ \lambda\end{pmatrix} = \begin{pmatrix}\mathbf{f}_1 \\ \mathbf{g}_1 \\ f_2 \\ g_2 \\ h\end{pmatrix}$$

is well-posed provided we have the strengthened inf-sup condition, and the assumption that that the original subproblems were well-posed. The strengthened inf-sup condition follows from the fact that the diagonal blocks $\frac{\mu}{K}(\mathbf{u}_p, \mathbf{v}_p)$, $\mu(\nabla\mathbf{u}_f, \nabla\mathbf{v}_f)$ are coercive on all of $\ker\nabla\cdot$, not merely on $\ker\nabla \cap \ker T_n$. In the

[14]Note that strictly speaking, the resulting system lacks a boundary condition for the Stokes problem. For ease of exposition, we shall ignore it, effectively setting $\alpha_{\mathrm{BJS}} = 0$. Even though this is nonphysical, including the condition does not substantially change the analysis, and the reader is referred to Paper 3 for details.

[15]Here, we have abused notation slightly - strictly speaking, $T_1 \neq T_2$, as their domains and codomains are different.

case of the Darcy problem, this is trivial, while in the case of the Stokes problem we require our hypothesis of there being a nonempty Dirichlet domain, so that the Poincarè inequality can be applied. Note, by Lemma 1.4.1, that this is equivalent to the multiplier-less subproblems being well-posed.

1.6 Solution techniques for the coupled multiphysics problem

In this section, we briefly sketch some approaches for solving a coupled problem with unknowns u_1, u_2 coupled with some condition $T_1 u_1 = T_2 u_2$. Heuristically, when assessing strengths of the methods, we shall take for granted that we understand the subproblems well, and that we have good methods for solving them. Broadly speaking, such methods can be divided into monolithic approaches, where the whole coupled problem is solved "all at once", and separated or domain decomposition methods, where the coupled problem is in some way decomposed into subproblems. The subproblems involved in the decomposition are commonly modified variants of the original subproblems the coupled problem were built out of.

The main challenge involved in a monolithic approach is that an element which work wells well for one subproblem is not guaranteed to work for the other subproblem. As an example, for the coupled Darcy-Stokes problem, the P2-P1 element is stable for the Stokes problem, but not for the Darcy problem. Indeed, many elements are stable for the Darcy problem precisely because they emulate properties of $H(\mathrm{div})$. As discussed in Section 1.5.3, this is the natural choice of space for the Darcy problem, but not for the Stokes problem.

Accordingly, if we want to use the same element across the entire domain in a so-called unified approach, care must be taken to choose one which works well for both subproblems. Although by design this will then result in a stable problem, the use of novel elements means existing solvers cannot always be reused. Often a custom element must be desigend for the particular coupling under consideration, with e.g. [69, 85] doing so for the Darcy-Stokes problem. In the case of the Darcy-Stokes problem, elements which are stable for the Brinkman problem can also be used for a unified approach, with [68, 111] developing suitable elements. [53] develops a hybrid DG (HDG) method including unknowns also on the mesh facets for the Brinkman problem robust in the material parameters.

An alternative to developing a custom element for the coupling being studied is to use a standard element. As this will then be unstable, one needs to include stabilization terms to have a well-posed problem. Provided one ensures these do not artificially affect the solution, this can be a good way to obtain a well-posed problem, with [98, 114] utilizing stabilized methods to solve the Darcy-Stokes problem. [8] develops a stabilized formulation for the Darcy-Stokes problem that they prove to be stable for all conforming elements, while [54] gives a stable hybrid DG (HDG) formulation with stabilization terms on the Stokes facets. For other multiphysics problems, [24] uses a P1-P1 element with interior penalty stabilization to solve a fluid-structure interaction in modeling of the aortic valve.

Alternatively, a non-unified monolithic approach may be used. In such approaches, the problem is still solved monolithically, but differing discretizations are used in the differing domains. This frees us from designing new elements, and properly handling of the coupling at the interface becomes the main challenge, with a result proving something analogous to Theorem 1.5.3 generally being required. A common technique for enforcing coupling conditions applicable to many different kind of couplings is the use of Lagrange multipliers, with [77],[55] being influential examples of how this can be done for the coupled Darcy-Stokes problem, with [36] developing a parameter-robust GMRES preconditioner. See also [7] for a scheme for solving the Navier-Stokes-Darcy problem, and [1] for a method solving a fluid-structure interaction problem coupling the Stokes and the Biot problems via a Lagrange multiplier.

A formulation of the Stokes-Biot problem without an interface multiplier is studied in [10], with both monolithic and domain decomposition solvers considered. In [32], a domain decomposition technique is developed for the Stokes-Biot problem, and investigated both as a solver in its own right and as a preconditioner for a monolithic formulation, and shown numerically to be robust in the problem parameters. The solver is then applied to physiological modeling in [31]. In [109], Lagrange multipliers are used to enforce coupling conditions in a problem from contact mechanics, and an AMG preconditioner for the resulting problem is studied, with special aggregation operators defined to ensure the multipliers are aggregated in a way consistent with the displacement variables. [106] considers an essentially arbitrary abstract monolithically coupled problem, and defines a general AMG block preconditioner for it, showing experimentally that it performs well over several widely different problems.

A technique deserving of mention in the context of monolithic approaches is the introduction of auxiliary unknowns to obtain a "better-conditioned" formulation, which can be illustrated in the context of the Biot and multiple-network poroelastic (MPET) problems[16] in poroelasticity. The weak form of the Biot problem can be posed in a 2-field formulation, where only the fluid pressure and the solid displacement are used as unknowns. However, in [78], this formulation is found to be less parameter robust than a 3-field formulation in which an additional[17] unknown called the total pressure representing a weighted sum of

[16]The Biot problem models porous fluid flow in a porous medium and its elastic deformation, with the two phenomena coupled together. An example physical situation it describes is squeezing the water out of a water-filled sponge, where deformation of the porous solid affects the flow of fluid in it. The MPET problem can be viewed as a generalized version of the Biot problem, where the solid can now contain multiple fluid networks. It can be used to build macro-scale models of complex systems like brain tissue, where fluid networks like arterial blood, venous blood, CSF and so forth can be modeled as separate fluid networks, each of which can exchange water with each other and interact with the elastic deformation of the brain tissue in which they are embedded.

[17]The fact that the introduction of an auxilliary unknown should make the system easier to solve, and not harder, seeing as we are adding more degrees of freedom, might surprise the reader. A digression, but hopefully also an illuminating analogy for how this can happen, is our treatment of the abstract multiphysics problems in Section 1.5.4. There, we also found that the introduction of an auxilliary unknown λ could "mediate" the interaction of two other unknowns u_1, u_2 coupling together in a challenging way by replacing the $u_1 \leftrightarrow u_2$ interaction

the fluid pressure and the "solid pressure", i.e. the divergence of the displacement. Using this formulation, a robust solver and preconditioner is developed for the Biot problem, and is shown to be viable in parameter regimes closer to the incompressible limit than the 2-field formulation.

A similar technique is used in [79], where the introduction of an additional unknown defined in a similar manner as in the above, and the resulting formulation is shown to be parameter robust in a way the original formulation is not. An alternative 3-field formulation for the Biot problem where the Darcy velocity is introduced as a third unknown is used in [64] to develop an operator preconditioner with the displacement, fluid pressure and fluid velocity living in, respectively, (parameter-weighted) $H^1 \cap H\,(\mathrm{div})$, $L^2 \cap H\,(\mathrm{div})$ and L^2 norms, with the same authors extending their approach to the MPET problem in [65]. In [75], a 4-field formulation utilizing both of these unknowns is developed, yielding a particularly simple set of stable spaces, with the displacement, fluid pressure, fluid velocity and total pressure living in H^1, L^2, $H\,(\mathrm{div})$, L^2 respectively. The introduction of additional unknowns to obtain a more pleasant formulation is in no way limited to problems in poroelasticity. For instance, the Darcy-Stokes formulation used in Paper 3 of this thesis can be viewed as a modification of the one considered in Paper 4 to sidestep the difficulties therein. Other examples abound in the literature, with [105] including the curl of the flow to obtain a formulation for the Brinkman problem which yields a block diagonal operator preconditioner, and [56, 57] analyzing a formulation of the Darcy-Stokes problem where the Stokes (psuedo-)stress is introduced as an auxiliary unknown in order to obtain a stable discretization.

Subdomain decomposition is an iterative method in which we start with an initial guess for the coupled solution, which is successively refined by alternatingly solving the first and second subproblems in isolation. At each step, the condition coupling the subproblem we are solving to the other subproblem is replaced by a boundary condition derived from the current guess for the solution to other subproblem. The resulting procedure yields an iterative method which in the limit must satisfy the coupling condition. For an example application of this method to the Darcy-Stokes problem with proof that the resulting method converges, see [47], who compares it to an alternative scheme where each iteration involves only computation at the interface, or see [96] for an example where the Navier-Stokes and Darcy problem are coupled together with a Lagrange multiplier. [44] considers a domain decomposition method for the Darcy-Stokes problem where a stabilized formulation is used to obtain parameter robustness, and gives some biological applications. Other authors using domain decomposition techniques to solve the Darcy-(Navier-)Stokes system are [37, 38, 87]. [104] uses analogous techniques to different ends, developing a parallel algorithm for solving the coupled Stokes-Darcy problem by partitioning the full domain into numerous "fictious" subdomains used for computation, and then

with the two interactions $u_1 \leftrightarrow \lambda$, $u_2 \leftrightarrow \lambda$. This decomposition also, in principle, permits the use of Theorem 1.5.3 to show stability of the coupled problem, with the total pressure playing the role of the Lagrange multiplier, although we emphasize that this is not the approach taken in [78].

using domain decomposition to solve the resulting problem iteratively in parallel and doing extensive experiments to assess its conditioning.

The clear advantage of this method, as identified by [87], is that often, very little modification is required to the subproblem solvers, greatly simplifying the implementation and permitting reuse of good subproblem solver code. In this way, a good preconditioner for the subproblems can be reused for the coupled problem. However, in return, we are required to solve each subproblem multiple times, which increases the computational cost. An analysis of the convergence properties of the iterative method must also be performed, which can be challenging and require tuning auxilaiary parameters. However, this difficulty can be overcome to obtain modified systems where the number of required subproblem solves is guaranteed to remain constant, as in [34, 37].

We also mention the somewhat similar two-grid method introduced in [112], and applied in [35] for the coupled Darcy-Navier-Stokes problem, where the coupled problem is solved on a coarse grid to obtain a coarse global solution, which is then refined by solving two decoupled problems, yielding an algorithm for the coupled problem which is shown not to lose order of approximation. See also [42, 117] for other applications of this method to the coupled Darcy-Stokes problem.

1.6.1 Unfitted mesh methods

It has long been recognized that several important classes of coupled problems pose challenges for mesh generation. According to a 2003 survey [6], "Mesh generation is delicate in many situations, for instance, when the domain has complicated geometry; when the mesh changes with time, as in crack propagation, and remeshing is required at each timestep; when a Lagrangian formulation is employed, especially with nonlinear PDEs.". Another class of problems where mesh generation becomes a challenge are fluid-structure interaction problems. Indeed, accurately modeling the time-varying domain is frequently a primary motivation for studying these problems in the first place, precluding standard approaches which do not face this challenge head-on. Several methods for facing this challenge directly by considering a more general, flexible or time-varying notion of spatial discretization have been extensively studied, and it is our aim to give an overview of some of the larger directions.

The immersed boundary method has roots dating back to [93], where it was developed for a finite difference method, and handles fluid-structure interaction problems by modeling the fluid in a "background" Eulerian framework, meshing the fluid domain without any special accounting of the solid domain. The solid boundary is tracked via Eulerian variables on a fixed mesh. The fluid and solid problems are generally solved separately, and the fact that a fixed mesh is used dispenses with the need for time-varying conforming meshes for the two domains. As the coupling between the structure and the fluid is included directly in the flow equation, an accurate treatment of the interaction force term is a central challenge for the analysis of immersed boundary methods, with [21] developing and rigorously analysiing a finite element method.

The immersed finite element or immersed domain method, analysed in e.g. [115], includes a mesh of the entire solid domain. This is still tracked on the fluid mesh by Lagrangian variables, but the modeling of the interior of the solid enables more sophisticated modeling of its elasticity. A later generalization where the interfacial forces used as boundary conditions for the solid problem are evaluated implicitly is given in [107]. We refer the reader to [58, 70] for reviews of recent advanced in immersed boundary and immersed domain methods, to [66] for a review focusing particularly on such methods as they pertain to fluid-structure interaction problems, and to [4] for an example application investigating time-splitting schemes for fluid-structure interaction problems.

Because the solid mesh does not conform to the fluid mesh, it will intersect the fluid mesh partially, cutting across mesh elements. As they then have very small contributions to the system matrix, the resulting problem can be poorly conditioned, as discussed in [80]. This makes the use of iterative methods expensive, and several authors (e.g. [63], [76], [49]) have opted to use direct solvers. In order to make iterative methods feasible, some form of stabilization or penalty term must be incorporated, or a preconditioner must be developed.

Virtual element methods (VEM) and hybrid higher-order (HHO) methods all involve the generalization of the finite element method to irregular mesh elements. Frequently, the polygons/polyhedra can be almost arbitrary, with a number of vertices varying along the mesh. Though different, both VEM and HHO methods carry out this generalization in analogous ways [81]. In VEM, the local function spaces on each cell are only implicitly defined, and the contributions of a degree of freedom d to the system matrix is a function $\pi(d)$, where π need not be easily invertible. Accordingly, the resulting element can be thought of as only existing "virtually".

In HHO methods, the problem is posed in terms of degrees of freedom living on mesh cells and facets. On each cell, local reconstruction operators are defined which map facet degrees of freedom to a local high-order function in the cell, and the variational problem can then be posed in a local high order space. See [13] or [89] for reviews of the virtual element method, and [46] for a review of HHO methods.

VEM and HHO methods are of particular interest in applications with complex geometries, where the added flexibility in what elements can be used for meshing is important, and the computational overhead of a more complex method can be a price worth paying for not having to deal with an ill-conditioned tetrahedral mesh.

However, care must still be taken to ensure the resulting discrete problem is well-posed. Poor choices of reconstruction operators in HHO methods can result in an ill-posed problem. [27] develops a robust discretisation for the Brinkman problem by defining the reconstruction operators as a sum of Stokes reconstruction operators, which create functions similar to the regular Lagrangian basis functions stable for the Stokes problem, and Darcy reconstruction operators, which create functions similar to the Raviart–Thomas–Nédélec finite element stable for the Darcy problem. Analogously, [43] defines families of $H(\mathrm{div})$ and $H(\mathrm{curl})$ conforming elements for the Virtual Element Method by generalizing

the classical BDM, Nédélec and RT elements.

VEM and HHO methods also have appealing benefits for multigrid preconditioners and adaptive solvers, because the ability to handle irregular elements means mesh coarsening can be carried out by agglomeration, or simply taking the union of adjacent cells. The fact that the resulting mesh conforms perfectly to the original mesh simplifies construction of the restriction and extension operators used for multigrid, which is used to develop an agglomerative multigrid preconditioner in [28] for elliptic problems, and in [26] for the Stokes and Navier-Stokes problems. Though one gets better accuracy by reassembling the fine grid problem on the coarse grid, the latter paper also tries simply restricting the fine grid problem to the coarse grid for an approximation which trades some accuracy for computational cost. In [12], the same simplicity of coarsening the mesh is used to define an adaptive solver for a discontinuous Galerkin discretization of the Poisson problem, and proves the resulting scheme stable. See [39] for a review and an axiomatization of adaptive finite element methods in general, and [22] for the development of a technique for computing adaptive error indicators over clusters of eigenvalues, and results establishing that this yields better performance than just computing it over a single eigenvalue.

Given the significant interest in unfitted mesh methods, it is perhaps no surprise that there is also significant interest in developing preconditioners for them. Some of the techniques are applicable to multiple unfitted mesh methods. [17] proves (under fairly general assumptions) that the inf-sup constant for the divergence operator, viewed as a map $H^1 \to L^2$, depends continuously on the domain and the degree or number of elements of the discrete function spaces. In [61], this is applied to show an inf-sup property for a wider class of unfitted Stokes elements. [92] develops domain decomposition preconditioners for the Stokes problem and the mixed formulation of the linear elasticity suitable for nonconforming mesh methods.

The literature on unfitted mesh methods and their preconditioning is extensive, with numerous approaches not explicitly discussed above. In the context of multiphysics problems, mortar methods are of particular relevance, providing a way to couple subproblems together on a common interface via Lagrange multipliers, without requiring the interface meshes to match. See [16] for a description, [52] for a comparison to the Nitsche method, and [55, 67, 100, 103] for some example applications. We refer to [99] for a review of the finite cell method, to [94, 95] for discussions of precondition of the extended (XFEM) and generalized finite element methods, and to [116] for an application of XFEM to heart valve modeling developing a block preconditioner for the resulting FSI problem. [9, 11] discusses conditioning issues arising from cut cells in unfitted mesh methods in general, as well as an aggregated finite element method for elliptic problems in general and the Stokes problem in particular.

1.6.2 Dimensionality reduction techniques

Dimensionality reduction techniques have seen particular use in modeling of the cardiovascular system, and involve simplifying a 3D model of e.g. a blood vessel

to a 1D model, where only axial flow in the vessel is assumed to be relevant. Dimensionality can even be reduced further to a 0D (lumped parameter) model, where blood flow through the vessel is assumed to be simply proportional to the pressure difference across it. Such models are particularly simple to implement, and due to the significantly reduced computational cost, can often be written entirely in high-level languages like Python [84]. The low computational cost of the models also makes it feasible to model large networks of vessels, or even the entire cardiovascular system ([91], [18]) by coupling together lower-dimensional models. See [97] for a recent review, and [23] for a systematic review of numerical methods for 1D models.

Of note is that 0D models result in models analogous to simple electric circuits consisting of resistors, inductors and capaticors. In this analogy, the fluid flow through a vessel is the current through a circuit, and the pressure difference across the vessel is the potential difference across the circuit. Despite its simplicity, this analogy and model extends to several types of fluid behavior. The fluid viscosity is modeled by a resistor, the inertia of the fluid is modeled by an inductor, and the interaction between the fluid and the elastic vessel wall is modeled by a capacitor. Interaction between the blood vessel and elastic surrounding tissue may also be incorporated in 0D models through an additional capacitor representing the elasticity of the surrounding tissue.

A related class of models are ones which couple models of reduced dimensionality to a standard 3D finite element model, such as is done in this thesis' Paper II. Such an approach is appealing because it permits incorporating the effects of larger parts of the cardiovascular system into the model, while still spending as much of our "complexity budget" on a particular region of interest as possible. A challenge in the development of such models is, evidently, handling this nonstandard coupling properly, with several of the same challenges as in this thesis' Papers II, III and IV. See [51], [110] for reviews of the 3D-1D coupling in hemodynamics in particular, and [101] for a survey of some applications. For an investigation of fractional Sobolev preconditioners for trace coupled $3D$-$1D$ systems, see [72, 73], and [74] for $2D - 1D$ coupled saddle point problems.

Examples of finite element models of coupled problems where one problem has been approximated by a lower dimensional domain abound also outside hemodynamics. [30] numerically investigates a block diagonal preconditioner for a model of the coupling between a poroelastic medium containing a fluid-filled fracture by coupling together the Biot and the Brinkman problems, where the fracture is modeled as a codimension 1 domain. Porous flow in a domain with cracks of differing dimensionalities is considered in [25].

References

[1] Ambartsumyan, I. et al. "A Lagrange multiplier method for a Stokes–Biot fluid–poroelastic structure interaction model". In: *Numerische Mathematik* vol. 140, no. 2 (2018), pp. 513–553.

[2] Amestoy, P. R. et al. “MUMPS: a general purpose distributed memory sparse solver”. In: *International Workshop on Applied Parallel Computing*. Springer. 2000, pp. 121–130.

[3] Amestoy, P. et al. “On the complexity of the block low-rank multifrontal factorization”. In: *SIAM Journal on Scientific Computing* vol. 39, no. 4 (2017), A1710–A1740.

[4] Annese, M., Fernández, M. A., and Gastaldi, L. “Splitting schemes for a Lagrange multiplier formulation of FSI with immersed thin-walled structure: stability and convergence analysis”. In: *arXiv preprint arXiv:2007.04699* (2020).

[5] Babuška, I. “Error-bounds for finite element method”. In: *Numerische Mathematik* vol. 16, no. 4 (1971), pp. 322–333.

[6] Babuška, I., Banerjee, U., and Osborn, J. E. “Survey of meshless and generalized finite element methods: a unified approach”. In: *Acta Numerica* vol. 12 (2003), pp. 1–125.

[7] Badea, L., Discacciati, M., and Quarteroni, A. “Numerical analysis of the Navier–Stokes/Darcy coupling”. In: *Numerische Mathematik* vol. 115, no. 2 (2010), pp. 195–227.

[8] Badia, S. and Codina, R. “Unified stabilized finite element formulations for the Stokes and the Darcy problems”. In: *SIAM journal on Numerical Analysis* vol. 47, no. 3 (2009), pp. 1971–2000.

[9] Badia, S., Martin, A. F., and Verdugo, F. “Mixed aggregated finite element methods for the unfitted discretization of the Stokes problem”. In: *SIAM journal on scientific computing* vol. 40, no. 6 (2018), B1541–B1576.

[10] Badia, S., Quaini, A., and Quarteroni, A. “Coupling Biot and Navier–Stokes equations for modelling fluid–poroelastic media interaction”. In: *Journal of Computational Physics* vol. 228, no. 21 (2009), pp. 7986–8014.

[11] Badia, S., Verdugo, F., and Martin, A. F. “The aggregated unfitted finite element method for elliptic problems”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 336 (2018), pp. 533–553.

[12] Bassi, F. et al. “On the flexibility of agglomeration based physical space discontinuous Galerkin discretizations”. In: *Journal of Computational Physics* vol. 231, no. 1 (2012), pp. 45–65.

[13] Beirão da Veiga, L. et al. “The hitchhiker’s guide to the virtual element method”. In: *Mathematical models and methods in applied sciences* vol. 24, no. 08 (2014), pp. 1541–1573.

[14] Benzi, M. “Some uses of the field of values in numerical analysis”. In: *Bollettino dell’Unione Matematica Italiana* (2020), pp. 1–19.

[15] Benzi, M. and Olshanskii, M. A. “Field-of-values convergence analysis of augmented Lagrangian preconditioners for the linearized Navier–Stokes problem”. In: *SIAM Journal on Numerical Analysis* vol. 49, no. 2 (2011), pp. 770–788.

[16] Bernardi, C., Maday, Y., and Rapetti, F. "Basics and some applications of the mortar element method". In: *GAMM-Mitteilungen* vol. 28, no. 2 (2005), pp. 97–123.

[17] Bernardi, C. et al. "Continuity properties of the inf-sup constant for the divergence". In: *SIAM Journal on Mathematical Analysis* vol. 48, no. 2 (2016), pp. 1250–1271.

[18] Blanco, P. J. et al. "An anatomically detailed arterial network model for one-dimensional computational hemodynamics". In: *IEEE Transactions on biomedical engineering* vol. 62, no. 2 (2014), pp. 736–753.

[19] Boffi, D., Gardini, F., and Gastaldi, L. "Approximation of PDE eigenvalue problems involving parameter dependent matrices". In: *arXiv preprint arXiv:2001.01304* (2020).

[20] Boffi, D., Brezzi, F., and Gastaldi, L. "On the convergence of eigenvalues for mixed formulations". In: *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* vol. 25, no. 1-2 (1997), pp. 131–154.

[21] Boffi, D. and Gastaldi, L. "A finite element approach for the immersed boundary method". In: *Computers & structures* vol. 81, no. 8-11 (2003), pp. 491–501.

[22] Boffi, D. et al. "Optimal convergence of adaptive FEM for eigenvalue clusters in mixed form". In: *Mathematics of Computation* vol. 86, no. 307 (2017), pp. 2213–2237.

[23] Boileau, E. et al. "A benchmark study of numerical schemes for one-dimensional arterial blood flow modelling". In: *International journal for numerical methods in biomedical engineering* vol. 31, no. 10 (2015), e02732.

[24] Bonomi, D. et al. "Influence of the aortic valve leaflets on the fluid-dynamics in aorta in presence of a normally functioning bicuspid valve". In: *Biomechanics and modeling in mechanobiology* vol. 14, no. 6 (2015), pp. 1349–1361.

[25] Boon, W. M., Nordbotten, J. M., and Yotov, I. "Robust discretization of flow in fractured porous media". In: *SIAM Journal on Numerical Analysis* vol. 56, no. 4 (2018), pp. 2203–2233.

[26] Botti, L., Colombo, A., and Bassi, F. "h-multigrid agglomeration based solution strategies for discontinuous Galerkin discretizations of incompressible flow problems". In: *Journal of Computational Physics* vol. 347 (2017), pp. 382–415.

[27] Botti, L., Di Pietro, D. A., and Droniou, J. "A Hybrid High-Order discretisation of the Brinkman problem robust in the Darcy and Stokes limits". In: *Computer Methods in Applied Mechanics and Engineering* vol. 341 (2018), pp. 278–310.

[28] Botti, L. et al. "hp-hp-Multilevel discontinuous Galerkin solution strategies for elliptic operators". In: *International Journal of Computational Fluid Dynamics* vol. 33, no. 9 (2019), pp. 362–370.

[29] Brezzi, F. “On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers”. In: *Publications mathématiques et informatique de Rennes*, no. S4 (1974), pp. 1–26.

[30] Bukač, M., Yotov, I., and Zunino, P. “Dimensional model reduction for flow through fractures in poroelastic media”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* vol. 51, no. 4 (2017), pp. 1429–1471.

[31] Bukac, M. et al. “Effects of poroelasticity on fluid-structure interaction in arteries: A mputational sensitivity study”. In: *Modeling the heart and the circulatory system*. Springer, 2015, pp. 197–220.

[32] Bukač, M. et al. “Partitioning strategies for the interaction of a fluid with a poroelastic material based on a Nitsche’s coupling approach”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 292 (2015), pp. 138–170.

[33] Buttari, A. “Scalability of parallel sparse direct solvers: methods, memory and performance”. PhD thesis. 2018.

[34] Cai, M., Huang, P., and Mu, M. “Some multilevel decoupled algorithms for a mixed navier-stokes/darcy model”. In: *Advances in Computational Mathematics* vol. 44, no. 1 (2018), pp. 115–145.

[35] Cai, M., Mu, M., and Xu, J. “Numerical solution to a mixed Navier–Stokes/Darcy model by the two-grid approach”. In: *SIAM Journal on Numerical Analysis* vol. 47, no. 5 (2009), pp. 3325–3338.

[36] Cai, M., Mu, M., and Xu, J. “Preconditioning techniques for a mixed Stokes/Darcy model in porous media applications”. In: *Journal of computational and applied mathematics* vol. 233, no. 2 (2009), pp. 346–355.

[37] Cao, Y. et al. “Parallel, non-iterative, multi-physics domain decomposition methods for time-dependent Stokes-Darcy systems”. In: *Mathematics of Computation* vol. 83, no. 288 (2014), pp. 1617–1644.

[38] Cao, Y. et al. “Decoupling the stationary Navier-Stokes-Darcy system with the Beavers-Joseph-Saffman interface condition”. In: *Abstract and Applied Analysis*. Vol. 2013. Hindawi. 2013.

[39] Carstensen, C. et al. “Axioms of adaptivity”. In: *Computers & Mathematics with Applications* vol. 67, no. 6 (2014), pp. 1195–1253.

[40] Chandler-Wilde, S. N., Hewett, D. P., and Moiola, A. “Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples”. In: *Mathematika* vol. 61, no. 2 (2015), pp. 414–443.

[41] Chidyagwai, P., Ladenheim, S., and Szyld, D. B. “Constraint preconditioning for the coupled Stokes–Darcy system”. In: *SIAM Journal on Scientific Computing* vol. 38, no. 2 (2016), A668–A690.

[42] Chidyagwai, P. and Rivière, B. “A two-grid method for coupled free flow with porous media flow”. In: *Advances in Water Resources* vol. 34, no. 9 (2011), pp. 1113–1123.

[43] Da Veiga, L. B. et al. “H(div) and H(curl)-conforming virtual element methods”. In: *Numerische Mathematik* vol. 133, no. 2 (2016), pp. 303–332.

[44] D’Angelo, C. and Zunino, P. “Robust numerical approximation of coupled Stokes’ and Darcy’s flows applied to vascular hemodynamics and biochemical transport”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* vol. 45, no. 3 (2011), pp. 447–476.

[45] Di Nezza, E., Palatucci, G., and Valdinoci, E. “Hitchhiker’s guide to the fractional Sobolev spaces”. In: *arXiv preprint arXiv:1104.4345* (2011).

[46] Di Pietro, D. A., Ern, A., and Lemaire, S. “A review of hybrid high-order methods: formulations, computational aspects, comparison with other methods”. In: *Building bridges: connections and challenges in modern approaches to numerical partial differential equations*. Springer, 2016, pp. 205–236.

[47] Discacciati, M. and Quarteroni, A. “Convergence analysis of a subdomain iterative method for the finite element approximation of the coupling of Stokes and Darcy equations”. In: *Computing and Visualization in Science* vol. 6, no. 2-3 (2004), pp. 93–103.

[48] Durán, R. G. and Muschietti, M. A. “An explicit right inverse of the divergence operator which is continuous in weighted norms”. In: *Studia Mathematica* vol. 148 (2001).

[49] Elhaddad, M. et al. “Multi-level hp-finite cell method for embedded interface problems with application in biomechanics”. In: *International journal for numerical methods in biomedical engineering* vol. 34, no. 4 (2018), e2951.

[50] Evans, L. C. *Partial Differential Equations*. AMS, 2010.

[51] Formaggia, L., Quarteroni, A., and Vergara, C. “On the physical consistency between three-dimensional and one-dimensional models in haemodynamics”. In: *Journal of Computational Physics* vol. 244 (2013), pp. 97–112.

[52] Fritz, A., Hüeber, S., and Wohlmuth, B. I. “A comparison of mortar and Nitsche techniques for linear elasticity”. In: *Calcolo* vol. 41, no. 3 (2004), pp. 115–137.

[53] Fu, G., Jin, Y., and Qiu, W. “Parameter-free superconvergent H (div)-conforming HDG methods for the Brinkman equations”. In: *IMA Journal of Numerical Analysis* vol. 39, no. 2 (2019), pp. 957–982.

[54] Fu, G. and Lehrenfeld, C. “A strongly conservative hybrid DG/mixed FEM for the coupling of Stokes and Darcy flow”. In: *Journal of Scientific Computing* vol. 77, no. 3 (2018), pp. 1605–1620.

[55] Galvis, J. and Sarkis, M. “Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations”. In: *Electron. Trans. Numer. Anal* vol. 26, no. 20 (2007), p. 07.

[56] Gatica, G. N., Oyarzúa, R., and Sayas, F.-J. "A residual-based a posteriori error estimator for a fully-mixed formulation of the Stokes–Darcy coupled problem". In: *Computer Methods in Applied Mechanics and Engineering* vol. 200, no. 21-22 (2011), pp. 1877–1891.

[57] Gatica, G., Oyarzúa, R., and Sayas, F.-J. "Analysis of fully-mixed finite element methods for the Stokes-Darcy coupled problem". In: *Mathematics of Computation* vol. 80, no. 276 (2011), pp. 1911–1948.

[58] Griffith, B. E. and Patankar, N. A. "Immersed methods for fluid–structure interaction". In: *Annual Review of Fluid Mechanics* vol. 52 (2020), pp. 421–448.

[59] Grisvard, P. *Elliptic problems in nonsmooth domains*. SIAM, 2011.

[60] Gustafson, K. "The Toeplitz-Hausdorff theorem for linear operators". In: *Proceedings of the American Mathematical Society* vol. 25, no. 1 (1970), pp. 203–204.

[61] Guzmán, J. and Olshanskii, M. "Inf-sup stability of geometrically unfitted Stokes finite elements". In: *Mathematics of Computation* vol. 87, no. 313 (2018), pp. 2091–2112.

[62] Hansbo, P. "Nitsche's method for interface problems in computa-tional mechanics". In: *GAMM-Mitteilungen* vol. 28, no. 2 (2005), pp. 183–206.

[63] Heltai, L., Roy, S., and Costanzo, F. "A fully coupled immersed finite element method for fluid structure interaction via the Deal. II library". In: *arXiv preprint arXiv:1209.2811* (2012).

[64] Hong, Q. and Kraus, J. "Parameter-robust stability of classical three-field formulation of Biot's consolidation model". In: *arXiv preprint arXiv:1706.00724* (2017).

[65] Hong, Q. et al. "Conservative discretizations and parameter-robust preconditioners for Biot and multiple-network flux-based poroelasticity models". In: *Numerical Linear Algebra with Applications* vol. 26, no. 4 (2019), e2242.

[66] Hou, G., Wang, J., and Layton, A. "Numerical methods for fluid-structure interaction—a review". In: *Communications in Computational Physics* vol. 12, no. 2 (2012), pp. 337–377.

[67] Huang, P., Chen, J., and Cai, M. "A Mortar Method Using Nonconforming and Mixed Finite Elements for the Coupled Stokes-Darcy Model". In: *Advances in Applied Mathematics and Mechanics* vol. 9, no. 3 (2017), pp. 596–620.

[68] Johnny, G. and Michael, N. "A family of nonconforming elements for the Brinkman problem". In: *IMA Journal of Numerical Analysis* vol. 32, no. 4 (2012), pp. 1484–1508.

[69] Karper, T., Mardal, K.-A., and Winther, R. "Unified finite element discretizations of coupled Darcy–Stokes flow". In: *Numerical Methods for Partial Differential Equations: An International Journal* vol. 25, no. 2 (2009), pp. 311–326.

[70] Kim, W. and Choi, H. "Immersed boundary methods for fluid-structure interaction: A review". In: *International Journal of Heat and Fluid Flow* vol. 75 (2019), pp. 301–309.

[71] Kirby, R. C. "From functional analysis to iterative methods". In: *SIAM review* vol. 52, no. 2 (2010), pp. 269–293.

[72] Kuchta, M., Mardal, K.-A., and Mortensen, M. "Preconditioning trace coupled 3d-1d systems using fractional Laplacian". In: *Numerical Methods for Partial Differential Equations* vol. 35, no. 1 (2019), pp. 375–393.

[73] Kuchta, M. et al. "Analysis and approximation of mixed-dimensional PDEs on 3D-1D domains coupled with Lagrange multipliers". In: *arXiv preprint arXiv:2004.02722* (2020).

[74] Kuchta, M. et al. "Preconditioners for saddle point systems with trace constraints coupling 2d and 1d domains". In: *SIAM Journal on Scientific Computing* vol. 38, no. 6 (2016), B962–B987.

[75] Kumar, S. et al. "Conservative discontinuous finite volume and mixed schemes for a new four-field formulation in poroelasticity". In: *ESAIM: Mathematical Modelling and Numerical Analysis* vol. 54, no. 1 (2020), pp. 273–299.

[76] Lācis, U., Taira, K., and Bagheri, S. "A stable fluid–structure-interaction solver for low-density rigid bodies using the immersed boundary projection method". In: *Journal of Computational Physics* vol. 305 (2016), pp. 300–318.

[77] Layton, W. J., Schieweck, F., and Yotov, I. "Coupling fluid flow with porous media flow". In: *SIAM Journal on Numerical Analysis* vol. 40, no. 6 (2002), pp. 2195–2218.

[78] Lee, J. J., Mardal, K.-A., and Winther, R. "Parameter-robust discretization and preconditioning of Biot's consolidation model". In: *SIAM Journal on Scientific Computing* vol. 39, no. 1 (2017), A1–A24.

[79] Lee, J. J. et al. "A mixed finite element method for nearly incompressible multiple-network poroelasticity". In: *SIAM Journal on Scientific Computing* vol. 41, no. 2 (2019), A722–A747.

[80] Lehrenfeld, C. and Reusken, A. "Optimal preconditioners for Nitsche-XFEM discretizations of interface problems". In: *Numerische Mathematik* vol. 135, no. 2 (2017), pp. 313–332.

[81] Lemaire, S. "Bridging the hybrid high-order and virtual element methods". In: *IMA Journal of Numerical Analysis* (2019).

[82] Liesen, J. and Tichý, P. "Convergence analysis of Krylov subspace methods". In: *GAMM-Mitteilungen* vol. 27, no. 2 (2004), pp. 153–173.

[83] Logg, A., Mardal, K., and Wells, G. *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*. Lecture Notes in Computational Science and Engineering. Springer Berlin Heidelberg, 2012.

[84] Manini, S. et al. "pyNS: an open-source framework for 0D haemodynamic modelling". In: *Annals of biomedical engineering* vol. 43, no. 6 (2015), pp. 1461–1473.

[85] Mardal, K. A., Tai, X.-C., and Winther, R. "A robust finite element method for Darcy–Stokes flow". In: *SIAM Journal on Numerical Analysis* vol. 40, no. 5 (2002), pp. 1605–1631.

[86] Mardal, K.-A. and Winther, R. "Preconditioning discretizations of systems of partial differential equations". In: *Numerical Linear Algebra with Applications* vol. 18, no. 1 (2011), pp. 1–40.

[87] Márquez, A., Meddahi, S., and Sayas, F.-J. "A decoupled preconditioning technique for a mixed Stokes–Darcy model". In: *Journal of Scientific Computing* vol. 57, no. 1 (2013), pp. 174–192.

[88] Mary, T. "Block Low-Rank multifrontal solvers: complexity, performance, and scalability". PhD thesis. 2017.

[89] Mengolini, M., Benedetto, M. F., and Aragón, A. M. "An engineering perspective to the virtual element method and its interplay with the standard finite element method". In: *Computer Methods in Applied Mechanics and Engineering* vol. 350 (2019), pp. 995–1023.

[90] Mikelic, A. and Jäger, W. "On the interface boundary condition of Beavers, Joseph, and Saffman". In: *SIAM Journal on Applied Mathematics* vol. 60, no. 4 (2000), pp. 1111–1127.

[91] Mynard, J. P. and Smolich, J. J. "One-dimensional haemodynamic modeling and wave dynamics in the entire adult circulation". In: *Annals of biomedical engineering* vol. 43, no. 6 (2015), pp. 1443–1460.

[92] Pavarino, L. and Scacchi, S. "Isogeometric block FETI-DP preconditioners for the Stokes and mixed linear elasticity systems". In: *Computer Methods in Applied Mechanics and Engineering* vol. 310 (2016), pp. 694–710.

[93] Peskin, C. S. "Numerical analysis of blood flow in the heart". In: *Journal of computational physics* vol. 25, no. 3 (1977), pp. 220–252.

[94] Prenter, F. de, Verhoosel, C., and Brummelen, E. van. "Preconditioning immersed isogeometric finite element methods with application to flow problems". In: *Computer Methods in Applied Mechanics and Engineering* vol. 348 (2019), pp. 604–631.

[95] Prenter, F. de et al. "Condition number analysis and preconditioning of the finite cell method". In: *Computer Methods in Applied Mechanics and Engineering* vol. 316 (2017), pp. 297–327.

[96] Qiu, C. et al. “A domain decomposition method for the time-dependent Navier-Stokes-Darcy model with Beavers-Joseph interface condition and defective boundary condition”. In: *Journal of Computational Physics* (2020), p. 109400.

[97] Quarteroni, A., Veneziani, A., and Vergara, C. “Geometric multiscale modeling of the cardiovascular system, between theory and practice”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 302 (2016), pp. 193–252.

[98] Rui, H. and Zhang, R. “A unified stabilized mixed finite element method for coupling Stokes and Darcy flows”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 198, no. 33-36 (2009), pp. 2692–2699.

[99] Schillinger, D. and Ruess, M. “The Finite Cell Method: A review in the context of higher-order structural analysis of CAD and image-based geometric models”. In: *Archives of Computational Methods in Engineering* vol. 22, no. 3 (2015), pp. 391–455.

[100] Seitz, A. et al. “Isogeometric dual mortar methods for computational contact mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 301 (2016), pp. 259–280.

[101] Shi, Y., Lawford, P., and Hose, R. “Review of zero-D and 1-D models of blood flow in the cardiovascular system”. In: *Biomedical engineering online* vol. 10, no. 1 (2011), p. 33.

[102] Szyld, D. B. “The many proofs of an identity on the norm of oblique projections”. In: *Numerical Algorithms* vol. 42, no. 3-4 (2006), pp. 309–323.

[103] Triebenbacher, S. et al. “Applications of the mortar finite element method in vibroacoustics and flow induced noise computations”. In: *Acta Acustica united with Acustica* vol. 96, no. 3 (2010), pp. 536–553.

[104] Vassilev, D., Wang, C., and Yotov, I. “Domain decomposition for coupled Stokes and Darcy flows”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 268 (2014), pp. 264–283.

[105] Vassilevski, P. S. and Villa, U. “A block-diagonal algebraic multigrid preconditioner for the Brinkman problem”. In: *SIAM Journal on Scientific Computing* vol. 35, no. 5 (2013), S3–S17.

[106] Verdugo, F. and Wall, W. A. “Unified computational framework for the efficient solution of n-field coupled problems with monolithic schemes”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 310 (2016), pp. 335–366.

[107] Wang, X. and Zhang, L. T. “Modified immersed finite element method for fully-coupled fluid–structure interactions”. In: *Computer methods in applied mechanics and engineering* vol. 267 (2013), pp. 150–169.

[108] Wathen, A. J. “Preconditioning”. In: *Acta Numerica* vol. 24 (2015).

[109] Wiesner, T. et al. “Algebraic multigrid methods for dual mortar finite element formulations in contact mechanics”. In: *International Journal for Numerical Methods in Engineering* vol. 114, no. 4 (2018), pp. 399–430.

[110] Xiao, N., Alastruey, J., and Alberto Figueroa, C. “A systematic comparison between 1-D and 3-D hemodynamics in compliant arterial models”. In: *International journal for numerical methods in biomedical engineering* vol. 30, no. 2 (2014), pp. 204–231.

[111] Xie, X., Xu, J., and Xue, G. “Uniformly-stable finite element methods for Darcy-Stokes-Brinkman models”. In: *Journal of Computational Mathematics* (2008), pp. 437–455.

[112] Xu, J. “A novel two-grid method for semilinear elliptic equations”. In: *SIAM Journal on Scientific Computing* vol. 15, no. 1 (1994), pp. 231–237.

[113] Xu, J. and Zikatanov, L. “Some observations on Babuška and Brezzi theories”. In: *Numerische Mathematik* vol. 94, no. 1 (2003), pp. 195–202.

[114] Yu, J. et al. “Nitsche’s type stabilized finite element method for the fully mixed Stokes–Darcy problem with Beavers–Joseph conditions”. In: *Applied Mathematics Letters* vol. 110 (2020), p. 106588.

[115] Zhang, L. et al. “Immersed finite element method”. In: *Computer Methods in Applied Mechanics and Engineering* vol. 193, no. 21-22 (2004), pp. 2051–2067.

[116] Zonca, S., Vergara, C., and Formaggia, L. “An unfitted formulation for the interaction of an incompressible fluid with a thick structure via an XFEM/DG approach”. In: *SIAM Journal on Scientific Computing* vol. 40, no. 1 (2018), B59–B84.

[117] Zuo, L. and Hou, Y. “A decoupling two-grid algorithm for the mixed Stokes-Darcy model with the Beavers-Joseph interface condition”. In: *Numerical Methods for Partial Differential Equations* vol. 30, no. 3 (2014), pp. 1066–1082.

Chapter 2

Summary of papers

2.1 Paper I

In this paper, we apply a computational model to assess the plausibility of microscale intervascular convective flow in the brain, and estimate the permeability of brain tissue. This work was motivated by the recent interest in the glymphatic system, a proposed mechanism of cerebral waste clearance thought to involve a convective flow between cerebral blood vessels illustrated in Figure 2.1.

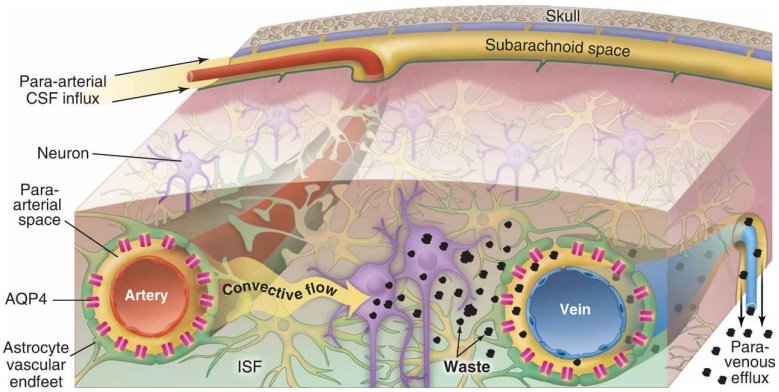

Figure 2.1: Illustration of the proposed mechanism of waste clearance from [5]. Given a pressure gradient between the high-pressure arteries/arterioles and the low-pressure veins/venules, one would observe the efflux of water marked 'Convective flow' in the figure. The magnitude of the water flux would be proportional to the pressure difference, and to the permeability of the extravascular space.

We construct a mesh of the extracellular space using a reconstruction of an $\approx (4\mu\text{m})^3$ portion of rat neuropil generated from electron microscopy data in [2]. As waste products are constrained to the extracellular space, we simulate viscous flow in the extracellular space (ECS), and determine necessary pressure gradients to explain the rates of waste clearance seen experimentally. As these are very large, we conclude that convective transport through the glymphatic pathway is unlikely to be the main cause of cerebral waste clearance. In order to ensure the effects of some parameters involved in the ECS reconstruction do not significantly impact our conclusions, we repeat the experiment for different reconstruction parameter choices and obtain substantially similar results.

Mathematically, our computational model consists of the Stokes equations for viscous flow. Due to the large mesh size required to capture the complex geometry

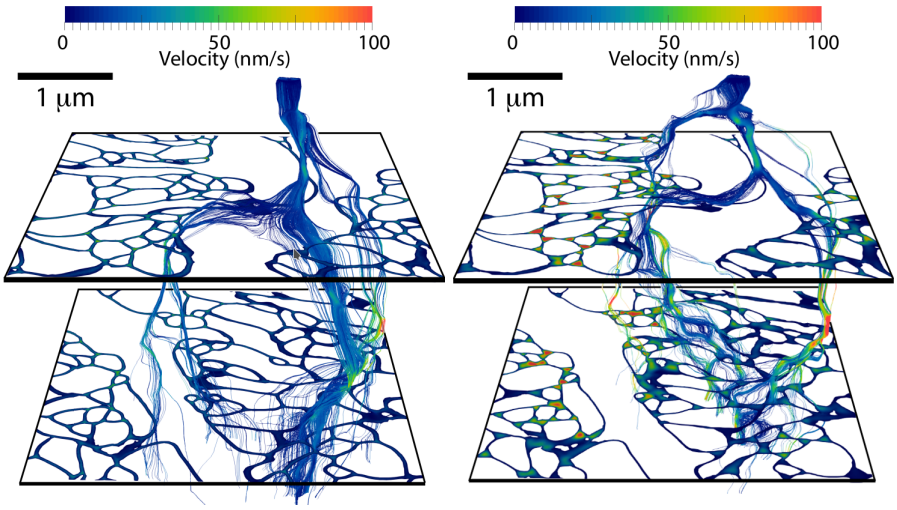

Figure 2.2: Partial visualization of the extracellcular flow field from Paper I for two different ECS reconstruction. In each reconstruction, two cross-sections of the ECS are visible, showing its tortuous structure.

of the extracellular space, we opted to use the unstable $P1 - P1$ elements with a stabilization term instead of the more standard stable $P2 - P1$ element, as this would have come at a higher computational cost. As our mathematical methods are relatively standard, the primary interest of this paper is in the significance of its results for physiology. Viewed in connection with Paper II, however, it nevertheless serves as a good illustration of the immense variation in parameter- and length scales relevant for biomechanical modeling.

2.2 Paper II

In this paper, we investigate numerical methods for the solution of a coupled 3d-1d diffusion problem using `fenics`$_{\text{ii}}$ [3], and apply them to model diffusion of a bloodborne MRI tracer. This is to some extent related to the problem considered in the previous paper, which modeled convective solute transport only in a microscale part of the interstitium. In this paper, the interstitium is "homogenized", in that the intricate geometry of the ECS shown in Figure 2.2 is reduced to homogenous space, its effects on fluid flow instead incorporated in the model through a material parameter. As this frees us from needing a quite so detailed mesh, we are able to model a length scale of millimeters instead of microns, meaning the interstitial space can be coupled to the vasculature, which we include as a highly detailed 1D mesh.

The coupling between the 3D and the 1D diffusion problem is accomplished by the use of an averaging operator π between the 3D and 1D spaces, yielding

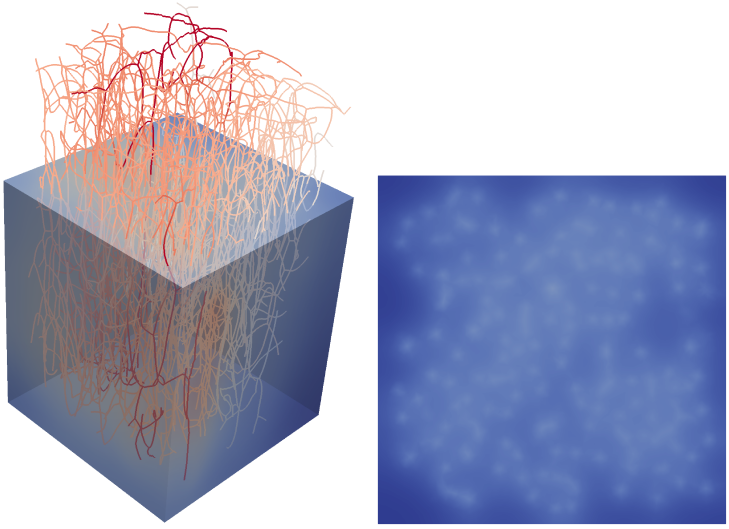

Figure 2.3: Example results shown on a (left) clip and (right) on a slice of the 3D domain. In the left, the 1D mesh of the cerebral vasculature is visible, showing the high level of detail we are able to use in our model of the 1D domain. In the right, the 'halos' of increased concentration immediately around the vessels occur because tracer enters the 3D domain via the 1D domain.

the following problem:

$$\begin{bmatrix} I - kD_\Omega\Delta_\Omega & 0 & k\beta\Pi^* \\ 0 & I - kD_\Gamma\Delta_\Gamma & k\beta I \\ k\beta\Pi & k\beta I & -k. \end{bmatrix} \begin{bmatrix} p_{3\mathrm{D}} \\ p_{1\mathrm{D}} \\ \lambda \end{bmatrix} = \begin{bmatrix} f \\ g \\ h \end{bmatrix} \tag{2.1}$$

Here, π is an averaging operator mapping the 3D space to the 1D space, studied further in [4]. Deciding the appropriate Sobolev spaces to use so that this trace operator bounded and right invertible is theoretically nontrivial, complicating our use of the framework of operator preconditioning to find an appropriate preconditioner for the coupled problem. The choice of space is therefore investigated numerically by assessing the preconditioner arising from various choices of spaces. As the resulting solution method is quite fast, and the addition of the multiplier inexpensive, the resulting method seems likely to be faster than a standard domain decomposition method.

Seeing as the model is capable of efficiently estimating solute transport from the blood vessels to the interstital space, it is particularly suitable for applications. Possible future work in this area might therefore involve applying the model to study phenomena where local variations in the vasculature are of interest. One examples of such a phenomena is the effect of local plaque buildup (reducing local blood throughput in the vasculature) on oxygen delivery in that region.

Another direction for future work involves extending the model to incorporate convection, as it currently does not properly include convection.

2.3 Paper III

In this paper, we study the general form of a multiphysics problem consisting of two coupled "single-physics" problems, where the coupling is enforced using a Lagrange multiplier. We show theoretically that under the assumption that the problems are individually well-posed in a slightly stronger sense than usual, so is the coupled problem, with no modification of either problem required beyond modifying the norm of the multiplier space. This result also extends to the discrete corollary that if both discrete systems are individually stable, so is the coupled system, given in its abstract form below.

$$\mathcal{A}\begin{pmatrix} u_1 \\ u_2 \\ p_1 \\ p_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} A_i & & B_i* & & T_1* \\ & A_2 & & B_2* & T_2* \\ B_1 & & & & \\ & B_2 & & & \\ T_1 & T_2 & & & \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ p_1 \\ p_2 \\ \lambda \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ g_1 \\ g_2 \\ h \end{pmatrix} \in \begin{pmatrix} V_1* \\ V_2* \\ Q_1* \\ Q_2* \\ \Lambda *_1 + \Lambda *_2 \end{pmatrix} \tag{2.2}$$

From our results, we derive preconditioners in the operator preconditioning framework for the coupled Darcy-Stokes and Stokes-linear elasticity problems, which are proven to be stable and robust in the material parameters under the assumption that the Stokes, Darcy and linear elasticity problems with Lagrange multiplier are individually well-posed on their respective domains. All assumptions and results are justified by numerical experiments, demonstrating the suitability of our approach for applications. Needing this assumption is, to some extent, a drawback, but in return we are able to prove a quite general result, as our proof applies whenever the assumption holds. The sting of having to make the assumption is also somewhat lessened by the fact that the subproblems we consider have previously been studied in the literature, see e.g. [1], meaning the required assumptions have in fact been proven in several cases.

The effects of the intersections of the interface with the Dirichlet domain of the subproblems is thoroughly investigated, and is found to affect which norm we have to use for the multiplier space in a manner related to the work of [1]. Additionally, we demonstrate that failing to use parameter weighted norms at the interface makes the natural operator preconditioner quite poorly-conditioned, showing that the methods we develop are necessary.

A clear direction for future work is to more clearly establish what is required for the subproblem stability assumptions to hold. Here, we briefly comment on a sense in which the presence of Lagrange multipliers do not cause any additional difficulties. By Lemma 1.4.1, we know that provided the interfacial trace map used in the problem formulation has a bounded right inverse, the subproblem where the interface condition is enforced with a Lagrange multiplier is well-posed

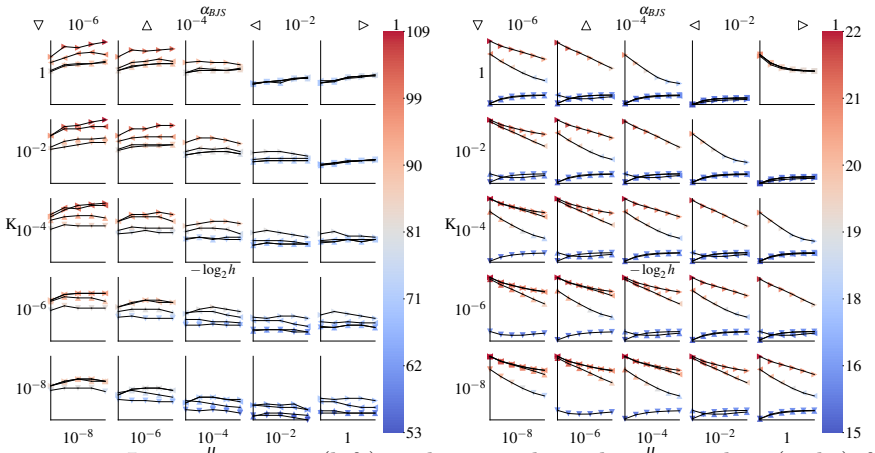

Figure 2.4: Iteration count (left) and spectral condition number (right) for the Darcy-Stokes problem with homogeneous Dirichlet boundary conditions preconditioned by Riesz map preconditioner developed in Paper III. Each subplot plots the (logarithm of) the discretization parameter against the iteration count / spectral condition number for the indicated fixed value of the material parameters K, μ so that the system size grows from left to right, with the Beaver-Joseph-Saffman parameter α_{BJS} indicated by the line marker. The resulting plot is somewhat busy, but clearly demonstrates the stability of the preconditioner with respect to discretization and three different material parameters.

if and only if the subproblem where the interface condition is enforced strongly is well-posed.

Applying this to the Darcy subproblem of Assumption III.4.2, for example, and keeping with its notation, we get that the problem of finding $(\mathbf{u}_p, p_p, \lambda) \in W$ such that $a((\mathbf{u}_p, p_p, \lambda), (\mathbf{v}_p, q_p, w)) = L((\mathbf{v}_p, q_p, w))$ for all $(\mathbf{v}_p, q_p, w) \in W$ is well-posed if and only if the problem of finding $(\mathbf{u}_p, p_p) \in W_0$ such that $a_0((\mathbf{u}_p, p_p), (\mathbf{v}_p, q_p)) = L_0((\mathbf{v}_p, q_p))$ for all $(\mathbf{v}_p, q_p) \in W_0$ is well-posed, where the equivalence is contingent on the map $T_n : \mathbf{V} \to \Lambda^*$ being bounded, surjective (and linear), meaning it has a bounded right inverse. Here,

$$
\begin{aligned}
a_0((\mathbf{u}_p, p_p), (\mathbf{v}_p, q_p)) &= K^{-1}(\mathbf{u}_p, \mathbf{v}_p) + (p, \nabla \cdot \mathbf{v}_p) + (\nabla \cdot \mathbf{u}_p, q), \\
L_0((\mathbf{v}_p, q_p)) &= (\mathbf{f}, \mathbf{v}_p) + (g, q_p) \\
a((\mathbf{u}_p, p_p, \lambda), (\mathbf{v}_p, q_p, w)) &= a_0((\mathbf{u}_p, p_p), (\mathbf{v}_p, q_p)) + (T_n \mathbf{u}_p, w)_\Gamma + (\lambda, T_n \mathbf{v}_p)_\Gamma, \\
L((\mathbf{v}_p, q_p, w)) &= L_0((\mathbf{v}_p, q_p)) + (h, w)_\Gamma
\end{aligned}
$$

$W = \mathbf{V} \times Q \times \Lambda$, $W_0 = \mathbf{V}_0 \times Q \times \Lambda$, and $\mathbf{V} = \frac{1}{\sqrt{K}} \mathbf{H}_{0,D}(\mathrm{div}, \Omega_p)$, $\mathbf{V}_0 = \frac{1}{\sqrt{K}} \mathbf{H}_{0,D \cup \Gamma}(\mathrm{div}, \Omega_p)$ $Q = \sqrt{K} L^2(\Omega_p)$, and $\Lambda = \sqrt{K} H^{\frac{1}{2}}_{00}(\Gamma)$. (Observe that $\mathbf{V}_0$ is simply the kernel of T_n.)

Hence, we see that provided the interfacial trace map we have used to define

our subproblem is continuous and surjective[1], we may replace Assumption III.4.2 by an equivalent assumption of stability of a Darcy problem without Lagrange multiplier, and instead a Dirichlet condition also on Γ.

We remark that the above reasoning can also be applied to any conforming discretization $\mathbf{V}^h \times Q^h \times \Lambda^h \subset \mathbf{V} \times Q \times \Lambda$, showing that if the discretization is such that $\mathbf{V}_0^h \times Q^h$ is a stable discretization of the Darcy problem without Lagrange multiplier, then $\mathbf{V}^h \times Q^h \times \Lambda^h$ is a stable discretization of the Darcy problem with Lagrange multiplier, provided the restriction of T_n to a map $\mathbf{V}^h \times \Lambda^h$ is surjective and bounded. This gives a method by which we can obtain a stable discretization for the subproblems with Lagrange multiplier studied in Paper III given stable discretizations of their versions without Lagrange multiplier.

2.4 Paper IV

In this paper, we use the methods of Paper III to study the coupled Darcy-Stokes system with the Darcy problem in primal form, e.g. without a separate variable for the fluid velocity. This means the Darcy pressure must now be coupled directly to the Stokes velocity, as opposed the case of Paper III, where the Darcy pressure and Stokes velocity were both coupled to the Darcy velocity.

The advantage of this formulation is that the resulting system has smaller dimension than the one with the Darcy problem in mixed form, making it of relevance for applications where compuational cost is a concern, especially if the Darcy domain is significantly larger than the Stokes domain. However, our use of the primal formulation means we lose the mass conservation enjoyed by the mixed formulation. Additionally, because the normal derivative trace operator used to couple the problems cannot be defined on all of the H^1 space we use for the Darcy pressure, putting the resulting formulation on a firm mathematical foundation is nontrivial.

We assume and partially motivate the existence of an alternative coupling operator for coupling the problems, and prove well-posedness of the continuous problem under this assumption using the methods of Paper III. Using the operator preconditioning framework, we then obtain a robust preconditioner. The robustness of our preconditioner and our assumption of the existence of a coupling operator is supported by numerical experiments. In particular, we support the claim made that our results extend to the discrete case by experimentally showing that each choice of stable subproblem elements results in a stable discretization of the coupled problem.

[1] Note that this is nontrivial. Indeed, as discussed in Section 1.4.1, for insufficiently regular $\partial\Omega, \Gamma$, the trace operators we use are not guaranteed to exist.

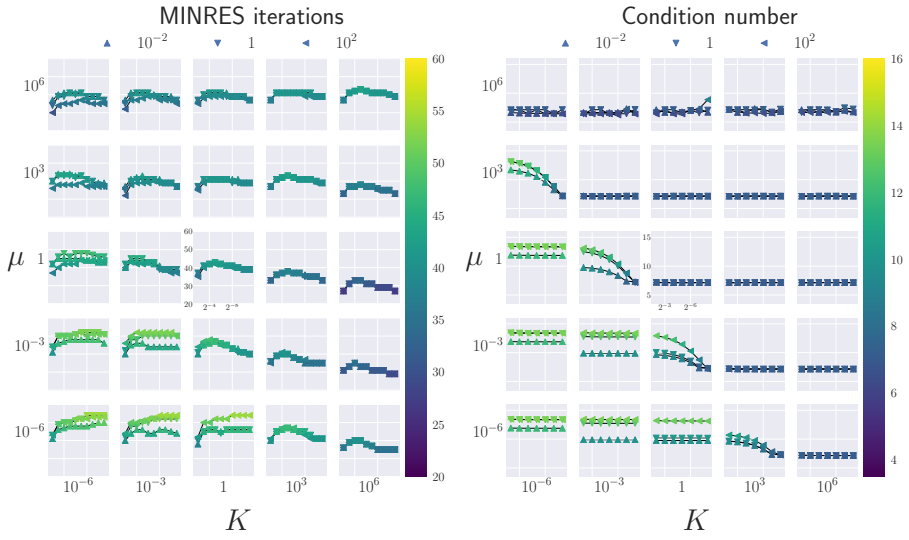

Figure 2.5: Mesh refinement vs. iteration counts (left) and condition numbers (right) for the Stokes-primal Darcy problem using the Riesz map preconditioner developed in Paper III. The plot is made in the same style as Figure 2.4, again demonstrating the robustness of our preconditioner. Note that the triangular markers showing the value of α_{BJS} are quite close, making them look like squares.

References

[1] Galvis, J. and Sarkis, M. "Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations". In: *Electron. Trans. Numer. Anal* vol. 26, no. 20 (2007), p. 07.

[2] Kinney, J. P. et al. "Extracellular sheets and tunnels modulate glutamate diffusion in hippocampal neuropil". In: *The Journal of comparative neurology* vol. 521, no. 2 (Dec. 2012), pp. 448–464.

[3] Kuchta, M. "Assembly of multiscale linear PDE operators". In: *arXiv preprint arXiv:1912.09319* (2019).

[4] Kuchta, M. et al. "Analysis and approximation of mixed-dimensional PDEs on 3D-1D domains coupled with Lagrange multipliers". In: *arXiv preprint arXiv:2004.02722* (2020).

[5] Nedergaard, M. "Garbage truck of the brain". In: *Science* vol. 340, no. 6140 (2013), pp. 1529–1530.

Papers

Paper I

Interstitial solute transport in 3D reconstructed neuropil occurs by diffusion rather than bulk flow

Karl Erik Holter, Benjamin Kehlet, Anna Devor, Terrence J. Sejnowski, Anders M. Dale, Stig W. Omholt, Ole Petter Ottersen, Erlend Arnulf Nagelhus, Kent-André Mardal, Klas H. Pettersen

"Interstitial solute transport in 3D reconstructed neuropil occurs by diffusion rather than bulk flow." Published in *Proceedings of the National Academy of Sciences*, September 2017, volume 114, issue 37, pp. 9894–9899. DOI: 10.1073/pnas.1706942114.

Abstract

The brain lacks lymph vessels and must rely on other mechanisms for clearance of waste products, including amyloid β that may form pathological aggregates if not effectively cleared. It has been proposed that flow of interstitial fluid through the brain's interstitial space provides a mechanism for waste clearance. Here we compute the permeability and simulate pressure-mediated bulk flow through 3D electron microscope (EM) reconstructions of interstitial space. The space was divided into sheets (i.e., space between two parallel membranes) and tunnels (where three or more membranes meet). Simulation results indicate that even for larger extracellular volume fractions than what is reported for sleep and for geometries with a high tunnel volume fraction, the permeability was too low to allow for any substantial bulk flow at physiological hydrostatic pressure gradients. For two different geometries with the same extracellular volume fraction the geometry with the most tunnel volume had 36% higher permeability, but the bulk flow was still insignificant. These simulation results suggest that even large molecule solutes would be more easily cleared from the brain interstitium by diffusion than by bulk flow. Thus, diffusion within the interstitial space combined with advection along vessels is likely to substitute for the lymphatic drainage system in other organs.

Transport of nutrients and waste within the brain's parenchyma is paramount to healthy brain function.

Although lymphatic vessels occur within the meninges [6, 22], they are absent from the brain's parenchyma. This raises the question of how waste products are

cleared from the brain [1, 5, 13, 17, 30, 34]. There is an urgent need to resolve this question, given the fact that several neurological disorders are associated with accumulation of toxic debris and molecules in the brain interstitium [21]. Most notably, insufficient clearance may contribute to the development of Alzheimer's disease and multiple sclerosis [15, 21].

Recently the "glymphatic" hypothesis [15] was launched. This hypothesis holds that the brain is endowed with a waste clearance system driven by bulk flow of interstitial fluid through the interstitium, from para-arterial to paravenous spaces, facilitated by astrocytic aquaporin-4 (AQP4). Further, it was proposed that cerebral arterial pulsation [16] and respiration [19] drive paravascular fluid movement and cerebrospinal (CSF)-interstitial fluid (ISF) exchange. Here, bulk flow is defined as the movement of fluid down the pressure gradient, advection is the transport of a substance by bulk flow, and convection is transport by a combination of advection and diffusion.

There is strong evidence for paravascular advection [10, 13, 29, 33], although the details of influx and efflux pathways and the underlying driving forces are debated [9, 15, 23, 33]. There are, however, controversies regarding the relative importance of advective versus diffusive transport within the interstitial space [1, 5, 13, 30], and the idea that a hydrostatic pressure gradient can cause an advective flow within the interstitium has been questioned [5, 17, 30].

The recent generation of 3D reconstructions of brain neuropil together with representative extracellular space volume estimates have now finally opened for realistic simulations of solute transport in brain. Though the convoluted and very fine structure of the interstitial space makes such simulations challenging, we were able to simulate the flow and estimate the permeability for EM reconstructions from Kinney et al. [18] by meshing the interstitial space into almost 100 million tetrahedrons and describing the relevant physics in each tetrahedron by differential equations.

By simulating bulk flow in two versions of the EM reconstruction we find that the permeability is too low to allow for any substantial bulk flow for realistic hydrostatic pressure gradients. The results imply that diffusion prevails. Besides advancing understanding of waste clearance in brain, our results also elucidate how drugs distribute within brain neuropil after having permeated the blood-brain barrier.

I.1 Results

We used publicly available reconstructions [18] to simulate bulk flow through the interstitial space. The reconstructions were based on electron microscopy of serial sections of rat CA1 hippocampal neuropil. To correct for the volume changes known to occur during tissue preparation and embedding, Kinney et al. [18] adjusted the interstitial volume fraction from 8% in the original EM reconstruction to more physiologically realistic volume fractions of about 20 % [32].

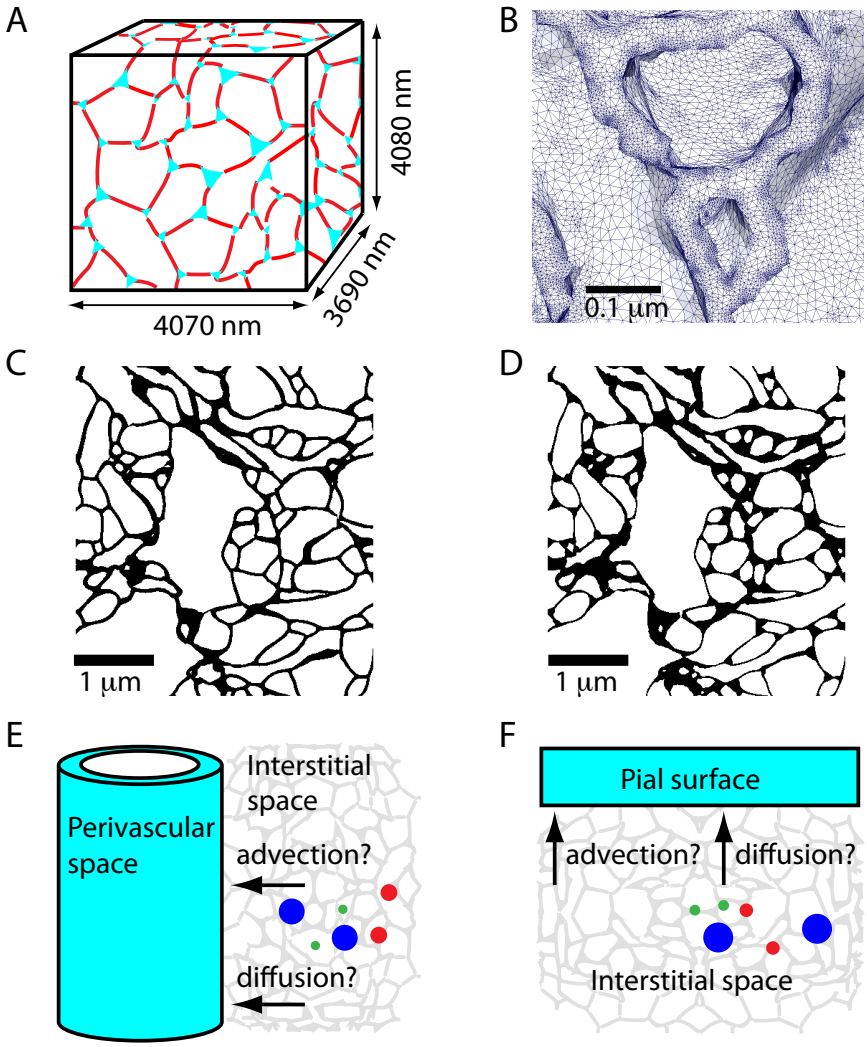

Figure I.1: Model systems and microscopic structure of the extracellular volume. (*A*) Schematic illustration of the EM reconstruction. Tunnels in cyan, sheets in red. (*B*) Sub-micrometer partition of the EM reconstruction showing typical sizes of the 84 million tetrahedrons used in the simulation. (*C*) EM reconstruction from Kinney et al. [18] with a small tunnel volume fraction and (*D*) with a larger tunnel volume fraction. Both *C* and *D* have extracellular volume fractions of about 20% (20.1% and 20.7%, respectively). (*E*) Schematic illustration showing the cylinder model of the paravascular space and solutes (filled circles) in the surrounding interstitial space. (*F*) Schematic illustration showing the pial surface model.

Kinney et al. [18] grouped the interstitial volume into tunnels or sheets. Sheets are the volumes between two adjacent membranes, typically 10 nm to 40 nm wide, tunnels are the wider, interconnected structures found at the junction of three or more cells, about 40 nm to 80 nm wide. In Fig. I.1*A* tunnels are colored in cyan, sheets in red. Kinney et al. [18] used different volume scaling procedures, some adding volume mainly to the tunnels, some adding volume to the sheets. We simulated interstitial bulk flow and computed the permeabilities from two different realizations of the EM reconstruction, both having approximately the same total interstitial volume fraction, but with different relative tunnel volume fractions. We also simulated bulk flow and permeability for smaller subvolumes with interstitial volume fractions up to 32.1 %.

Example sections from the two realizations are shown in Fig. I.1, *C* and *D*, where *C* has the smallest relative tunnel fraction (33 %), *D* has the largest (63 %). As described in Methods, the two tissue realizations were divided into 84 million and 25 million tetrahedrons, respectively, the smallest tetrahedrons with sides less than 1 nm, see Fig. I.1*B*. The flow and permeability were estimated by solving the Stokes equations in the FEniCS simulator [20] for a pressure gradient of 1 mmHg/mm applied between opposite sides of the tissue cube, assuming nonelastic and impermeable obstacles. The pressure gradient of 1 mmHg/mm is considered as an absolute upper estimate of the assumed pressure gradient within brain tissue (see Discussion), and the flow velocities and Péclet numbers shown here should therefore be considered as upper estimates. Note that there is a linear relationship between pressure gradient and flow velocity, implying that a pressure gradient different from the 1 mmHg/mm used here will change the velocities with the same factor. In contrast, the estimated permeabilities will be preserved.

Based on the estimated permeabilities from the EM reconstructions we created two simplified model systems to compare the effect of solute clearance by diffusion versus advection. In Fig. I.1, *E* and *F*, schematic illustrations of the two models are shown. Fig. I.1*E* illustrates clearance towards the paravascular space, Fig. I.1*F* illustrates clearance towards the pial surface. Three solutes with different diffusion constants were studied, the smallest corresponding to the effective diffusion coefficient of potassium ions ($D^* = 77 \times 10^{-7}\,\mathrm{cm}^2/\mathrm{s}$ [12]), the medium sized corresponding to 3 kDa Texas Red Dextran ($D^* = 5.3 \times 10^{-7}\,\mathrm{cm}^2/\mathrm{s}$ [32]), while the largest had a diffusion constant corresponding to 70 kDa Dextran ($D^* = 0.84 \times 10^{-7}\,\mathrm{cm}^2/\mathrm{s}$ [32]).

I.1.1 Flow and permeability in reconstructed neuropil

The intrinsic hydrodynamic permeability, κ, is defined by Darcy's law, $q = -\frac{\kappa}{\mu}\nabla p$, which states that there is a proportionality between the flux, q (discharge per unit area, with units of length per time), and the pressure gradient, ∇p, with μ denoting the viscosity. For the geometry with the smallest tunnel fraction (Fig. I.1*C*) we estimated the permeability to be $10.9\,\mathrm{nm}^2$, $10.3\,\mathrm{nm}^2$ and $11.0\,\mathrm{nm}^2$ (mean $10.7\,\mathrm{nm}^2$) along the three orthogonal axes perpendicular to the sides of the rectangular tissue cuboid. For the geometry with a larger tunnel fraction

(Fig. I.1*D*) the permeability was estimated to be 16.6 nm^2, 14.4 nm^2 and 13.1 nm^2 (mean 14.7 nm^2) along the three orthogonal axes. Thus, the anisotropy was maximum 6 % for the geometry with a low tunnel fraction and maximum 26 % for the geometry with a high tunnel fraction.

The geometry with a high tunnel fraction had a 36 % higher mean permeability than the geometry with a lower tunnel fraction [18], even though the extracellular volume fraction was approximately the same. In Fig. I.2 the maximal velocities in *A*–C are substantially lower than the maximal velocities in panels *D*–*F*, where the former corresponds to the geometry with a low tunnel fraction and the latter corresponds to the geometry with a higher tunnel fraction. Further, the cross-sections show that the velocities are highest within the centers of the larger tunnels (Fig. I.2, *A* and *D*). For all plots we have assumed a pressure gradient of 1 mmHg/mm. This assumption should be considered as an upper estimate (see Discussion). The average extracellular velocities are 8.95 nm/s and 12.2 nm/s, corresponding to permeabilities of 10.7 nm^2 and 14.7 nm^2, respectively. Note, however, that our convergence tests (see Methods) revealed that the permeabilities and velocities may have been underestimated with as much as 30%. Thus, an upper estimate of the permeabilities would be 14 nm^2 and 19 nm^2, and corresponding mean velocities of 12 nm/s and 16 nm/s, respectively.

For both geometries it takes several hundred minutes before 50 % of the fluid has traveled more than 100 μm (Fig. I.2, *C* and *F*). For comparison, Xie et al. [35] show that 3 kDa Texas Red Dextran typically penetrated 100 μm in about 20 minutes in sleeping and in anesthetized mice, a much shorter time interval than what could have been achieved for advection based tracer penetration from the cortical surface. However, Xie et al. show that a substantial part of the tracer (administered intrathecally) first travels along vessels before it starts penetrating laterally into the interstitial space. Although this could explain the short time scale for tracer penetration seen in Xie et al., Figs. I.3–I.5 show that interstitial diffusion predominates over interstitial advection, also when the tracer originates from paravascular spaces. We find that diffusion is compatible with the time scale seen in the tracer experiments in Xie et al. (Fig. I.4), and the estimated permeabilities were too low to allow for any significant advection. Even when we simulated flow and permeabilities for subvolumes with a much larger extracellular volume fraction than would be realistic for any physiological situation, we still estimated permeabilities incompatible with tracer velocites from Xie et al. (subvolumes with extracellular volume fractions of 27.9 % and 32.1 % gave permeabilites of 33 nm^2 and 70 nm^2, respectively). Table I.1 shows that our estimated permeabilites are about two orders of magnitudes lower than what is typically found in the literature.

I.1.2 Advection versus diffusion

Using the above estimated permeabilities we found that the bulk flow velocities are low also when we assume an arterial source and a venous sink. In this model the vessels are assumed to be surrounded by a medium with homogeneous permeability and an extracellular volume fraction of 20 %. Fig. I.3 shows that

Table I.1: Comparison of permeabilities from the literature

A viscosity of 0.8 mPa s was assumed when the permeability was converted from hydraulic conductance.

Permeability (nm^2)	Reference
10 to 20	this study, geometry from [18]
1280	[31], estimated from [24]
2480	[31], estimated from [8]
1360	[31], estimated from [28]
720	[17]
4000	[7]
1600	[5]

except for the volume just outside the vessels, where the pressure gradient is steepest, the flow velocities would typically be less than 10 nm/s for our assumed pressure differences of 1 mmHg/mm, even for the permeability value from the geometry with the higher permeability.

The typical time scale for diffusion is much smaller than the time scale for advection and comparable to typical time scales seen in tracer recordings (Fig. I.4). Fig. I.4 shows clearance of an interstitial solute, i.e., we assume the concentration to be higher inside the parenchyma than at the pial surface or within the paravascular spaces. For concentration gradients in the opposite direction, as after intrathecal tracer infusion, the y-axes would be symmetrically inverted.

In Fig. I.4, A and B, we show the concentration profile of different substances at three time instances after we decrease the concentration by Δc at the boundary, which is either the paravascular space (A), or the pial surface (B). The light substance (green) with an effective diffusion constant corresponding to ions such as potassium, shows a prominent decay already after 5 seconds (broken line), even at distances as far as 100 μm from the vessel (A) or the cortical surface (B). For larger solutes diffusion takes much longer time. The red lines correspond to effective diffusion constants for 3 kDa Texas Red Dextran and the blue lines correspond to 70 kDa Dextran. However, even for 70 kDa Dextran the concentration is seen to be substantially reduced at a time scale of minutes, both around vessels (A) and as a function of distance from the cortical surface (B).

Diffusion is seen to reduce the concentration at a distance 100 μm from a vessel (Fig. I.4C) and 100 μm from the cortical surface (Fig. I.4D) substantially within an hour, even for the very heavy 70 kDa Dextran. Note that here we have only assumed efflux from one vessel. If more vessels were assumed, the concentrations would have been decreased substantially in Fig. I.4, A and C.

A more direct way to compare advection to diffusion is to compare the size of the advection term to the size of the diffusion term in the diffusion-convection equation by use of the the Péclet number (Pe), $\mathrm{Pe} = Lv/D^*$. This number is plotted for a series of solutes of different sizes in Fig. I.5. L is the typical size of the system, here taken to be the average distance between the surfaces of an arteriole-venule pair (238 μm), $v = 12.2$ nm/s is the advection velocity, here

taken to be the average velocity for the geometry with the highest permeability, and D^* is the effective diffusion constant of the different solutes in brain tissue. For Pe $\ll 1$ diffusion predominates, and in Fig. I.5 we see that even for the most heavy solutes, such as 70 kDa Dextran and ovalbumin, the Péclet number is substantially lower than one for the assumed pressure gradient of 1 mmHg/mm. Hence, diffusion predominates over advection, even for large molecules. For illustrational purposes we have added a pressure gradient of 2 mmHg/mm in Fig. I.5. Even for this pressure gradient most solutes have Péclet numbers well below one, although 70 kDa Dextran is seen to be approaching one (0.69).

I.2 Discussion

Surprisingly little is known about the mechanisms that govern the movement of molecules between brain cells. As the brain interstitial space is particularly narrow and tortuous, the complexity of this space has so far defied any attempts to realistically simulate solute movement within it. New opportunities for such simulations arose with the recent generation of 3D representations that faithfully describe the interstitial space [18]. Here we take advantage of these representations—and of recent development in computer hardware, processing power and software tools—to show that interstitial permeability is much lower and solute movement is much more constrained than previously assumed. Movement occurs by diffusion rather than being driven by bulk flow. This conclusion holds even in simulations with an abnormally high extracellular volume fraction (32.1 %).

The existence of a bulk flow of interstitial fluid has been debated for decades. Syková and Nicholson [32] concluded that such flow is restricted to the paravascular spaces rather than taking place throughout the ECS. However, on introducing the glymphatic concept Nedergaard and coworkers expressed the view that waste products are cleared by bulk flow through the interstitium. The present data compel us to revise the concept of the glymphatic system [15]. The key idea embedded in the term "glymphatic" is that waste is cleared from the brain by a glia-dependent mechanism, analogous to the lymphatic system in other organs [25, 26]. The critical experiment in support of this concept showed that amyloid β and other compounds were cleared less efficiently in AQP4-deficient mice than in wild-type [15]. AQP4 is strongly expressed in glia, more specifically in the astrocytic endfeet that surround brain vessels [27]. In terms of involvement of glia in waste removal the glymphatic concept is not challenged by our results. However, according to the glymphatic concept as originally described, paraarterial and paravenous spaces connect through convective flow in the neuropil. Our findings strongly suggest that this is untenable and that diffusion prevails in the interstitial space.

The present findings have pronounced implications for future research. The idea of there being an advection in the interstitial space directed attention to mechanisms underlying the control of extracellular volume and hydrostatic pressure gradients within brain tissue. On the other hand, if diffusion predominates—

as the present data suggest—future research efforts should aim at understanding how concentration gradients are established and maintained. Attention should then be directed to transport processes at the brain-blood interface, and to the nature and scale of advection along brain vessels. Paravascular advection is required to effectively maintain the concentration gradients that are prerequisites for diffusion through neuropil. AQP4 could facilitate paravascular advection, which in turn could explain why appropriate clearance may depend on the presence of this water channel.

The major premise for our conclusion is that the permeability of the interstitial space is so low that it effectively precludes advection through brain neuropil at realistic pressure gradients. The question is why our permeability estimates differ by order of magnitudes from those of previous studies. The other high permeabilities reported in Table I.1 are either based on simultaneous fluid infusion and pressure recordings [5, 7, 8, 24, 28, 31], or simulated by the use of simplified geometries [17]. Combined infusion and pressure recordings may lead to overestimated permeabilities due to tissue displacement and because fluid is escaping along high-permeability paths such as the paravascular spaces. Simulations are, on the other hand, critically dependent on the right dimensions of the interstitial space. For a given extracellular volume fraction the dimension of the extracellular space is a function of the obstacle size. The 3D reconstructions used in our simulations indicate a mean obstacle size of far less than 1 μm, and we end up with a relatively low permeability. By comparison, Jin et al. [17] assume an extracellular volume fraction similar to what is used here (20 %), but their simulations are based on artificially created 2D obstacles with a much larger mean obstacle size of 5 μm, and they arrive at a much larger permeability. However, even with such a large obstacle size they end up with a conclusion that is in line with ours: diffusion predominates when it comes to solute movement through the extracellular space. The same conclusion is also reached by Asgari et al. [5], using a simplified model. We show that this conclusion holds in a realistic 3D model and even for very large molecules such as ovalbumin.

As stated above, our simulation precludes advection through brain neuropil at realistic pressure gradients. What are realistic pressure gradients in this context? Through a cardiac cycle the peak-to-peak intracranial pressure amounts to less than 10 mmHg. However, the pulsatility is almost synchronous throughout the brain, and the minute differences seen in simultaneous recordings of intracranial pressures give rise to much smaller pressure gradients than the 1 mmHg/mm assumed here [11]. Pressure gradients within the brain and/or CSF is typically reported to be less than 0.01 mmHg/mm [4, 11]. Thus, our assumption that these gradients are 1 mmHg/mm should be seen as an upper estimate. Unfortunately, technologies are not available for direct measurements of pressure gradients between neighboring brain vessels, i.e, those gradients that drive advection, if any, through brain neuropil.

We conclude that diffusion through interstitial space combined with paravascular advection substitutes for the lymphatic drainage system in other organs. This has profound implications for our understanding of how waste products are cleared from brain and of how drugs, nutrients and signal molecules permeate

brain neuropil.

I.3 Methods

I.3.1 Finite element simulations

ISF is assumed to be incompressible Newtonian fluid, and the flow is modeled by the Stokes equations: $\mu\nabla^2\mathbf{v}+\nabla p = \mathbf{0}$ and $\nabla\cdot\mathbf{v} = \mathbf{0}$. Here $\mathbf{v}$ is velocity vector and p the pressure within ISF. The viscosity μ of the ISF is assumed to be 0.8 mPa s. As we use linear elements for both velocity and pressure, a stabilization term $0.2\,h^2\nabla^2 p$ is added to the second equation [14], with h denoting the element size. To drive flow, a pressure gradient of 1 mmHg/mm is applied in one direction. This is enforced as a Neumann boundary condition, i.e., constant pressure at the inflow and outflow surfaces. On the remaining exterior boundary we used a symmetry assumption ($\mathbf{v}\cdot\mathbf{n}_\perp = 0$, where $\mathbf{n}_\perp$ is the unit normal vector for the outer surface), and at the interior cell surface boundaries we use the no-slip condition, $\mathbf{v} = 0$.

The resulting PDE system is solved in FEniCS [20]. Post processing of the data, including computation of total flux and visualization, was carried out using Paraview [3].

The meshes on which the computations are performed are generated using the CGAL backend of FEniCS' mesh generation submodule `mshr`. For the largest simulation the mesh consisted of 84 million tetrahedrons and more than 1000 CPU hours was needed to simulate the flow (279 minutes on 224 Intel E5-2670 processors).

A highly detailed mesh is required to adequately resolve the intricate geometry of the interstitial space. In order to test whether the mesh is sufficiently fine, the ideal test would be to refine it once, repeat all computations and check that the results do not significantly change. However, because the number of mesh elements is already very large, this is not computationally feasible.

Instead we used a less strict test. For both geometries we performed the simulation on a smaller volume measuring $0.52\,\mu\text{m} \times 0.52\,\mu\text{m} \times 0.45\,\mu\text{m}$. For the default simulation we found the baseline permeability for these subvolumes by applying a mesh size similar to what was used in the full simulations. We then refined the mesh by increasing the number of tetrahedrons 7 times. After three refinements (resulting in a total of 343 times as many volume elements as the original mesh) we reached the upper limit for what was computational feasible. For the reduced geometry with a high tunnel fraction (Fig. I.1*D*) the extracellular volume fraction was 32.1 % and refinements gave the following permeability series for the subvolume, listed from the default value to the most refined value: $54.26\,\text{nm}^2$, $61.91\,\text{nm}^2$, $65.59\,\text{nm}^2$ and $67.74\,\text{nm}^2$. For the reduced geometry with a low tunnel fraction (Fig. I.1*C*) the extracellular volume fraction was 27.9 % and the corresponding series was $25.22\,\text{nm}^2$, $29.49\,\text{nm}^2$, $31.24\,\text{nm}^2$ and $32.14\,\text{nm}^2$. These trends predict that the series should converge for about $70\,\text{nm}^2$ and $33\,\text{nm}^2$, respectively. This is seen by fitting each series to a permeability model $\kappa = \kappa_\infty - a/(x-b)$, where κ_∞ is the asymptotic permeability for small

mesh sizes, and x is the reciprocal of the mesh size. As the volumes of the tetrahedrons are reduced by 7 times for each refinement, the typical mesh size is correspondingly reduced by $7^{1/3} \approx 1.9$ times. Thus, $x = x_j = 7^{j/3}$ denotes the x-values fitted for $j = 0, 1, 2, 3$, with corresponding permeabilities κ_j at the y-axes. κ_∞, a and b are parameters fitted in the model. A non-linear least square fit gives the asymptotic permeabilities $69.8\,\mathrm{nm}^2$ and $33.0\,\mathrm{nm}^2$, corresponding to a 28.7% and a 30.8% increase from the baseline permeabilities, respectively. When the mesh size becomes infinitesimally small we therefore expect the permeabilities to increase by about 30 % also for the full geometries.

I.3.2 Interstitial flow from arteriole to venule

The interstitial flow velocities for the arteriole-venule geometry plotted in Fig. I.3 were found analytically, see Supplemental Material.

I.3.3 Diffusion from the cortical surface

If we assume the cortical surface to be perfectly planar and the lateral concentration to be constant, the one dimensional diffusion equation describes the system. A constant concentration $c(z,t) = c_0$ at time $t < 0$ followed by an abrupt decrease in concentration at the boundary (cortical surface), $c(z = 0, t \geq 0) = 0$, has the solution $c(z,t) = c_0 \,\mathrm{erf}(z/\sqrt{4D^*t})$, where $\mathrm{erf}(x)$ is the standard error function, z is distance from the cortical surface and D^* is the effective diffusion coefficient.

I.3.4 Diffusion from the paravascular space

The diffusion equation was solved in polar coordinates with a commercial software package (MATLAB 8.6, R2015b, The MathWorks Inc., Natick, MA, 2015). The outer surface of the paravascular space was assumed to have the shape of an infinitely long cylinder with an outer radius a, and the solute was allowed to diffuse throughout the interstitial space defined by $a < r < R$. Similarly to the planar diffusion, a constant concentration $c(r,t) = c_0$ was assumed at time $t < 0$ followed by an abrupt decrease in concentration at the boundary (paravascular space), $c(r = a, t \geq 0) = 0$. The concentration was kept constant at the distal boundary $c(r = R, t > 0) = c_0$, where R is much larger than the distances plotted in Fig. I.4 ($R \gg 0.2$ mm).

I.3.5 Acknowledgements

This work was funded by the Research Council of Norway (grants #226696 and #240476); the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 601055, the Molecular Life Science Initiative at the University of Oslo, Simula-UCSD-University of Oslo Research and PhD training (SUURPh) program; and the Letten Foundation. The simulations were run on the Abel Cluster (project

NN9279K), owned by the University of Oslo and the Norwegian metacenter for High Performance Computing (NOTUR), and operated by the Department for Research Computing at USIT, the University of Oslo IT-department, http://www.hpc.uio.no/. An approximate total of 60 thousand CPU hours were spent on the simulations for this study.

References

[1] Abbott, N. J. "Evidence for bulk flow of brain interstitial fluid: significance for physiology and pathology." In: *Neurochemistry international* vol. 45, no. 4 (Sept. 2004), pp. 545–552.

[2] Adams, D. L. et al. "Vascular Supply of the Cerebral Cortex is Specialized for Cell Layers but Not Columns." In: *Cerebral Cortex* vol. 25, no. 10 (Oct. 2015), pp. 3673–3681.

[3] Ahrens, J., Geveci, B., and Law, C. "ParaView: An End-User Tool for Large Data Visualization". In: *Visualization Handbook, Elsevier, ISBN-13: 978-0123875822* (Jan. 2005).

[4] Alperin, N. J. et al. "MR-Intracranial pressure (ICP): a method to measure intracranial elastance and pressure noninvasively by means of MR imaging: baboon and human study." In: *Radiology* vol. 217, no. 3 (Dec. 2000), pp. 877–885.

[5] Asgari, M., Zélicourt, D. de, and Kurtcuoglu, V. "Glymphatic solute transport does not require bulk flow." In: *Scientific Reports* vol. 6, no. 1 (Dec. 2016), p. 38635.

[6] Aspelund, A. et al. "A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules". In: *Journal of Experimental Medicine* vol. 212, no. 7 (June 2015), pp. 991–999.

[7] Basser, P. J. "Interstitial pressure, volume, and flow during infusion into brain tissue." In: *Microvascular Research* vol. 44, no. 2 (Sept. 1992), pp. 143–165.

[8] Bobo, R. H. et al. "Convection-enhanced delivery of macromolecules in the brain." In: *Proceedings of the National Academy of Sciences of the United States of America* vol. 91, no. 6 (Mar. 1994), pp. 2076–2080.

[9] Carare, R. O. et al. "Solutes, but not cells, drain from the brain parenchyma along basement membranes of capillaries and arteries: significance for cerebral amyloid angiopathy and neuroimmunology." In: *Neuropathology and applied neurobiology* vol. 34, no. 2 (Apr. 2008), pp. 131–144.

[10] Cserr, H. F. and Ostrach, L. H. "Bulk flow of interstitial fluid after intracranial injection of blue dextran 2000." In: *Experimental neurology* vol. 45, no. 1 (Oct. 1974), pp. 50–60.

[11] Eide, P. K. and Sæhle, T. "Is ventriculomegaly in idiopathic normal pressure hydrocephalus associated with a transmantle gradient in pulsatile intracranial pressure?" In: *Acta Neurochirurgica* vol. 152, no. 6 (Feb. 2010), pp. 989–995.

[12] Halnes, G. et al. "Electrodiffusive Model for Astrocytic and Neuronal Ion Concentration Dynamics". In: *PLoS Computational Biology* vol. 9, no. 12 (Dec. 2013), e1003386.

[13] Hladky, S. B. and Barrand, M. A. "Mechanisms of fluid movement into, through and out of the brain: evaluation of the evidence". In: *Fluids and Barriers of the CNS* (2014).

[14] Hughes, T. J. R., Franca, L. P., and Balestra, M. "A new finite element formulation for computational fluid dynamics: V. Circumventing the babuška-brezzi condition: a stable Petrov-Galerkin formulation of the stokes problem accommodating equal-order interpolations". In: *Computer Methods in Applied Mechanics and Engineering* vol. 59, no. 1 (Nov. 1986), pp. 85–99.

[15] Iliff, J. J. et al. "A Paravascular Pathway Facilitates CSF Flow Through the Brain Parenchyma and the Clearance of Interstitial Solutes, Including Amyloid β." In: *Science Translational Medicine* vol. 4, no. 147 (Aug. 2012), 147ra111–147ra111.

[16] Iliff, J. J. et al. "Cerebral arterial pulsation drives paravascular CSF-interstitial fluid exchange in the murine brain." In: *Journal of Neuroscience* vol. 33, no. 46 (Nov. 2013), pp. 18190–18199.

[17] Jin, B.-J., Smith, A. J., and Verkman, A. S. "Spatial model of convective solute transport in brain extracellular space does not support a "glymphatic" mechanism." In: *The Journal of general physiology* vol. 148, no. 6 (Dec. 2016), pp. 489–501.

[18] Kinney, J. P. et al. "Extracellular sheets and tunnels modulate glutamate diffusion in hippocampal neuropil". In: *The Journal of comparative neurology* vol. 521, no. 2 (Dec. 2012), pp. 448–464.

[19] Kiviniemi, V. et al. "Ultra-fast magnetic resonance encephalography of physiological brain activity - Glymphatic pulsation mechanisms?" In: *Journal of Cerebral Blood Flow and Metabolism* vol. 36, no. 6 (May 2016), pp. 1033–1045.

[20] Logg, A., Mardal, K.-A., and Wells, G. *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*. Ed. by Logg, A., Mardal, K.-A., and Wells, G. Vol. 84. Lecture Notes in Computational Science and Engineering. Berlin: Springer, 2012.

[21] Louveau, A., Da Mesquita, S., and Kipnis, J. "Lymphatics in Neurological Disorders: A Neuro-Lympho-Vascular Component of Multiple Sclerosis and Alzheimer's Disease?" In: *Neuron* vol. 91, no. 5 (Sept. 2016), pp. 957–973.

[22] Louveau, A. et al. “Structural and functional features of central nervous system lymphatic vessels.” In: *Nature* vol. 523, no. 7560 (July 2015), pp. 337–341.

[23] Morris, A. W. J. et al. “Vascular basement membranes as pathways for the passage of fluid into and out of the brain.” In: *Acta neuropathologica* vol. 131, no. 5 (May 2016), pp. 725–736.

[24] Morrison, P. F. et al. “High-flow microinfusion: tissue penetration and pharmacodynamics.” In: *The American journal of physiology* vol. 266, no. 1 Pt 2 (Jan. 1994), R292–305.

[25] Nagelhus, E. A. and Ottersen, O. P. “Physiological roles of aquaporin-4 in brain.” In: *Physiological reviews* vol. 93, no. 4 (Oct. 2013), pp. 1543–1562.

[26] Nedergaard, M. “Garbage Truck of the Brain”. In: *Science* vol. 340, no. 6140 (June 2013), pp. 1529–1530.

[27] Nielsen, S. et al. “Specialized membrane domains for water transport in glial cells: high-resolution immunogold cytochemistry of aquaporin-4 in rat brain.” In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* vol. 17, no. 1 (Jan. 1997), pp. 171–180.

[28] Prabhu, S. S. et al. “Distribution of macromolecular dyes in brain using positive pressure infusion: a model for direct controlled delivery of therapeutic agents.” In: *Surgical neurology* vol. 50, no. 4 (Oct. 1998), 367–75–discussion 375.

[29] Rennels, M. L. et al. “Evidence for a ’paravascular’ fluid circulation in the mammalian central nervous system, provided by the rapid distribution of tracer protein throughout the brain from the subarachnoid space.” In: *Brain research* vol. 326, no. 1 (Feb. 1985), pp. 47–63.

[30] Smith, A. J., Jin, B.-J., and Verkman, A. S. “Muddying the water in brain edema?” In: *Trends in neurosciences* vol. 38, no. 6 (June 2015), pp. 331–332.

[31] Smith, J. H. and Humphrey, J. A. C. “Interstitial transport and transvascular fluid exchange during infusion into brain and tumor tissue”. In: *Microvascular Research* vol. 73, no. 1 (Jan. 2007), pp. 58–73.

[32] Syková, E. and Nicholson, C. “Diffusion in brain extracellular space.” In: *Physiological reviews* vol. 88, no. 4 (Oct. 2008), pp. 1277–1340.

[33] Tarasoff-Conway, J. M. et al. “Clearance systems in the brain-implications for Alzheimer disease.” In: *Nature reviews. Neurology* vol. 11, no. 8 (Aug. 2015), pp. 457–470.

[34] Thrane, A. S. et al. “Filtering the muddied waters of brain edema.” In: *Trends in neurosciences* vol. 38, no. 6 (June 2015), pp. 333–335.

[35] Xie, L. et al. “Sleep drives metabolite clearance from the adult brain.” In: *Science* vol. 342, no. 6156 (Oct. 2013), pp. 373–377.

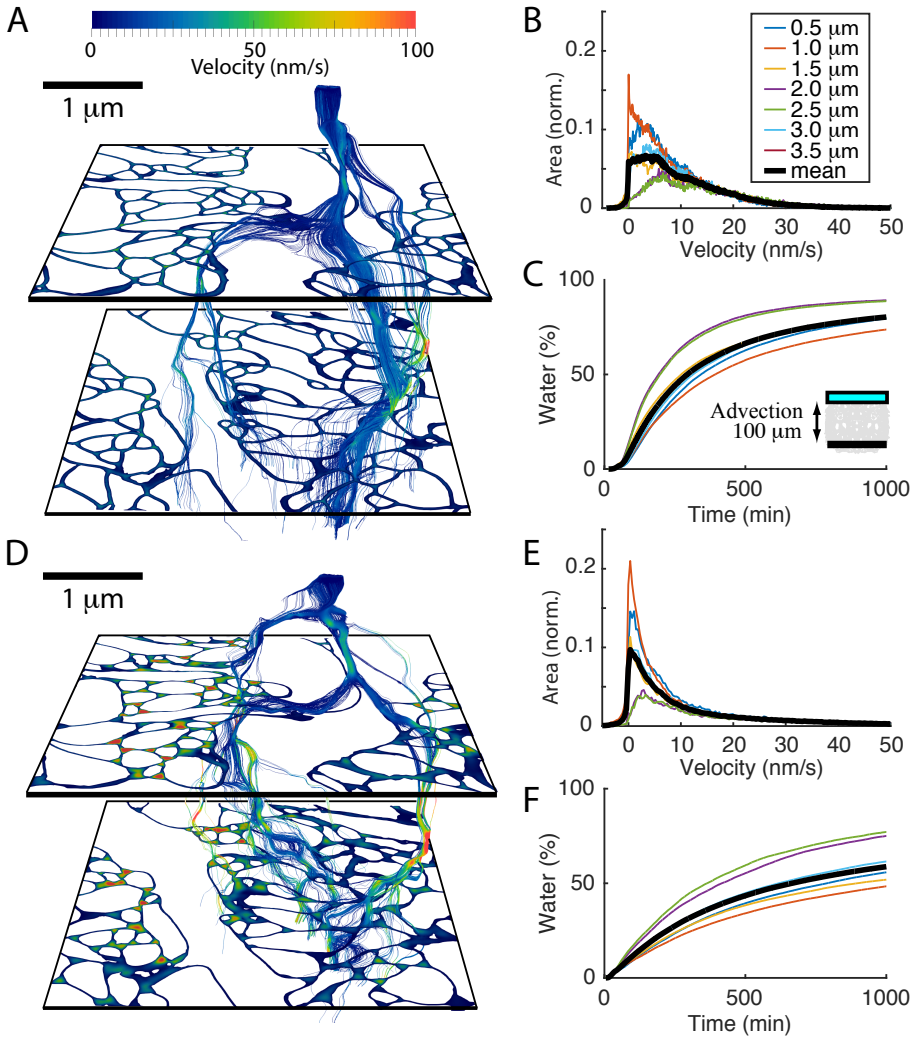

Figure I.2: Bulk flow velocity through the EM reconstruction from Kinney et al. [18]. A pressure gradient of 1 mmHg/mm is applied in the vertical (z) direction. (A) The geometry with a low tunnel volume fraction. The cross sections are at depth $z = 1.5\,\mu$m and $z = 3.5\,\mu$m. For clarity only streamlines originating from a small circle with radius $0.1\,\mu$m at $z = 0$ are shown. (B) Distribution of the z-component of flow velocities through different cross-sectional extracellular areas of the geometry in A, with the corresponding depth of the plane expressed in the legend. All traces are normalized to the mean extracellular cross sectional area. The mean distribution is shown in black. (C) The percentage of water which has reached $100\,\mu$m as a function of time (see inset), assuming each streamline to be straight, along the z-axis and with a constant velocity given by the velocity distribution in B. (D–F) Same as A–C for the EM reconstruction with a higher tunnel volume fraction, but approximately the same extracellular volume fraction.

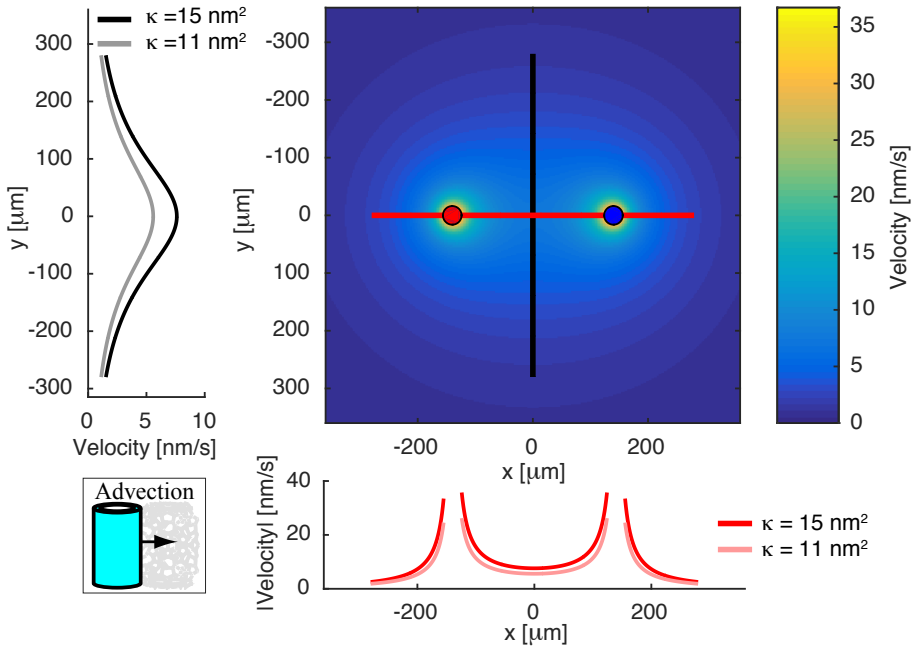

Figure I.3: Color plot showing velocity for the bulk flow from arteriole (red, filled circle) to venule (blue, filled circle) for the highest permeability, $\kappa = 14.69\,\mathrm{nm}^2$, assumed viscosity $\mu = 0.8\,\mathrm{mPa\,s}$ and extracellular volume fraction of 20 %. Diameter is 30 μm for both arteriole and venule, their center to center distance is 280 μm [2, 17]. The line plots in red/pink and black/gray correspond to the absolute value of the velocity profiles along the red (x-axis) and black (y-axis) lines in the color plot, and the two colors correspond to the two different permeabilities derived from the geometries with high tunnel volume fraction and low tunnel fraction. The pressure difference between the two vessels surfaces facing each other is 1 mmHg/mm. Lower left inset illustrates the cylindrical geometry of the vessels.

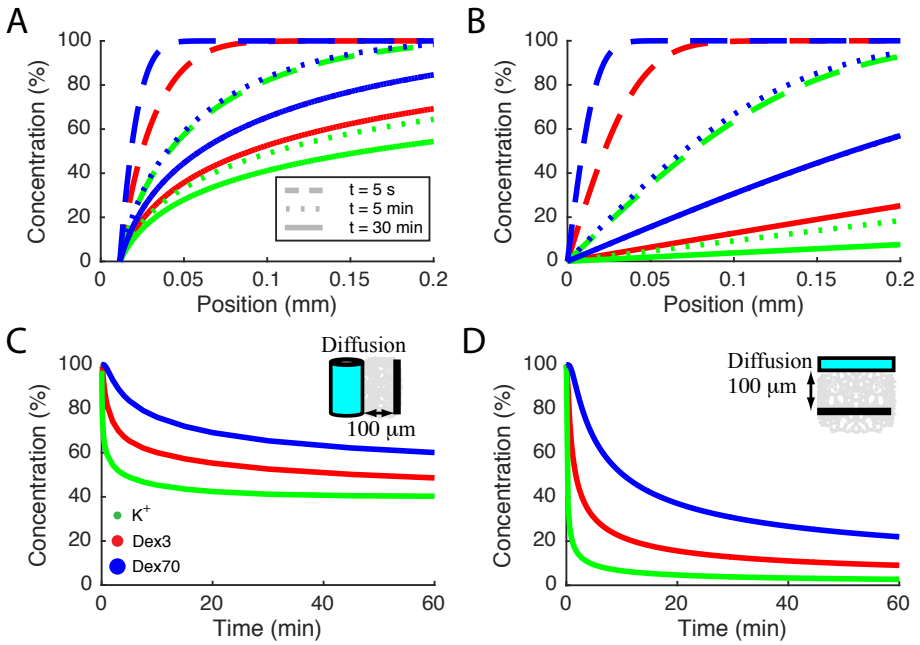

Figure I.4: Diffusion from neuropil towards (A and C) a cylindric vessel (see inset in C) and (B and D) the cortical surface (see inset in D). At time $t = 0$ the solute is assumed to be evenly spread throughout the interstitial space, and the cortical surface/cylinder is assumed to have zero concentration of the solute. The different colors correspond to effective diffusion coefficients for potassium ions (green), 3 kDa Texas Red Dextran (red) and 70 kDa dextran (blue). (A) Concentration profile around a vessel for three time instances. (B) Concentration profile below the cortical surface for three time instances. (C) Concentration of the three solutes as a function of time at a distance 100 μm from the cylinder center. (D) Concentration of the three solutes as a function of time 100 μm below the cortical surface.

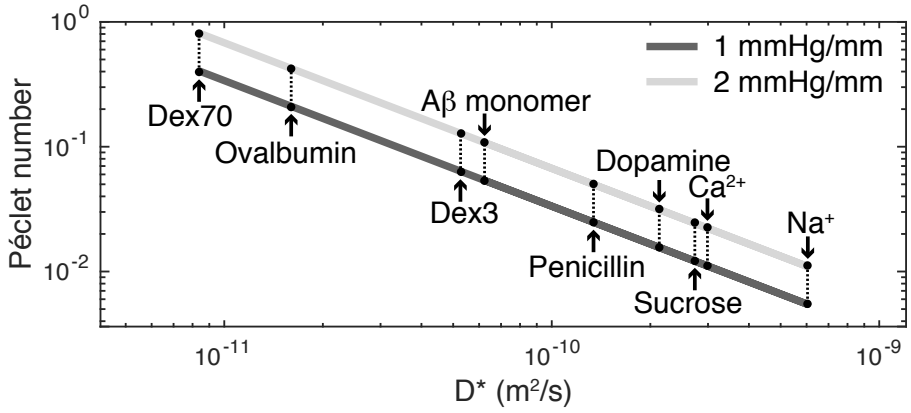

Figure I.5: Péclet numbers. Effective diffusion coefficients (D^*) from Syková and Nicholson [32].

Paper II

Sub-voxel perfusion modeling in terms of coupled 3d-1d problem

Karl Erik Holter, Miroslav Kuchta, Kent-André Mardal

arXiv: 1803.04896. *ENUMATH 2017 Proceedings.*

Abstract

We study perfusion by a multiscale model coupling diffusion in the tissue and diffusion along the one-dimensional segments representing the vasculature. We propose a block-diagonal preconditioner for the model equations and demonstrate its robustness by numerical experiments. We compare our model to a macroscale model by [P. Tofts, Modelling in DCE MRI, 2012].

II.1 Introduction

The micro-circulation is altered in diseases such as cancer and Alzheimer's disease, as demonstrated with modern perfusion MRI. In cancer, the so-called enhanced permeability and retention (EPR) effect describes the fact that the smaller vessels in a tumor are leaky, highly permeable vessels that enable the tumor cells to grow quicker than normal cells.

In Alzheimer's disease (AD), the opposite is alleged to happen. According to [8], hypoperfusion is a precursor to AD, and the cause of the pathological cell-level changes occuring in AD. This could also explain why various kinds of heart disease are risk factors for AD, as changes in the blood pressure would affect the perfusion of the brain.

The vasculature (e.g. idealized as a system of pipes) and the surrounding tissue are clearly three-dimensional. However, the fact that in many applications the radii of the vessels are negligible compared to their lengths, permits reducing their governing equations to models prescribed on (one dimensional) curves where the physical radius R enters as a parameter. In this paper, we consider a coupled $3d$-$1d$ system

$$\begin{aligned} \frac{\partial u}{\partial t} &= D_\Omega \Delta_\Omega u + \delta_\Gamma \beta \lambda && \text{in } \Omega, \\ \frac{\partial \hat{u}}{\partial t} &= D_\Gamma \Delta_\Gamma \hat{u} - \beta \lambda && \text{in } \Gamma, \\ \lambda &= \beta(\Pi_R u - \hat{u}) && \text{in } \Gamma. \end{aligned} \tag{II.1}$$

Here, δ_Γ is the Dirac measure on Γ; u, $\hat{u}$ are the concentrations in the tissue domain Ω and the one-dimensional vasculature representation Γ; and D_Ω, D_Γ are conductivities on the respective domains. The last equation is then a generalized Starling's law; relating $\hat{u}$ and the value $\Pi_R u$ of u averaged over the idealized cylindrical vessel surface centered around Γ[1]. The equation thus represents a coupling between the domains.

Variants of the system (II.1) have been used to study coupling between tissue and vasculature flow in numerous applications. In [7], a steady state limit of the system is used to numerically investigate oxygen supply of the skeletal muscles. Finite differences were used for the discretization. Using the method of Green's functions, [17], [18] and [3] studied oxygen transport in the brain and tumors.

Transport of oxygen inside the brain was also investigated by [12] using the finite volume method (FVM) and [6] using the finite element method (FEM) for the $3d$ diffusion problem and FVM elsewhere. In these studies the $1d$ problem was transient. More recently, coupled models discretized entirely by FEM were applied to study cancer therapies, see e.g. [15] and references therein. The mathematical foundations of these works are rooted in the seminal contributions of [5], [4] where well-posedness of the following problem is analyzed

$$\begin{aligned} -\Delta_\Omega u + (\hat{u} - \Pi_R u)\delta_\Gamma &= f\delta_\Gamma \quad \text{in } \Omega \\ -\Delta_\Gamma \hat{u} - (\hat{u} - \Pi_R u) &= g \quad \text{in } \Gamma. \end{aligned} \tag{II.2}$$

The presence of the measure term in (II.2) requires the use of non-standard spaces in the analysis. In [5], the variational formulation of the problem is proven to be well-posed using weighted Sobolev spaces. In particular, the solution u is sought in $H^1_\alpha(\Omega)$, $\alpha > 0$ while the test functions v for the first equation in (II.2) are taken from $H^1_{-\alpha}(\Omega)$. With this choice, the right-hand side $\langle f\delta_\Gamma, v\rangle$ as well as the trace operator $v \mapsto v|_\Gamma$ are well defined, while the reduced regularity of u is sufficient to for the average $\Pi_R u$ to make sense. As shown in [4], use of FEM for the formulation in weighted spaces yields optimal rates if the computational mesh is gradually refined towards Γ (graded meshes).

Another approach to the analysis of (II.2) has recently been suggested in the numerical study [2]. Building on the analysis of [9] for the elliptic problem with a 0 dimensional Dirac right-hand side, the wellposedness of the problem was shown with trial spaces $W^{1,p}(\Omega)$, $p = 3 - \frac{d}{2}$ and test spaces $W^{1,q}(\Omega)$, $p^{-1} + q^{-1} = 1$, and quasi-optimal error estimates for FEM shown in the norms which excluded a fixed neighborhood of Γ of radius R.

In studying AD or EPR, the physical parameters may vary across several orders of magnitude while small or large time steps can be desirable depending on the time scales of interest. The solution algorithm for the employed model

[1] Let $x \in \Gamma$ and $C_R(x)$ be a circular crossection of the vessel surface with a plane $\left\{y \in \mathbb{R}^3, (y-x)\cdot \frac{\mathrm{d}\Gamma}{\mathrm{d}s}(x) = 0\right\}$ defined by the tangent vector of Γ at x. The surface avarage $\Pi_R u$ of u is then defined by

$$(\Pi_R u)(x) = |C_R(x)|^{-1} \int_{C_R(x)} u(y)\,\mathrm{d}y.$$

equations is thus required to be robust with respect to these parameters as well independent of the discretization. For (the transient version of) (II.2) the construction of such algorithms is complicated by the non-standard spaces on the domain Ω.

A potential remedy for the problem can be introduction of a Lagrange multiplier which enforces the coupling between the domains with the goal of confining the non-standard spaces to the smaller domain Γ. This idea has been used by [11] to analyze robust preconditioners for $2d$-$1d$ coupled problems based on operators in fractional Sobolev spaces, in particular, $(-\Delta_\Gamma)^{-\frac{1}{2}}$, while numerical experiments reported in [10] suggest that for suitable exponents $(-\Delta_\Gamma)^s$ defines a preconditioner for the Schur complement of a $3d$-$1d$ coupled system with a *trace* constraint[2]. We note that in both cases off-the-shelf methods were used as preconditioners for the operators on Ω.

The system (II.1) includes an additional variable λ for the coupling constraint, cf. (II.2). We therefore aim to apply the techiques of [10], [11] to construct a mesh-independent preconditioner for the problem, while the ideas of operator preconditioning [13] are used to ensure robustness with respect to the physical parameters and the time-stepping.

The rest of the paper is organized as follows. Section II.2 identifies the structure of the preconditioner. In §II.3 we discuss discretization of the proposed operator and report numerical experiments which demonstrate the robust properties. In §II.4 the system (II.1) is used to model tissue perfusion using a realistic geometry of the rat cortex. Conclusions are finally summarized in §II.5.

II.2 Preconditioner for the coupled problem

In the following we let Ω be a bounded domain in $\mathbb{R}^d$, $d = 2$ or 3 and Γ be a subdomain of Ω of dimension 1. By $L_2(D)$ we denote the space of square-integrable functions over D and $H^1(D)$ is the space of functions with first order derivatives in $L^2(D)$.

Discretizing (II.1) in time by backward-Euler discretization the problem to be solved at each temporal level is of the form

$$\mathcal{A}u = f$$

with

$$\mathcal{A} = \begin{bmatrix} I - kD_\Omega\Delta_\Omega & 0 & k\beta\Pi_R^* \\ 0 & I - kD_\Gamma\Delta_\Gamma & k\beta I \\ k\beta\Pi_R & k\beta I & -k. \end{bmatrix} \tag{II.3}$$

and k being the time step size. Note that in order to obtain a symmetric problem the operator $\mathcal{A}$ uses the adjoint Π_R^* of the *averaging* operator Π_R instead of the *trace*, cf. (II.1). The choice results in modeling error of order $\mathcal{O}(R)$.

[2] Note that in (II.2) and (II.1) the constraint/coupling is defined in terms of a surface averaging operator Π_R.

To motivate the structure of the preconditioner let us consider a 3×3 matrix

$$A = \begin{bmatrix} 1+\alpha_1 & 0 & \beta_1 \\ 0 & 1+\alpha_2 & \beta_2 \\ \beta_1 & \beta_2 & -\gamma \end{bmatrix}$$

where α_1, α_2, β_1, β_2, and γ are assumed to be positive. It can then be shown that with

$$B = \begin{bmatrix} (1+\alpha_1)^{-1} & 0 & 0 \\ 0 & (1+\alpha_2)^{-1} & 0 \\ 0 & 0 & (\gamma+\beta_1^2+\beta_2^2)^{-1} + (\gamma+\beta_1^2/\alpha_1+\beta_2^2/\alpha_2)^{-1} \end{bmatrix}$$

the condition number of BA is bounded independent of the parameters. With this in mind we propose that $\mathcal{A}$ can be preconditioned by a block-diagonal operator

$$\mathcal{B} = \begin{bmatrix} (I - kD_\Omega \Delta_\Omega)^{-1} & 0 & 0 \\ 0 & (I - kD_\Gamma \Delta_\Gamma)^{-1} & 0 \\ 0 & 0 & S \end{bmatrix}, \tag{II.4}$$

where

$$S = S_1^{-1} + S_2^{-1} \tag{II.5}$$

and

$$\begin{aligned} S_1 &= \gamma I + (k\beta)^2 \Pi_R \Pi_R^* + (k\beta)^2 I, \\ S_2 &= \gamma I + (k\beta)^2 \Pi_R (-kD_\Omega \Delta_\Omega)^{-1} \Pi_R^* + (k\beta)^2 (-kD_\Gamma \Delta_\Gamma)^{-1}. \end{aligned} \tag{II.6}$$

The operator $\mathcal{B}$ could be rigorously derived within operator preconditioning [13] as a Riesz map preconditioner for $\mathcal{A}$ viewed as an isomorphism from $V(\Omega) \times \hat{V}(\Gamma) \times Q(\Gamma)$ to its dual space. In [11] the framework was applied to a system of two elliptic problems coupled by a $2d$-$1d$ constraint. For (II.1) extension of the analysis to parabolic problems would be required. In [14] robust preconditioners for time-dependent Stokes problem were analyzed as operators between sums of (parameter) weighted Sobolev spaces. Similarly, the structure of (II.5) suggests that $Q = Q_1 + Q_2$ with Q_1, Q_2 being suitable interpolation spaces. However, here we shall not justify Q_1 and Q_2 (and the preconditioner $\mathcal{B}$) theoretically. Instead, Q_1, Q_2 are characterized and robustness of $\mathcal{B}$ is demonstrated by numerical experiments.

II.3 Discrete preconditioner

Considering (II.4), both $(I - kD_\Omega \Delta_\Omega)^{-1}$ and $(I - kD_\Gamma \Delta_\Gamma)^{-1}$ can be realized with off-the-shelf multilevel algorithms and the crucial question is thus how to construct S efficiently. Note that assembling S_1 and in particular S_2 might be too costly or even prohibitive, cf. $(-\Delta_\Omega)^{-1}$ in S_2. However, following (II.6) the preconditioner can be realized if operators spectrally equivalent to $\Pi_R(-\Delta_\Omega)^{-1}\Pi_R^*$ and $\Pi_R \Pi_R^*$ are known and if the inverse (action) of the resulting approximations to S_1 and S_2 is inexpensive to compute.

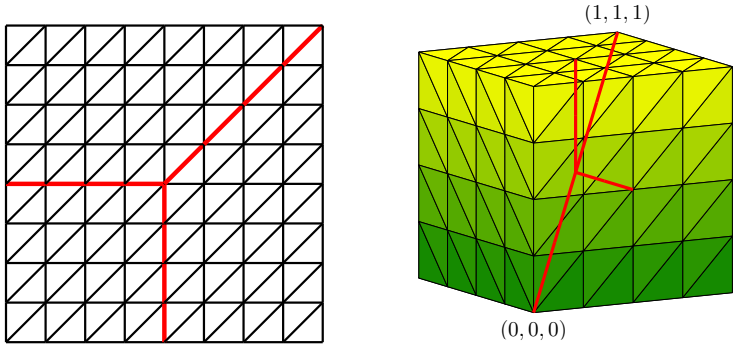

Figure II.1: Geometries used in preconditioning numerical experiments. The domain is $\Omega = [0,1]^d$ while in order to prevent symmetries Γ (pictured in red) always features a branching point. Triangulation of Γ is made up of edges of the cells that triangulate Ω.

II.3.1 Auxiliary operators

If $\Omega \subset \mathbb{R}^2$ and Π_R is understood as the trace operator, the mapping properties of trace as a bounded surjective operator $H^1(\Omega) \to H^{\frac{1}{2}}(\Gamma)$ can be used to show that $\Pi_R(-\Delta_\Omega)^{-1}\Pi_R^*$ is spectrally equivalent with $(-\Delta_\Gamma)^{-\frac{1}{2}}$. At the same time $\Pi_R\Pi_R^* = \Pi_R(-\Delta)^0_\Omega\Pi_R^*$ requires characterizing the space of traces of functions in $L^2(\Omega)$. We shall demonstrate by a numerical experiment that a spectrally equivalent operator is here provided by an operator $h^{-1}I_h$.

Let Ω_h, Γ_h be triangulations of Ω and Γ such that Γ_h consists of a subset of edges of the elements Ω_h, cf. Figure II.1. Further let V_h, Q_h be finite element spaces of continuous linear Lagrange elements on Ω_h and Γ_h respectively. Finally we consider the eigenvalue problem: Find $\lambda \in \mathbb{R}$, $u \in V_h$, $p \in Q_h$ such that

$$\begin{aligned} \int_\Omega uv + \int_\Gamma p\Pi_R v &= \lambda \int_\Omega uv \qquad && \forall v \in V_h, \\ \int_\Gamma q\Pi_R u &= \lambda \int_\Gamma h^{-1}pq \qquad && \forall q \in Q_h. \end{aligned} \tag{II.7}$$

Table II.1 shows the spectral condition number $\kappa = \max|\lambda| / \min|\lambda|$ of the linear systems (II.7). For all the considered resolutions h the value of κ is bounded.

As the mapping properties of $\Pi_R\Pi_R^*$ and $\Pi_R(-\Delta_\Omega)^{-1}\Pi_R^*$ in case $\Omega \subset \mathbb{R}^3$ are not trivially obtained from the continuous analysis, we again resort to finding the suitable approximations by numerical experiments. Similar to the two dimensional case, the first operator with Π_R having the constant radius $R = 0.02$ is found to be spectrally equivalent with $h^{-1}I_h$. Following [10], an approximation to $\Pi_R(-\Delta_\Omega)^{-1}\Pi_R^*$ is searched for as a suitable power $s < 0$ of $(-\Delta_\Gamma + I_\Gamma)$. More precisely, we look for the exponent yielding the most h-stable condition number

	$\Omega \subset \mathbb{R}^2$						$\Omega \subset \mathbb{R}^3$					
	$1/h$						$1/h$					
	32	64	128	256	512	1024	4	8	16	32	64	128
(II.7)	4.36	4.33	4.35	4.36	4.36	4.35	4.26	5.22	5.58	4.44	4.85	4.85
(II.8)	8.80	8.84	8.84	8.86	8.88	8.87	11.04	8.25	7.44	9.02	10.30	11.25

Table II.1: Spectral condition numbers of the eigenvalue problems related to approximations of $\Pi_R \Pi_R^*$(eq (II.7)) and $\Pi_R(-\Delta_\Omega)\Pi_R^*$(eq (II.8)). In the two-dimensional case (II.8) uses $s = -\frac{1}{2}$ in agreement with the mapping properties of the continuous trace operator. Results for $s = -0.55$ are reported in the three dimensional case. On the finest triangulation $\dim V_h \sim 10^6$ and $\dim Q_h \sim 10^3$ when $d = 2$ and $\dim Q_h \sim 10^2$ for $d = 3$.

of the eigenvalue problem: Find $\lambda \in \mathbb{R}$, $u \in V_h$, $p \in Q_h$ such that

$$\begin{aligned} \int_\Omega \nabla u \cdot \nabla v + \int_\Gamma p \Pi_R v &= \lambda \int_\Omega \nabla u \nabla v + uv \qquad && \forall v \in V_h, \\ \int_\Gamma q \Pi_R u &= \lambda \int_\Gamma p(-\Delta_\Gamma + I_\Gamma)^s q \qquad && \forall q \in Q_h. \end{aligned} \tag{II.8}$$

In (II.8) the powers are computed using the spectral decomposition of the operator $-\Delta_\Gamma + I_\Gamma$.

We shall not present here the results for the entire optimization problem and only report on the optimum which is found to be $s = -0.55$. For this value Table II.1 shows the history of the condition numbers of system (II.8) using again $R = 0.02$. There is a slight growth by a constant increment on the finer meshes, however, the final condition number is comparable with that obtained on the coarsest mesh. We note that we have not investigated the behaviour of the exponent on different curves or with variable radius. This subject is left for future works.

II.3.2 Discrete preconditioner for the coupled problem

Applying the proposed preconditioner (II.4) of the coupled $3d$-$1d$ problem (II.3) requires evaluating the inverse of operators S_1 and S_2 in (II.6). The former is readily computed since, due to the suggested equivalence of $\Pi_R \Pi_R^*$ and $h^{-1} I_h$, the matrix representation of S_1 is essentialy a rescaled mass matrix. For S_2 we show that if $(-\Delta_\Gamma + I_\Gamma)^s$ is computed from the spectral decomposition then the inverse S_2^{-1} can be computed in a closed form.

Let A, M be the $n \times n$ matrix representations of Galerkin approximations of $-\Delta_\Gamma + I_\Gamma$ and I in the space $Q_h \subset H^1(\Gamma)$. Following [11] the matrix representation of $(-\Delta_\Gamma + I_\Gamma)_h^s$ is $H_s = MU\Lambda^s(MU)'$ where the matrices U, Λ solve the generalized eigenvalue problem $AU = MU\Lambda$ such that $U'MU = I$. Using H_s it is easily established that the matrix representation of S_2 is

$$\gamma H_0 + (k\beta)^2 (kD_\Omega)^{-1} H_{-\frac{1}{2}} + (k\beta)^2 (kD_\Gamma)^{-1} H_{-1}.$$

As $H_s^{-1} = U\Lambda^{-s}U'$ the inverse of the S_2 matrix is given by

$$U\left(\gamma\Lambda^0 + (k\beta)^2(kD_\Omega)^{-1}\Lambda^{-\frac{1}{2}} + (k\beta)^2(kD_\Gamma)^{-1}\Lambda^{-1}\right)^{-1}U'. \qquad \text{(II.9)}$$

Using the spectral decomposition, the cost of setting up the preconditioner S is determined by the cost of solving the generalized eigenvalue problem for U and Λ. This practically limits the construction to systems where $\dim Q_h \sim 10^3$. However, for the problems considered further, this limitation does not present an issue. In particular, the preconditioner can be setup on the discretization of vasculature of the cortex tissue used in §II.4 which contains approximately twenty thousand vertices.

With a discrete approximation of S we can finally address the question of $\mathcal{B}$ being a robust preconditioner for (II.3). Motivated by (II.1) we do not vary all the paramaters, and instead, set $\gamma = 1$ in $\mathcal{A}$. Morever, the conductivity on Ω is taken as unity and only variable $D_\Gamma > 1$ is considered mimicking the expected faster propagation along the one-dimensional domain. Finally, the time step k and the coupling constant β shall take values between 10^{-4} and 10^{-8}.

For a fixed choice of parameters D_Γ, β, k, the preconditioned problem $\mathcal{B}\mathcal{A}x = \mathcal{B}f$ is considered on geometries from Figure II.1 and discretized with continuous linear Lagrange elements. The resulting linear system is then solved with the MinRes method where the iterations are stopped once the preconditioned residual norm is less than 10^{-10} in magnitude.

The observed iteration counts are reported in Table II.2. For both $2d$-$1d$ and $3d$-$1d$ coupled problems the iterations can be seen to be bounded in the discretization parameter. Moreover the preconditioner performs almost uniformly in the considered ranges of D_Γ and β while there is a clear boundedness in k as well. We note that these conclusions are not significantly altered if the ranges for k and β are extended to 1.

Having demonstrated the numerical stability of our model, we next test it on the same problem as [19], namely a bloodborne tracer perfusing and later being cleared from tissue. While [19] considers this problem on a macroscopic scale, we model it on the micro-scale, where individual blood vessels can be resolved as part of our $1d$ domain.

II.4 Perfusion experiment

In [16], a 0.7mm × 0.7mm × 0.7mm piece of mouse brain microvasculature was imaged using two-photon microscopy. To obtain a realistic geometry for our model, we used this data to generate a $3d$ mesh of the extravascular space in which vessel segments corresponded to $1d$ mesh edges. The radius of the blood vessels is used as the radius R in the definition of the averaging operator Π_R, and ranged between 1 and 15 micron.

To model a small region of tissue being perfused by a bloodborne tracer, we use initial conditions of $u, \hat{u} = 0$, and a boundary condition of $\hat{u} = 1\ \frac{\text{mol}}{\text{L}}$ on the part of the boundary corresponding to inlet vessels. To model clearance, the

D_Γ	β	k	$2d$-$1d$ $1/h$						$3d$-$1d$ $1/h$					
			32	64	128	256	512	1024	4	8	16	32	64	128
10^0	10^{-8}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	16	11	12	13	13	14	12	16	18	16	13	14
	10^{-6}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	13	14	12	16	18	16	13	14
	10^{-4}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	16	11	12	13	13	14	12	16	18	16	13	14
10^2	10^{-8}	10^{-8}	11	10	10	9	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	13	13	11	16	18	13	13	14
	10^{-6}	10^{-8}	11	10	10	9	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	13	13	11	16	18	13	13	14
	10^{-4}	10^{-8}	11	10	10	9	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	13	13	12	16	18	13	13	14
10^4	10^{-8}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	12	10	11	16	18	13	13	14
	10^{-6}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	12	10	11	16	18	13	13	14
	10^{-4}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	15	11	12	13	12	10	11	16	18	13	13	14
10^6	10^{-8}	10^{-8}	11	10	10	8	7	7	8	13	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	14	10	10	9	8	5	11	15	18	13	11	11
	10^{-6}	10^{-8}	11	10	10	8	7	7	8	12	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	14	10	10	9	8	5	11	15	18	13	10	11
	10^{-4}	10^{-8}	11	10	10	8	7	7	8	12	12	11	11	10
		10^{-6}	15	13	12	10	8	8	10	14	16	16	15	13
		10^{-4}	14	10	10	9	8	5	11	16	18	13	10	11

Table II.2: Number of iterations of MinRes method on (II.3) using (II.4) as preconditioner with S approximated using (II.9). (Left) $2d$-$1d$ coupled problem and (right) $3d$-$1d$ coupled problem from Figure II.1 are considered.

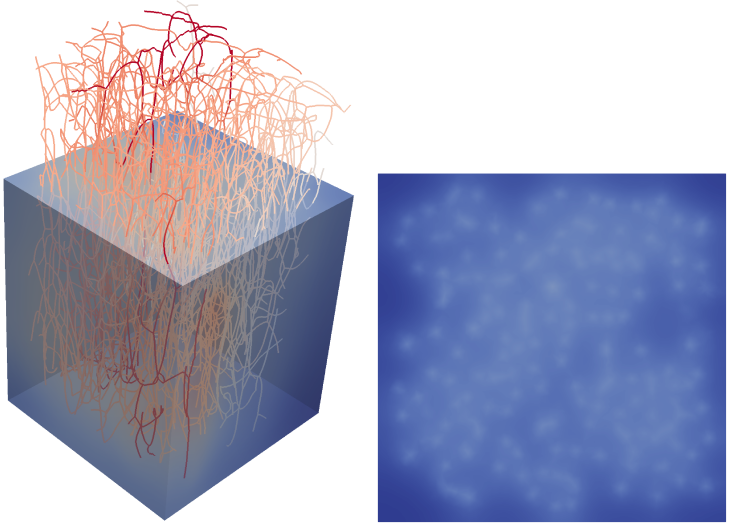

Figure II.2: Example results shown on a (left) clip of the $3d$ domain and (right) on a slice. Notice the 'halos' of increased concentration immediately around the vessels.

inlet boundary condition was swapped from $\hat{u} = 1$ to $\hat{u} = 0$ after a third of the simulation time had passed.

As parameters, we use $D_\Gamma = 6.926 \times 10^7\ \mu\text{m}^2/\text{s}$, $D_\Omega = 1.87 \times 10^2\ \mu\text{m}^2/\text{s}$ and $\beta = 50\ \mu\text{m/s}$. We used $k = 1\text{s}$ after verifying that reducing it did not significantly affect our results. In this experiment, it was unneccessary to use the preconditioner described in §II.3 since the problem size was small enough to allow for use of a direct solver.

Tofts [19] assumes a relation

$$\frac{\partial C_t}{\partial t} = \frac{K_{\text{trans}}}{\nu}(C_v - C_t) \tag{II.10}$$

between the pixel tissue concentration C_t and the pixel vessel concentration C_v for some constant K_{trans}. Here, ν is the vascular volume fraction, which in our geometry is about 0.76%. Our geometry is of a size comparable to a single pixel in [19], so C_t and C_v correspond to the normalized averages

$$C_t = \frac{\int_\Omega u}{\int_\Omega 1} \quad \text{and} \quad C_v = \frac{\int_\Gamma \pi R^2 \hat{u}}{\int_\Gamma \pi R^2}.$$

We computed K_{trans} by solving for $u, \hat{u}$ using our model, and then computing C_t, C_v as given above, and defining K_{trans} such that equation (II.10) holds at each time point. This makes K_{trans} a function of time, with units seconds^{-1}.

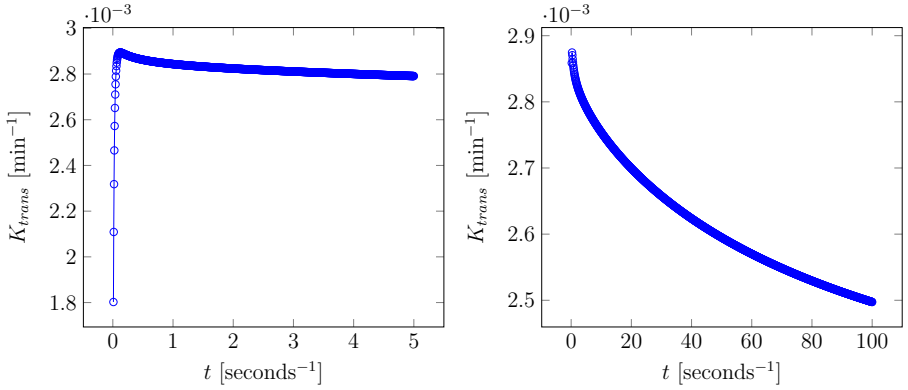

Figure II.3: Behavior of K_{trans} on short and long time scales.

II.4.1 Discussion of perfusion experiment

Our value for K_{trans} is not entirely constant. There is a small variation in time, as perfusion seems to be somewhat faster when the extravascular space is completely empty of the tracer. As it starts getting saturated, perfusion slows down somewhat. This translates into K_{trans} decreasing by about 20% over a period of about 5 minutes, from 0.0028 min^{-1} to 0.0023 min^{-1}.

In [20], K_{trans} was estimated in healthy and cancerous human brain tissue from MRI scans. In healthy tissue, they estimate K_{trans} to be between 0.003 min^{-1} and 0.005 min^{-1}, that is, slightly higher than our results. There are several possible explanations for this difference. One might be that in our model, vascular transport is modeled as exceptionally fast diffusion for convenience, whereas in reality it occurs by convection. However, in both cases $1d$ transport is very fast compared to the $3d$ transport and the $1d$-$3d$ exchange. Further, K_{trans} is defined in terms of the $1d$-$3d$ exchange alone, so non-extreme variations in the $1d$ transport seem unlikely to be relevant.

Another possibility might be that the data of [20] are taken from human brain tissue, while our vasculature is taken from a mouse brain tissue, likely from a different region of the brain. A third reason might be our diffusion constants not exactly matching the tracer used by [20].

In further work, it would be interesting to incorporate convective transport into the model and see if better agreement with the experimental data is observed. A suitable starting point here is [1], who derive a convection-diffusion type system (equations (3a), (3b)) by assuming that the blood flow $\hat{q}$ in a segment is laminar and follows Poiseuille's law

$$\hat{q} = R^4 C \nabla \hat{p}.$$

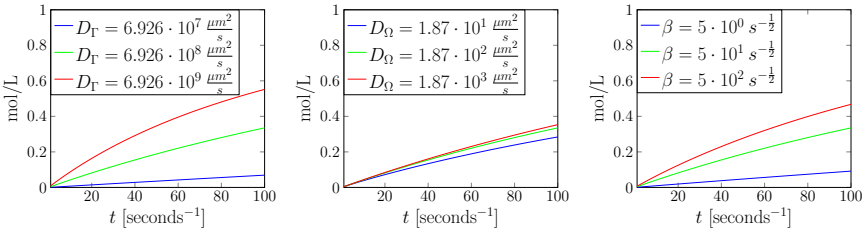

Figure II.4: Plots of variation in C_t when different parameters are varied.

II.4.2 Parameter sensitivity analysis

We carried out a rudimentary parameter sensitivity analysis by varying the three parameters $D_\Gamma, D_\Omega, \beta$ by a factor 10 and seeing how that affected the tissue concentration C_t. Specifically, we started from a baseline of $D_\Gamma = 6.926 \cdot 10^8 \frac{\mu m^2}{s}, D_\Omega = 1.87 \cdot 10^2 \frac{\mu m^2}{s}, \beta = 5 \cdot 10^1\, \mathrm{s}^{-\frac{1}{2}}$, and for each parameter, increased or decreased it by a factor 10.

The results are shown in Figure II.4. They indicate that C_t depends more strongly on β and D_Γ than on D_Ω for the set of parameters considered here.

II.5 Conclusions

A coupled $3d$-$1d$ system with an additional unknown enforcing the coupling between the domains was used as a model of tissue perfusion. For the system we proposed a robust preconditioner and demonstrated its properties through numerical experiments. Further, we have shown that the model can be applied to a physiological problem with reasonable results.

References

[1] Cattaneo, L. and Zunino, P. "A computational model of drug delivery through microcirculation to compare different tumor treatments". In: *International journal for numerical methods in biomedical engineering* vol. 30, no. 11 (2014), pp. 1347–1371.

[2] Cattaneo, L. and Zunino, P. "Numerical Investigation of Convergence Rates for the FEM Approximation of 3D-1D Coupled Problems". In: *Numerical Mathematics and Advanced Applications-ENUMATH 2013*. Springer, 2015, pp. 727–734.

[3] Chapman, S. J., Shipley, R. J., and Jawad, R. "Multiscale modeling of fluid transport in tumors". In: *Bulletin of Mathematical Biology* vol. 70, no. 8 (2008), p. 2334.

[4] D'Angelo, C. "Finite element approximation of elliptic problems with Dirac measure terms in weighted spaces: applications to one-and three-dimensional coupled problems". In: *SIAM Journal on Numerical Analysis* vol. 50, no. 1 (2012), pp. 194–215.

[5] D'Angelo, C. and Quarteroni, A. "On the coupling of 1D and 3D diffusion-reaction equations: application to tissue perfusion problems". In: *Mathematical Models and Methods in Applied Sciences* vol. 18, no. 08 (2008), pp. 1481–1504.

[6] Fang, Q. et al. "Oxygen advection and diffusion in a three dimensional vascular anatomical network". In: *Optics express* vol. 16, no. 22 (2008), p. 17530.

[7] Goldman, D. and Popel, A. S. "A computational study of the effect of capillary network anastomoses and tortuosity on oxygen transport". In: *Journal of Theoretical Biology* vol. 206, no. 2 (2000), pp. 181–194.

[8] Jack, C. "Cerebrovascular and cardiovascular pathology in Alzheimer's disease". In: *International Review of Neurobiology* vol. 84 (2009), pp. 35–48.

[9] Koppl, T. and Wohlmuth, B. "Optimal a priori error estimates for an elliptic problem with Dirac right-hand side". In: *SIAM Journal on Numerical Analysis* vol. 52, no. 4 (2014), pp. 1753–1769.

[10] Kuchta, M., Mardal, K.-A., and Mortensen, M. "On preconditioning saddle point systems with trace constraints coupling 3D and 1D domains–applications to matching and nonmatching FEM discretizations". In: *arXiv preprint arXiv:1612.03574* (2016).

[11] Kuchta, M. et al. "Preconditioners for saddle point systems with trace constraints coupling 2d and 1d domains". In: *SIAM Journal on Scientific Computing* vol. 38, no. 6 (2016), B962–B987.

[12] Linninger, A. et al. "Cerebral microcirculation and oxygen tension in the human secondary cortex". In: *Annals of biomedical engineering* vol. 41, no. 11 (2013), pp. 2264–2284.

[13] Mardal, K.-A. and Winther, R. "Preconditioning discretizations of systems of partial differential equations". In: *Numerical Linear Algebra with Applications* vol. 18, no. 1 (2011), pp. 1–40.

[14] Mardal, K.-A. and Winther, R. "Uniform preconditioners for the time dependent Stokes problem". In: *Numerische Mathematik* vol. 98, no. 2 (2004), pp. 305–327.

[15] Nabil, M., Decuzzi, P., and Zunino, P. "Modelling mass and heat transfer in nano-based cancer hyperthermia". In: *Royal Society open science* vol. 2, no. 10 (2015), p. 150447.

[16] Sakadžić, S. et al. "Large arteriolar component of oxygen delivery implies a safe margin of oxygen supply to cerebral tissue". In: *Nature communications* vol. 5 (2014), p. 5734.

[17] Secomb, T. W. et al. “Green’s function methods for analysis of oxygen delivery to tissue by microvascular networks”. In: *Annals of biomedical engineering* vol. 32, no. 11 (2004), pp. 1519–1529.

[18] Secomb, T. et al. “Theoretical simulation of oxygen transport to brain by networks of microvessels: effects of oxygen supply and demand on tissue hypoxia”. In: *Microcirculation* vol. 7, no. 4 (2000), pp. 237–247.

[19] Tofts, P. S. “T1-weighted DCE imaging concepts: modelling, acquisition and analysis”. In: *Signal* vol. 500, no. 450 (2010), p. 400.

[20] Zhang, N. et al. “Correlation of volume transfer coefficient K_{trans} with histopathologic grades of gliomas”. In: *Journal of Magnetic Resonance Imaging* vol. 36, no. 2 (2012), pp. 355–363.

Paper IV

Robust preconditioning for coupled Stokes-Darcy problems with the Darcy problem in primal form

Karl Erik Holter, Miroslav Kuchta, Kent-André Mardal

arXiv: 2001.05529 To appear in *Computers & Mathematics with Applications.*

Abstract

The coupled Darcy-Stokes problem is widely used for modeling fluid transport in physical systems consisting of a porous part and a free part. In this work we consider preconditioners for monolitic solution algorithms of the coupled Darcy-Stokes problem, where the Darcy problem is in primal form. We employ the operator preconditioning framework and utilize a fractional solver at the interface between the problems to obtain order optimal schemes that are robust with respect to the material parameters, i.e. the permeability, viscosity and Beavers-Joseph-Saffman condition. Our approach is similar to that of [19], but since the Darcy problem is in primal form, expressing mass conservation at the interface involves the normal derivative, which introduces some mathematical challenges. These challenges will be specifically addressed in this paper, in particular we will employ fractional Laplacians at the interface. Numerical experiments illustrating the performance are provided. The preconditioner is posed in non-standard Sobolev spaces which may be perceived as an obstacle for its use in applications. However, we detail the implementational aspects and show that the preconditioner is quite feasible to realize in practice.

IV.1 Introduction

Let $\Omega = \Omega_f \cup \Omega_p$, where Ω_f is the domain of the viscous flow, Ω_p is the domain of the porous media and Γ their common interface. Further let the domain boundaries be decomposed as $\partial\Omega_f = \Gamma \cup \partial\Omega_{f,D} \cup \partial\Omega_{f,N}$ and $\partial\Omega_p = \Gamma \cup \partial\Omega_{p,D} \cup \partial\Omega_{p,N}$, where subscripts D, N signify respectively that Dirichlet and Neumann boundary conditions are prescribed on the part of the boundary. The boundary of Γ, i.e., the intersection of Γ and $\partial\Omega$ is denoted by $\partial\Gamma$. An illustration is given in Figure IV.1.

The Stokes problem reads:

$$\mu\Delta\mathbf{u}_f - \nabla p_f = \mathbf{f}_f \text{ in } \Omega_f, \quad \text{(IV.1)}$$
$$\nabla \cdot \mathbf{u}_f = 0 \text{ in } \Omega_f, \quad \text{(IV.2)}$$

while the Darcy problem in primal form reads:

$$-K\Delta p_p = g \text{ in } \Omega_p. \quad \text{(IV.3)}$$

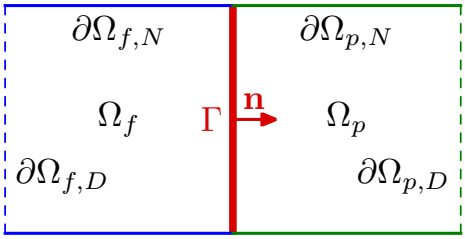

Figure IV.1: Schematic domain of Darcy-Stokes problem. Dirichlet conditions shown in dashed line, and interface in red.

Here, $\mathbf{u}_f, p_f$ are the unknown velocity and pressure for the Stokes problem (IV.1)-(IV.2) in Ω_f, p_p is the unknown pressure of the Darcy problem (IV.3) in Ω_p. The material parameters are the fluid viscosity μ and the permeability K. Here we shall consider the problem with the Dirichlet boundary conditions

$$\mathbf{u}_f = \mathbf{u}_f^0 \text{ on } \partial\Omega_{f,D}, \quad p_p = p_p^0 \text{ on } \partial\Omega_{p,D}$$

and Neumann conditions

$$(\mu\nabla\mathbf{u}_f - p_f I)\cdot\mathbf{n}_f = \mathbf{h} \text{ on } \partial\Omega_{f,N}, \quad \nabla p_p \cdot \mathbf{n}_p = h_p \text{ on } \partial\Omega_{p,N},$$

where $\mathbf{n}_f$, $\mathbf{n}_p$ are the outer unit normals of the respective subdomains. In particular we assume that $|\partial\Omega_{i,D}| > 0$ and $|\partial\Omega_{i,N}| > 0$ for $i = p, f$. Moreover, the coupled problem must be equipped with interface conditions expressing the continuity of stress as well as mass balance. We postpone their description until we describe the weak formulation of the problem.

The discretization of the coupled Darcy-Stokes problem with the Darcy problem in a mixed form is challenging since the Darcy and Stokes problems, respectively, call for different schemes. For example, typical finite element methods for the Darcy problem, like the Raviart-Thomas or Brezzi-Douglas-Marini elements, are not stable for Stokes problem as the discretization of the flux specifically targets the properties of $H(\mathrm{div})$ rather than H^1 which is natural for Stokes discretizations. For this reason, a wide range of methods have been proposed over the last decade that address this particular challenge. For example, new elements robust for both the Darcy and Stokes problem have been proposed in [1, 23, 24, 29, 35]. Alternatively, stabilization or modifications of standard methods may be used as in [7, 14, 20, 21, 34]. In this work we will consider the coupled problem with the Darcy equation in a primal form and a Lagrange multiplier to couple the Stokes and Darcy problems. Standard elements in both the Darcy and the Stokes domain can then be used, but a main problem with such schemes is the stability of the discretization at the interface.

The well-posedness of the Darcy-Stokes problem coupled together through the use of a Lagrange multiplier is well-known when the Darcy problem is in mixed form [15, 27], where both the continuous setting and various discretizations were proposed. Other solution and discretization algorithms for the coupled

problem are presented in e.g. [16, 33], see [11, 27] for an overview. For the mixed formulation we have, in our previous work [19], developed monolithic solvers that are robust with respect to all material parameters by utilizing fractional solvers on the interface. Here, we continue with the same type of approach, but address the difficulty of the Darcy problem in primal form, where the main concern from a mathematical point of view is the normal gradient at the interface. However, as the interface is of lower dimension, the number of degrees of freedom at the interface is typically small compared to the overall problem and preconditioning blocks at the interface based on fractional Laplacians are hence feasible to realise without sacrificing performance. Furthermore, multilevel solvers are available for fractional Laplacians [3, 5] if the interface requires more degrees of freedom than can be solved for by direct methods. We remark that the problem to be studied further is symmetric and includes an explicit variable, the Lagrange multiplier, on Γ. In this respect it differs from the more common primal formulation, which leads to a non-symmetric system to be solved for $\mathbf{u}_f$, p_f and p_p. Well-posedness of the latter problem was established in [11] with efficient solvers proposed and analyzed e.g. in [8, 10, 12].

An outline of the paper is as follows: Section IV.2 describes the notation, introduces the symmetric primal Darcy-Stokes problem and illustrates the difficulties in its preconditioning. The main challenge for the solver construction, i.e. the proper posing of the coupling operator, is addressed in Section IV.3. Parameter robust preconditioners are then established in Section IV.4.

IV.2 Preliminaries

Let Ω be a bounded Lipschitz domain in $\mathbb{R}^n$, $n = 2$ or 3, and denote its boundary by $\partial\Omega$. We denote by $L^2(\Omega)$ the Lebesgue space of square integrable functions, with the norm $\|u\|^2_{L^2(\Omega)} = \int_\Omega |u|^2\, dx$, and by $H^1(\Omega)$ the Sobolev space of functions with first derivative in $L^2(\Omega)$ with norm $\|u\|^2_{H^1(\Omega)} = \|u\|^2_{L^2(\Omega)} + \|\nabla u\|^2_{L^2(\Omega)}$. Note that the spaces are both Hilbert spaces, with the standard inner products. These spaces are defined in the same way when u is a vector field, in which case we will write $\mathbf{u}$ in boldface. We also define the subspace $H^1_0(\Omega)$ to be the completion in $\|\cdot\|_{H^1(\Omega)}$ of $C^\infty_0(\Omega)$, the space of smooth functions on Ω whose restriction to $\partial\Omega$ is zero.

For a Lipschitz domain Ω with $\Gamma \subset \partial\Omega$, we can define a trace operator T by $Tu = u|_\Gamma$ for smooth u. This can be extended to a bounded, surjective and right-invertible operator $H^1(\Omega) \to H^{\frac{1}{2}}(\Gamma)$ (cf. e.g. [9]), where the space $H^{\frac{1}{2}}(\Gamma)$ will be defined later. Given a subset $\partial\Omega_D$ of $\partial\Omega$, we let $H^1_{0,\partial\Omega_D}(\Omega)$, or for readability just $H^1_{0,D}(\Omega)$, be the subspace of $H^1(\Omega)$ for which the restriction to $\partial\Omega_D$ is zero, where the restriction is defined in terms of the trace operator. Typically, $\partial\Omega_D$ will be the subset of $\partial\Omega$ on which Dirichlet conditions are prescribed. We also define the semi-norm $L^2_\tau(\Gamma)$ on $H^1(\Omega)$ to be the $L^2(\Gamma)$ norm of the tangential component of $\mathbf{u}$ at Γ. In 2D, this is just $\|\mathbf{u}|_\Gamma \cdot \tau\|_{L^2(\Gamma)}$ where τ is a tangent unit vector, while in 3D it is more conveniently written as $\|\mathbf{u}|_\Gamma - (\mathbf{u}|_\Gamma \cdot \mathbf{n})\mathbf{n}\|_{L^2(\Gamma)}$.

For any inner product space X, we let $(\cdot,\cdot)_X$ denote its inner product. When $X = L^2(\Omega)$, we will omit the subscript if there is no cause for confusion. We write the space of continuous linear operators from X to Y as $\mathcal{L}(X,Y)$, or just as $\mathcal{L}(X)$ if $Y = X$. For any two Sobolev spaces X, Y both contained in a common ambient space, we define the intersection and sum spaces $X \cap Y$ and $X + Y$ in terms of the norms

$$\|u\|^2_{X\cap Y} = \|u\|^2_X + \|u\|^2_Y \quad \text{and} \quad \|u\|^2_{X+Y} = \inf_{\substack{x+y=u \\ x\in X, y\in Y}} \|x\|^2_X + \|y\|^2_Y.$$

For any $c > 0$, we define the scaled space cX to be just X as a set, but with the inner product $(u,v)_X = c(u,v)_X$. Its norm is trivially equivalent to $\|\cdot\|_X$, but because the equivalence constant depends on c, the distinction between the two norms becomes important when we need to establish the independence of bounds with respect to problem parameters.

We define the fractional space $H^s(\Gamma)$ following [26]. Let $S \in \mathcal{L}(H^1(\Gamma))$ be the operator such that $(Su,v)_{H^1} = (S(I-\Delta)u,v) = (u,v)_{L^2}$ for all $v \in H^1(\Gamma)$. We can then find a basis of $H^1(\Gamma)$ of orthonormal eigenfunctions e_i of S with eigenvalues $\lambda_i > 0$. Writing $u = \sum_i c_i e_i$ in this basis, we define the norm $\|u\|^2_{H^s(\Gamma)} = \sum c_i^2 \lambda_i^{-s}$ for any $s \in [-1,1]$. Further, let the space $H^s(\Gamma)$ be the completion of $C^\infty(\Gamma)$ with respect to $\|\cdot\|_{H^s(\Gamma)}$. We also define the space $H^s_{00}(\Gamma)$ in the same manner, except that we then apply Dirichlet boundary conditions by choosing S in $\mathcal{L}\left(H^1_0(\Gamma)\right)$. Furthermore, $H^s_{00}(\Gamma)$ is the completion of $C^\infty_0(\Gamma)$ rather than $C^\infty(\Gamma)$.

For the sake of completeness we review here the construction of a matrix realization of fractional operators given in [26]. To this end let $V_h \subset H^1(\Gamma)$, $n = \dim V_h$, be a finite dimensional finite element subspace with basis functions ϕ_i, $i = 1,\dots,n$ and A, $\mathsf{M} \in \mathbb{R}^{n\times n}$ be the symmetric positive definite (stiffness and mass) matrices such that

$$\mathsf{A}_{ij} = (\nabla\phi_j, \nabla\phi_i) \quad \text{and} \quad \mathsf{M}_{ij} = (\phi_j, \phi_i).$$

In case $V_h \not\subset H^1(\Gamma)$ and piecewise constant (P0) discretization is used we let

$$\mathsf{A}_{ij} = \sum_{\nu\in\mathcal{N}} \{\{h\}\}_\nu^{-1} (\llbracket\phi_j\rrbracket_\nu, \llbracket\phi_i\rrbracket_\nu)_\nu,$$

where $\mathcal{N}$ is a set of all the facets of the finite element mesh. Further the (facet) average and jump operators are defined as $\{\{u\}\}_\nu = \frac{1}{2}(u|_{K^+} + u|_{K^-})$, $\llbracket u\rrbracket_\nu = u|_{K^+} - u|_{K^-}$ with K^+ and K^- the two cells sharing facet ν. When ν is an exterior facet, we define $\llbracket u\rrbracket_\nu = \{\{u\}\}_\nu = u|_K$, where K is the unique cell with ν as facet.

It follows that the generalized eigenvalue problem $(\mathsf{A}+\mathsf{M})\mathsf{U} = \mathsf{M}\mathsf{U}\mathsf{\Lambda}$ has only positive eigenvalues and a complete set of eigenvectors that form the basis of $\mathbb{R}^n$ so that the powers of $\mathsf{S} = \mathsf{U}\mathsf{\Lambda}(\mathsf{M}\mathsf{U})^T$ are well defined. For $s \in [-1,1]$ we then

set $\mathsf{H}(s) = \mathsf{MS}^s$. Letting u be the vector of degrees of freedom of $u_h \in V_h$, i.e. $u_h = \sum_i^n (\mathsf{u})_i \phi_i$, we finally have

$$\|u_h\|_{H^s} = \sqrt{\sum_{i,j=1}^{n} \mathsf{u}_i \mathsf{H}_{ij}(s) \mathsf{u}_j}.$$

When $\mathbf{u}$ is a vector function, we define the normal trace $T_{\mathbf{n}}\mathbf{u} = \mathbf{u}|_\Gamma \cdot \mathbf{n}$ using the trace operator T component-wise. As such $T_{\mathbf{n}}$ is a continuous map $H^1(\Omega) \to H^{\frac{1}{2}}(\Gamma)$. Moreover, we let T_τ be the tangential trace operator. We remark that in 2D and 3D the operator maps to scalar, respectively vector fields. The normal derivative, $\partial_{\mathbf{n}} u = \nabla u \cdot \mathbf{n}|_\Gamma$, is more challenging to define properly in this context. Let us therefore briefly sketch an approach, which at least in the authors' opinion at first glance seems like a natural starting point. However, as we will show, the approach does not yield robust preconditioners in our context. First, notice that if we impose additional regularity on u and require that $\Delta u \in L^2$ then $\partial_{\mathbf{n}}$ is well defined. In detail, let $w \in H^{1/2}(\partial\Omega)$ and $E : H^{1/2}(\partial\Omega) \to H^1(\Omega)$ be a (harmonic) extension operator. Then $\partial_{\mathbf{n}} u$ clearly lies in $H^{-1/2}(\partial\Omega)$ because

$$\int_{\partial\Omega} \partial_{\mathbf{n}} u \cdot \mathbf{w}\, ds = \int_\Omega \Delta u \cdot E\mathbf{w}\, \mathrm{d}x + \int_\Omega \nabla \cdot (E\mathbf{w}) \cdot \nabla u\, \mathrm{d}x \leq \infty.$$

This extra regularity assumption is, however, hard to express in the operator preconditioning framework. In particular, to the author's knowledge, there are no *standard* finite elements that would enable us to exploit the extra regularity. A possible approach could be NURBS [22] or C^1 discretizations developed for fourth order problems. However, the latter often show poor performance for second order problems [32].

Alternatively, we may attempt to define $\partial_{\mathbf{n}}$ as a composition of the first order derivative operator, ∇, with the $1/2$ order normal trace operator $T_{\mathbf{n}}$. The composition $\partial_{\mathbf{n}}$ could then be expected to be a $3/2$ operator $\partial_{\mathbf{n}} : H^1(\Omega) \to H^{-1/2}(\partial\Omega)$. From an operator preconditioning point of view, this would be feasible to realize, as we will see below. However, as we will demonstrate, robustness will not be obtained if we realize $\partial_{\mathbf{n}}$ as a $3/2$ operator. In fact, robustness is only obtained if $\partial_{\mathbf{n}}$ is a first order operator, $\partial_{\mathbf{n}} : H^1(\Omega) \to L^2(\partial\Omega)$. We remark here that while the operator in a continuous setting is $\partial_{\mathbf{n}} : H^1(\Omega) \to L^2(\partial\Omega)$, in the discrete setting we will include a scaling parameter, i.e. the mesh size, because we use the finite element method. To see that this is reasonable, notice that for finite elements, the mass matrix, as representation of the identity, is differently scaled in different dimensions. In Example IV.3.2 we detail the scaling in a simplified example.

In order to demonstrate why posing the $\partial_{\mathbf{n}}$ operator properly is required, let us now formulate the coupled Darcy-Stokes problem, where the Darcy problem is in a primal form. As a starting point, let the Lagrangian of the coupled problem

be,

$$\begin{aligned} L(\mathbf{u}_f, p_f, p_p, \lambda) = &\int_{\Omega_f} \frac{1}{2}\left(\mu(\nabla \mathbf{u}_f)^2 - \mathbf{f}_f \cdot \mathbf{u}_f\right) \mathrm{d}x + \int_\Gamma \frac{1}{2} D(\mathbf{u}_f \cdot \tau)^2 \,\mathrm{d}s \\ &+ \int_{\Omega_p} \frac{1}{2} K\left((\nabla p_p)^2 - g\, p_p\right) \mathrm{d}x + \int_{\Omega_f} \nabla \cdot \mathbf{u}_f \, p_f \,\mathrm{d}x \\ &+ \int_\Gamma (T_{\mathbf{n}} \mathbf{u}_f - K \partial_{\mathbf{n}} p_p) \lambda \,\mathrm{d}s. \end{aligned}$$

Note that the sign of p_f has been changed from (IV.1). Here, the Lagrange multiplier λ in $\int_\Gamma (T_{\mathbf{n}} \mathbf{u}_f - K\partial_{\mathbf{n}} p_p)\lambda \,\mathrm{d}s$ is used to ensure mass conservation, while the extra term $\int_\Gamma \frac{1}{2} D(\mathbf{u}_f \cdot \tau)^2 \,\mathrm{d}s$, where $D = \alpha_{\mathrm{BJS}} \sqrt{\frac{\mu}{K}}$, corresponds to the Beavers-Joseph-Saffman condition [31].

The corresponding weak formulation is obtained by the first order optimality conditions of the Lagrangian, that is; $\frac{\partial L}{\partial \mathbf{u}_f} = 0$, $\frac{\partial L}{\partial p_f} = 0$, $\frac{\partial L}{\partial p_p} = 0$, and $\frac{\partial L}{\partial \lambda} = 0$. A variational formulation hence reads: Find $(\mathbf{u}_f, p_p, p_f, \lambda)$ such that

$$\begin{aligned} a((\mathbf{u}_f, p_p), (\mathbf{v}_f, q_p)) + b((\mathbf{v}_f, q_p), (p_f, \lambda)) &= f((\mathbf{v}_f, q_p)) \quad \forall (\mathbf{v}_f, q_p), \\ b((\mathbf{u}_f, p_p), (q_f, w)) &= g((q_f, w)) \qquad \forall (q_f, w), \end{aligned} \tag{IV.4}$$

where the bilinear forms a, b are defined as

$$\begin{aligned} a((\mathbf{u}_f, p_p), (\mathbf{v}_f, q_p)) &= \mu(\nabla \mathbf{u}_f, \nabla \mathbf{v}_f)_{\Omega_f} + D(\mathbf{u}_f \cdot \tau, \mathbf{v}_f \cdot \tau)_\Gamma + K(\nabla p_p, \nabla q_p)_{\Omega_p}, \\ b((\mathbf{u}_f, p_p), (q_f, w)) &= (\nabla \cdot \mathbf{u}_f, q_f)_{\Omega_f} + (T_{\mathbf{n}} \mathbf{u}_f, w)_\Gamma - K(\partial_{\mathbf{n}} p_p, w)_\Gamma. \end{aligned} \tag{IV.5}$$

We shall refer to (IV.4) as the (primal) Darcy-Stokes problem. Note that the resulting formulation is symmetric.

While appropriate function spaces are readily available for $\mathbf{u}_f, p_p, p_f$ and their corresponding test functions, it is less clear what the appropriate requirements are for w and λ. This will be addressed below.

Example IV.2.1 (Preconditioner for coupled Darcy-Stokes problem assuming $\partial_{\mathbf{n}} : H^1 \to H^{-1/2}$)**.** Let us assume that $\partial_{\mathbf{n}}$ is a $3/2$ operator so that $K\partial_{\mathbf{n}} p_p \in \frac{1}{\sqrt{K}} H^{-1/2}$ for $p_p \in \sqrt{K} H^1_{0,D}(\Omega_p)$. Next, observe that since $\mathbf{u}_f \in \sqrt{\mu} H^1_{0,D}(\Omega_f) \cap \sqrt{D} L^2_\tau(\Gamma)$ then $T_{\mathbf{n}} \mathbf{u}_f \in \sqrt{\mu} H^{1/2}$. Per assumption the coupling term $T_{\mathbf{n}} \mathbf{u}_f - K\partial_{\mathbf{n}} p_p$ belongs to $\sqrt{\mu} H^{1/2} + \frac{1}{\sqrt{K}} H^{-1/2}$ so that the dual variable $w \in \frac{1}{\sqrt{\mu}} H^{-1/2} \cap \sqrt{K} H^{1/2}$. In turn, we consider the following weak formulation: Find $\mathbf{u}_f, p_p, p_f, \lambda \in \sqrt{\mu} H^1_{0,D}(\Omega_f) \cap \sqrt{D} L^2_\tau(\Gamma), \sqrt{K} H^1_{0,D}(\Omega_p), \frac{1}{\sqrt{\mu}} L^2(\Omega_f), \frac{1}{\sqrt{\mu}} H^{-1/2} \cap \sqrt{K} H^{1/2}$ such that

$$\begin{aligned} a((\mathbf{u}_f, p_p), (\mathbf{v}_f, q_p)) + b((\mathbf{v}_f, q_p), (p_f, \lambda)) &= f((\mathbf{v}_f, q_p)) \quad \forall (v_f, q_p), \\ b((\mathbf{u}_f, p_p), (q_f, w)) &= g((q_f, w)) \qquad \forall (q_f, w). \end{aligned} \tag{IV.6}$$

The coefficient matrix associated with (IV.4) reads

$$\mathcal{A} = \left(\begin{array}{cc|cc} -\mu\Delta + DT_\tau' T_\tau & & (\nabla\cdot)' & T_{\mathbf{n}}' \\ & -K\Delta & & -K\partial_{\mathbf{n}}' \\ \hline \nabla\cdot & & & \\ T_{\mathbf{n}} & -K\partial_{\mathbf{n}} & & \end{array}\right). \tag{IV.7}$$

Assuming that the proposed spaces indeed lead to well-posed operator $\mathcal{A}$, the operator preconditioning framework [30] yields as a preconditioner the Riesz mapping

$$\mathcal{B} = \left(\begin{array}{cc|c} -\mu\Delta + DT_\tau' T_\tau & & \\ & -K\Delta & \\ \hline & & \frac{1}{\mu} I \\ & & \qquad \frac{1}{\mu}\left(I+\Delta\right)^{-1/2} + K\left(I+\Delta\right)^{1/2} \end{array}\right)^{-1}. \tag{IV.8}$$

In order to test the preconditioner, we solve problem (IV.6) on $\Omega = [0,2]\times[0,1]$, where $\Omega_f = [0,1]\times[0,1]$ and $\Omega_p = [1,2]\times[0,1]$ and the Dirichlet boundary domains are $\partial\Omega_{f,D} = \{(x,y)\in\partial\Omega_f, x = 0\}$ and $\partial\Omega_{p,D} = \{(x,y)\in\partial\Omega_p, x = 2\}$, cf. Figure IV.1. The mesh is a uniform triangular mesh, consisting of $4N^2$ equally sized isosceles triangles. To discretize (IV.4), we use lowest order (P2-P1) Taylor-Hood elements for the Stokes velocity and pressure, while piecewise quadratic elements (P2) were used for the Darcy pressure and piecewise constant elements (P0) for the Lagrange multiplier. Discretization is carried out in the FEniCS library [28], with coupling maps between the interface and domains and the fractional Laplacians being implemented by the extension $\text{FEniCS}_{\text{ii}}$ [25].

Approximation of the preconditioner (IV.8) is constructed by using single sweep of V-cycle of algebraic multigrid BoomerAMG from the Hypre library [13] for all the blocks except for the interface block, which is inverted exactly. Starting from a random initial vector, we count the number of iterations required to solve the preconditioned linear system using the MINRES solver from the PETSc library [4] with convergence criterion based on relative tolerance of 10^{-8} and absolute tolerance of 10^{-10}. Additionally, the condition numbers of $\mathcal{B}^{-1}\mathcal{A}$ are computed using an iterative solver from the SLEPc library [17]. In the condition number computations the operator $\mathcal{B}$ is computed exactly, that is, all the blocks are inverted by LU. We remark that this solver setup is used also in the subsequent examples.

The results of the experiment are plotted in Figure IV.2. By the failure of the iteration counts to stabilize, we see that using $\frac{1}{\mu}\left(I+\Delta\right)^{-1/2} + K\left(I+\Delta\right)^{1/2}$ as multiplier space does not lead to a robust preconditioner over the whole parameter range. Note, however, that in the regime where μ is significantly smaller than K (i.e. the lower left region of the plots in Figure IV.2), iteration counts and condition numbers appear to be stable as the mesh is refined. In this regime, the norm of the multiplier space is dominated by the part from $\frac{1}{\sqrt{\mu}} H^{-1/2}$, which is determined by posing of the trace operator. This suggests

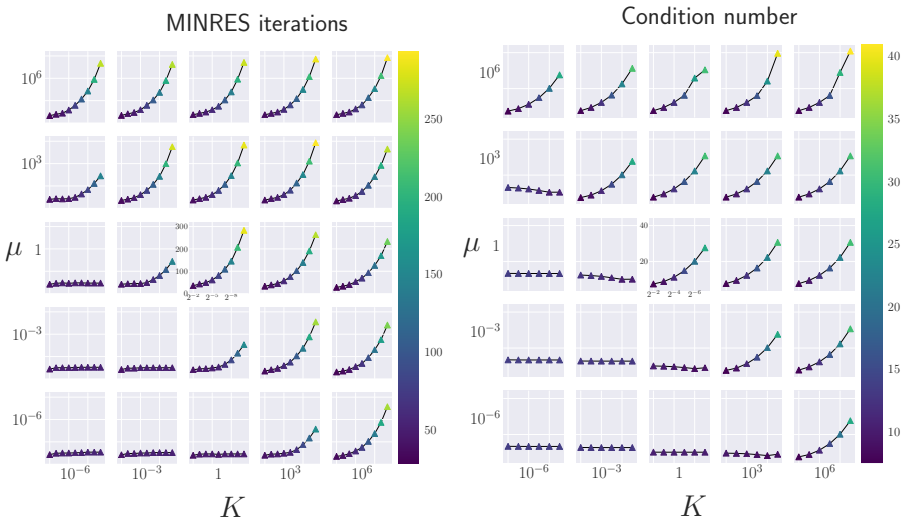

Figure IV.2: Mesh refinement vs. iteration counts (left) and condition numbers (right) for Example IV.2.1. All subplots share x- and y-axes. For fixed μ, K the x-axis range in the iterations subplot extends from (mesh size) $h = 2^{-2}$ to $h = 2^{-10}$. In the condition number plots the range is from $h = 2^{-2}$ to $h = 2^{-8}$. In all cases, $\alpha_{BJS} = 1$.

that the choice of $\sqrt{K}H^{1/2}$, i.e. wrong posing of the $\partial_{\mathbf{n}}$ operator, is responsible for the lack of boundedness.

IV.3 Approximating the trace normal gradient operator

A crucial step in the analysis of the Darcy-Stokes problem will be the mapping properties of the operator $\partial_{\mathbf{n}}$. As a computationally practical choice of space for the Darcy pressure is $\sqrt{K}H^1$, we immediately run into the problem discussed in the preliminaries because $\partial_{\mathbf{n}}$ cannot be defined on all of H^1. This necessitates either an assumption of extra regularity or an alternative approach.

Motivated by the observation in [18], that in a discrete finite element setting the trace operator is stable as a map $L^2(\Omega) \to L^2(\partial\Omega)$, we propose an alternative approach to construct the preconditioners. We start off by outlining the construction of an operator $\partial_{\mathbf{n},\epsilon} : H^1(\Omega_p) \to L^2(\Gamma)$ which will be an approximation to $\partial_{\mathbf{n}}$. Suppose Γ is a sufficiently regular subset of $\partial\Omega_p$, and that Γ is of co-dimension 1 in Ω_p. The ϵ-thick envelope $\Gamma_\epsilon = \{y \in \Omega_p, \mathrm{dist}(y, \Gamma) < \epsilon\}$ is a higher-dimensional approximation of Γ. For any $v \in H^1(\Omega_p)$,

$$\frac{1}{\epsilon}\int_{\Gamma_\epsilon} v\,\phi\,\mathrm{d}x \to \int_\Gamma Tv\,T\phi\,ds \text{ as } \epsilon \to 0, \tag{IV.9}$$

where ϕ is a test function in $H^1(\Omega_p)$.

Note that although the integral over Γ is not well-defined for a general $v \in L^2(\Omega_p)$, the integral over Γ_ϵ is. Provided Γ is sufficiently regular and ϵ sufficiently small, we assume that there exists a vector field $\mathbf{n}_{\Gamma_\epsilon}$ on Γ_ϵ which approximates the normal vector $\mathbf{n}_\Gamma$ of Γ at Γ. Using $\mathbf{n}_{\Gamma_\epsilon}$, we further assume that we can define a bounded extension $E_\epsilon : L^2(\Gamma) \to L^2(\Gamma_\epsilon)$ along $\mathbf{n}_{\Gamma_\epsilon}$ for which $\int_\Gamma w\, ds \approx \frac{1}{\epsilon}\int_{\Gamma_\epsilon} E_\epsilon w\, \mathrm{d}x$ for any $w \in L^2(\Gamma)$. Provided $\mathbf{n}_{\Gamma_\epsilon}$ and E_ϵ can be defined, then for any $u \in H^1(\Omega_p)$ we can define $\partial_{\mathbf{n},\epsilon} u$ by

$$\int_\Gamma \partial_{\mathbf{n},\epsilon} u \cdot w\, ds = \frac{1}{\epsilon}\int_{\Gamma_\epsilon} \nabla u \cdot \mathbf{n}_{\Gamma_\epsilon} E_\epsilon w\, dx$$

for any $w \in L^2(\Gamma)$, thus defining the required map $\partial_{\mathbf{n},\epsilon} : H^1(\Omega_p) \to L^2(\Gamma_\epsilon)$ approximating $\partial_{\mathbf{n}}$. We assume that the resulting operator $\partial_{\mathbf{n},\epsilon}$ is both surjective and bounded, with $\|\partial_{\mathbf{n},\epsilon} u\|_{L^2(\Gamma)} \leq C\|u\|_{H^1(\Omega_p)}$, and that $\partial_{\mathbf{n},\epsilon}$ has a bounded right inverse.

We emphasize that $\partial_{\mathbf{n},\epsilon}$ is just an analytical tool constructed for the analysis in the continuous setting and that ϵ is not related to the mesh size h. In fact, we can choose ϵ far smaller than the mesh size and for any practical purposes in computations we assume that $\partial_{\mathbf{n},\epsilon}$ will be practically identical to $\partial_{\mathbf{n}}$. We summarize the assumption as follows:

Assumption IV.3.1. *Given a sufficiently regular $\Gamma \subset \partial\Omega_p$, $\partial_{\mathbf{n},\epsilon} : H^1(\Omega_p) \to L^2(\Gamma)$ is a bounded surjection which approximates $\partial_{\mathbf{n}}$ on the subspace of H^1 on which $\partial_{\mathbf{n}}$ can be defined. Further, $\partial_{\mathbf{n},\epsilon}$ has a bounded right inverse.*

Although characterizing the conditions under which IV.3.1 holds is beyond the scope of this paper, we motivate the existence of the required constructions $E_\epsilon, \mathbf{n}_\epsilon$ in a few simple examples below.

Example IV.3.1. Let Γ be the $y-$axis, and Ω_p be the positive half-plane. The construction of $E_\epsilon, \mathbf{n}_{\Gamma_\epsilon}$ is then given by $\mathbf{n}_{\Gamma_\epsilon} = \mathbf{n}_\Gamma = (-1, 0)$ and for $w(y) \in C^1(\Gamma)$ we let $(E_\epsilon w)(x, y) = w(y)$. This continuously extends to all of $L^2(\Gamma)$. Clearly $\partial_{\mathbf{n},\epsilon} u \to \partial_{\mathbf{n}} u$ as $\epsilon \to 0$ for $u \in C^1$. Given any $w(y) \in C^0(\Gamma)$, define u by $u(x, y) = -xw(y)$. Then the map $w \to u$ continuously extends to a right inverse of $\partial_{\mathbf{n},\epsilon}$, as by linearity $\partial_{\mathbf{n},\epsilon} u = \partial_{\mathbf{n}} u = w$.

Next, suppose Ω_p is the unit disk, and Γ its boundary. By parametrizing Γ with e.g. polar coordinates, this case can be effectively translated to the above. $\mathbf{n}_{\Gamma_\epsilon}$ is now the unit radial vector $\mathbf{i}_r$, and for any $w(\theta) \in C^1(\Gamma)$, $(E_\epsilon w)(r, \theta) = w(\theta)$. Again, this definition of E_ϵ extends to all of $L^2(\Gamma)$. Because $\frac{1}{\epsilon}\int_{\Gamma_\epsilon} \nabla u \cdot \mathbf{n}_{\Gamma_\epsilon} E_\epsilon w\, dx = \int_0^{2\pi} w(\theta) \cdot \int_{1-\epsilon}^{1} \frac{1}{\epsilon}\frac{\partial u}{\partial r}\, r dr\, d\theta$ and $\frac{1}{\epsilon}\int_{1-\epsilon}^{1} f(r)\, dr \to f(1)$ as $\epsilon \to 0$, we again have $\partial_{\mathbf{n},\epsilon} u \to \partial_{\mathbf{n}} u$ as $\epsilon \to 0$ for $u \in C^1$. Analogously to the previous case, a right inverse can be defined by sending any $w(\theta) \in C^0(\Gamma)$ to $u(r, \theta) = rw(\theta)$.

Before considering the Darcy-Stokes problem, we justify Assumption IV.3.1. First we consider a simplified example in order to illustrate how the scaling of mass matrices in different dimensions affects preconditioners constructed

via the application of trace operators. Then, in Example IV.3.3 we construct preconditioners for a Poisson problem with a $\partial_{\mathbf{n}}$-constraint which is to be enforced by a Lagrange multiplier, cf. the Babuška problem [2] involving the trace operator.

Example IV.3.2 (Trace constrained L^2 projection)**.** Let Ω be a bounded domain with $\Gamma \subseteq \partial\Omega$ and $V = H^1(\Omega)$. We then consider the problem

$$\min_{u\in V} \int_\Omega u^2 \,\mathrm{d}x - 2\int_\Omega fu \,\mathrm{d}x \quad \text{subject to} \quad \int_\Gamma (Tu - g)\, p \,\mathrm{d}s = 0. \tag{IV.10}$$

Letting p denote the Lagrange multiplier associated with the boundary constraint, the extrema $u \in V$, $p \in Q = L^2(\Gamma)$ of the Lagrangian of (IV.10) satisfy the variational problem: Find $u \in V$ and $p \in Q$ such that

$$\begin{aligned} \int_\Omega uv \,\mathrm{d}x + \int_\Gamma pTv \,\mathrm{d}s &= \int_\Omega fv \,\mathrm{d}x \quad \forall v \in V, \\ \int_\Gamma qTu \,\mathrm{d}s &= \int_\Gamma gq \,\mathrm{d}s \quad \forall q \in Q. \end{aligned} \tag{IV.11}$$

The operator of the preconditioned continuous problem then reads

$$\mathcal{BA} = \begin{pmatrix} I & \\ & S \end{pmatrix}^{-1} \begin{pmatrix} I & T' \\ T & \end{pmatrix}, \tag{IV.12}$$

where S is to be constructed such that the condition number of the discrete problems is bounded in the discretization parameter h. Here we shall consider three constructions. We remark that when using the finite element realization of the identity operator, we mean the mass matrix and denoted it by I. The mass matrix has eigenvalues such that both the smallest and the largest eigenvalues scale as h^d on uniform mesh. First we consider $S = I$, with eigenvalues $\approx h$ because Γ is a 1D manifold. Then, following [18], we let $S = h^{-1}I$, i.e., a matrix with eigenvalues ≈ 1. Finally, the choice of $S = (-\Delta + I)^{-1/2}$ is included to show that the relevant trace space in (IV.12) is not (by viewing the trace as an order 1/2 operator) $H^{1/2}$ so that dual variable would reside in $H^{-1/2}$.

We remark that the first two operators are in practical computations assembled as weighted mass matrices where the weights for the respective operators are 1 and inverse cell volume. Recalling §IV.2 the matrix representation of the fractional operator is $\mathsf{H}(1/2)$.

To compare the three preconditioners, we let Ω be a unit square, $\Gamma = \{(x, y) \in \partial\Omega, x = 0\}$. Further, the domain shall be discretized uniformly into $4N^2$ isosceles triangles with size $h = 1/N$, see Figure IV.3. Such disretization shall be in the following referred to as uniformly structured (us). Considering finite element discretization by P2-P1 elements Table IV.3.1 lists spectral condition numbers of (IV.12). It can be seen that only the $S = h^{-1}I$ preconditioner leads to results independent of h.

The growth of the condition number in Table IV.3.1 due to the preconditioner with $-1/2$ power indeed confirms that $H^{1/2}$ is not appropriate in our setting.

Alternatively, an attempt to establish the trace space could be based on viewing the trace as an $1/2$ operator. Starting from L^2 a formal calculation then leads to the space $H^{-1/2}$ and $H^{1/2}$ as the multipler space. While we do not include here the results for $S = (-\Delta + I)^{1/2}$ we remark that the condition number behaves practically identically to $S = I$.

h	I	$(-\Delta + I)^{-1/2}$	$h^{-1}I$
2^{-2}	8.72	24.08	4.63
2^{-3}	12.11	47.84	4.63
2^{-4}	16.91	95.15	4.63
2^{-5}	23.70	189.9	4.63
2^{-6}	33.31	379.2	4.63
2^{-7}	46.90	758.0	4.63
2^{-8}	66.12	1515	4.63

Table IV.3.1: Condition numbers of (IV.12) with different preconditioners and discretization by P2-P1 elements on (us) mesh from Figure IV.3. Boundedness is obtained with the Schur complement preconditioner $h^{-1}I$.

l	P2-P1			P2-P0		
	(us)	(uu)	(nu)	(us)	(uu)	(nu)
1	4.63	4.63	4.10	4.63	4.63	3.98
2	4.63	4.06	4.32	4.63	4.07	4.33
3	4.63	4.20	4.28	4.63	4.20	4.31
4	4.63	4.29	4.31	4.63	4.32	4.34
5	4.63	4.45	4.50	4.63	4.43	4.45
6	4.63	4.25	4.28	4.63	4.32	4.37
7	4.63	4.25	4.36	4.63	4.28	4.39

Table IV.3.2: Condition numbers of (IV.12) with preconditioner using $S = h^{-1}I$. Boundedness with different types of triangulations, cf. Figure IV.3, and discretizations can be observed.

In order to verify that the properties of the $h^{-1}I$ preconditioner are not due to the uniformly structured (us) mesh, we also consider refinements of two additional discretizations of Ω shown in Figure IV.3, which we refer to as uniform unstructured (uu) and non-uniform unstructured (nu) respectively. In the uniform unstructured mesh, the vertices of the mesh are no longer on a uniform square grid, but the mesh elements are still close to uniform in size. In the non-uniform unstructured mesh, the vertices are not on a square grid, and the mesh elements close to Γ are finer than the ones further away.

Using these triangulations, problem (IV.12) shall be discretizated by P2-P1 elements as well P2-P0 elements to provide more evidence for the preconditioner construction. Indeed, Table IV.3.2 shows that the condition numbers of (IV.12) are bounded irrespective of the underlying mesh and the finite element discretization considered.

Example IV.3.3 (Babuška problem with Neumann boundary conditions)**.** Let Ω be a bounded domain with the boundary partitioned into non-overlapping subdomains $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N \cup \Gamma$ such that $|\partial\Omega_D| > 0$ and $|\Gamma| > 0$. We will consider both the case that $\partial\Omega_D \cap \Gamma = \emptyset$ and later the case that $\partial\Omega_N \cap \Gamma = \emptyset$. Let $V = H^1_{0,\partial\Omega_D}(\Omega)$ and consider the problem

$$\min_{u\in V} \int_\Omega |\nabla u|^2 \,\mathrm{d}x - 2\int_\Omega f u \,\mathrm{d}x \quad \text{subject to} \quad \int_\Gamma (\partial_{\mathbf{n}} u - g)\, p \,\mathrm{d}s = 0. \tag{IV.13}$$

With p the Lagrange multiplier associated with $\partial_{\mathbf{n}}$-constraint (IV.13) leads to a

variational problem: Find $u \in V$ and $p \in Q = L^2(\Gamma)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_\Gamma p \partial_{\mathbf{n}} v \,\mathrm{d}s = \int_\Omega f v \,\mathrm{d}x \qquad \forall v \in V, \tag{IV.14}$$

$$\int_\Gamma q \partial_{\mathbf{n}} u \,\mathrm{d}s = \int_\Gamma g q \,\mathrm{d}s \qquad \forall q \in Q. \tag{IV.15}$$

The preconditioned continuous problem then reads

$$\mathcal{BA} = \begin{pmatrix} -\Delta & \\ & S \end{pmatrix}^{-1} \begin{pmatrix} -\Delta & \partial_{\mathbf{n}}{}' \\ \partial_{\mathbf{n}} & \end{pmatrix}. \tag{IV.16}$$

Following the preliminaries where $\partial_{\mathbf{n}}$ was regarded as a $3/2$ operator we let $S = (-\Delta + I)^{1/2}$. Alternatively, $S = h^{-1}I$ is set following the Assumption IV.3.1. Finally $S = I$ is considered. Matrix realization of the S operators shall be identical to Example IV.3.2. We shall also use the tessellations described in Example IV.3.2 as well as identical eigenvalue solvers.

To compare the three preconditioners we let Ω be a unit square and $\Gamma = \{(x, y) \in \partial\Omega, x = 0\}$ (marked in red in Figure IV.3) and we consider first the (Neumann) case where $\partial\Omega_N = \{(x, y) \in \partial\Omega, y = 0 \text{ or } y = 1\}$, i.e. where the multiplier domain intersects the part of boundary with Neumann boundary conditions. On $\partial\Omega_D$, we have the Dirichlet boundary condition $u = g$, and on $\partial\Omega_N$ homogeneous Neumann condition $\nabla u \cdot \mathbf{n} = 0$ is assumed.

Using the uniform meshes (marked as (us) in Figure IV.3) and P2-P1 elements, Table IV.3.3 shows the spectral condition numbers of (IV.16). As in Example IV.3.2 only $S = h^{-1}I$ preconditioner (based on Assumption IV.3.1) leads to results independent of h.

h	$(-\Delta + I)^{1/2}$	I	$h^{-1}I$
2^{-2}	11.99	6.70	4.88
2^{-3}	14.55	9.27	4.88
2^{-4}	18.47	12.89	4.88
2^{-5}	24.44	18.01	4.88
2^{-6}	33.25	25.26	4.88
2^{-7}	45.96	35.52	4.88
2^{-8}	64.10	50.02	4.88

Table IV.3.3: Condition numbers of (IV.16) discretized by P2-P1 elements on uniform refinements of (us) mesh in Figure IV.3. Boundednes in discretization is obtained only with $S = h^{-1}I$.

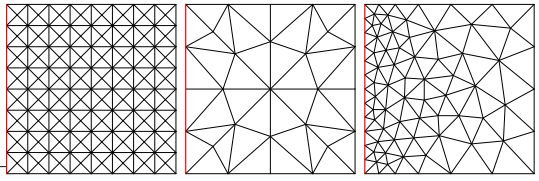

Figure IV.3: Parent meshes for uniform refinement. From left to right: uniform structured(us), uniform unstructured(uu), non-uniform unstructured(nu). Non-uniform mesh has finer (by factor 3) mesh size close to Γ.

Table IV.3.4 shows that the performance of $h^{-1}I$ in (IV.16) remains robust if different tessellations and finite element discretizations are used.

l	P2-P1			P2-P0			l	P2-P1			P2-P0		
	(us)	(uu)	(nu)	(us)	(uu)	(nu)		(us)	(uu)	(nu)	(us)	(uu)	(nu)
1	4.88	4.77	6.64	3.49	3.49	3.06	1	5.34	5.25	6.67	3.48	3.45	3.04
2	4.88	5.98	6.56	3.49	3.04	3.37	2	5.34	6.25	6.67	3.49	2.99	3.37
3	4.88	5.78	5.67	3.49	3.24	3.36	3	5.34	5.94	5.84	3.49	3.24	3.36
4	4.88	6.31	6.67	3.49	3.40	3.40	4	5.34	6.52	6.93	3.49	3.40	3.40
5	4.88	5.25	5.68	3.49	3.44	3.48	5	5.34	5.56	6.07	3.49	3.44	3.48
6	4.88	5.71	5.89	3.49	3.41	3.44	6	5.34	5.91	6.17	3.49	3.41	3.44
7	4.88	6.14	6.61	3.49	3.35	3.47	7	5.34	6.40	6.85	3.49	3.35	3.47

Table IV.3.4: Condition numbers of (IV.16) using $S = h^{-1}I$ preconditioner discretized on uniform refinements of parent meshes in Figure IV.3 using two element types. Refinement level is indicated by l. (Left) Γ intersects $\partial\Omega_N$. (Right) Γ intersects $\partial\Omega_D$.

In the context of multiscale problems, compatibility of boundary conditions of the multiplier space and the boundary conditions prescribed on the domain intesecting Γ is known to present an issue, cf. e.g. [15]. Here, we address this problem by considering (IV.16) with $|\partial_N\Omega| = 0$, i.e. we let Γ intersect only the Dirichlet boundary. We remark that until this point only intersection with Neumann boundary was considered.

In Table IV.3.4 the Dirichlet problem is considered with an *unmodified* $h^{-1}I$ preconditioner. In particular, with P2-P1 discretization we impose *no* boundary conditions on the multiplier space. Using this construction the condition numbers can be seen to remain bounded on all the meshes and with both finite element discretizations.

We remark that the $h^{-1}I$ preconditioner is equally unaffected by the Dirichlet boundary conditions on $\partial\Omega_D = \partial\Omega \setminus \Gamma$ in the trace-constrained L^2 projection problem (IV.10) with $V = H^1_{0,\partial\Omega_D}(\Omega)$, cf. Example IV.3.2. This is in contrast to the H^1 problems considered in [19], where the appropriate preconditioner was $H^{-\frac{1}{2}}_{00}$ or $H^{-\frac{1}{2}}$ depending on whether the interface intersected the Dirichlet boundary or not. Let us note that in the continuous setting boundary values have measure zero and this may then be perceived as support for our assumption that the L^2 space is the correct one in our discrete setting. Of course, the counterargument in the continuous setting is that then the trace cannot be defined. However, in the discrete setting, this can be done.

Without including the simulation results we comment here that the condition numbers of the Dirichlet problem are practically identical to those presented in Tables IV.3.1 and IV.3.2. In addition, with the two preconditioners $S = I$ and $S = (-\Delta + I)^{1/2}$ on the unstructured meshes a growth of condition numbers with h is observed similar to Table IV.3.3.

We remark that the stability of the preconditioner $h^{-1}I$ in Example IV.3.3 provides numerical evidence for well-posedness of (IV.14), i.e. the Darcy subproblem in the coupled Darcy-Stokes system (IV.4).

IV.4 Robust Preconditioners for the Darcy–Stokes system

In Example IV.2.1, we showed that the efficiency of the preconditioner (IV.8) for the primal Darcy–Stokes problem (IV.7) varied substantially with the material parameters even though the Stokes block and the Darcy block were preconditioned with appropriate preconditioners, and argued that the reason was a poor preconditioner at the interface. In this section we demonstrate that robustness with respect to mesh resolution and variations in material parameters can be obtained by posing the Lagrange multiplier in properly weighted fractional spaces, namely the intersection space $X_\Gamma = \sqrt{K}L^2(\Gamma) \cap \frac{1}{\sqrt{\mu}}H^{-1/2}(\Gamma)$. No modifications of the velocity or pressure space norms will be required. Our analysis is closely related to [19], and based on Assumption IV.3.1 along with an assumption of stability for the Stokes problem. We remark that although Assumption IV.3.1 is motivated by the discrete problem, our analysis is carried out in a continuous setting.

Let $\partial\Omega_i = \partial\Omega_{i,D} \cup \partial\Omega_{i,N} \cup \Gamma$ for $i = f, p$ such that $\partial\Omega_{f,D} \cap \Gamma = \emptyset$. We shall prove well-posedness of the coupled Darcy-Stokes problem (IV.4) with spaces

$$\begin{aligned} V_f = \sqrt{\mu}H^1_{0,D}(\Omega_f) \cap \sqrt{D}L^2_\tau(\Gamma),\ Q_f = \frac{1}{\sqrt{\mu}}L^2(\Omega_f), \\ Q_p = \sqrt{K}H^1_{0,D}(\Omega_p),\ X_\Gamma = \sqrt{K}L^2(\Gamma) \cap \frac{1}{\sqrt{\mu}}H^{-1/2}(\Gamma). \end{aligned} \tag{IV.17}$$

Note that in case Γ intersects only the Dirichlet boundary $\partial\Omega_{f,D}$ the space $H^{-1/2}$ needs to be modified to reflect $H^{1/2}_{00}$ as the appropriate trace space of V_f. We refer to [19] for a thorough discussion of the subject.

As a prerequisite for the coupled problem to be well-posed, we require that each subproblem is well-posed. For the Stokes subproblem the property has been demonstrated by numerical experiments in [19]. Here we state the result without proof.

Assumption IV.4.1. *Let Ω_f be such that $\partial\Omega_f = \partial\Omega_{f,D} \cup \partial\Omega_{f,N} \cup \Gamma$, $|\partial\Omega_{f,D}| > 0$ and $\partial\Omega_{f,D} \cap \Gamma = \emptyset$. We define $V_S = \sqrt{\mu}H^1_{0,D}(\Omega_f) \cap \sqrt{D}L^2_\tau(\Gamma) \times \frac{1}{\sqrt{\mu}}L^2(\Omega_f) \times \frac{1}{\sqrt{\mu}}H^{-1/2}(\Gamma)$ and the forms*

$$\begin{aligned} a_S((\mathbf{u}_f, p_f, \lambda), (\mathbf{v}_f, q_f, w)) =& \mu(\nabla\mathbf{u}_f, \nabla\mathbf{v}_f) + D(\mathbf{u}_f \cdot \tau, \mathbf{v}_f \cdot \tau)_\Gamma + (p_f, \nabla \cdot \mathbf{v}_f) \\ &+ (\nabla \cdot \mathbf{u}_f, q_f) + (T_{\mathbf{n}}\mathbf{u}_f, w)_\Gamma + (\lambda, T_{\mathbf{n}}\mathbf{v}_f)_\Gamma, \\ L_S((\mathbf{v}_f, q_f, w)) =& (\mathbf{f}, \mathbf{v}_f) + (g, q_f) + (h_D, w)_\Gamma, \end{aligned}$$

where $\mathbf{f} \in \frac{1}{\sqrt{\mu}}H^{-1}(\Omega_f), g \in \sqrt{\mu}L^2(\Omega_f), h_D \in \sqrt{\mu}H^{\frac{1}{2}}(\Gamma)$ are arbitrary. Then we assume that the Stokes problem: Find $(\mathbf{u}_f, q_f, \lambda) \in V_s$ such that

$$a_S((\mathbf{u}_f, p_f, \lambda), (\mathbf{v}_f, q_f, w)) = L_S((\mathbf{v}_f, q_f, w)) \quad \forall (\mathbf{v}_f, q_f, w) \in V_S$$

satisfies the Brezzi conditions and hence has a unique solution $(\mathbf{u}_f, p_f, \lambda) \in V_S$ and with constant C depending only on Ω_f, $\partial\Omega_{f,D}$ and Γ, we have

$$\|(\mathbf{u}_f, p_f, \lambda)\|_{V_S} \leq C \left(\|\mathbf{f}\|^2_{\frac{1}{\sqrt{\mu}}H^{-1}(\Omega_f)} + \|g\|^2_{\sqrt{\mu}L^2(\Omega_f)} + \|h_D\|^2_{\sqrt{\mu}H^{\frac{1}{2}}(\Gamma)} \right)^{\frac{1}{2}}.$$

Corresponding well-posedness of the Darcy problem with $\partial_{\mathbf{n}}$-constraint was demonstrated numerically for $K = 1$ in Example IV.3.3. Here, we analyze the general case.

Lemma IV.4.1. *Suppose Ω_p, Γ are such that Assumption IV.3.1 holds and $|\partial\Omega_{p,D}| > 0$. Then for any $f \in \frac{1}{\sqrt{K}}H^{-1}(\Omega_p)$, $h \in \frac{1}{\sqrt{K}}L^2(\Gamma)$, the problem of finding $(p_p, \lambda) \in \sqrt{K}H^1_{0,D}(\Omega_p) \times \sqrt{K}L^2(\Gamma)$ so that*

$$\begin{aligned} K(\nabla p_p, \nabla q_p)_{\Omega_p} + K(\lambda, \partial_{\mathbf{n},\epsilon} q_p)_\Gamma &= (f, q_p) \quad \forall q_p \in \sqrt{K}H^1_{0,D}(\Omega_p), \\ K(\partial_{\mathbf{n},\epsilon} p_p, w)_\Gamma &= (h, w)_\Gamma \quad \forall w \in \sqrt{K}L^2(\Gamma) \end{aligned} \tag{IV.18}$$

has a unique solution satisfying

$$\|p_p\|_{\sqrt{K}H^1_{0,D}(\Omega_p)} \leq C \left(\|h\|^2_{\frac{1}{\sqrt{K}}L^2(\Gamma)} + \|f\|^2_{\frac{1}{\sqrt{K}}H^{-1}(\Omega_p)} \right)^{\frac{1}{2}},$$

where C is a constant depending only on Ω_p.

Proof. Let $V = \sqrt{K}H^1_{0,D}(\Omega_p)$, $Q = \sqrt{K}L^2(\Gamma)$. We consider the left-hand side of (IV.3) as an operator

$$\begin{pmatrix} A & B' \\ B & \end{pmatrix} : V \times Q \to V' \times Q', \tag{IV.19}$$

where $(Ap_p, q_p) = K(\nabla p_p, \nabla q_p)_{\Omega_p}$ and $(Bp_p, w) = K(\partial_{\mathbf{n},\epsilon} p_p, w)_\Gamma$.

The statement of the theorem follows from the Brezzi theory [6] once the Brezzi conditions are verified. That is, we must show that A, B are bounded, A is coercive on $\ker B$ and that the inf-sup condition $\inf_{q\in Q} \sup_{v\in V}(Bv, q) \geq \beta\|v\|\|q\|$ holds for some constant $\beta > 0$.

Here the boundedness of A and the coercivity on V are evident. For the latter we recall that $|\partial\Omega_{p,D}| > 0$ is assumed and invoke the Poincare inequality. Assumption IV.3.1 is needed to show the properties of B. Because

$$K(\lambda, \partial_{\mathbf{n},\epsilon} q_p)_\Gamma \leq K\|\lambda\|_{\sqrt{K}L^2(\Gamma)}\|\partial_{\mathbf{n},\epsilon} q_p\|_{\frac{1}{\sqrt{K}}L_2(\Gamma)} \leq \|\partial_{\mathbf{n},\epsilon}\|\|\lambda\|_{\sqrt{K}L^2(\Gamma)}\|q_p\|_{\sqrt{K}H^1(\Omega_p)},$$

we have boundedness with constant $\|\partial_{\mathbf{n},\epsilon}\|$. For the inf-sup condition, we recall the bounded right inverse E of $\partial_{\mathbf{n},\epsilon}$. Letting $p_p^* = E(\lambda)$, we have $K(\lambda, \partial_{\mathbf{n},\epsilon} p_p^*) = \|\lambda\|^2_{\sqrt{K}L^2(\Gamma)}$ and $\|p_p^*\|_{\sqrt{K}H^1_{0,D}(\Omega_p)} \leq \|E\|\|\lambda\|_{\sqrt{K}L_2(\Gamma)}$ so that

$$\begin{aligned} \sup_{p_p \in \sqrt{K}H^1_{0,D}(\Omega_p)} \frac{K(\lambda, \partial_{\mathbf{n},\epsilon} p_p)}{\|p_p\|_{\sqrt{K}H^1_{0,D}(\Omega_p)}} &\geq \frac{K(\lambda, \partial_{\mathbf{n},\epsilon} p_p^*)}{\|p_p^*\|_{\sqrt{K}H^1_{0,D}(\Omega_p)}} = \frac{\|\lambda\|^2_{\sqrt{K}L^2(\Gamma)}}{\|p_p^*\|_{\sqrt{K}H^1_{0,D}(\Omega_p)}} \\ &\geq \frac{\|\lambda\|^2_{\sqrt{K}L^2(\Gamma)}}{\|E\|\|\lambda\|_{\sqrt{K}L^2(\Gamma)}} = \frac{1}{\|E\|}\|\lambda\|_{\sqrt{K}L^2(\Gamma)}. \end{aligned}$$

This proves all the Brezzi conditions. ∎

Having discussed well-posedness of the Stokes and Darcy subproblems our main result concerning the coupled Darcy-Stokes problem (IV.4) is given in Theorem IV.4.2. We remark that given two well-posed subproblems the coupled system could be analyzed with the framework of [19]. Here we provide a standalone proof.

Theorem IV.4.2. *Let Ω_f, Ω_p be as defined in Assumption IV.3.1 and Assumption IV.4.1. Further let*

$$V_f = \sqrt{\mu}H^1_{0,D}(\Omega_f) \cap \sqrt{D}L^2_\tau(\Gamma),\ Q_f = \frac{1}{\sqrt{\mu}}L^2(\Omega_f),$$
$$Q_p = \sqrt{K}H^1_{0,D}(\Omega_p),\ X_\Gamma = \sqrt{K}L^2(\Gamma) \cap \frac{1}{\sqrt{\mu}}H^{-1/2}(\Gamma).$$

Then the operator $\mathcal{A}$ in (IV.7) *is an isomorphism mapping W to its dual space W' such that $\|\mathcal{A}\|_{\mathcal{L}(W,W')} \leq C$ and $\|\mathcal{A}^{-1}\|_{\mathcal{L}(W',W)} \leq \frac{1}{C}$, where C is independent of μ, K, and D.*

Proof of Theorem IV.4.2. We aim to apply Brezzi theory [6] to the Darcy-Stokes operator (IV.7) in the abstract form (IV.19). To this end let $V = V_f \times Q_p$ and $Q = Q_f \times X_\Gamma$ where for brevity $X_f = \frac{1}{\sqrt{\mu}}H^{-1/2}(\Gamma)$, $X_p = \sqrt{K}L^2(\Gamma)$ and we let the operators A, B be defined in terms of bilinear forms from (IV.5) as

$$(A(\mathbf{u}_f, p_p), (\mathbf{v}_f, q_p)) = \mu(\nabla\mathbf{u}_f, \nabla\mathbf{v}_f)_{\Omega_f} + D(\mathbf{u}_f \cdot \tau, \mathbf{v}_f \cdot \tau)_\Gamma + K(\nabla p_p, \nabla q_p)_{\Omega_p},$$
$$(B(\mathbf{u}_f, p_p), (q_f, w)) = (\nabla \cdot \mathbf{u}_f, q_f)_\Gamma + (T_\mathbf{n}\mathbf{u}_f, w)_\Gamma - K(\partial_{\mathbf{n},\epsilon}p_p, w)_\Gamma.$$

We proceed to verify the Brezzi conditions. Note that by assumption $|\partial\Omega_{i,D}| > 0$, $i = p, f$ so that by Poincare inequality on both subdomains A is coercive. For boundedness of A observe that $\mu(\nabla\mathbf{u}_f, \nabla\mathbf{v}_f) + D(\mathbf{u}_f \cdot \tau, \mathbf{v}_f \cdot \tau)_\Gamma < \|\mathbf{u}_f\|_{V_f}\|\mathbf{v}_f\|_{V_f}$ by Cauchy Schwarz inequality. Moreover, following Lemma IV.4.1, we have $K(\nabla p_p, \nabla q_p) \leq \|p_p\|_{Q_p}\|q_p\|_{Q_p}$. Combining the two bounds and applying the Cauchy-Schwarz inequality,

$$(A(\mathbf{u}_f, p_p), (\mathbf{v}_f, q_p)) \leq \|\mathbf{u}_f\|_{V_f}\|\mathbf{v}_f\|_{V_f} + \|p_p\|_{Q_p}\|q_p\|_{Q_p} \leq \|(\mathbf{u}_f, p_p)\|_V\|(\mathbf{v}_f, q_p)\|_V.$$

To show boundedness of B we recall that $\|\nabla \cdot \mathbf{u}_f\|_{Q_f'} = \|\nabla \cdot \mathbf{u}_f\|_{\sqrt{\mu}L^2(\Omega_f)} \leq C\|\mathbf{u}_f\|_{V_f}$, where C depends on dimensionality of Ω_f. Further, by the trace inequality $\|T_\mathbf{n}\mathbf{u}_f\|_{X_f'} = \|T_\mathbf{n}\mathbf{u}_f\|_{\sqrt{\mu}H^{\frac{1}{2}}(\Gamma)} \leq \|T_\mathbf{n}\|\|\mathbf{u}_f\|_{V_f}$, and by Assumption IV.3.1 $\|K\partial_{\mathbf{n},\epsilon}p_p\|_{X_p'} = \|K\partial_{\mathbf{n},\epsilon}p_p\|_{\frac{1}{\sqrt{K}}L^2(\Gamma)} = \|\partial_{\mathbf{n},\epsilon}p_p\|_{\sqrt{K}L^2(\Gamma)} \leq \|\partial_{\mathbf{n},\epsilon}\|\|p_p\|_{Q_p}$. Hence, per definition of dual norms,

$$\begin{aligned}(\nabla \cdot \mathbf{u}_f, q_f)_\Gamma + (T_\mathbf{n}\mathbf{u}_f, w)_\Gamma &\leq \|\nabla \cdot \mathbf{u}_f\|_{Q_f'}\|q_f\|_{Q_f} + \|T_\mathbf{n}\mathbf{u}_f\|_{X_f'}\|w\|_{X_f} \\ &\leq \max(1, C)\|\mathbf{u}_f\|_{V_f}\left(\|q_f\|_{Q_f} + \|T_\mathbf{n}\|\|w\|_{X_f}\right)\end{aligned}$$

and

$$K(\partial_{\mathbf{n},\epsilon}p_p, w)_\Gamma \leq \|K\partial_{\mathbf{n},\epsilon}p_p\|_{X_p'}\|w\|_{X_p} \leq \|\partial_{\mathbf{n},\epsilon}\|\|p_p\|_{Q_p}\|w\|_{X_p}.$$

Combining the two estimates we show boundedness of B

$$(B(\mathbf{u}_f, p_p), (q_f, w)) \leq 2\max(1, C, \|T_{\mathbf{n}}\|, \|\partial_{\mathbf{n},\epsilon}\|\|)(\mathbf{u}_f, p_p)\|_v \|(q_f, w)\|_Q.$$

Finally, we turn to the inf-sup condition. Let $R_{Q_f}^{-1}, R_{X_f}^{-1}, R_{X_p}^{-1}$ be the inverse Riesz maps of their respective spaces, so that $R_V u = (u, \cdot)_V \in V^{'}$. Let $(q_f, w) \in Q_f \times X$ be arbitrary. We first define two extensions by using the two subproblems. Recalling the notation of Assumption IV.4.1, let $(\mathbf{u}_f^*, p_f^*, \lambda*)$ be the solution of

$$a_S((\mathbf{u}_f^*, p_f^*, \lambda^*), (\mathbf{v}_f', q_f', w')) = (R_{Q_f}^{-1} q_f, q_f') + (R_{X_f}^{-1} w, w')_\Gamma \text{ for all } (\mathbf{v}_f', q_f', w') \in V_S.$$

Per assumption, there is a constant C_f so that we have the bound

$$\|\mathbf{u}_f^*\|_{V_f} \leq C_f \left(\|R_{Q_f}^{-1} q_f\|^2_{\sqrt{\mu} L^2(\Omega_f)} + \|R_{X_f}^{-1} w\|^2_{\sqrt{\mu} H^{\frac{1}{2}}(\Gamma)} \right)^{\frac{1}{2}} = C_f \left(\|q_f\|^2_{Q_f} + \|w\|^2_{X_f} \right)^{\frac{1}{2}}, \tag{IV.20}$$

where the right equality follows from the fact that $\sqrt{\mu} L^2(\Omega_f) = Q_f^{'}$, $\sqrt{\mu} H^{\frac{1}{2}}(\Gamma) = X_f^{'}$ and that the Riesz map is an isometry. Similarly, let p_p^*, λ_2^* be the solution of

$$\begin{aligned} K(\nabla p_p^*, \nabla q_p')_{\Omega_p} + K(\lambda_2^*, \partial_{\mathbf{n},\epsilon} q_p')_\Gamma + K(\partial_{\mathbf{n},\epsilon} p_p^*, w')_\Gamma = (R_{X_p}^{-1} w, w')_\Gamma \\ \text{for all} (q_p', w') \in \sqrt{K} H^1_{0,D}(\Omega_p) \times \sqrt{K} L^2(\Gamma). \end{aligned}$$

By Lemma IV.4.1, we then have the bound

$$\|p_p^*\|_{Q_p} \leq C_p \|R_{X_p}^{-1} w\|_{\frac{1}{\sqrt{K}} L^2(\Gamma)} = C_p \|w\|_{X_p} \tag{IV.21}$$

for a constant C_p. Observe now that by our definitions of $\mathbf{u}_f^*, p_p^*$,

$$\begin{aligned} &(\nabla \cdot \mathbf{u}_f^*, q_f)_\Gamma + (T_{\mathbf{n}} \mathbf{u}_f^*, w)_\Gamma - K(\partial_{\mathbf{n},\epsilon} p_p^*, w)_\Gamma \\ =&(R_{Q_f}^{-1} q_f, q_f)_\Gamma + (R_{X_f}^{-1} w, w)_\Gamma - (R_{X_p}^{-1} w, w)_\Gamma \\ =&\|q_f\|^2_{Q_f} + \|w\|^2_{X_f} + \|w\|^2_{X_p} = \|q_f\|^2 + \|w\|^2_X. \end{aligned}$$

Using (IV.20), (IV.21)

$$\begin{aligned} \|(\mathbf{u}_f^*, p_p^*)\|_V &= \left(\|\mathbf{u}_f^*\|^2_{V_f} + \|p_p^*)\|^2_{Q_p} \right)^{\frac{1}{2}} \\ &\leq C \left(\|q_f\|^2_{Q_f} + \|w\|^2_{X_f} + \|w\|^2_{X_p} \right)^{\frac{1}{2}} = C \left(\|q_f\|^2_{Q_f} + \|w\|^2_X \right)^{\frac{1}{2}}, \end{aligned}$$

where $C = \max(C_f, C_p)$. Putting this together, we can prove the inf-sup condition:

$$\begin{aligned} \sup_{(\mathbf{u}_f, p_p) \in V} &\frac{(\nabla \cdot \mathbf{u}_f, q_f)_\Gamma + (T_{\mathbf{n}} \mathbf{u}_f, w)_\Gamma - K(\partial_{\mathbf{n},\epsilon} p_p, w)_\Gamma}{\|(\mathbf{u}_f, p_p)\|_V} \\ \geq &\frac{(\nabla \cdot \mathbf{u}_f^*, q_f)_\Gamma + (T_{\mathbf{n}} \mathbf{u}_f^*, w)_\Gamma - K(\partial_{\mathbf{n},\epsilon} p_p^*, w)_\Gamma}{\|(\mathbf{u}_f^*, p_p^*)\|_V} \\ \geq &\frac{1}{C} \frac{\|q_f\|^2 + \|w\|^2_X}{\left(\|q_f\|^2_{Q_f} + \|w\|^2_X \right)^{\frac{1}{2}}} = \frac{1}{C} \|(q_f, w)\|_Q. \end{aligned}$$

Hence, the inf-sup condition holds with $\beta = \frac{1}{C}$. By Brezzi theory, Theorem IV.4.2 follows, and the problem is well-posed. As in the argument of [19], we note that due to our use of parameter weighted spaces, all constants are in fact independent of the problem parameters, and the operator preconditioner is therefore robust to parameter variations.

■

Using operator preconditioning and Theorem IV.4.2 a suitable preconditioner for the primal Darcy-Stokes problem (IV.4) is a Riesz map with respect to the inner product of W in (IV.17), that is, the operator

$$\mathcal{B} = \begin{pmatrix} -\mu\Delta + DT_\tau' T_\tau & & & \\ & K\Delta & & \\ & & \frac{1}{\mu} I & \\ & & & -\mu\,(I+\Delta)_\Gamma^{-1/2} + \frac{K}{h} I \end{pmatrix}^{-1}. \tag{IV.22}$$

We remark that all of the components of the preconditioner can be realized in an efficient, order optimal manner with multilevel schemes. In particular, the only non-standard component here is the multilevel scheme for the fractional operator which, however, has been established in [3].

Example IV.4.3 (Robust Darcy-Stokes preconditioning)**.** We consider the setup from Example IV.2.1, i.e., a P2-P1-P2-P0 discretization, while using the operator (IV.22) as preconditioner. As before, the leading blocks of the preconditioner are realized using single algebraic multigrid V-cycle. The multiplier block is then assembled using the eigenvalue decomposition and its inverse is computed by a direct solver.

The obtained iteration and condition numbers are plotted in Figure IV.4. It can be seen that both quantities are bounded in mesh size N as well as the physical parameters μ, κ and α_{BJS}.

Example IV.4.4 (Alternative choices of function spaces for the coupled problem)**.** As demonstrated in Example IV.3.3, the Darcy subproblem also appears to be stable when discretized by P2-P1 elements. According to the reasoning given in [19], we would therefore expect our preconditioner for the coupled problem to remain robust if we instead use a P1 element for the multiplier. Additionally, because the P2-P0 element is known to be stable for the Stokes problem, we would also expect that our preconditioner would remain robust if the P2-P0 element was used to discretize the Stokes subproblem instead.

Altogether, this suggests three new discretizations for the coupled problem: $V_f \times Q_f \times Q_p \times X$ = P2-P0-P2-P0, P2-P0-P2-P1, or P2-P1-P2-P1. We repeat the experiment performed in Example IV.4.3 for all three discretizations. The results for $\alpha_{\mathrm{BJS}} = 1$ are shown in Figure IV.5. It can be seen that the preconditioner appears robust both in the mesh size and in the physical parameters for all three choices of discretizations. We remark that experiments were also carried out with

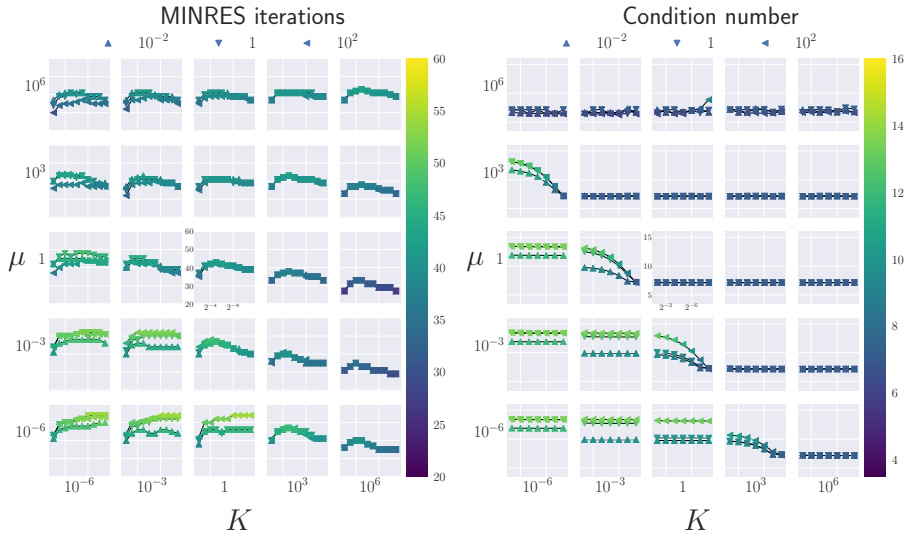

Figure IV.4: Mesh refinement vs. iteration counts (left) and condition numbers (right) for Example IV.4.3 using the preconditioner (IV.22). All subplots share x- and y-axes. For fixed μ, K the x-axis range in the iterations subplot extends from $h = 2^{-2}$ to $h = 2^{-11}$. In conditioning plots the range is from $h = 2^{-2}$ to $h = 2^{-8}$. The value of α_{BJS} is indicated by the line marker. Triangular markers on top of each other look like squares.

$\alpha_{\text{BJS}} = 10^{-2}$ and 10^{2}. The results were substantially similar to Example IV.4.3, and for legibility are therefore not shown.

Remark IV.4.5. Below we consider the validity of Assumption IV.3.1 in a continuous and discrete setting. Clearly, in a continuous setting it is easy to find a function that violates the assumption. Consider the case where $\epsilon \ll h$ while Ω and Γ are both unit sized. Further, let $u \in H^1(\Omega_p)$ be a function which is zero in $\Omega \backslash \Gamma_\epsilon$ and has a gradient of 1 in Γ_ϵ. Recalling our definition of the operator $\partial_{\mathbf{n},\epsilon} : H^1(\Omega_p) \to L^2(\Gamma_\epsilon)$ by

$$\int_\Gamma \partial_{\mathbf{n},\epsilon} u \cdot w \, ds = \frac{1}{\epsilon} \int_{\Gamma_\epsilon} \nabla u \cdot \mathbf{n}_{\Gamma_\epsilon} E_\epsilon w \, dx,$$

we see that $\partial_{\mathbf{n},\epsilon} u \in L^2(\Gamma_\epsilon)$ is the unit constant function whereas $\|u\|_1 \approx \sqrt{\epsilon}$. Hence $\|\partial_{\mathbf{n},\epsilon}\| \geq \frac{\|\partial_{\mathbf{n},\epsilon} u\|_{L^2(\Gamma_\epsilon)}}{\|u\|_{H^1(\Omega_p)}} \approx \frac{1}{\sqrt{\epsilon}}$, which is very large for small ϵ. Clearly, this function violates Assumption IV.3.1.

The above construction of a function that violates the assumption is however clearly not relevant in our discrete setting as these functions are below the

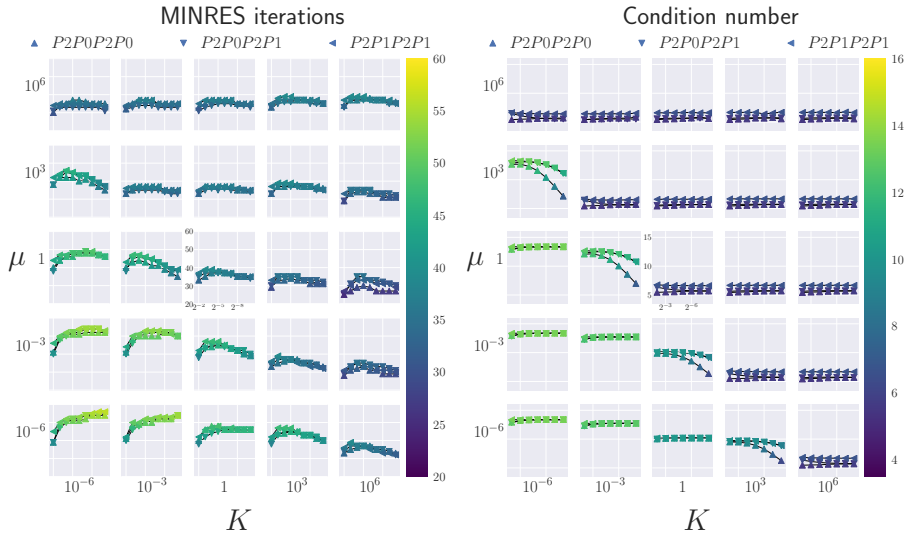

Figure IV.5: Mesh refinement vs. iteration counts (left) and condition numbers (right) for alternative discretizations. The line marker indicates the discretization used. All subplots share x- and y-axes and have $\alpha_{\mathrm{BJS}} = 1$. For fixed μ, K the x-axis range in the iterations subplot extends from $h = 2^{-2}$ to $h = 2^{-10}$. In conditioning plots the range is from $h = 2^{-2}$ to $h = 2^{-8}$.

resolution of our finite element mesh. Indeed, in our numerical experiments, we use discrete subspaces of $H^1(\Omega_p)$, so that any function whose gradient is nonzero on Γ also has nonzero gradient at distance h from Γ. This means that if ϵ is chosen smaller than h, functions like u above which are zero immediately outside of Γ_ϵ are not admissible.

For a relevant finite element function u_h, constructed as above, i.e., such that u_h is zero everywhere except having a gradient of 1 on the finite elements with facets on Γ, assuming that $\epsilon \ll h$, we have $\frac{\|\partial_{\mathbf{n},\epsilon} u_h\|_{L^2(\Gamma_\epsilon)}}{\|u_h\|_{H^1(\Omega_p)}} \approx \frac{1}{\sqrt{h}}$. Indeed this estimate corresponds to the scaling shown in the Examples IV.3.2 and IV.3.3.

Acknowledgments

Karl Erik Holter is a doctoral fellow in the Simula-UCSD-University of Oslo Research and PhD training (SUURPh) program, an international collaboration in computational biology and medicine funded by the Norwegian Ministry of Education and Research. Miroslav Kuchta acknowledges support by the Research Council of Norway (NFR) grant 280709. Kent-Andre Mardal acknowledges support by the Research Council of Norway (NFR) grant 280709 and 301013.

References

[1] Arbogast, T. and Brunson, D. S. “A computational method for approximating a Darcy–Stokes system governing a vuggy porous medium”. In: *Computational geosciences* vol. 11, no. 3 (2007), pp. 207–218.

[2] Babuška, I. “The finite element method with Lagrangian multipliers”. In: *Numerische Mathematik* vol. 20, no. 3 (1973), pp. 179–192.

[3] Bærland, T., Kuchta, M., and Mardal, K.-A. “Multigrid Methods for Discrete Fractional Sobolev Spaces”. In: *SIAM Journal on Scientific Computing* vol. 41, no. 2 (2019), A948–A972.

[4] Balay, S. et al. *PETSc Web page*. https://www.mcs.anl.gov/petsc. 2019.

[5] Bramble, J., Pasciak, J., and Vassilevski, P. “Computational scales of Sobolev norms with application to preconditioning”. In: *Mathematics of Computation* vol. 69, no. 230 (2000), pp. 463–480.

[6] Brezzi, F. “On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers”. In: *Publications mathématiques et informatique de Rennes*, no. S4 (1974), pp. 1–26.

[7] Burman, E. and Hansbo, P. “A unified stabilized method for Stokes’ and Darcy’s equations”. In: *Journal of Computational and Applied Mathematics* vol. 198, no. 1 (2007), pp. 35–51.

[8] Cai, M., Mu, M., and Xu, J. “Preconditioning techniques for a mixed Stokes/Darcy model in porous media applications”. In: *Journal of computational and applied mathematics* vol. 233, no. 2 (2009), pp. 346–355.

[9] Ding, Z. “A proof of the trace theorem of Sobolev spaces on Lipschitz domains”. In: *Proceedings of the American Mathematical Society* vol. 124, no. 2 (1996), pp. 591–600.

[10] Discacciati, M. and Quarteroni, A. “Analysis of a domain decomposition method for the coupling of Stokes and Darcy equations”. In: *Numerical mathematics and advanced applications*. Springer, 2003, pp. 3–20.

[11] Discacciati, M. and Quarteroni, A. “Navier-Stokes/Darcy coupling: modeling, analysis, and numerical approximation”. In: (2009).

[12] Discacciati, M., Quarteroni, A., and Valli, A. “Robin–Robin domain decomposition methods for the Stokes–Darcy coupling”. In: *SIAM Journal on Numerical Analysis* vol. 45, no. 3 (2007), pp. 1246–1268.

[13] Falgout, R. D. and Yang, U. M. “hypre: A library of high performance preconditioners”. In: *International Conference on Computational Science*. Springer. 2002, pp. 632–641.

[14] Feng, M.-f. et al. “Stabilized Crouzeix-Raviart element for the coupled Stokes and Darcy problem”. In: *Applied Mathematics and Mechanics* vol. 31, no. 3 (2010), pp. 393–404.

[15] Galvis, J. and Sarkis, M. "Non-matching mortar discretization analysis for the coupling Stokes-Darcy equations". In: *Electron. Trans. Numer. Anal* vol. 26, no. 20 (2007), p. 07.

[16] Gatica, G. N., Meddahi, S., and Oyarzúa, R. "A conforming mixed finite-element method for the coupling of fluid flow with porous media flow". In: *IMA Journal of Numerical Analysis* vol. 29, no. 1 (2008), pp. 86–108.

[17] Hernandez, V., Roman, J. E., and Vidal, V. "SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems". In: *ACM Trans. Math. Software* vol. 31, no. 3 (2005), pp. 351–362.

[18] Holter, K. E., Kuchta, M., and Mardal, K.-A. "Sub-voxel perfusion modeling in terms of coupled 3d-1d problem". In: *European Conference on Numerical Mathematics and Advanced Applications*. Springer. 2017, pp. 35–47.

[19] Holter Kuchta, M. "Robust preconditioning of monolithically coupled multiphysics problems". In: *ArXiv* (2020).

[20] Hong, Q. and Kraus, J. "Uniformly stable discontinuous Galerkin discretization and robust iterative solution methods for the Brinkman problem". In: *SIAM Journal on Numerical Analysis* vol. 54, no. 5 (2016), pp. 2750–2774.

[21] Hong, Q. et al. "A robust multigrid method for discontinuous Galerkin discretizations of Stokes and linear elasticity equations". In: *Numerische Mathematik* vol. 132, no. 1 (2016), pp. 23–49.

[22] Hughes, T. J., Cottrell, J. A., and Bazilevs, Y. "Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement". In: *Computer methods in applied mechanics and engineering* vol. 194, no. 39-41 (2005), pp. 4135–4195.

[23] Johnny, G. and Michael, N. "A family of nonconforming elements for the Brinkman problem". In: *IMA Journal of Numerical Analysis* vol. 32, no. 4 (2012), pp. 1484–1508.

[24] Karper, T., Mardal, K.-A., and Winther, R. "Unified finite element discretizations of coupled Darcy–Stokes flow". In: *Numerical Methods for Partial Differential Equations: An International Journal* vol. 25, no. 2 (2009), pp. 311–326.

[25] Kuchta, M. "Assembly of multiscale linear PDE operators". In: *arXiv preprint arXiv:1912.09319* (2019).

[26] Kuchta, M. et al. "Preconditioners for saddle point systems with trace constraints coupling 2d and 1d domains". In: *SIAM Journal on Scientific Computing* vol. 38, no. 6 (2016), B962–B987.

[27] Layton, W. J., Schieweck, F., and Yotov, I. "Coupling fluid flow with porous media flow". In: *SIAM Journal on Numerical Analysis* vol. 40, no. 6 (2002), pp. 2195–2218.

[28] Logg, A., Mardal, K., and Wells, G. *Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book*. Lecture Notes in Computational Science and Engineering. Springer Berlin Heidelberg, 2012.

[29] Mardal, K. A., Tai, X.-C., and Winther, R. "A robust finite element method for Darcy–Stokes flow". In: *SIAM Journal on Numerical Analysis* vol. 40, no. 5 (2002), pp. 1605–1631.

[30] Mardal, K.-A. and Winther, R. "Preconditioning discretizations of systems of partial differential equations". In: *Numerical Linear Algebra with Applications* vol. 18, no. 1 (2011), pp. 1–40.

[31] Mikelic, A. and Jäger, W. "On the interface boundary condition of Beavers, Joseph, and Saffman". In: *SIAM Journal on Applied Mathematics* vol. 60, no. 4 (2000), pp. 1111–1127.

[32] Nilssen, T., Tai, X.-C., and Winther, R. "A robust nonconforming H^2-element". In: *Mathematics of Computation* vol. 70, no. 234 (2001), pp. 489–505.

[33] Rivière, B. and Yotov, I. "Locally conservative coupling of Stokes and Darcy flows". In: *SIAM Journal on Numerical Analysis* vol. 42, no. 5 (2005), pp. 1959–1977.

[34] Xie, X., Xu, J., and Xue, G. "Uniformly-stable finite element methods for Darcy-Stokes-Brinkman models". In: *Journal of Computational Mathematics* (2008), pp. 437–455.

[35] Zhang, S., Xie, X., and Chen, Y. "Low order nonconforming rectangular finite element methods for Darcy-Stokes problems". In: *Journal of Computational Mathematics* (2009), pp. 400–424.