STABILIZED FINITE ELEMENT METHODS FOR COUPLED FLOW AND
GEOMECHANICS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
CIVIL AND ENVIRONMENTAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Joshua A. White
August 2009

UMI Number: 3382958

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy
submitted.   Broken or indistinct print, colored or poor quality illustrations
and photographs, print bleed-through, substandard margins, and improper
alignment can adversely affect reproduction.
In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted.   Also, if unauthorized
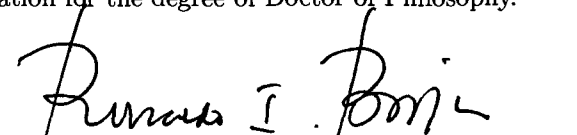copyright material had to be removed, a note will indicate the deletion.

UMI®

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Ronaldo I. Borja)   Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Peter K. Kitanidis)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.
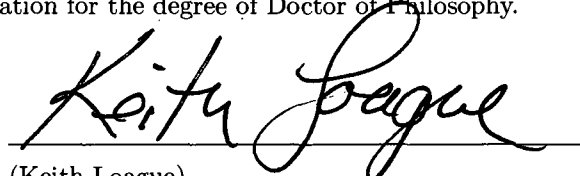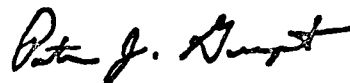
_____

(Keith Loague)

Approved for the University Committee on Graduate Studies.

iii

Abstract

In this work, we present a finite element formulation for variably-saturated porous geomaterials undergoing elastoplastic deformations. The deforming body is treated as a multiphase continuum, and the governing mass and momentum balance equations are solved in a fully-coupled manner. It is well-known, however, that mixed formulations of the type examined here may lead to unstable approximations unless the spaces chosen for the pressure and displacement interpolation satisfy stringent stability restrictions. Failure to choose a stable pair typically leads to spurious pressure oscillations and poor convergence behavior. Unfortunately, many seemingly natural combinations— including equal-order interpolation for all field variables—do not satisfy the necessary requirements. In this work, we propose a stabilized formulation, based on a minor modification of the variational equations, which allows one to circumvent these restrictions and employ equal-order mixed elements. Several numerical examples are used to demonstrate the computationally appealing features of this alternative formulation.

The resulting implicit, nonlinear algebraic systems are then solved using an inexact Newton algorithm. We discuss methods for solving the linearized systems using memory-efficient iterative solvers, both on serial and parallel computing platforms. In order to deal with inherent ill-conditioning, we propose a block-structured, multilevel preconditioner that both accelerates the convergence of the Krylov solver and exhibits excellent scaling properties as the number of unknowns and number of processors increase.

To demonstrate the effectiveness of these approaches, the analysis framework is applied to modeling hydrologically-driven slope failure. This analysis is motivated by a recent landslide that occurred at a steep experimental catchment (CB1) near Coos Bay, Oregon. Simulations are used to quantify the rainfall-induced slope deformation and assess the failure potential. Results of parametric studies suggest that for a steep hillside slope underlain by shallow bedrock similar to the CB1 site, failure would occur by a multiple slide block mechanism, with progressive failure surfaces forming at the bedrock interface and then propagating to the slope surface. A key observation is that significant computational resources are required to capture these complex solid/fluid interaction mechanisms at sufficient resolution, further justifying the use of the proposed approaches over conventional methods.

Acknowledgements

This work would not have been possible without the support of many people.

First and foremost, I would like to thank my advisor, Ronnie Borja. His support and encouragement over the past five years have been invaluable. I would also like to thank Joanne Fredrich, who hosted me for several summers and with whom it has always been a pleasure to collaborate. I wish to thank the members of my dissertation committee—Profs. Kitanidis, Loague, Nur, and Pollard—for their insight and their helpful comments on my work. This dissertation would also not have been possible without funding support through a William K. Bowes Jr. Stanford Graduate Fellowship and a National Science Foundation Graduate Research Fellowship.

I want to thank my family—Mom, Dad, Billy, and Charlotte—who have always encouraged me in my endeavours, whatever they may be. I also want to thank all of my friends at Stanford, who have made the past few years fly by. Finally, I want to thank Dashi, who helped in countless ways—both large and small—and who has always been there for me.

Contents

List of Tables

List of Figures

1 Introduction

The focus of this dissertation is the development of efficient numerical methods for modeling variably-saturated geomaterials. In particular, we examine fully-coupled models that simultaneously capture the deformation of the solid matrix and fluid flow through the pore space. In many geotechnical and geoscientific applications, it is necessary to model this fluid-structure interaction in order to make meaningful predictions. Unfortunately, these multiphysics models are challenging both from a theoretical and numerical point of view, and their application in day-to-day practice has been extremely limited. The goal of this work is to address some of these challenges.

This introductory chapter has four basic goals:

- First, we introduce a few representative applications of fluid-structure interaction, to motivate the usefulness of fully-coupled simulations. These applications are chosen from a variety of disciplines, though they are biased toward geotechnical and geoscientific applications.

- Second, we present a grain-scale description of porous media, to highlight some of the microscale mechanisms we wish to capture in our macroscale (continuum) models.

- Third, we discuss the basic challenges to developing meaningful numerical models. These challenges can loosely be grouped as theoretical, experimental, and computational challenges. While this dissertation focuses heavily on numerical methods, theory and experiment obviously play a crucial role.

- Finally, we outline the remainder of the dissertation, so the reader may know what to expect in each chapter and understand the unifying themes.

1.1 Applications

We now consider a few representative applications of fully-coupled flow and geomechanics. Fluid-filled porous media appear in many disciplines, however, and so the applicable scope of these models is much broader than the applications listed here.

1.1.1 Civil Engineering

An understanding of coupled solid-fluid flow behavior is crucial for assessing the safety and serviceability of many geotechnical projects—e.g. foundations, excavations, cuts, embankments, and dams.

Table 1.1: Major landslides and debris flows in recent decades

Year	Location	Rainfall Observations	Lives Lost
2006	Guinsaugon, Philippines	2.0 m / 10 days	1126
1999	Vargas, Venezuela	0.9 m / 3 days	10,000[1,2]
1991	Antofagasta, Chile	Brief torrential rain, peak rate 60 mm/h	101
1988	Rio de Janeiro, Brazil	3 weeks of tropical storm conditions	320
1987	Rio Limon, Venezuela	0.2 m / 5 hours	210
1985	Mameyes, Puerto Rico	0.6 m / 24 hours	129

[1] Includes losses due to flooding.
[2] Low estimate. Actual losses were difficult to determine due to poor census data.

Numerical simulations provide a useful tool in the design stage of new projects, as well as for the analysis of already completed works.

Unfortunately, many dramatic geotechnical failures have been attributed to fluid-structure interaction. A classic example is the effect of the 1971 San Fernando Earthquake on the lower and upper San Fernando Dams. The lower dam experienced a major flow failure, while the upper dam experienced significant but limited downstream movement. Given the close proximity of the two dams and similar construction materials, the question naturally arises as to what led to the differing performance. The investigation by Seed et al. [1973] immediately after the event indicated that liquefaction and weakening of the hydraulic fill had occurred in both dams. In the upper dam, however, the upstream and downstream shells remained relatively strong, while in the lower dam a large part of the embankment near the base of the dam had liquefied and led to flow failure. In this context, coupled finite element analyses are useful for understanding the liquefaction mechanisms and designing remediation measures—see, e.g., the coupled analysis in [Ming and Li 2003].

1.1.2 Landslides and Debris Flows

Hydrologically-driven slope instability threatens lives and infrastructure worldwide. Table 1.1 provides a representative list of six major events in recent decades [Jibson 1992; Lagmay et al. 2006; Martinez 2000; Schuster et al. 2002; USAID 2000; Van Sint Jan and Talloni 1993]. A recent case is the 2006 Guinsaugon, Philippines slides: heavy and persistent rainfall over a ten day period triggered several slides, killing 1126 people and destroying vital infrastructure.

Despite decades of extensive slope stability model development, the fundamental controls connecting the hydrologic and geotechnical processes triggering slope failure are still not well quantified. This is evident from the La Conchita landslide of January 10, 2005 in southern California (see Figures 1.1 and 1.2). The slide occurred without warning in a populated area—yet the site was well-known to be susceptible to landsliding from previous events [Jibson 2005].

This lack of understanding is partially the result of the simplified physics in current models, with
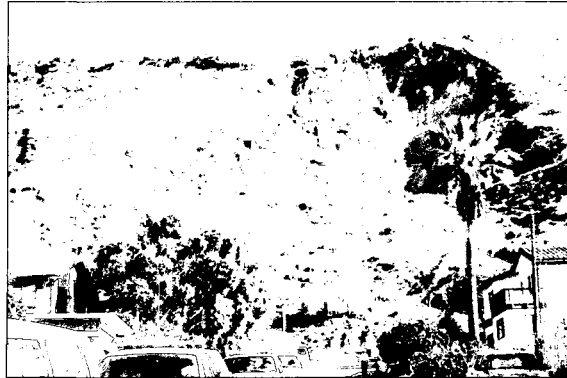
Figure 1.1: Photograph of the 2005 La Conchita landslide (from [Jibson 2005]).
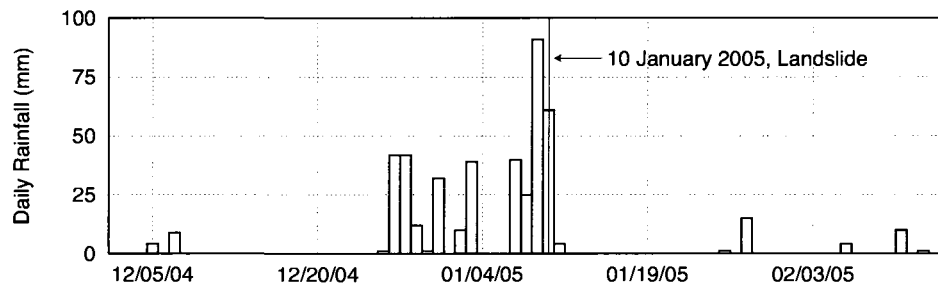
Figure 1.2: Rainfall record in the weeks leading up to the La Concita landslide.

the omission of the effect of partial saturation from slope stability calculations. It is known that increasing the degree of saturation decreases the capillary pressure, which in turn weakens the slope. Despite the expected significant impact, this interplay between fluid flow, increases in saturation, and loss of strength is typically not well-accounted for.

1.1.3 Fault-zone Processes

Solid-fluid coupling is of particular interest when studying fault zone processes, and is central to many open questions about fault behavior. The presence of fluids might explain why some faults, such as the San Andreas, are weaker than expected [Hardebeck and Hauksson 1999; Sleep and Blanpied 1992]. Increases in pore pressure may tend to weaken faults by reducing the effective normal stress, and trigger seismic activity. If the overpressures are too large, however, the fault could experience stable, rather than unstable, sliding [Segall and Rice 1995]. Dilatancy or compaction within the fault zone will also play a crucial role, as well as the degree to which fluid exchange is allowed to occur between the fault and its surroundings. Numerical simulations are especially useful for testing hypotheses in these cases since field observation is necessarily limited.

1.1.4 Reservoir Engineering

Reservoir engineering also represents another broad area where coupled geomechanical simulations can prove useful.

Our current energy economy, for example, is vitally dependent on subsurface hydrocarbon reservoirs. The ability to accurately model the behavior of oil and natural gas therefore has large commercial consequences. Coupled simulations can be used to assess the optimal placement and orientation of wellbores, both for maximizing fluid recovery and to ensure wellbore stability in the prevailing stress field. When callibrated with historical performance data, a numerical model can also be used to assess current and future reservoir performance, accounting for compaction drive and similar effects. Finally, as a result of deep-water exploration, significant investments have been made is seafloor pipelines and infrastructure. This infrastructure is susceptible to submarine landslides and wave-loading failures. Numerical models therefore provide a useful tool for seafloor hazard assessment.

In the future, a societal priority will be maintaining access to clean drinking water. Therefore, the sustainable management of our subsurface aquifers is critically important. Also, as we look beyond our current energy mix to a more sustainable model, coupled subsurface simulations can prove invaluable. Technologies likes subsurface CO_2 sequestration, in-situ coal gasification, and enhanced geothermal energy all require detailed understanding of coupled behavior. Many of these processes include other couplings as well (e.g. thermal and reactive-transport) and so there are significant research needs in the area of efficient numerical methods for multiphysics models.

1.2 Micromechanics of Porous Media

In studying the coupled behavior of saturated and unsaturated porous media, it is useful to begin with a grain-scale description. Before we start, however, we remark that while an understanding of the micromechanical mechanisms is crucial to successful numerical modeling, in this work we are ultimately interested in field scale applications. As a result, we will eventually introduce a macroscale (continuum) mathematical description of the problem. For the remainder of the dissertation we will then work within this continuum framework. While not the focus of this thesis, however, micromechanical methods can be tremendously useful in their own right—particularly for estimating constitutive properties to be later inserted into continuum-scale models. See, for example, [Fredrich et al. 2006; White et al. 2006].

Let us consider a typical three-phase medium with solid, water, and air constituents (Figure 1.3). The solid, granular phase is packed together to form a structural skeleton, while the fluid and gas phases reside in the pore space. As a result of capillary suction in this unsaturated material, the wetting phase (water) is drawn into the smaller crevices and pores. In the limit as the pore water fills all of the voids–except, perhaps, for some small, irreducible air saturation—we obtain a saturated material. In this work we are concerned with soils and rocks exhibiting the entire range from irreducible water saturation to irreducible air saturation, and so will generically refer to them as
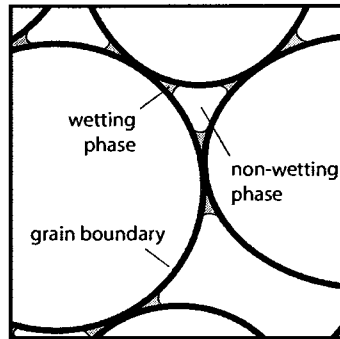
Figure 1.3: Schematic illustration of the microstructure of a three-phase porous medium.

variably-saturated porous media. We will use this simple, three-phase model to illustrate many of the key features we will be concerned with in this work. In our discussion, let us loosely subdivide these aspects into solid mechanisms, fluid mechanisms, and fluid-structure interaction mechanisms.

First, consider the behavior of the porous skeleton in the absence of any pore fluid interaction. In this case, the microstructure has a granular composition for which we would like to know the stress and deformation characteristics. Consider the behavior of the medium under increasing levels of volumetric and deviatoric stress. At low stresses, the structure may have sufficient inter-grain contact forces (due to pure friction or some cementation) for the structure to remain intact without any inter-grain slip. At larger stresses, however, we expect to observe grain rearrangement as grains roll and slide past one another to acommodate increasing stresses and deformations. At very large stresses, we may enter a regime where the grains themselves begin to fracture and crush. Note that as the microstructural fabric changes, we also have a corresponding change in the pore network topology. The macroscale, observed specimen response is likely to be highly nonlinear—particularly at large stresses.

Due to grain rearrangement, the volumetric and deviatoric responses in granular materials are strongly coupled. Consider the volumetric behavior of the medium as it is sheared. If the medium originally had a loose packing, it will likely compact when sheared, with a corresponding decrease in void ratio. If, on the other hand, the original packing was quite dense, then the material may actually dilate as grains shift and roll over one another to accommodate the deformation. Furthermore, a commonly observed feature of geomaterials is that, regardless of the initial density, under large shear strains a steady "critical state" is reached where increasing shear deformations can be accomodated at constant volume—i.e. as isochoric response. Figure 1.4 is a schematic illustration of this behavior for typical loose and dense granular materials. These basic ideas forms the foundation of Critical State Soil Mechanics [Roscoe and Burland 1968; Roscoe et al. 1958; Schofield and Wroth 1968], a conceptual framework that has led to the development of many useful material models.

Even the simple skeleton drawn in Figure 1.3 will show a decidedly complex constitutive behavior. We should also bear in mind that real porous materials are geometrically more complex than our simple drawing suggests. For comparison, the left image in Figure 1.5 is a segmented
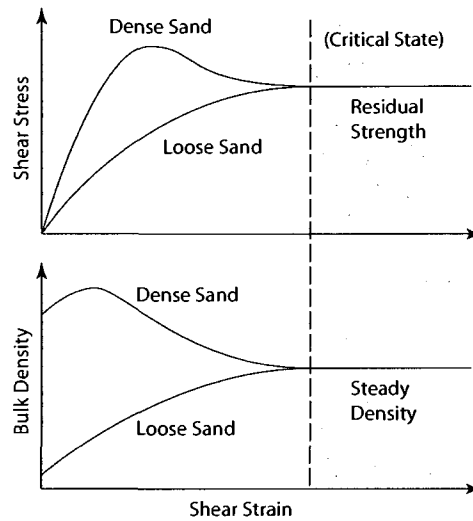
Figure 1.4: Schematic illustration of the behavior of loose and dense sands when sheared (after [Iverson et al. 1997]).

microtomographic image of the pore-space of a Castlegate sandstone [Fredrich et al. 2006]. Note that this image is inverted, so that voids are opaque and the solid phase is transparent. We can immediately see the striking topological complexity of real geomaterials, and naturally expect that they will behave in a highly nonlinear manner.

Now, consider the behavior of the fluid phases as they flow through the porous skeleton. For this discussion, we assume that the skeleton remains rigid and undeformable. As a result of gravitational and pressure potentials, the fluids will migrate from one region to another. Because of the topological complexity of the pore network, however, these flow paths become tremendously complex. The right image in Figure 1.5 shows the velocity magnitude of a single-phase fluid flowing through the Castlegate pore network. The velocity field was computed using a lattice-Boltzmann simulation on a computational grid mapped directly from the microtomographic data set [Fredrich et al. 2006; White and Fredrich 2007]. Even a cursory analysis of these flow paths reveals their tortuous nature. Furthermore, we observe a few, well-connected, high-flux pathways surrounded by significant regions of stagnant fluid. If we consider the possibility of multiple fluid phases in the pore space, then the story becomes even more complex as we must account for relative flow behavior of the multiple fluids, as well as a host of more sophisticated interaction mechanisms.

As we have seen, the separate solid and fluid problems are intrinsically challenging. Clearly, however, the solid deformation and fluid flow will also have mutual interaction effects. Let us briefly describe two of these interaction mechanisms, though there are many more we might consider.

The first interaction mechanism appears in the form of deformation constraints. As described earlier, geomaterials show a complex volumetric response. If the there are fluids in the pore space, they will play a role in controlling this dilation and compaction. Consider, for example, a saturated
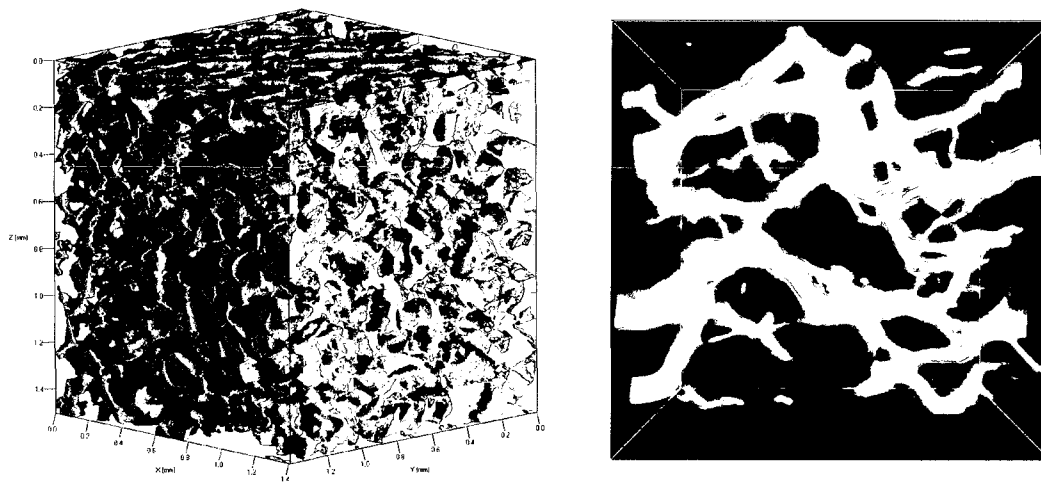
Figure 1.5: Segmented microtomographic image of the pore space of Castlegate sandstone [Fredrich et al. 2006], and a lattice-Boltzmann simulation of single-phase flow paths through the pore network [White and Fredrich 2007]. Note that the figures are not to scale. The microtomographic volume is $1.416 \times 1.416 \times 1.503$ mm^3, while the simulation image is of a $0.668 \times 0.688 \times 0.688$ mm^3 subvolume.

porous soil or rock whose permeability is very low in comparison to the time scale of loading. In this case, as the skeleton deforms the pore fluid cannot drain away to other regions. Since the bulk modulus of the fluid is typically very large (for water, $K_w \approx 2.2$ GPa), the trapped fluid will prevent the porous skeleton from undergoing any significant volumetric deformation. Thus, the mixture as a whole may act as an incompressible material. The numerical challenge of dealing with this constraint is the subject of Chapter 3.

Second, the presence of either positive or negative pore water pressure has a significant effect on the apparent strength of the geomaterials. It is commonly observed, for example, that the shear capacity of a specimen increases with effective confining pressure. The simplest model of this effect is the Mohr-Coulomb failure criterion, in which the shear capacity is a linear function of confining pressure. Now, if significant positive pore water pressure is present, it will tend to relieve the effective stress that would otherwise be carried by the solid matrix. It will decrease the effective hydrostatic stress in the skeleton, while leaving the deviator stress unaffected. As a result, positive pressure can have a destabilizing effect, reducing the intrinsic shear capacity of the material. For soils with little to no cohesion, the pore water pressure may even become so large that the soil liquefies—as was the case with the lower San Fernando dam. In a similar vein, consider the behavior of an unsaturated soil. In this case, there is capillary suction in the pore space due to the presence of both air and water phases. This suction will tend to pull grains together, and has a stabilizing effect on the porous medium. That is, the material in an unsaturated state exhibits a larger apparent strength than in the fully-saturated state. The classic example of this fact is sand-castle building. Completely dry or fully-saturated sand is almost useless as a building material, but at moderate saturations the sand can be quite strong. These effects will play a central role in Chapter 5. There, we consider

rainfall-induced landslides, an application in which loss of strength due to saturation increases can have dramatic and sometimes devastating effects.

In this dissertation we are concerned with modeling field-scale behavior. At this point we will therefore move away from our grain-scale description and adopt a macroscale (continuum) perspective. Nevertheless, the previous microscale discussion is useful in two respects. First, it provides an important conceptual framework for understanding the macroscale phenomena we would like to capture. Second, it illustrates the fundamental challenge of trying to represent very complex microscale interactions with macroscale, phenomenological models.

Finally, we also wish to emphasize that the macroscale/microscale division we have adopted here is somewhat artificial. In reality there is a continuum of scales, with physical processes and heterogeneities at many levels playing a role in the global system response. Ultimately, the success and failure of geomechanical models lies in the ability to capture the important effects at every scale in a meaningful (and tractable) manner.

1.3 Challenges and Research Needs

We now turn our attention to some of the basic challenges associated with modeling coupled systems. To ground the discussion, we introduce the simplest model: a saturated, quasi-static material with incompressible solid and fluid constituent phases. In this case, the governing equations are

$$\nabla \cdot \boldsymbol{\sigma}' - \nabla p + \rho \boldsymbol{g} = \boldsymbol{0} \tag{1.1}$$

$$\nabla \cdot \dot{\boldsymbol{u}} + \nabla \cdot \boldsymbol{w} = 0 \tag{1.2}$$

We will discuss coupled formulations in great detail in Chapter 2, and only introduce the model here to fix some ideas. The first equation is a linear momentum balance, relating the effective stress in the solid matrix $\boldsymbol{\sigma}'$, the pore water pressure p, and a body force due to the mixture density ρ. The second equation is a mass balance relating the displacement velocity $\dot{\boldsymbol{u}}$ and the seepage velocity $\boldsymbol{w}$. This model must be supplemented with suitable initial and boundary conditions, and closed with two constitutive relationships. The first relates the seepage velocity to the pressure, for which we use a generalized Darcy's law,

$$\boldsymbol{w} = -\frac{\boldsymbol{k}}{\eta} \left(\nabla p - \rho_w \boldsymbol{g} \right) \tag{1.3}$$

Here, we have introduced the intrinsic permeability $\boldsymbol{k}$ (units of m^2), the dynamic fluid viscosity η (units of Pa·s), and the fluid density ρ_w. The second constitutive relation relates the displacement $\boldsymbol{u}$ to the effective stress $\boldsymbol{\sigma}'$. Under a small-strain assumption, most hyperelastic and elastoplastic material models can be written in rate form as

$$\dot{\boldsymbol{\sigma}}' = \mathbb{C} : \dot{\boldsymbol{\epsilon}}, \qquad \dot{\boldsymbol{\epsilon}} = \frac{\nabla \dot{\boldsymbol{u}} + \nabla^\mathsf{T} \dot{\boldsymbol{u}}}{2} \tag{1.4}$$

Here, the strain rate $\dot{\epsilon}$ is the symmetric gradient of the displacement velocity $\dot{u}$, $\mathbb{C}$ is a non-constant fourth-order tensor of tangential moduli, and the colon operator : denotes a double contraction. In general, this coupled model can only be tackled analytically for simple geometries and material models. In most cases, a numerical method is required.

Equations (1.1) and (1.2) can be derived from the continuum theory of mixtures, and provide a physics-based rational for modeling coupled geomaterial behavior. There are three important steps to go through, however, to get from this mathematical framework to a working model:

- First, we need numerical methods for solving (1.1) and (1.2). These methods must be both accurate and inexpensive.

- Second, there is a constitutive modeling effort—that is, equations (1.3) and (1.4). The chosen models should be predictive, theoretically grounded, and able to be constrained with available field data.

- Finally, there is the application phase, in which one defines the geometry, initial and boundary conditions, and other features specific to the given problem. This phase likely also includes validation and revision efforts.

None of these phases is trivial, and demand significant research attention. Furthermore, the construction of meaningful numerical models is clearly more than an exercise in numerics: theory and experiment play a central (and inescapable) role. Thus, while this thesis is focused on some of the numerical challenges associated with coupled problems, we also need to address these other issues. This fact will be most apparent in Chapter 5, where we look at a specific application—studying hydrologically-driven landslides.

With respect to designing efficient methods (Step 1) there are three basic challenges that we attempt to address in this work. The first challenge is the obvious one: coupled problems are simply "larger" because we must simultaneously model multiple physical mechanisms and multiple unknown fields. There are other challenges, though, that are more subtle. First, the mixed finite element formulations we consider are subject to certain mathematical stability restrictions. These restrictions limit the available choice of interpolations one can use and still obtain good, convergent approximations. Unfortunately, interpolations that do "work" are often inconvenient and expensive. Fortunately, there are methods for circumventing these stability restrictions, and they will be the focus of Chapter 3. The second issue is that coupled problems are inherently stiff and ill-conditioned. This ill-conditioning derives from the fact that the individual governing equations have widely-differing magnitudes and mathematical character. The focus of Chapter 4 is the development of intelligent solution methods that recognize the coupled nature of the problem and exploit it in the solution process. We will then tie these methods together in Chapter 5 to look at a the specific application of hydrologically-driven landslides.

1.4   Outline

With this background in mind, the remainder of this dissertation is organized as follows:

- **Chapter 2: Multiphase Mechanics.** This chapter introduces a mathematical framework for modeling multiphase materials based on continuum theory of mixtures. We discuss the general formulation, as well as simplified models that are useful in practice. We also present the basics of the numerical discretization using finite elements.

- **Chapter 3: Stabilized FEM.** In this chapter we address the inherent stability restrictions embedded in mixed finite element formulations of coupled behavior. These restrictions are one of the major hurdles to computational efficiency. Standard methods that satisfy these restrictions are quite expensive, and limit the size and resolution of coupled models. In this work, we propose a "stabilized" finite element method that circumvents these restrictions, and allows for the successful use of efficient, low-order elements. We compare the accuracy and the performance of the intrinsically-stable and new stabilized methods with several numerical examples.

- **Chapter 4: Solution Methods.** In this chapter, we explore solution methods for the coupled model that address the size, stiffness, and ill-conditioning of coupled systems. We propose a block-structured preconditioning approach that speeds the convergence of iterative solvers and shows excellent scaling properties as the problem size grows. We also explore improvements to the standard Newton iteration that provide additional boosts in performance. When combined with the stabilized finite element formulation, we are able to solve three-dimensional, fully-coupled problems with hundreds of thousands to millions of unknowns on relatively modest computational platforms.

- **Chapter 5: Hydrologically-Driven Slope Instability.** In this chapter, we apply numerical simulations to study hydrologically-driven slope failure. The motivation for this work is the rainfall-induced failure that occurred at a steep experimental catchment (CB1) near Coos-Bay, Oregon, in 1996. Through several numerical examples we quantify the deformation and stability of the slope, assess it failure potential, and discuss the kinematics of the likely failure mechanisms. Through parametric studies we also try to identify the key uncertainties that need to be constrained in future modeling efforts.

- **Chapter 6: Conclusions.** At the end, we will identify the unifying themes and key contributions of this dissertation, as well as lay out important areas that should be addressed in future work.

2 Multiphase Mechanics

This chapter introduces a mathematical framework for modeling multiphase materials based on continuum theory of mixtures. We discuss a very general formulation, as well as simplified models that are useful in practice. We also present the basics of the numerical discretization using finite elements.

2.1 Conservation Laws

Consider a representative elementary volume (REV) of a three phase mixture containing solid, water, and air. We have chosen these three phases since they are the ones commonly present in geotechnical applications. Most of what follows, however, can be readily extended to multiphase mixtures with other fluid and gas constituents.

The total volume of the REV is $dv = dv_s + dv_w + dv_a$. The volume fraction of the π-phase ($\pi = s, w, a$) is denoted by $\phi^\pi = dv_\pi / dv$, with the restriction

$$\phi^s + \phi^w + \phi^a = 1. \tag{2.1}$$

The void fractions for the water and air phases are given by

$$\psi^w = \frac{\phi^w}{1 - \phi^s} \qquad \text{and} \qquad \psi^a = 1 - \psi^w. \tag{2.2}$$

Note that the void fraction ψ^w is also commonly referred to as the the water saturation, and ψ^a as the air saturation. The intrinsic mass density of each phase is ρ_π, while the averaged (or partial) density of the phase within the REV is given by $\rho^\pi = \phi^\pi \rho_\pi$. The total mass density of the mixture is therefore

$$\rho = \rho^s + \rho^w + \rho^a. \tag{2.3}$$

As a general notation, we will use phase designations in the superscript for average or partial quantities, and in the subscript for intrinsic or true quantities.

2.1.1 Mass Balance

For simplicity, we assume that no mass is exchanged between phases. In this case, we can write the balance of mass for each phase as

$$\frac{d^\pi \rho^\pi}{dt} + \rho^\pi \, \nabla \cdot \boldsymbol{v}_\pi = 0, \qquad \pi = s, w, a. \tag{2.4}$$

Here, the total derivative $d^\pi(\circ)/dt$ is reckoned with respect to the motion of the π-phase. In practice, however, it is convenient to reckon all quantities with respect to the motion of a single phase, typically the solid. We can do so using the transformation,

$$\frac{d^s(\circ)}{dt} = \frac{d^\pi(\circ)}{dt} + \nabla(\circ) \cdot \boldsymbol{v}_{\pi s}, \qquad \boldsymbol{v}_{\pi s} = \boldsymbol{v}_\pi - \boldsymbol{v}_s, \tag{2.5}$$

where $\boldsymbol{v}_{\pi s}$ is the relative velocity of the π-phase with respect to the solid. Given the central importance of the solid motion, from now on we drop the superscript on $d^s(\circ)/dt$ and simply write $d(\circ)/dt$. Similarly, we will refer to $\boldsymbol{v}_s$ simply as $\boldsymbol{v}$, with the connection to the solid implicitly understood. With this new notation, we can rewrite the conservation equations as

$$\frac{d\rho^s}{dt} + \rho^s \, \nabla \cdot \boldsymbol{v} = 0, \tag{2.6}$$

$$\frac{d\rho^\pi}{dt} + \rho^\pi \, \nabla \cdot \boldsymbol{v} + \nabla \cdot (\rho^\pi \boldsymbol{v}_{\pi s}) = 0, \qquad \pi = w, a. \tag{2.7}$$

We now introduced an equation of state for the density/pressure relation in each phase through a barotropic flow assumption: the intrinsic pressure $p_\pi = -\operatorname{tr}(\boldsymbol{\sigma}_\pi)/3$ is a solely a function of the density ρ_π and vice versa. In particular, let us introduce a bulk modulus K_π such that

$$K_\pi = \rho_\pi \frac{dp_\pi}{d\rho_\pi}. \tag{2.8}$$

Expanding the total time derivative of the partial density ρ^π,

$$\frac{d\rho^\pi}{dt} = \frac{d(\phi^\pi \rho_\pi)}{dt} \tag{2.9}$$

$$= \rho_\pi \frac{d\phi^\pi}{dt} + \phi^\pi \frac{d\rho_\pi}{dp_\pi} \frac{dp_\pi}{dt} \tag{2.10}$$

$$= \rho_\pi \left(\frac{d\phi^\pi}{dt} + \frac{\phi^\pi}{K_\pi} \frac{dp_\pi}{dt} \right). \tag{2.11}$$

Inserting this relationship into (2.6) and (2.7),

$$\frac{d\phi^s}{dt} + \frac{\phi^s}{K_s} \frac{dp_s}{dt} + \phi^s \, \nabla \cdot \boldsymbol{v} = 0, \tag{2.12}$$

$$\frac{d\phi^\pi}{dt} + \frac{\phi^\pi}{K_\pi} \frac{dp_\pi}{dt} + \phi^\pi \, \nabla \cdot \boldsymbol{v} + \frac{1}{\rho_\pi} \nabla \cdot (\rho^\pi \boldsymbol{v}_{\pi s}) = 0. \tag{2.13}$$

We now make use of the following identity:

$$\frac{d\phi^\pi}{dt} = \frac{d}{dt}\left[\psi^\pi(1-\phi^s)\right] \tag{2.14}$$

$$= (1-\phi^s)\frac{d\psi^\pi}{dt} - \psi^\pi\frac{d\phi^s}{dt}. \tag{2.15}$$

Combining equations (2.12), (2.13), and (2.15) leads to an expression for the π-phase mass balance ($\pi = w, a$) with the solid-phase balance embedded in it,

$$(1-\phi^s)\frac{d\psi^\pi}{dt} + \frac{\phi^\pi}{K_\pi}\frac{dp_\pi}{dt} + \frac{\psi^\pi\phi^s}{K_s}\frac{dp_s}{dt} + \psi^\pi\,\nabla\cdot\boldsymbol{v} + \frac{1}{\rho_\pi}\,\nabla\cdot(\rho^\pi\boldsymbol{v}_{\pi s}) = 0. \tag{2.16}$$

We now make use of the following relation for the bulk modulus of the porous skeleton, K,

$$\phi^s\frac{dp_s}{dt} = -K\,\nabla\cdot\boldsymbol{v}. \tag{2.17}$$

See Borja [2006c] for a discussion of this relationship and a derivation of an expression for the skeleton modulus based on a thermodynamic analysis. Equation (2.16) can therefore be rewritten as

$$(1-\phi^s)\frac{d\psi^\pi}{dt} + \frac{\phi^\pi}{K_\pi}\frac{dp_\pi}{dt} + \psi^\pi\left(1-\frac{K}{K_s}\right)\nabla\cdot\boldsymbol{v} + \frac{1}{\rho_\pi}\,\nabla\cdot(\rho^\pi\boldsymbol{v}_{\pi s}) = 0, \tag{2.18}$$

where we recognize the appearance of the Biot coefficient $b = 1 - K/K_s$. Our final manipulations concern the last term, containing the relative flow vector $\boldsymbol{v}_{\pi s}$,

$$\frac{1}{\rho_\pi}\,\nabla\cdot(\rho^\pi\boldsymbol{v}_{\pi s}) = \frac{1}{\rho_\pi}\,\nabla\cdot(\rho_\pi\phi^\pi\boldsymbol{v}_{\pi s}) \tag{2.19}$$

$$= \frac{1}{\rho_\pi}\nabla\rho_\pi\cdot(\phi^\pi\boldsymbol{v}_{\pi s}) + \nabla\cdot(\phi^\pi\boldsymbol{v}_{\pi s}) \tag{2.20}$$

$$= \frac{\phi^\pi}{K_\pi}\nabla p_\pi\cdot\boldsymbol{v}_{\pi s} + \nabla\cdot(\phi^\pi\boldsymbol{v}_{\pi s}). \tag{2.21}$$

We have again made use of the bulk modulus relationship (2.8). The final form for the conservation equation, valid for the solid/water or solid/air combination, is

$$(1-\phi^s)\frac{d\psi^\pi}{dt} + b\psi^\pi\,\nabla\cdot\boldsymbol{v} + \nabla\cdot(\phi^\pi\boldsymbol{v}_{\pi s}) + \frac{\phi^\pi}{K_\pi}\left(\frac{dp_\pi}{dt} + \nabla p_\pi\cdot\boldsymbol{v}_{\pi s}\right) = 0. \tag{2.22}$$

We note that the final term in parenthesis represents the total time derivative $d^\pi(p_\pi)/dt$ reckoned with respect to the solid motion. The term $\phi^\pi\boldsymbol{v}_{\pi s}$ represents the superficial (Darcy) velocity, which for convenience we will denote as $\boldsymbol{w}^\pi$.

To conclude, we recall that this formulation rests on the assumptions of isothermal deformation, no mass exchanges between phases, and the existence of bulk moduli K_π and K. Otherwise, the formulation remains quite general and can be readily expanded to handle an arbitrary number of

compressible components. In some cases, however, an even more general formulation may be desired. See, e.g., Gawin et al. [1995] for a formulation that includes thermal effects.

2.1.2 Phase Compressibility

In an isothermal system of the type considered here, the bulk modulus of an ideal gas is equal to the absolute pressure of the gas. Therefore, if p_a is the gauge pressure in the air, then we define the bulk modulus for the air as

$$K_a = p_a + p_{atm}. \tag{2.23}$$

The atmospheric pressure for air at STP conditions is 101.3 kPa. In contrast, the bulk modulus K_w for water is 2.2×10^6 kPa and the bulk modulus K_s for sand or clay is typically $\mathcal{O}(10^7 - 10^8)$ kPa.

In many engineering cases of interest, reasonable approximations may neglect the intrinsic compressibilities of one or more phases. We will come back to this point as we discuss simplified formulations.

2.1.3 Effective Stress Relationship

The effective stress is that portion of the total stress that controls the stress/strain, volume change, and strength behavior of the solid matrix, independently of the magnitude of the fluid pressure [Schrefler and Gawin 1996]. In saturated soil mechanics the effective stress is most commonly written as

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + b p_w \mathbf{1}, \tag{2.24}$$

where b is the Biot coefficient introduced earlier. For many soils it is adequate to assume $b = 1$ ($K/K_s \approx 0$). For saturated rocks and concrete b is typically in the range 0.5–0.9.

In the unsaturated case, there is a lack of consensus on the best form for the effective stress relationship. In this work, we employ the following definition of the effective stress,

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + b \bar{p} \mathbf{1} \tag{2.25}$$

where $\bar{p}$ is the phase-averaged pressure of the fluid and gas surrounding the solid grains,

$$\bar{p} = \psi^w p_w + \psi^a p_a. \tag{2.26}$$

A convenient feature of this definition is that it naturally reduces to the definition (2.24) in the saturated limit, and so no special treatment is needed to handle variably-saturated conditions.

Borja [Borja 2006c; Borja and Koliji 2009] demonstrated that this expression for the constitutive stress tensor is energy-conjugate to the solid rate of deformation. Note that this definition is similar to the Bishop stress when $b = 0$ (incompressible solid grains) and ψ^w is replaced with the

experimentally determined Bishop parameter χ. We recover the Skempton [Skempton 1961] and Nur-Byerlee [Nur and Byerlee 1971] stress when $\psi^w = 1$; and the Terzaghi [Terzaghi 1943] stress when $b = 0$ and $\psi^w = 1$. The Bishop parameter χ can be determined experimentally, and the substitution of degree of saturation ψ^w in lieu of this parameter is a simplified approximation derived from the volume averaging over a representative elementary volume (REV).

The pressures are assumed to be relative pressures, measured with respect to atmospheric pressure or another suitable reference. We also define the matric suction

$$s = p_a - p_w \geq 0. \tag{2.27}$$

As we will see, the suction (also commonly referred to as capillary pressure) is a crucial state variable used in the definition of the necessary constitutive relationships. The lower bound follows from the fact that the medium is assumed to be fully saturated when $p_w > p_a$, and without a fluid/gas interface no capillary pressure can exist.

If the pore-space is well-connected and the formation of interest is at a shallow depth, it is frequently assumed that the pore gas pressure is in equilibrium with atmospheric pressure, or $p_a \approx 0$. This assumption leads to a "passive gas" formulation in which the only non-trivial balance equations are the linear momentum balance and the conservation law for the fluid phase. This assumption also implies a simplified definition of the suction, $s = -p_w$, and effective stress,

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + b(\psi^w p_w)\mathbf{1}. \tag{2.28}$$

2.1.4 Linear Momentum Balance

The local form of the balance of linear momentum for the mixture under quasi-static conditions is

$$\nabla \cdot \boldsymbol{\sigma} + \rho \boldsymbol{g} = \mathbf{0}. \tag{2.29}$$

Based on the discussion of the previous section, the momentum balance can be directly written in terms of the effective stress $\boldsymbol{\sigma}'$ and the phase-averaged pressure $\bar{p}$,

$$\nabla \cdot (\boldsymbol{\sigma}' - b\bar{p}\mathbf{1}) + \rho \boldsymbol{g} = \mathbf{0}. \tag{2.30}$$

The multiphase system under consideration has five unknowns—three components of displacement and two pressures. The momentum balance provides three governing equations, while the remaining two follow from the mass conservation equations. In order to close the model, it is necessary to define certain constitutive relationships. These relationships are the subject of the next section.

2.2 Constitutive Relations

In the following sections we discuss the constitutive relationships necessary to complete the definition of the multiphase model, as well as present a few representative models. The key observation is that the constitutive models are based on two state variables, the effective stress and the matric suction.

2.2.1 Saturation Relationship

In this work we will use saturation relationships of the form $\psi^\pi(s)$, where s is the suction pressure defined earlier. One model commonly employed in hydrology is the van Genuchten [1980] model,

$$\psi^w(s) = \psi_1 + (\psi_2 - \psi_1) \left[1 + (\frac{s}{s_a})^n \right]^{-m}.$$ (2.31)

The model contains four parameters: ψ_1 is the residual water saturation, ψ_2 is the maximum water saturation, s_a is a scaling pressure, and n and m are empirical constants defining the shape of the saturation curve. In the original formulation n and m are not independent constants, but are rather related to one another as

$$m = \frac{n-1}{n}.$$ (2.32)

Using this relationship reduces the number of empirical parameters to three. We remark that while the van Genuchten curve is a popular choice, many other models exist.

2.2.2 Generalized Darcy's Law

The quantity $w^\pi = \phi^\pi v_{\pi s}$ appearing in the mass balance equations is the superficial velocity of the π-phase (units of m/s). In this work we use a generalized Darcy's law to relate the superficial velocity field to the pressure and elevation potential in each phase,

$$w^\pi = -k_{r\pi} \frac{k_\pi}{\eta_\pi} \left(\nabla p_\pi - \rho_\pi g \right).$$ (2.33)

Here, k_π is a second-order tensor defining the saturated intrinsic permeability of the π-phase (units of m^2), η_π is the fluid viscosity (units of Pa·s), and $k_{r\pi}$ is a dimensionless relative conductivity factor that accounts for the effect of partial saturations.

The factor $k_{r\pi}$ is typically taken as a function of the saturation ψ^w, though it can also be related directly to the suction s. From an experimental perspective, the former is often easier to measure than the latter. The van Genuchten [1980] model mentioned earlier provides such relationships. Defining a scaled saturation θ as

$$\theta = \frac{\psi^w - \psi_1}{\psi_2 - \psi_1},$$ (2.34)

the fluid and gas relative permeabilities are given by

$$k_{rw}(\theta) = \theta^{1/2} \left[1 - \left(1 - \theta^{1/m}\right)^m\right]^2 , \qquad (2.35)$$

$$k_{ra}(\theta) = (1 - \theta)^{1/3} \left(1 - \theta^{1/m}\right)^{2m} . \qquad (2.36)$$

2.2.3 Elastoplastic Models

In this work we will use a variety of material models—linear elastic, hyperelastic, and elastoplastic. For simplicity of presentation, we will use a small-strain assumption throughout and define the strain measure ϵ as the symmetric gradient of the displacement,

$$\epsilon = \nabla^s u = \frac{1}{2}\left(\nabla u + \nabla^\mathsf{T} u\right) \qquad (2.37)$$

For saturated materials, the effective stress/strain relationship can be written in a general "rate form" as

$$\dot{\sigma}' = \mathbb{C} : \dot{\epsilon} \qquad (2.38)$$

where $\mathbb{C}$ is a fourth-order tensor of tangential moduli. For unsaturated materials, the story is slightly more complicated as some models include an explicit dependence of the stress on the suction s as well. See, for example, the Critical State Model proposed by Borja [2004]. For this case we need a more general relationship

$$\dot{\sigma}' = \mathbb{C} : \dot{\epsilon} + \mathbb{D}\dot{s} \qquad (2.39)$$

where $\mathbb{D}$ is a second-order tensor of tangential moduli. These general relationships are sufficient to formulate the problem, and so we will not narrow the discussion at this point to a specific material model. We will instead introduce models as necessary to describe specific applications.

2.3 Simplified Forms

The triphasic model just introduced is quite general, and can be extended to accomodate an arbitrary number of compressible constituent phases. In practice, however, simplifying assumptions can often be made. In this section we consider a few of these simplified forms.

In many "near-surface" applications it is reasonable to make the passive gas assumption introduce earlier, $p_a = 0$. This assumption eliminates the need for the air mass balance equation, and reduces to number of governing equations to two. To compare these formulations, we consider the one-dimensional drainage experiments by Liakopoulos [1964]. In the experiment, a column of sand was fed with water from a reservoir at the top and allowed to freely drain through the bottom. Once the sand was saturated and the pore water pressure in the column was approximately atmospheric (free
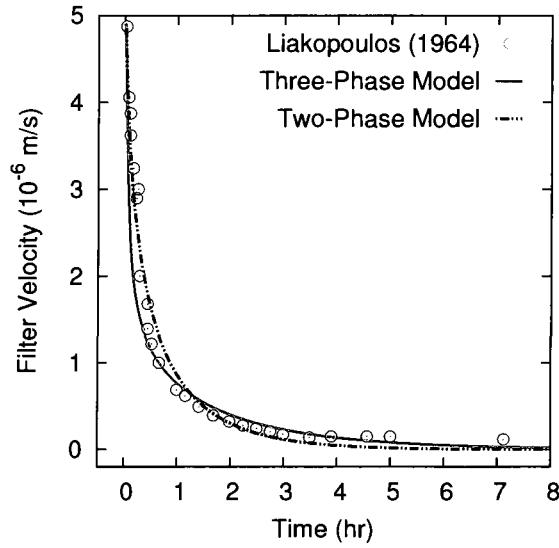
Figure 2.1: Comparison of two- and three-phase formulations with experimental data by Liakopoulos [1964].

drainage), the inlet at the top was closed. The column was then allowed to drain from a saturated to a final unsaturated state. Figure (2.1) presents experimental measurements of the fluid velocity at the lower outlet filter, along with simulations of this experiment using both a three-phase and the simpler two-phase models. We see that for this case, both approximations work quite well. The simpler two-phase model tends to underestimate the time to steady-state, but otherwise provides a reasonable approximation.

In many applications, it is often reasonable to ignore certain phase compressibilities. For example, at low confining stresses such as those encountered in geotechnical applications, it is reasonable to assume the that bulk modulus of the water and solid phases are essentially infinite. Note that we still assume that the solid skeleton modulus is finite, but that the Biot coefficient $b = 1$. Along with the passive gas assumption, this leads to the simpler formulation,

$$\nabla \cdot (\boldsymbol{\sigma}' - \psi^w p_w \mathbf{1}) + \rho \boldsymbol{g} = \mathbf{0} \tag{2.40}$$

$$(1 - \phi^s)\frac{d\psi^\pi}{dt} + \psi^w \nabla \cdot \dot{\boldsymbol{u}} + \nabla \cdot \boldsymbol{w} = 0 \tag{2.41}$$

In this limit of full saturation ($\psi^w = 1$), this model reduces to the even simpler form,

$$\nabla \cdot (\boldsymbol{\sigma}' - p_w \mathbf{1}) + \rho \boldsymbol{g} = \mathbf{0} \tag{2.42}$$

$$\nabla \cdot \dot{\boldsymbol{u}} + \nabla \cdot \boldsymbol{w} = 0 \tag{2.43}$$

Since these last two models are conveniently simple but are also very useful in practice, we will tend to focus on them in later chapters. It should be apparent, however, that much of what is said here

can be straightforwardly extended to the more general formulation when the additional physics is necessary.

2.4 Limiting Forms

An interesting feature of the coupled model is that it simplifies to other well known mathematical formulations under certain limiting assumptions. For example, consider the undrained limit as the permeability $k \rightarrow 0$. If we assume a linear elastic material model (parameterized by the Lamé parameters λ and μ), the resulting set of governing equations can be rewritten as

$$2\mu \, \nabla \cdot \epsilon(u) - \nabla p = f \tag{2.44}$$

$$\nabla \cdot u = 0 \tag{2.45}$$

These are the well-known Stokes equations. This model can be used to represent the behavior of an incompressible elastic solid. With a suitable reinterpretation of the variables, the Stokes model also describes the behavior of slow, viscosity-dominated fluid flows. For this case, u represents the velocity field (rather than displacements) and μ is the viscosity (rather than the shear modulus).

Now consider the other limiting case of the coupled model: the limit as the solid matrix becomes rigid. In this case, the governing equations reduce to the Darcy flow equations,

$$w + \frac{k}{\mu} \nabla p = f \tag{2.46}$$

$$\nabla \cdot w = 0 \tag{2.47}$$

An important feature to recognize in both the Stokes and Darcy models is that the second equation in each is a constraint equation. In the first case, there is a divergence constraint on the displacement field (no volumetric deformation) and in the second case there is a divergence constraint on the seepage velocity (solendoidal flow). As we briefly mentioned in the introduction, the coupled model is subject to certain stability restrictions. These restrictions stem from the constrained nature of the problem in certain limit states. We will discuss these issues extensively in the next chapter.

2.5 Weak Forms

We have presented the governing model in strong (differential) form. In the context of the finite element method, however, we transform the strong statement to a weak (integral) statement. In this section, we present this weak integral statement. To avoid a proliferation of terms, we will use the simplest, saturated model given by (2.42) and (2.43). For convenience, we will therefore drop the subscript w and write the water pressure simply as p. We will retain the subscript only in those cases where the phase designation is ambiguous.

The mixture occupies a domain Ω with boundary Γ. This boundary is suitably decomposed into

regions where boundary conditions are specified for both the solid and fluid. In particular, we define

$$\Gamma_u : \quad \text{solid displacement boundary} \tag{2.48}$$

$$\Gamma_t : \quad \text{solid traction boundary} \tag{2.49}$$

$$\Gamma_p : \quad \text{fluid pressure boundary} \tag{2.50}$$

$$\Gamma_q : \quad \text{fluid flux boundary} \tag{2.51}$$

with the restrictions

$$\Gamma = \Gamma_u \cup \Gamma_t = \Gamma_p \cup \Gamma_q, \tag{2.52}$$

$$\Gamma_u \cap \Gamma_t = \Gamma_p \cap \Gamma_q = \varnothing. \tag{2.53}$$

The boundary conditions are given as

$$\boldsymbol{u} = \bar{\bar{\boldsymbol{u}}} \quad \text{on } \Gamma_u \quad \text{(specified displacement)}, \tag{2.54}$$

$$\boldsymbol{n} \cdot \boldsymbol{\sigma}' = \bar{\bar{\boldsymbol{t}}} \quad \text{on } \Gamma_t \quad \text{(specified traction)}, \tag{2.55}$$

$$p = \bar{\bar{p}} \quad \text{on } \Gamma_p \quad \text{(specified pressure)}, \tag{2.56}$$

$$-\boldsymbol{n} \cdot \boldsymbol{w} = \bar{\bar{q}} \quad \text{on } \Gamma_q \quad \text{(specified flux)}. \tag{2.57}$$

Finally, initial conditions at $t = 0$ are given as $\{\boldsymbol{u}_o, p_o\}$.

In performing the transformation from strong to weak, we are immediately presented with a choice: which variables to choose as our primary unknowns. In this work we will rely upon a two-field formulation in which the displacement and the water pressure are the primary variables. This leads to a so-called $\boldsymbol{u}/p$ formulation. In this case the mass and momentum balance equations are written in weak form to furnish two governing equations. Another option that is slightly more complicated, but also has nice approximation properties, is a three-field $(\boldsymbol{u}/p/\boldsymbol{w})$ approach.

For the two-field discretization, two spaces of trial functions are defined as

$$\mathcal{U} = \{\boldsymbol{u} : \Omega \to \mathbb{R}^3 \mid \boldsymbol{u} \in \boldsymbol{H}^1, \, \boldsymbol{u} = \bar{\bar{\boldsymbol{u}}} \text{ on } \Gamma_u\}, \tag{2.58}$$

$$\mathcal{P} = \{p : \Omega \to \mathbb{R} \mid p \in H^1, \, p = \bar{\bar{p}} \text{ on } \Gamma_p\}, \tag{2.59}$$

where H^1 denotes a Sobolev space of degree one. We also define the corresponding spaces of weighting functions, with homogeneous conditions on the essential boundaries,

$$\mathcal{U}_o = \{\boldsymbol{\eta} : \Omega \to \mathbb{R}^3 \mid \boldsymbol{\eta} \in \boldsymbol{H}^1, \, \boldsymbol{\eta} = \boldsymbol{0} \text{ on } \Gamma_u\}, \tag{2.60}$$

$$\mathcal{P}_o = \{\phi : \Omega \to \mathbb{R} \mid \phi \in H^1, \, \phi = 0 \text{ on } \Gamma_p\}. \tag{2.61}$$

The weak form of the problem is then to find $\{u, p\} \in \mathcal{U} \times \mathcal{P}$ such that for all $\{\eta, \phi\} \in \mathcal{U}_o \times \mathcal{P}_o$,

$$\mathcal{R}_{\mathrm{mom.}} = \int_\Omega \nabla^s \eta : \sigma' \, \mathrm{d}\Omega - \int_\Omega p \nabla \cdot \eta \, \mathrm{d}\Omega - \int_\Omega \eta \cdot \rho g \, \mathrm{d}\Omega - \int_\Gamma \eta \cdot \bar{t} \, \mathrm{d}\Gamma = 0 \tag{2.62}$$

$$\mathcal{R}_{\mathrm{mass}} = -\int_\Omega \phi \nabla \cdot \dot{u} \, \mathrm{d}\Omega + \int_\Omega \nabla\phi \cdot w \, \mathrm{d}\Omega + \int_\Gamma \phi \bar{\bar{q}} \, \mathrm{d}\Gamma = 0 \tag{2.63}$$

2.6 Numerical Discretization

With the weak form in hand, the construction of the numerical model follows the standard finite element recipe. For the spatial discretization, we introduce discrete interpolation spaces to represent u^h and p^h. Since the second equation is time-dependent, we use a finite-difference method (implicit Euler update or Crank-Nicolson) for the temporal discretization. Since these manipulations are standard, for the sake of brevity we will not report them here and refer the interested reader to [White and Borja 2008].

The most important point is that this discretization eventually leads to a vector of nonlinear residual equations $R(x) = 0$ that must be solved in each time step to advance the time-dependent solution $x(t)$. The solution vector has length $n_u + n_p$, where n_u is the number of displacement degrees of freedom, and n_p is the number of pressure degrees of freedom. It can be partitioned into u and p blocks as

$$x = \begin{bmatrix} u \\ p \end{bmatrix} \tag{2.64}$$

The residual vector R can also be partitioned into two blocks, an n_u block representing the discrete momentum balance, and a n_p block representing the discrete mass balance,

$$R(x) = \begin{bmatrix} R_{\mathrm{mom.}} \\ R_{\mathrm{mass}} \end{bmatrix} = 0 \tag{2.65}$$

Since the equations are nonlinear (in this case due to the nonlinear constitutive behavior of σ') we use a Newton iteration to drive them to zero. That is, at the kth iteration, given a current guess at the solution $x_k = (u_k, p_k)$, we build a linear approximation to the nonlinear residual about the current configuration,

$$R(x) \approx R(x_k) + J_k \, \Delta x = 0 \tag{2.66}$$

where $J_k = [\partial R / \partial x]$ is the $(n_u + n_p) \times (n_u + n_p)$ Jacobian of the system, evaluated at the current configuration x_k. This implies an update scheme in which we solve the linear system,

$$J_k \Delta x = -R(x_k) \tag{2.67}$$

for the update increment $\Delta x = (\Delta u, \Delta p)$, and set $x_{k+1} = x_k + \Delta x$. Since the residual vector has a

2 blocks, the Jacobian system has a 2×2 block structure,

$$\begin{bmatrix} A & B_1 \\ B_2 & C \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \end{bmatrix} = - \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \end{bmatrix}_k \tag{2.68}$$

where the A block is $n_u \times n_u$ and the C block is $n_p \times n_p$. Because the pressure is a scalar field, while the displacement is a vector field with as many components as there are spatial dimensions, the dimensions of the A block will typically be much larger than those of the C block. The A block contains coefficients that couple displacement degrees of freedom to displacement degrees of freedom (u/u), the C block couples pressure degrees of freedom to pressure degrees of freedom (p/p), and the off-diagonal blocks B_1 and B_2 govern the interaction of displacements with pressures (u/p) and vice-versa (p/u). Efficient methods for solving block-structured systems of this type are the focus of Chapter 4.

2.7 Conclusion

In this chapter, we have presented a very general formulation for modeling multiphase mixtures using a physics-based approach. We have shown how the general formulations reduces to a variety of simpler models commonly used in practice. We have presented both the strong and weak forms of the problem, and discussed the block structure of the resulting algebraic problems that must be solved in the context of a nonlinear Newton iteration.

In the next few chapters we consider some of the difficult numerical issues associated with solving these equations.

3 Stabilized FEM

*The key results of this chapter were originally published in J.A. White and R.I. Borja (2008),
"Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their ap-
plication to fault zone transients", Comp. Meth. Appl. Mech. Eng., 197(49–50):4353–4366.
Additional material is presented therein and we refer the interested reader to this work.*

As we have seen, in many applications it is necessary to build a fully-coupled model of solid de-
formation and fluid flow. Finite element simulations provide a natural tool for investigating such
processes. To do so, we employ a mixed formulation to solve for the solid displacements (u) and fluid
pressures (p). In comparison to pure displacement formulations, however, mixed schemes creates
additional challenges for the numerical analyst.

For example, consider the use of a mixed u/p formulation to solve for the displacements and
pressures in a saturated soil. In the limit of low permeability or fast loading rates, the pore fluid
can impose near or exact incompressibility on the deformation of the solid matrix. Since the fluid
itself has a large bulk modulus (for water, $K \approx 2$ GPa), it prevents the locally-undrained mixture
from undergoing any volumetric deformation. In the presence of such incompressibility constraints,
it is well known that only certain combinations of discrete spaces for the pressure and displacement
interpolation exhibit stable behavior. Failure to choose a stable pair can lead to poor results, typically
in the form of spurious oscillations in the pressure field and sub-optimal convergence behavior. The
same restrictions are found in many other constrained problems. Classic examples include mixed
finite element formulations for Stokes flow, Darcy flow, and incompressible elasticity. Recall that
these simpler, linear models can be recovered under certain limiting assumptions from the current
coupled formulation. Therefore, as we approach these limits we must address the same stability
restrictions.

The mathematical theory establishing the solvability and stability characteristics of mixed formu-
lations is well-developed. The key ingredients are an ellipticity requirement and the Ladyzhenskaya-
Babuška-Brezzi (LBB) condition [Arnold 1990; Brezzi 1990]. Unfortunately, many seemingly nat-
ural interpolation pairs—including equal-order interpolation for all field variables—do not satisfy
the necessary stability requirements. In practice, most analysts rely on "safe" elements such as the
Taylor-Hood family, in which the displacement interpolation is one-order higher than the pressure
interpolation. A variety of more sophisticated stable elements are also available, e.g. [Arnold et al.
1984].

An illustration of spurious pressure behavior is given in Figure 3.1. We have modeled a very
low permeability soil mass subject to a surface traction (representing a square footing load) using
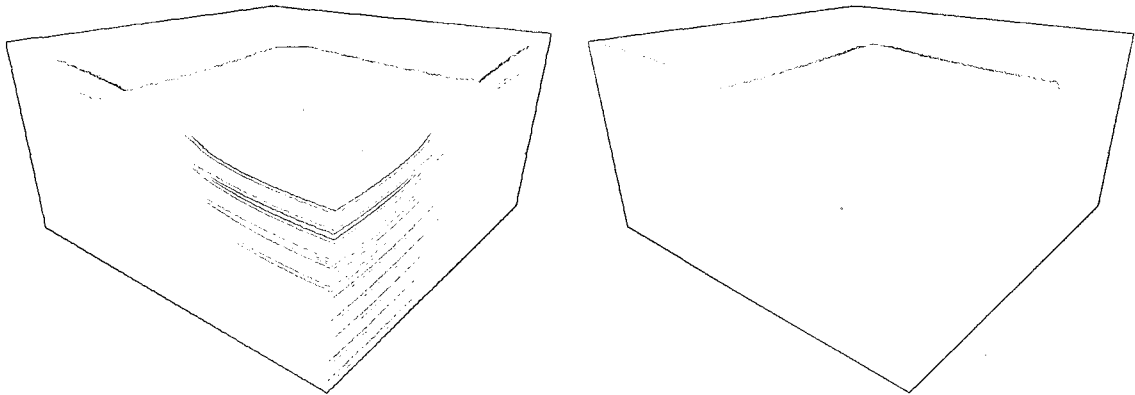
Figure 3.1: Predicted pressure field for the undrained footing problem using (left) unstable Q1Q1 elements and (right) intrinsically-stable Q2Q1 elements. The pressure solution for the Q1Q1 pair is ruined by spurious pressure oscillations.

two different interpolation pairs. In the left figure, equal order linear interpolation is used for both u and p (Q1Q1 elements) which is known to be an intrinsically unstable pair. In the right figure, an intrinsically stable pair is used (Taylor-Hood Q2Q1 elements). The spurious oscilliations in the left figure are clearly visible, while the right solution displays a smooth pressure field. Upon mesh refinement, the stable model displays convergent behavior, while in the unstable model spurious oscillations persist.

While stable combinations exist, they are often expensive to use. From an implementation point of view it would be appealing to circumvent the stability restrictions and employ a broader class of interpolation pairs. Over the years, many "stabilization" techniques have been proposed for doing precisely this, most extensively in the fluid dynamics community. The model equations used to study these schemes are typically the Stokes or Darcy equations, which despite their simplicity contain all of the salient features of a constrained problem. We mention the early Brezzi-Pitkäranta scheme [Brezzi and Pitkäranta 1984], the Galerkin Least-Squares (GLS) approach of Hughes et al. [Hughes et al. 1986], and the more recent Variational Multiscale Methods [Hughes et al. 1998]—but many others exist. See, for example, [Dohrmann and Bochev 2004; Douglas and Wang 1989; Masud and Hughes 2002]. In solid mechanics, a variety of schemes have also been developed for incompressible and quasi-incompressible elasticity in order to overcome volumetric locking associated with pure-displacement formulations. For example, in [Oñate et al. 2004], Oñate and co-workers proposed a formulation based on the concept of Finite Calculus. Masud and Xia in [Masud and Xia 2005, 2006] developed a formulation for both linear and nonlinear constitutive models based on a Variational Multiscale approach . Romero and Bischoff have recently proposed an interesting method for linear elasticity which involves enriching the finite element spaces with incompatible bubble functions [Romero and Bischoff 2007]. Of course, the above schemes are merely a representative sample of an extensive literature that has developed for each class of problems.

While it is difficult to classify all stabilization schemes in a unified framework, most frequently

the methods lead to a modified variational formulation in which additional terms are added to the mass balance equation, modifying the incompressibility constraint in such a way that stability of the mixed formulation is increased while maintaining a convergent method. In this way, meaningful results can be obtained when using otherwise unstable elements. The goal of this contribution is to extend the stabilization concept to coupled solid-deformation/fluid-diffusion problems. While stabilized methods are employed frequently in fluid and solid mechanics problems, their use in coupled geomechanical problems is limited. Nevertheless, some good work in this direction has begun. In [Wan 2002], Wan used the GLS approach to stabilize both a displacement-pressure (u/p) and a displacement-pressure-velocity ($u/p/w$) formulation. In [Truty 2001; Truty and Zimmermann 2006], Truty and Zimmerman compared three schemes: one based on the Brezzi-Pitkäranta stabilization and two based on the GLS approach. They also extended their formulation to account for partial saturations. In [Pastor et al. 2000, 1999], Pastor et al. proposed a stabilization scheme for dynamic problems using a fractional-step algorithm, incorporating the stabilization into the time-stepping scheme. In each case, the authors demonstrated that the stabilizations can successfully suppress instabilities and lead to good-quality solutions. Of course, each scheme has its own shortcomings. For example, the GLS method is based on adding the residual of the strong form of the governing equations. As such, second-order derivatives with respect to the displacements appear, and when using linear interpolation, these terms either vanish or are poorly approximated. A special technique must generally be employed to improve the accuracy of these calculations, introducing additional computational work. See [Wan 2002], for example, where Wan develops such a stress recovery technique. The GLS formulations also often lead to a non-symmetric modification of the system matrix. While this makes little difference if the original problem is non-symmetric, it would be appealing to preserve any symmetry if it does exist. Indeed, a key advantage of the methods of [Masud and Xia 2005, 2006; Romero and Bischoff 2007] is their symmetry-preserving property, but these schemes have only been employed for incompressible solids and have not been extended to coupled solid/fluid formulations. The Brezzi-Pitkäranta scheme does lead to a symmetric modification and can be cheaply implemented for equal-order linear interpolations. Unfortunately, the formulation cannot be extended to stabilize other unstable pairs such as linear-displacement/constant-pressure elements. The fractional-step method is primarily designed for dynamic problems, and may not be an efficient approach for quasi-static models. It also leads to a conditionally stable time-integration scheme even if the underlying algorithm is implicit, though recent improvements by the authors have significantly improved the stability restriction [Li et al. 2003].

In this chapter, we introduce a new stabilization scheme for coupled geomechanical problems based on the concept of Polynomial-Pressure-Projections. In this approach, the additional stabilizing terms use element-local projections of the pressure field to counteract the inherent instabilities in the chosen interpolation pair. The technique was recently proposed by Dohrmann, Bochev, and Gunzburger, and has been successfully employed for stabilizing the Stokes problem [Bochev et al. 2006; Dohrmann and Bochev 2004] and Darcy problem [Bochev and Dohrmann 2006]. An analysis of similar pressure projection methods, and a unifying framework for their analysis, has also been proposed by Burman [Burman 2007].
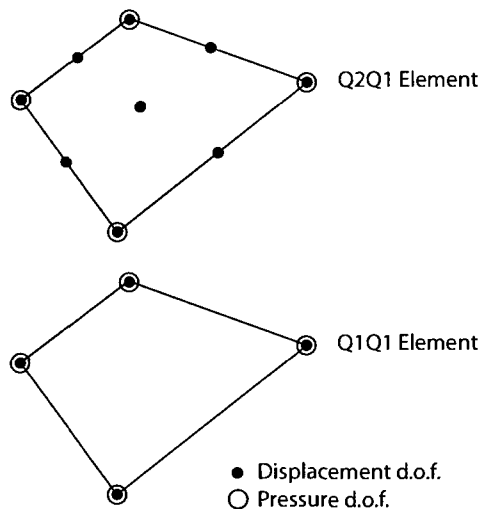
Figure 3.2: Q2Q1 and Q1Q1 mixed quadrilateral elements, showing the support points for displacement and pressure degrees of freedom.

In this work we employ pressure projections to address instabilities that arise in the geomechanical problems under consideration. The new stabilization has several appealing features. In particular, the additional stabilizing terms can be assembled locally on each element using standard shape function information, and no specialized subroutines are required. The scheme does not require the calculation of higher-order derivatives or special stress-recovery techniques. The method introduces minimal additional computational work, and can be readily implemented in a standard finite element code. The scheme also leads to a symmetric modification of the system matrix, preserving any symmetry that was inherent in the original variational formulation. The resulting method thus shares many of the positive features of the Brezzi-Pitkäranta stabilization, but can be used to stabilize a broader class of unstable pairs.

The primary motivation for using stabilization is computational efficiency. As an example, consider two meshes composed of an equal number of two-dimensional, quadrilateral elements. The first mesh employs continuous quadratic-displacement/linear-pressure interpolation (Q2Q1), while the second uses linear-displacement/linear-pressure interpolation (Q1Q1). Both elements are illustrated in Figure 3.2. The first element possesses 22 degrees of freedom and is known to be stable, while the second element has 12 degrees of freedom and is known to be unstable—unless a stabilized formulation is employed. The two elements are comparable in the sense that they produce the same order of pressure interpolation. The Q2Q1, however, leads to algebraic problems with many more degrees of freedom. As the number of elements in each mesh grows, a simple argument shows that the total number of unknowns in the two meshes quickly approachs a ratio of 3:1. If we consider the equivalent three-dimensional situation with hexahedral elements, this limit ratio is $6\frac{1}{4}$:1. Referring to Figure 3.1 as a typical example, the Q1Q1 mesh on the left contains 19,652 degrees of freedom,

while the Q2Q1 mesh on the right has 112,724 degrees of freedom. Since both meshes exhibit the same convergence behavior for the pressure solution, it seems preferable to use the much cheaper Q1Q1 interpolation.

Further computational savings can also be associated with the quadrature rule employed. The Q2Q1 element typically requires 3×3 Gauss-quadrature in order to accurately integrate the quadratic displacement field. In the Q1Q1 mesh, we only need 2×2 quadrature. If we consider an elastoplastic material in which a significant level of computation must be performed in the material subroutine at each Gauss point, the lower-order quadrature rule will lead to additional efficiency. The equal-order element can also somewhat simplify the code implementation, particularly when employing adaptive mesh refinement or a parallel decomposition of the domain. Finally we note that the introduction of stabilization terms can often improve the convergence behavior of iterative solvers. For extremely large problems, the memory-efficiency of iterative solvers makes them a more attractive choice than sparse direct solvers. We will come back to schemes for solving large linear systems arising from mixed formulations in Chapter 4.

The rest of this chapter is organized as follows. First, we briefly review the origins of instability in mixed formulations, and the necessary conditions for stability. Next, we present a stabilizing modification to the standard u/p formulation that allows us to circumvent these restrictions. In particular, we will see how we can successfully use the appealing equal-order Q1Q1 interpolation. We then demonstrate the performance of the new scheme on several numerical examples. We start with the classic model equations of Stokes flow and Darcy flow, and then examine coupled flow/deformation applications. The examples are mostly motivated by geotechnical applications. In all cases we compare the performance of a reference stable element (Q2Q1), an unstable element (Q1Q1), and the same unstable element with stabilization (denoted Q1Q1s). For completeness, we also attempt to highlight certain situations where the stabilization scheme must be employed with caution.

The primary focus of this chapter is on stabilizing the mixed u/p formulation—the most common formulation used in geotechnical practice. At several points, however, we will attempt to draw connections with $u/p/w$ formulations, which provide an interesting counterpoint. We also confine ourselves in this chapter to discussing fully-saturated applications, as these are the most critical with respect to instability issues. It is straightforward to apply the proposed scheme to unsaturated models as well, as demonstrated in later chapters. At the end we will conclude with a discussion of issues to be addressed and future avenues of research with respect to stabilized methods.

3.1   Stabilized Formulation

The goal in this section is to develop a stabilized formulation that allows for the successful use of Q1Q1 elements in coupled deformation-diffusion problems. We first begin with a few preliminary observations, and then present a stabilized formulation based on the Polynomial-Pressure-Projection technique.

Recall that in each Newton iteration k of each time step we need to solve block-structured linear

systems of the form

$$\begin{bmatrix} A & B \\ B^\mathsf{T} & C \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \end{bmatrix} = \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \end{bmatrix}_k \tag{3.1}$$

For the saturated model the system is block-symmetric. The C block contains coefficients coupling pressure degrees of freedom to other pressure degrees of freedom, and is a function of the permeability k. In the undrained limit, as $k \to 0$, $C \to 0$, leading to a linear system of the form

$$\begin{bmatrix} A & B \\ B^\mathsf{T} & 0 \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \end{bmatrix} = \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \end{bmatrix}_k \tag{3.2}$$

Although (3.2) can be thought of as a single problem, it is helpful rather to think of it as a series of algebraic problems parameterized by the element diameter h. The goal is to ensure that the approximate solution converges to the exact solution at optimal convergence rates as $h \to 0$.

Algebraic matrices in this form, with a zero (2,2) block, are common in mixed finite element formulations. In order for this system to display stable, convergent behavior, it is necessary for the spaces S_u^h and S_p^h chosen for the displacement and pressure interpolation to satisfy the discrete LBB condition [Brezzi 1990],

$$\sup_{v^h \in S_u^h} \frac{\int_\Omega q^h \nabla \cdot v^h \, d\Omega}{\|v^h\|_1} \geq c\|q^h\|_0 \qquad \forall q^h \in S_p^h, \tag{3.3}$$

with constant $c > 0$ independent of h. Unfortunately, the spaces S_u^h and S_p^h associated with linear-pressure/linear-displacement interpolations do not satisfy this condition and lead to unstable approximations. In [Bochev et al. 2006], however, the authors proved that this pair does satisfy a weaker condition. Consider a projection operator $\Pi : L_2(\Omega) \to R_0$, where R_0 is the space of piecewise constants. The authors showed that the discrete spaces S_u^h and S_p^h satisfy

$$\sup_{v^h \in S_u^h} \frac{\int_\Omega q^h \nabla \cdot v^h \, d\Omega}{\|v^h\|_1} \geq c_1\|q^h\|_0 - c_2\|q^h - \Pi q^h\|_0 \qquad \forall q^h \in S_p^h, \tag{3.4}$$

with $c_1 > 0$ and $c_2 > 0$ independent of h. Comparing this result with the discrete LBB condition, we see that the term $c_2\|q^h - \Pi q^h\|_0$ quantifies the inherent deficiency in the Q1Q1 pair. The stabilization methodology is therefore to add stabilizing terms to the variational equations to penalize this deficiency. Schemes for the Stokes equations and Darcy equations can be found in [Bochev and Dohrmann 2006; Dohrmann and Bochev 2004]. We now take the same approach for the coupled deformation-diffusion problems under consideration. We first need to define a projection operator with a suitable range—that of piecewise constants. Therefore, let

$$\Pi p^h \Big|_{\Omega^e} = \frac{1}{V^e} \int_{\Omega^e} p^h \, d\Omega. \tag{3.5}$$

Here, V^e is the volume of the element. The value of the projected field within each element is simply equal to the element average of p^h. We then modify the discrete variational equation for the balance of mass, $\mathcal{H}^h = 0$, to include an additional term,

$$\mathcal{H}^h - \mathcal{H}^{\text{stab}} = 0, \qquad \mathcal{H}^{\text{stab}} = \int_\Omega \frac{\tau}{2\mu}(\psi^h - \Pi\psi^h)\overline{(\dot{p}^h - \Pi\dot{p}^h)}\, d\Omega. \tag{3.6}$$

where the superposed dot denotes a time derivative. The temporally-discrete version is

$$\mathcal{H}_{n+1}^{\text{stab}} = \int_\Omega \frac{\tau}{2\mu_{n+\theta}\Delta t}(\psi^h - \Pi\psi^h)(p_{n+1}^h - \Pi p_{n+1}^h - p_n^h + \Pi p_n^h)\, d\Omega. \tag{3.7}$$

Here, μ is the shear modulus and $\tau > 0$ is a constant multiplier. Typically this parameter is $\mathcal{O}(1)$, but it can be used to "tune" the level of stabilization. We note that stabilization terms has a diffusive effect, and choosing a large τ will tend to smooth sharp gradients if they are present in the solution. There is thus a balance between choosing τ large enough to provide stability while not choosing it so large that the solution becomes overly diffusive. In most practical applications, however, we have not found this balance to be difficult to achieve.

In equation (3.7) the shear modulus appears as a coefficient on the stabilizing term. In many hyperelastic and elastoplastic models, including those examined in this work, the shear modulus is a function of the current state of stress and will therefore evolve with the configuration. To be fully consistent in such cases, we evaluate

$$\mu_{n+\theta} = \theta\mu_{n+1} + (1 - \theta)\mu_n. \tag{3.8}$$

where θ is the time-integration weighting parameter. The linearization of the stabilizing term about an intermediate configuration $(\boldsymbol{u}_k, p_k)$ is taken as

$$\Delta\mathcal{H}_k^{\text{stab}} = \int_\Omega \frac{\tau}{2\mu_k\Delta t}(\psi_h - \Pi\psi_h)[\Delta p_k^h - \Pi(\Delta p_k^h)]\, d\Omega. \tag{3.9}$$

We note that additional terms associated with the linearization of the coefficient $1/\mu_k$ have been omitted. Unless the modulus changes dramatically over an increment, these additional contributions are minor and can be ignored without losing quadratic convergence behavior in the global Newton iterations.

The additional quantities associated with the stabilization scheme can be readily assembled into the matrix problem using standard shape function information. Noting that $\Pi(\psi^h) = \Pi(\boldsymbol{N}^p\bar{c}) = \Pi(\boldsymbol{N}^p)\bar{c}$, the stabilized version of equation (3.1) becomes

$$\begin{bmatrix} A & B \\ B^{\mathsf{T}} & C - M \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \end{bmatrix} = \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} + H \end{bmatrix}_k \tag{3.10}$$

where the two stabilizing terms are given by

$$H = \int_{\Omega} \frac{\tau}{2\mu_{n+\theta}} \left[N^p - \Pi(N^p) \right] \left[p_{n+1}^h - \Pi p_{n+1}^h - p_n^h + \Pi p_n^h \right] d\Omega, \tag{3.11}$$

$$M = \int_{\Omega} \frac{\tau}{2\mu_{n+\theta}} \left[N^p - \Pi(N^p) \right] \left[N^p - \Pi(N^p) \right] d\Omega. \tag{3.12}$$

We see that the result of the scheme is to introduce a stabilizing sub-matrix into the (2,2) position, eliminating the zero-block that would otherwise appear in the incompressible case. Because the stabilization is based on the original shape functions, however, the sparsity pattern of this block remains unchanged and no fill-in occurs. Also note that modifications are only made to the pressure-pressure coupling block—the other sub-matrices remain unchanged. This is a key contrast with stabilization schemes based on adding residual equations to the variational form, e.g. [Truty 2001; Truty and Zimmermann 2006; Wan 2002], where other sub-matrices are modified as well. Finally, unlike subgrid scale methods, the stabilization does not require additional basis functions or element level condensation.

3.2 Numerical Examples

We now present a few numerical examples to test the performance of the stabilization technique. The examples were implemented using the DEAL.II Finite Element Library—a collaborative, open source project focused on developing a toolbox of common algorithms and data structures for use in object-oriented finite element codes [Bangerth et al. 2007]. For the linear solvers, we use several components from the TRILINOS Project—a growing collection of algorithms and technologies for solving large-scale scientific and engineering problems on serial and parallel computing platforms [Heroux et al. 2005]. Note that the main focus of Chapter 4 is efficient solution techniques for block-structured linear systems of the type encountered here. We therefore postpone the discussion of the particular solution methodology until then.

3.2.1 Stokes Flow

The most important requirement of any stabilization scheme is that it displays good convergence properties. As a first test of the performance of the stabilization method, we consider the Stokes equations,

$$2\mu \, \nabla \cdot \epsilon(u) - \nabla p = f \tag{3.13}$$

$$\nabla \cdot u = 0 \tag{3.14}$$

As already mentioned, the nonlinear poroplastic model reduces to Stokes equations in the linear elastic and undrained limit ($k \to 0$). With a suitable reinterpretation of the variables, the Stokes
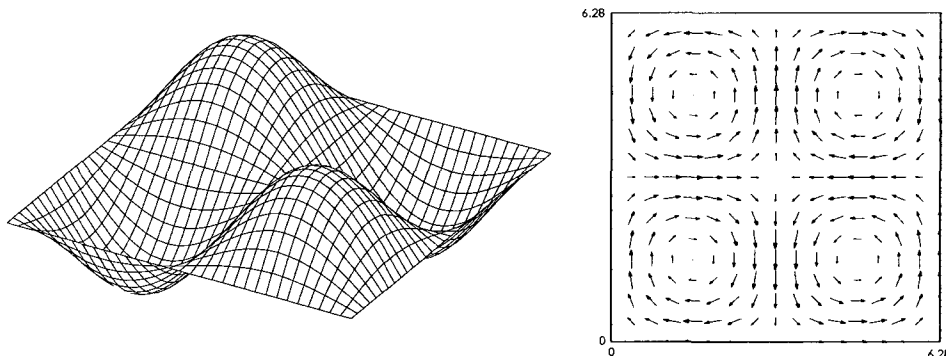
Figure 3.3: Pressure (left) and velocity (right) for the Stokes and Darcy examples.

model describes the behavior of slow, viscosity-dominated fluid flows. For this case, u represents the velocity field (rather than displacements) and μ is the viscosity (rather than the shear modulus). In weak form, we seek (u, p) such that for all (η, ϕ)

$$(\epsilon(\eta), 2\mu\epsilon(u)) - (\nabla \cdot \eta, p) = (\eta, f) - (n \cdot \eta, p)_\Gamma \tag{3.15}$$

$$(q, \nabla \cdot u) = 0 \tag{3.16}$$

For this model, $u \in H^1$ and $p \in L_2$. The requirements on the pressure space have been relaxed since, in the undrained limit, the derivatives on the pressure variable disappear. Thus, it is possible to use finite element spaces with discontinuous pressures, though we continue to examine continuous pressure elements in the following.

Due to the simplicity of the Stokes equations, it is easy to construct analytical solutions with which which we can test our numerical model. For example, consider a unit square $\Omega = [0, 2\pi] \times [0, 2\pi]$ in which the pressure and velocity fields are (somewhat arbitrarily) chosen as

$$\begin{bmatrix} p \\ u \\ v \end{bmatrix} = \begin{bmatrix} \sin(x)\cos(y) + c \\ -\cos(x)\sin(y) \\ -\sin(x)\sin(y) \end{bmatrix} \tag{3.17}$$

where c is a constant (Figure 3.3). Clearly the velocity field satisfies the divergence constraint $\nabla \cdot u$. Plugging these (u, p) fields into the left-hand side of the momentum equation (3.13), we can compute the body force f that leads to this solution. In the finite element discretization, this body force is applied along with Dirichlet boundary conditions on the normal components of the velocity field. Since the pressure is only determined to an arbitrary constant, the mean value pressure is subtracted from the solution so that $c = 0$.

Table 3.1 illustrates the behavior of the L_2-error in the pressure and velocity solutions using both the stabilized Q1Q1s element and the LBB-stable Q2Q1 element. We see that both elements converge at order $k = l + 1$, where l is the interpolation order. For the Q2Q1 element, $k = 3$ for

Table 3.1: Comparison of observed convergence behavior in the Stokes example for the stabilized Q1Q1s (with $\tau = 0.001$) and LBB-stable Q2Q1. (Top) L_2 pressure error and estimated convergence order. (Bottom) L_2 velocity error and estimated convergence order.

Mesh	Q2Q1 $\|e_p\|$	k	Q1Q1s $\|e_p\|$	k
4×4	5.33×10^{-1}	–	5.83×10^{-1}	–
8×8	1.10×10^{-1}	2.28	1.10×10^{-1}	2.40
16×16	2.60×10^{-2}	2.08	2.60×10^{-2}	2.09
32 ×32	6.41×10^{-3}	2.02	6.41×10^{-3}	2.02
64 ×64	1.60×10^{-3}	2.00	1.60×10^{-3}	2.01

Mesh	Q2Q1 $\|e_u\|$	k	Q1Q1s $\|e_u\|$	k
4×4	1.77×10^{-1}	–	1.37×10^{0}	–
8×8	1.82×10^{-2}	3.28	3.65×10^{-1}	1.91
16×16	2.21×10^{-3}	3.04	9.25×10^{-2}	1.98
32 ×32	2.74×10^{-4}	3.01	2.32×10^{-2}	1.99
64 ×64	3.42×10^{-5}	3.00	5.81×10^{-3}	2.00

the velocity field and $k = 2$ for the pressure field. For the Q1Q1s element $k = 2$ for both fields. For this example then, both the LBB-stable and stabilized element show optimal convergence behavior. The Q1Q1 element without stabilization fails to converge.

3.2.2 Darcy Flow

We now consider the other limiting case of the coupled model: the limit as the solid matrix becomes rigid. In this case, the governing equations reduce to the Darcy flow equations,

$$w + \frac{k}{\mu}\nabla p = f \tag{3.18}$$

$$\nabla \cdot w = 0 \tag{3.19}$$

Assuming that the permeability tensor k is invertible, we may solve the equations in mixed weak form as

$$\left(\eta, \left(\frac{k}{\mu}\right)^{-1} w\right) - (\nabla \cdot \eta, p) = (\eta, f) - (n \cdot \eta, p)_{\Gamma} \tag{3.20}$$

$$(q, \nabla \cdot w) = 0 \tag{3.21}$$

Table 3.2: Comparison of observed convergence behavior in the Darcy example for the stabilized Q1Q1s (with $\tau = 0.001$) and LBB-stable Q2Q1. (Top) L_2 pressure error and estimated convergence order. (Bottom) L_2 velocity error and estimated convergence order.

Mesh	Q2Q1 $\|e_p\|$	k	Q1Q1s $\|e_p\|$	k
4×4	4.92×10^{-1}	–	5.38×10^{-1}	–
8×8	1.11×10^{-1}	2.15	1.10×10^{-1}	2.29
16×16	2.61×10^{-2}	2.09	2.60×10^{-2}	2.08
32×32	6.42×10^{-3}	2.02	6.41×10^{-3}	2.02
64×64	1.60×10^{-3}	2.01	1.60×10^{-3}	2.01

Mesh	Q2Q1 $\|e_w\|$	k	Q1Q1s $\|e_w\|$	k
4×4	1.02×10^{0}	–	7.48×10^{1}	–
8×8	3.37×10^{-1}	1.60	1.55×10^{-1}	2.27
16×16	9.55×10^{-1}	1.82	3.68×10^{-2}	2.08
32×32	2.52×10^{-2}	1.92	9.07×10^{-3}	2.02
64×64	6.45×10^{-3}	1.97	2.26×10^{-3}	2.01

For this model, we require $w \in H(\text{div})$ and $p \in L_2$. Note that this formulation is similar to the $u/p/w$ approach of treating Darcy's law in weak, rather than strong, form.

The presence of the divergence constraint implies that the resulting discrete system is again subject to stability restrictions. Proceeding just as before, let the exact solution be given by (3.17). We can solve for the required body force f from the momentum equation, and apply this force in the finite element discretization, along with the necessary boundary conditions.

Table 3.2 presents the convergence results for this example problem. For the pressure fields, the rate of convergence for both elements is the same as in the Stokes example, with $k = 2$. The behavior with respect to the velocity field is a different story, however. This time, *both* elements show $k = 2$ convergence for the velocity field. In fact, the stabilized Q1Q1s slightly outperforms the stable Q2Q1 in terms of overall accuracy. Even though the Q2Q1 element possesses a higher interpolation order, the convergence of the velocity field in this example seems to be limited by the linear pressure approximation. For this example, the use of equal order interpolation thus seems particularly compelling.
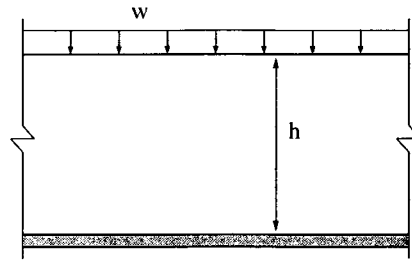
Figure 3.4: Geometry for Terzaghi's one-dimensional consolidation problem.

3.2.3 One-Dimensional Consolidation

The previous numerical examples represent limit states in which the porous medium is either impermeable but deformable (Stokes) or permeable but rigid (Darcy). The most common cases of interest, however, are porous media that are both permeable and non-rigid. Developing analytical solutions in this regime is more challenging, though several well-known examples exist [Cryer 1963; Terzaghi and Peck 1967].

In 1923, Terzaghi proposed the first theory to consider the one-dimensional consolidation of a saturated soil layer as a result of a uniform surface traction (Figure 3.4). In particular, he considered a homogeneous layer of height h resting on a rigid, impermeable bed. The upper surface is freely drained, so that the pore pressure at the surface boundary remains at atmospheric conditions at all times. At some initial time, a uniform traction w is applied at the surface. The sudden application of the load leads to an instantaneous rise in pore pressure throughout the soil layer. Since fluid is allowed to freely drain through the surface, however, there is a gradual dissipation of this excess pressure with time, and an increase in effective stress. As a result, surface settlement is observed.

For the simulations, we use a linear elastic model for the soil constitutive behavior. Because of the one-dimensional nature of the problem, the domain is discretized using a single column of quadrilateral elements. On the left and right boundaries, displacements are fixed in the horizontal direction. On the lower boundary, displacements are fixed vertically. For the fluid flow, all boundaries are impermeable, except for the upper surface where a prescribed pressure ($p = 0$) is set. For the purposes of comparison, we consider two cases: a high-permeability case (with $k_1 = 10^{-11}$ m^2) and a low-permeability case (with $k_2 = 10^{-14}$ m^2). The remaining parameters used for the simulation are as follows: surface load $w = 1$ kPa, bulk modulus $k = 100$ kPa, Poisson ratio $\nu = 0.2$, dynamic viscosity of water $\mu = 10^{-6}$ kPa·s, time step $\Delta t = 0.25$ s, stabilization $\tau = 1.0$. The domain was discretized with 32 quadrilateral elements.

We begin with the low-permeability simulation. Figure 3.5 presents the pressure profiles after the initial time step. When the surface traction is applied, the pore pressures throughout the domain rise to be equal to the overburden stress ($p = w$), except for in a thin layer near the drainage boundary. The unstable Q1Q1 element, however, predicts wild oscillations in the pressure field. In contrast, the stable Q2Q1 and stabilized Q1Q1s perform well.
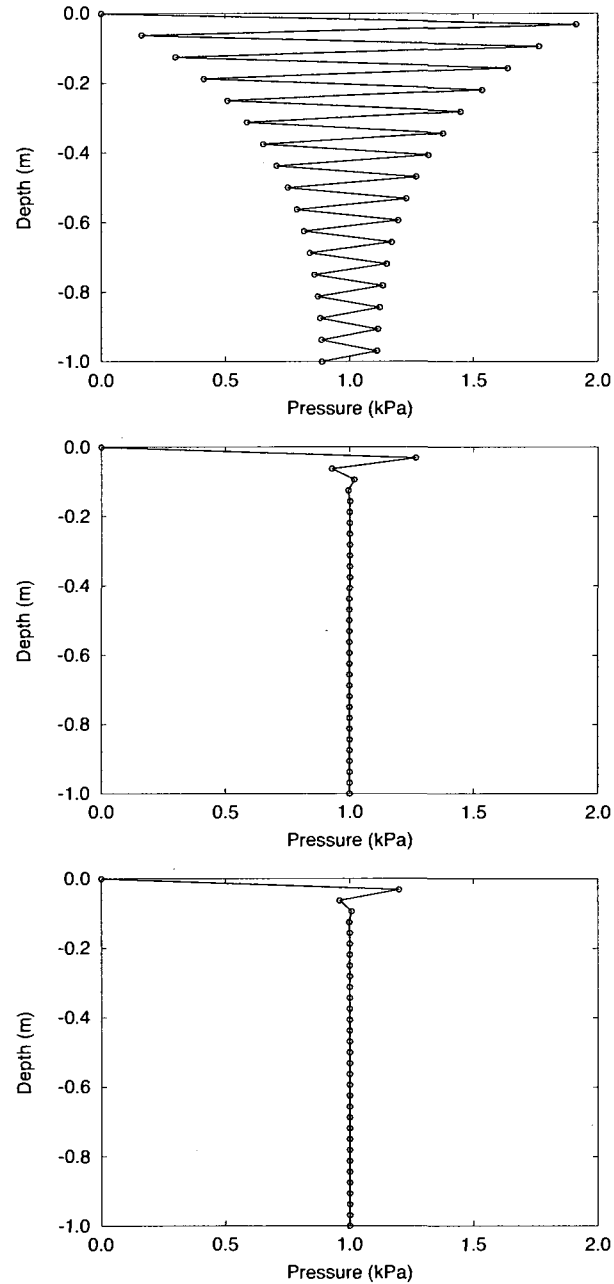
Figure 3.5: Comparison of excess pressure solutions in the first time step for Terzaghi's one-dimensional consolidation problem with $k = 10^{-14}$ m^2. Simulations results are presented for (top) unstable Q1Q1 elements, (middle) stable Q2Q1 elements, and (bottom) stabilized Q1Q1s elements.
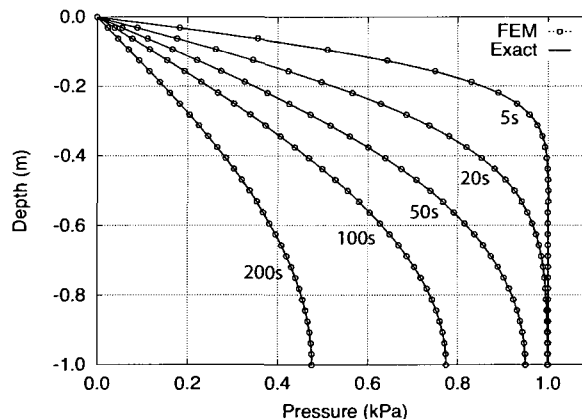
Figure 3.6: Comparison of the numerical and exact excess pressure solutions for the high-permeability case ($k = 10^{-11}$) using the stabilized formulation.

Note, however, that this simulation reveals another numerical difficulty that arises from the use of continuous pressure elements. Upon loading, the pressure throughout the domain is equal to w, but the condition $p = 0$ is strongly enforced at the surface. In reality, a thin boundary layer exists where there is a smooth transition from $p = 0$ to $p = w$, but the discrete mesh is unable to resolve this layer. As a result of the extremely sharp gradient, pressure overshoots are observed near the surface for early times. This effect is occasionally referred to as Gibbs' Phenomenon, in reference to a similar observation about the behavior of finite Fourier series interpolations near jump discontinuities [Gibbs 1898]. In this situation, a discontinuous pressure interpolation would help, as there is no hope of resolving a smooth transition. Alternatively, additional numerical diffusion could be added near the sharp gradient. Nevertheless, these minor oscillations do not propagate to the rest of the domain, and the performance of the two stable elements is far superior to the standard Q1Q1. We also note that with progressing time, the initial errors in the pressure field for the unstable Q1Q1 element slowly dissipate. A careful study of this effect can be found in [Murad and Loula 1994]. Clearly, however, the behavior at early times is unacceptable.

Before concluding this example, we examine the high permeability case. Due to the higher conductivity, drainage and settlement take place on a much shorter time scale. Figure 3.6 compares the finite element solution using the stabilized scheme with the exact analytical solution at several time steps, with good results. Upon refining the spatial and temporal discretizations, the stabilized scheme converges to the exact solution.

3.2.4 Three-dimensional Footing

Before concluding, we should emphasize that the same stabilization scheme can also be applied in three-dimensions. For our final example, we examine the three-dimensional footing previously introduced in Figure 3.1. There, strong pressure oscillations were observed using the unstable Q1Q1 hexahedra. Figure 3.7, in contrast, shows the results when using the proposed Q1Q1s formulation.
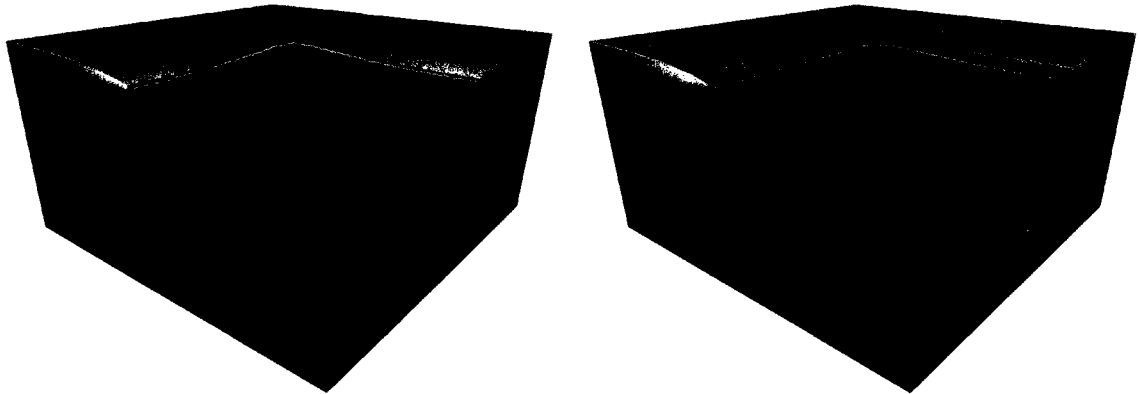
Figure 3.7: Predicted pressure field for the undrained footing problem using (left) stable Q2Q1 elements and (right) stabilized Q1Q1 elements. The unstable Q1Q1 element solution is presented in Figure 3.1

Upon comparison with the stable Q2Q1 solution, we find excellent agreement (both in the "eye norm" and quantitatively).

The three-dimensional example provides the most compelling computational argument for using a stabilized scheme. The stabilized Q1Q1s mesh has 19,652 degrees of freedom, while the Q2Q1 has 112,724 degrees of freedom. Furthermore, the Q1Q1s element requires 8 Gauss integration points per element for exact integration, while the Q2Q1 element requires 27. The total computational is therefore much reduced, while still obtaining good convergence behavior.

At this point, we will leave off with further numerical examples. We note that the simulations in later chapters use the same stabilization scheme, and illustrate its effectiveness for elastic, hyperelastic, and elastoplastic material models. The same formulation will also be applied to saturated, unsaturated, and variably-saturated media, with good results. Hopefully the broad array of applications will further serve to justify its robustness and generality.

3.3 Conclusion

Stabilized methods can offer tremendous computational advantages over standard approaches. In particular, one can employ meshes with fewer degrees of freedom, fewer Gauss points, and simpler data structures. The additional stabilization terms can also improve the convergence properties of iterative solvers. These factors become crucial when considering large-scale, coupled, three-dimensional problems.

In this work we have proposed a stabilization scheme to allow for the use of Q1Q1 elements in both 2D and 3D. The same scheme can also be applied to simplicial elements. The method employed has several appealing features. It requires only a minor modification of standard finite element codes, and adds little additional computational cost to the assembly routines. All necessary computations can be performed at the element level using standard shape-function information,

and no higher-order derivatives or stress-recovery techniques must be employed. It also leads to a symmetric modification of the system matrix, which is advantageous if the underlying problem is symmetric.

As the numerical examples have demonstrated, the stabilization scheme is robust and leads to high-quality solutions. In our opinion the key disadvantage is that the resulting solution may be overly diffusive in the presence of extremely sharp gradients. As we have indicated, however, this effect can be easily controlled using the stabilization parameter τ.

As far as future work is concerned, we believe additional investigation should be devoted to exploring the behavior of stabilization schemes near boundaries, as it is well known that many schemes have degraded accuracy in these regions [Becker and Braack 2001]. More detailed studies of the interaction of the stabilization scheme with the time integration method would also be of interest. The general conclusion, however, is that the computational advantages offered by stabilized methods are very appealing.

4 Solution Methods

We began this work by formulating a coupled model for the behavior of saturated and unsaturated porous geomaterials. We then designed a stabilized finite element method to discretize the coupled formulation with efficient, equal-order elements. From a computational point of view, this equal-order interpolation is far more appealing than the unequal interpolation commonly used in practice. Nevertheless, even the stabilized discretization can lead to very large algebraic problems, especially in three-dimensions. The size and ill-conditioning of these systems demands special attention, and is the focus of this chapter.

After the introduction of spatial and temporal discretizations, recall that we are left with a system of n nonlinear residual equations in the form

$$R(x) = 0 \tag{4.1}$$

that must be solved in each time step to advance the solution. For typical problems, n may be only a few thousand, or perhaps several million. There is therefore a clear need to develop robust, efficient, and scalable solution techniques for (4.1).

To deal with the nonlinearity inherent in our problem, we can use Newton's method (or a Newton-like method) to solve for the solution $\overset{*}{x}$ satisfying (4.1). A simple illustration of Newton's method in one-dimension is given in Figure 4.1. Given a trial configuration x_k, we construct a linear approximation to the governing equations about this configuration,

$$R(x) \approx R(x_k) + J_k \cdot \Delta x \tag{4.2}$$

where $J_k = [\partial R/\partial x]_k$ is the $n \times n$ Jacobian of the system. This leads to an iterative update scheme in which we solve

$$J_k \cdot \Delta x = -R(x_k) \tag{4.3}$$

for an increment Δx and update $x_{k+1} = x_k + \Delta x$. Given a good initial guess x_0 and some mild assumptions about $R(x)$, Newton's method will show quadratic convergence in a neighborhood of $\overset{*}{x}$. Clearly, however, a good initial guess is essential to success. Since little can be said about its global convergence properties, additional "globalization" procedures—e.g. linesearch and trust region methods—are typically included to improve robustness.
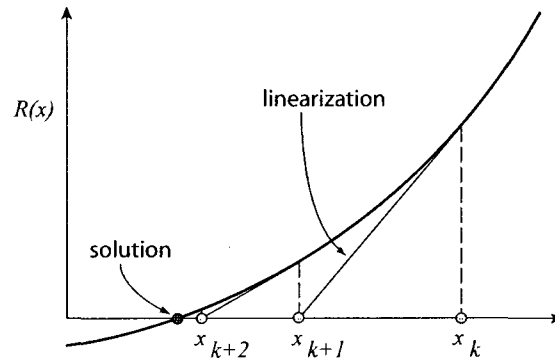
Figure 4.1: Illustration of Newton's method for a scalar function.

The development of practical methods for solving (4.1) relies on two key ingredients. First, we need a method for solving very-large systems of the form (4.3) as efficiently as possible. Second, the given linear solver must then be embedded within a nonlinear solver that actually implements Newton's method. Since the bulk of the computational expense in these simulations is spent solving linear systems, a goal of the nonlinear solver is to minimize the expense of each Newton update. At the same time, the nonlinear solver should include additional checks to ensure that only good update steps are taken. In this chapter we will explore both ingredients.

4.1 Linear Solvers

In this section, our focus is on efficient solution methods for the linear system (4.3). The solution of these systems form the basis for any Newton or Quasi-Newton algorithm, and the effectiveness of the linear solver is crucial to the performance of the nonlinear algorithm.

4.1.1 Direct and Iterative Solvers

Before focusing on the particular structure of our coupled problem, we remark that there are two broad classes of methods for solving linear systems of the form $Jx = b$, direct and iterative. The properties of each approach play a significant role in selecting the most appropriate solution algorithm, and so we provide a brief review of their primary features. Before we begin, however, we remark that these sorts of discussions are often phrased as direct *versus* iterative—with the implication that this is an either/or choice. Fortunately, the current spectrum of solution methods is much more complex, and the traditional advantages and disadvantages of the two methods have quickly blurred with the rise of more sophisticated implementations and even hybridized methods.

The first basic class encompasses direct solvers, in which one attempts to form an explicit factorization of the system matrix J. For example, a direct solver may attempt to form an LU or Incomplete-Cholesky factorization, depending on the symmetry properties of the system matrix.

There are several advantages to such an approach. First, the computational time for a direct factorization is typically fixed by the size of the matrix and number of nonzero entries, and is unaffected by the conditioning of J. As a result, for ill-conditioned matrices, direct factorizations are very appealing. Furthermore, once an explicit factorization is computed, the forward substitution phase required to compute the solution x for a given right hand side b requires a trivial amount of additional work. Therefore, if one needs to solve systems of the form $Jx_i = b_i$ with a constant J but many different right hands sides b_i, J need only be formed and factored once. The cost of the factorization can then be amortized over many right hand sides. As this situation arises frequently in practice, this is another key appeal of direct methods. Indeed, in our later discussion of Schur-complement reduction we will run into precisely this situation.

The key difficulty in developing scalable algorithms using sparse direct solvers, however, is their memory usage. While J is sparse, its factors are typically much denser. Thus, if one computes an LU factorization of J, large amounts of additional memory must be allocated to account for fill-in of entries within L and U. Careful re-ordering of the unknowns can minimize this fill-in, but by only so much. These re-ordering strategies are also expensive in and of themselves.

For 2D problems, this drawback is usually not severe enough to present a serious bottleneck. Also, even large 2D problems rarely have more than a few hundred-thousand unknowns. Direct factorizations of even modest 3D problems, however, can quickly overrun the available memory on a typical workstation. As a result, direct factorizations may be infeasible for large, 3D simulations unless substantial amounts of memory are available. This fact has prompted a significant effort at developing parallel sparse direct solvers, which maintain the advantages of a direct solve while using the much larger memory resources of shared- and distributed-memory platforms. Many direct-solvers also exploit out-of-core memory management to minimize memory bottlenecks.

In summary, direct solvers are typically robust and easy to use, but memory-intensive. Iterative solvers, loosely speaking, are the reverse. Instead of explicitly factoring J, iterative solvers attempt to form a sequence $\{x_m\}$ that converges to the desired solution. Since no factors must be stored, the additional memory-requirements associated with dense fill-in are avoided. The major downside to iterative methods is that they are very sensitive to the spectral properties of J. For ill-conditioned matrices, preconditioning techniques are therefore essential to achieving quick convergence. For example, instead of solving the original system $Jx = b$, we can introduce a (left) preconditioning matrix P and solve the preconditioned system

$$(P^{-1}J)x = P^{-1}b \tag{4.4}$$

If the matrix $(P^{-1}J)$ has better spectral properties than J itself, then the iterative solver will converge in fewer iterations. Of course, one must balance this decrease in iterations with the expense of computing P^{-1}. For example, neither $P^{-1} = I$ or $P^{-1} = J^{-1}$ are good preconditioning choices. In the first case, the preconditioner is trivial to apply, but leads to no convergence improvement at all. In the second case, the resulting system is perfectly conditioned and trivial to solve, but it is as expensive to compute the preconditioner as to solve the original system itself.

Unfortunately, since preconditioning is often problem-specific and requires careful tuning, iterative solvers are not as robust or easy to use as direct solvers. Thus, most commercial finite element codes provide a direct solver as their default choice, though iterative options are becoming increasingly frequent. Indeed, many good *black-box* preconditioners have been developed in recent years that work well for a wide-range of problems, with little or no tweaking on the part of the user. Typically, these are algebraic preconditioners that can be constructed from the matrix entries in J with no knowledge of the underlying problem being discretized. Unfortunately, these general-purpose preconditioners typically work best for algebraic problems that arise when solving scalar PDEs, or from single-physics applications. They are typically less successful when applied to multi-physics applications, like the coupled solid-deformation/fluid-flow problems considered in this work. In these situations, much of the ill-conditioning arises from attempting to couple governing equations which have widely different magnitudes and mathematical character. In these situations, physics-based preconditioning is often more appropriate. The preconditioner is endowed with knowledge of the underlying problem being discretized. By being problem specific, however, physics-based preconditioners are less widely applicable than their algebraic counterparts. Again, however, middle ground exists. In this work, we will explore preconditioning techniques that are aware of the coupled nature of the governing equations, but still use black-box preconditioning techniques to their best advantage.

A final advantage of iterative solvers is that the user specifies the accuracy with which a given linear system should be solved. In many cases, it is sufficient to only find a rough approximation to the solution—perhaps we only need to know x to one or two significant digits—and so computational effort can be saved by specifying a high tolerance for convergence. A direct solver always provides an "exact" solution, at constant computational cost, and so we cannot save effort when only an approximate solution is necessary. We will explore this feature as we discuss inexact Newton methods.

As stated earlier, the basic features of iterative and direct solvers cited here are generalizations, and current research is focused on overcoming these traditional disadvantages. These improvements may either come from modifying the original techniques, or combining direct and iterative methods into a hybrid method. For example, block iterative methods have been developed that are more efficient when solving with many right hand sides—a traditional strength of direct solvers [Saad 2003]. In multigrid methods, an iterative solver may be used to solve the full problem on the fine mesh, but the coarse-grid problems are sufficiently small to be handled by a direct solver. Similarly, when examining block-structured or domain-decomposed problems, a direct solver may not be able to handle the global problem, but can efficiently solve smaller sub-problems. An iterative phase can then be used to compute the global solution efficiently using the sub-problem solutions.

4.1.2 Krylov-Subspace Methods

Within the broader class of iterative methods, we are primarily interested in Krylov-subspace methods. This class encompasses a variety of approaches—e.g. Conjugate Gradients (CG) for symmetric systems and Generalized Minimum Residual (GMRES) and Stabilized Biconjugate Gradients (BICGSTAB) for non-symmetric systems, along with many others [Saad 2003]. To illustrate the basics of Krylov iterative methods, let the approximate solution at an iteration m be denoted by x_m, starting from an initial guess x_0. The residual at each iteration is $r_m = b - Jx_m$, with initial residual r_0. To avoid confusion, note here the use of a lowercase r_m for the linear residual at solver iteration m, distinct from the nonlinear residual R_k at Newton iteration k as introduced in equation (4.1). Now, let the mth-order Krylov subspace $\mathcal{K}_m$ be defined as

$$\mathcal{K}_m(J, r_0) = \mathrm{span}\{r_0, Jr_0, J^2 r_0, ..., J^{k-1} r_0\} \tag{4.5}$$

In general, a Krylov subspace method is a projection method in which one determines the approximate solution x_m from the affine subspace $x_0 + \mathcal{K}_m$ such that

$$r_m \perp \mathcal{C}_m \tag{4.6}$$

for a m-dimensional subspace $\mathcal{C}_m$ (the *constraints space*). Different versions of Krylov methods arise from different choices for $\mathcal{C}_m$. For example, in GMRES one chooses $\mathcal{C}_m = J\mathcal{K}_m$. In the resulting iteration, one can show that we approximate the exact solution to $Jx = b$ by the vector $x_m \in x_0 + \mathcal{K}_m$ that minimizes the norm of the residual, $\|r_m\|_2$.

A key observation is that, in the construction of the Krylov subspace $\mathcal{K}_m$, only matrix-vector products of the form $w = Jv$ are required. This fact has important algorithmic implications. In particular, when using Krylov methods we never need an explicit representation of J. All we need is a function that can return the matrix-vector product Jv given an input vector v. In many cases, J may be difficult or expensive to compute explicity, but we can develop cheap methods to compute matrix-vector products with it.

As a simple illustration, recall that the Jacobian matrix J_k is explicitly computed as

$$J(x_k) = \frac{\partial R}{\partial x}(x_k) \tag{4.7}$$

at some configuration x_k. Instead of explicitly forming and storing the system matrix J_k, the vector product $J_k v$ can often be accurately-approximated by a finite-difference directional derivative,

$$J_k v \approx \frac{R(x_k + \epsilon v) - R(x_k)}{\epsilon} \tag{4.8}$$

where ϵ is a scalar perturbation parameter. When using a Krylov-subspace method, it may then be possible to avoid evaluating the system matrix altogether, as the required matrix-vector products

| (a) Natural ordering | (b) Block sorted | (c) Block sorted, King ordering |
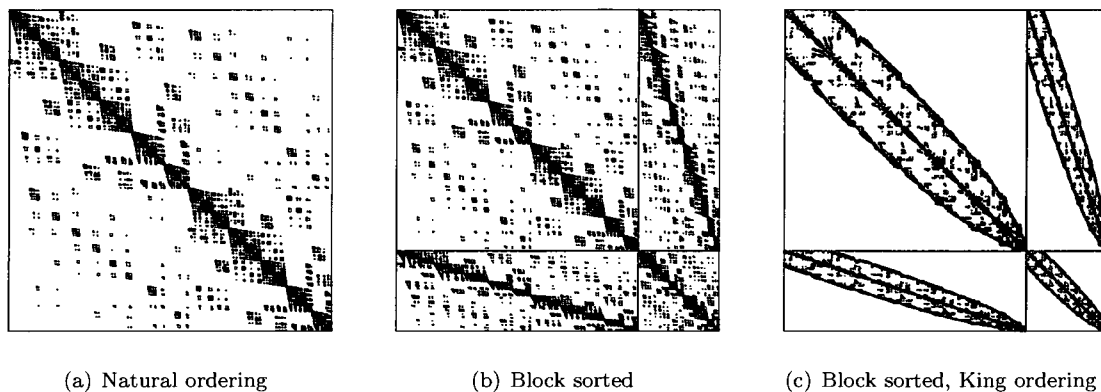
Figure 4.2: Sparsity patterns for the Jacobian J_k using various degree-of-freedom ordering strategies. Symbols represent non-zero entries in the matrix.

can be approximated instead with two residual-function evaluations. This observation leads to so-called *matrix-free* Newton-Krylov methods [Knoll and Keyes 2004]. Of course, in practice the story is not so simple, as J_k may be ill-conditioned. Often one ends up allocating some matrix storage for preconditioning purposes, leading to a "pseudo" matrix-free method. Nevertheless, the underlying principle is quite useful if J_k cannot be evaluated analytically, or if it is too expensive to compute and store.

As they are not the focus of the current work, we will not discuss matrix-free methods further. We will, however, exploit the matrix-vector properties of Krylov methods in a variety of other contexts, to implicitly represent quantities that cannot be explicitly formed. For the current case, evaluating the desired Jacobian analytically is not difficult, and so we will explicitly compute J_k. Further, when using complex elastoplastic material models, the additional time to evaluate the Jacobian and residual is not significantly greater than just evaluating the residual. In both cases an expensive stress integration routine must be called, and this may dominate the assembly time. From a memory point of view, however, the ability to avoid storing the Jacobian would be tremendously appealing. Unfortunately, Jacobian-free methods have not been well-explored for the class of problems considered in this work. They therefore represent an interesting avenue for future research.

4.1.3 Block Partitioning

With these general observations in mind, we now return to coupled problems, beginning with the u/p formulation. Because this FE model derives from two coupled governing equations, the discrete problem inherits an interesting 2×2 block-structure. In particular, the Jacobian system can be partitioned such that

$$\begin{bmatrix} A & B_1 \\ B_2 & C \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \end{bmatrix} = - \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \end{bmatrix}_k \tag{4.9}$$

where the A block is $n_u \times n_u$ and the C block is $n_p \times n_p$. Because the pressure is a scalar field, while the displacement is a vector field with as many components as there are spatial dimensions, the dimensions of the A block will typically be much larger than those of the C block. For an equal-order interpolation in three-dimensions (i.e. using a stabilized scheme), $n_u = 3n_p$. If a higher-order interpolation is used for the displacement field, then this size disparity is further exaggerated.

Figure 4.2 illustrates typical sparsity patterns for the Jacobian using various degree-of-freedom reordering strategies. The dark symbols denote non-zero entries in the system matrix, implying a coupling between the degrees of freedom for that row/column pair. Using a natural ordering, in which degrees of freedom are numbered in a more-or-less arbitrary way, no clear pattern is apparent. Pressure degrees of freedom are arbitrarily interlaced with displacement degrees of freedom. If however we first number all of the displacement degrees of freedom, only then followed by all of the pressure degrees of freedom, the block-structure expressed in (4.9) becomes readily visible. In this example, equal-order interpolation was used on a three-dimensional mesh, so the A block is three times as large as the C block. We can take this reordering process even further, renumbering degrees of freedom within each block. For example, we can use a Cuthill-McKee [Cuthill and McKee 1969] or King algorithm [King 1970] to attempt to minimize bandwidth, or a Minimum-Degree algorithm [George and Liu 1989] to attempt to minimize fill-in during a direct factorization.

Now, when attempting to solve such systems, one option is to ignore this inherent structure and simply treat the linear system monolithically. The alternative is to develop linear solution techniques that directly exploit the block-partitioning. As we shall see, these block-solution techniques can provide significant advantages in speed and memory usage in comparison to attacking the monolithic problem directly. The improvement in performance derives from the fact that we incorporate additional information about the underlying coupled PDEs into the solution procedure. In the next two sections we consider two block-based solution techniques. In the first section, we describe the classic technique of Schur-complement reduction. In the second section, we describe a block-preconditioning technique that circumvents some inherent bottlenecks in the Schur-complement approach. Having discussed the 2×2 problem, we will then extend the same framework to address the solution of linear systems arising from $u/p/w$ discretizations:

$$\begin{bmatrix} A & B_1 & \\ B_2 & C & D_1 \\ & D_2 & E \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \\ \Delta w \end{bmatrix} = - \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \\ R_{\text{darcy}} \end{bmatrix} \tag{4.10}$$

In this case, the system matrix has a 3×3 block-tridiagonal structure that can be put to good use.

As a preliminary observation, the 2×2 (u/p) system matrix admits a variety of block-factorizations [Benzi et al. 2005]. For example, a block LDU factorization is

$$J = \begin{bmatrix} I & 0 \\ B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & A^{-1}B_1 \\ 0 & I \end{bmatrix} \tag{4.11}$$

where $S = C - B_2 A^{-1} B_1$ is the *Schur-complement* (of A) for the system. Equivalent LU factorizations are

$$J = \begin{bmatrix} A & 0 \\ B_2 & S \end{bmatrix} \begin{bmatrix} I & A^{-1}B_1 \\ 0 & I \end{bmatrix} \tag{4.12}$$

and

$$J = \begin{bmatrix} I & 0 \\ B_2 A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B_1 \\ 0 & S \end{bmatrix} \tag{4.13}$$

These decompositions will prove useful in subsequent sections.

4.1.4 Schur-Complement Reduction

Schur-complement reduction is a common technique for solving block 2×2 systems of the form (4.9). Assuming that A is nonsingular, we can formally solve for Δu in terms of Δp using the first row of (4.9),

$$\Delta u = A^{-1} (R_{\text{mom.}} - B_1 \Delta p) \tag{4.14}$$

Substituting this expression for Δu in the second row, we arrive at a reduced $n_p \times n_p$ linear system

$$\left(C - B_2 A^{-1} B_1 \right) \Delta p = R_{\text{mass}} - B_2 A^{-1} R_{\text{mom.}} \tag{4.15}$$

where, again, $S = C - B_2 A^{-1} B_1$ is the Schur-complement. This leads to a two-step procedure for solving (4.9):

- Determine the pressure increment Δp by solving the reduced Schur system

$$S \Delta p = R_{\text{mass}} - B_2 A^{-1} R_{\text{mom.}} \tag{4.16}$$

- Determine the displacement increment Δu by solving

$$A \Delta u = R_{\text{mom.}} - B_1 \Delta p \tag{4.17}$$

Schur-reduction is simply block Gaussian elimination applied to J. On the surface, this two step approach appears to save quite a bit of work. The original system matrix was $(n_u + n_p) \times (n_u + n_p)$. Using the Schur-reduction approach, we need only solve a $n_p \times n_p$ system followed by a $n_u \times n_u$ system. The catch is that S contains a term involving A^{-1}, which in this case means that S is dense and expensive to form, let alone solve with.

In this case iterative methods provide a natural solution. To solve the system (4.16) using a Krylov-method, we never need an explicit representation of S, only the ability to form matrix-vector products with it. For example, given an input vector v_{in} we can compute the output vector

$v_{out} = Sv_{in}$ in a three step process:

- Set $y_1 = B_1 v_{in}$

- Solve $Ay_2 = y_1$

- Return $v_{out} = Cv_{in} - B_2^T y_2$

With this implicit representation, it is straightforward to solve the reduced Schur system without ever having to work directly with a dense S. The key downside to this approach, however, is that every time we need to multiply by S we must perform a solve with A. Using another iterative solver to deal with A establishes an "inner/outer" iterative scheme. At each outer iteration with S we perform inner iterations with A. As a result, we may end up solving with A dozens or even hundreds of times. The success of this method therefore hinges on the ability to develop good preconditioners P_A for A and P_S for S.

Fortunately, recall that the A sub-block is equivalent to the stiffness matrix one encounters in a simple solid-mechanics setting, without any coupled fluid flow. Solving the uncoupled, displacement problem is significantly easier than solving the fully-coupled solid/fluid problem. Many good preconditioners have been designed over the years for elastic and elastoplastic problems, and the Schur-reduction approach can directly incorporate these methods. Also, developing a good preconditioner for S in the current case is not too difficult. Therefore, since efficient methods exist for solving with A and S, the Schur-reduction approach can be quite practical—hence its popularity.

For two-dimensional and small three-dimensional problems, it may be possible to use a direct solver on A. Indeed, it is much easier to factor A alone than the entire coupled system J. It is also only necessary to factor A once, since it remains constant within each Newton iteration. Solving with many right hand sides can then be performed at trivial additional cost. The direct factorization avoids having to perform any inner iterations, and leads to a hybrid iterative/direct method that is tremendously efficient.

Unfortunately, this approach is limited to problems where the A block is sufficiently small that a direct solver can handle it. For very large problems, we must revert to an iterative inner solver. In the next section we consider an alternative approach that circumvents the outer/inner scheme and avoids the inner solver bottleneck.

4.1.5 Block u/p Preconditioning

Using the Schur-reduction approach, one breaks the coupled system J into two smaller, more manageable pieces. By breaking the problem into these pieces, we can use efficient methods that have been developed for the sub-problems. The downside is that we need to repeatedly solve with A, the largest sub-block. In this section an alternative approach is considered that leverages good solution methods for the sub-blocks, but avoids the inner/outer iteration bottleneck. To do so, we solve the full $(n_u + n_p) \times (n_u + n_p)$ system J all at once, but build a block-structured preconditioner P. In

particular, consider a block triangular preconditioner P_L of the form

$$P_L = \begin{bmatrix} A & 0 \\ B_2 & S \end{bmatrix} \quad \text{such that} \quad (P_L^{-1}J)x = (P_L^{-1}b). \tag{4.18}$$

The subscript L denotes that this preconditioner acts from the left. The equivalent right preconditioner is

$$P_R = \begin{bmatrix} A & B_1 \\ 0 & S \end{bmatrix} \quad \text{such that} \quad (JP_R^{-1})(P_Rx) = b. \tag{4.19}$$

Preconditioners of this type were first studied in [Bramble and Pasciak 1988], and have received significant attention in recent years in the context of solving the Stokes and Navier-Stokes equations [Elman et al. 2008]. Note that these choices of P_L and P_R are simply the lower and upper triangular factors in (4.12) and (4.13), respectively. Focusing on P_L for now, its inverse is also lower triangular,

$$P_L^{-1} = \begin{bmatrix} A^{-1} & 0 \\ S^{-1}B_2A^{-1} & S^{-1} \end{bmatrix} \tag{4.20}$$

Applying this preconditioner to the system matrix,

$$P_L^{-1}J = \begin{bmatrix} A^{-1} & 0 \\ S^{-1}B_2A^{-1} & S^{-1} \end{bmatrix} \begin{bmatrix} A & B_1 \\ B_2 & C \end{bmatrix} = \begin{bmatrix} I & A^{-1}B_1 \\ 0 & I \end{bmatrix} \tag{4.21}$$

The preconditioned matrix is block upper triangular, with identity blocks on the diagonal. It is straightforward to show that the preconditioned matrix has only one distinct eigenvalue ($\lambda = 1$), and a Krylov-based iteration on the preconditioned system would converge in at most two iterations [Benzi et al. 2005].

In practice, we should not form the "exact" preconditioner P_L indicated above. Note that P_L involves inverses of A and S, and so constructing P_L would be just as expensive as solving the original system J through the Schur-reduction technique. Instead, we will try to build a preconditioner $\tilde{P}_L$ that is a good approximation to P_L, but is much less expensive. Assuming the approximation is good, we can expect an iterative solver to converge in relatively few iterations (though likely more than two).

Also, note that the preconditioned matrix is nonsymmetric, even if the underlying problem had an original symmetry. In this form, we can no longer use a symmetric method such as the standard CG iteration, which is fast to apply on a per-iteration basis and has a short-recurrence relation. A nonsymmetric method like GMRES works, but is more expensive and needs to store previous basis vectors. In [Bramble and Pasciak 1988], the authors recovered a symmetric method by using a non-standard inner product inside the CG iterations. In practice, the use of such a symmetrized method may be unnecessary. If the preconditioner works well and the number of iterations is kept

small, the disadvantage to using a slower, nonsymmetric iteration such as GMRES will not be great. Further, in many cases (e.g unsaturated flow) the original system matrix is already nonsymmetric.

The form of $\tilde{P}_L$ involves two levels of approximation. Recall that S is explicitly dense. The first step is therefore to replace the dense S with a sparse approximation $\tilde{S}$. Then, instead of using explicit inverses A^{-1} and $\tilde{S}^{-1}$, these blocks are replaced with their own preconditioners, which by design are good inverse approximations. That is, we introduce sub-preconditioners P_A^{-1} for A^{-1} and P_S^{-1} for $\tilde{S}^{-1}$. With these component pieces, the final form of $\tilde{P}_L^{-1}$ is then

$$\tilde{P}_L^{-1} = \begin{bmatrix} P_A^{-1} & 0 \\ P_S^{-1} B_2 P_A^{-1} & P_S^{-1} \end{bmatrix} \tag{4.22}$$

We now consider the concrete implementation of $\tilde{P}_L^{-1}$ and the right-acting equivalent $\tilde{P}_R^{-1}$. For convenience, define the following four block matrices:

$$M_1 = \begin{bmatrix} I & \\ & P_S^{-1} \end{bmatrix} \quad M_2^L = \begin{bmatrix} I & \\ B_2 & I \end{bmatrix} \quad M_2^R = \begin{bmatrix} I & B_1 \\ & I \end{bmatrix} \quad M_3 = \begin{bmatrix} P_A^{-1} & \\ & I \end{bmatrix} \tag{4.23}$$

The the left and right acting inverses can be decomposed as

$$\tilde{P}_L^{-1} = M_1 M_2^L M_3 \quad \text{and} \quad \tilde{P}_R^{-1} = M_3 M_2^R M_1 \tag{4.24}$$

which furnish a ready way to compute the necessary matrix-vector multiply operations. For example, a vector-multiply operation with $\tilde{P}_R^{-1}$ and a partitioned input vector $\langle r_{in}^u, r_{in}^p \rangle$ can be performed piecewise as:

- Step one:

$$r_{out}^p = P_S^{-1} r_{in}^p \tag{4.25}$$

- Step two:

$$r_{out}^u = P_A^{-1}(r_{in}^u + B_1 r_{out}^p) \tag{4.26}$$

Once we have computed P_A^{-1} and P_S^{-1}, the application of the preconditioner in the context of a GMRES iteration is therefore straightforward.

Clearly the Schur-reduction and block preconditioning techniques are strongly connected. Note, for example, the formal similarity of the two-step preconditioner application (eqns. 4.25 and 4.26) and the two-step Schur-reduction solution process (eqns. 4.16 and 4.17). The key distinction is that the Schur-reduction technique forms implicit representations of the exact inverses A^{-1} and S^{-1}, while the block-preconditioning technique makes due with cheaper, approximate inverses P_A^{-1} and P_S^{-1}.

Before concluding this section, we make a final comment about software design. We observe
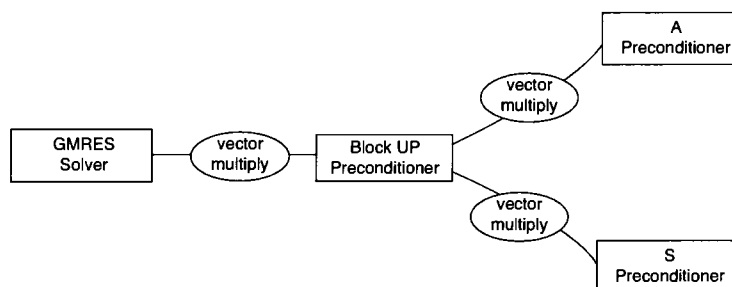
Figure 4.3: Schematic illustration of the interaction between the GMRES solver, block preconditioner, and sub-preconditioners for A and S. All interactions take place through vector-multiply interfaces, leading to a natural modularity in the code design.

that the block u/p preconditioning strategy has an inherent modularity built into it (Figure 4.3). The complete linear solver consists of a GMRES module, a block preconditioning module, and two sub-preconditioners. All interactions between these modules, however, only take place through vector-multiply operations. As a result, the inner workings of each module can be kept hidden from other modules, and need only provide a standardized vector-multiply interface. This approach therefore naturally lends itself to an object-oriented strategy. The modularity makes it very easy to modify or test new component pieces, without having to redesign the entire code. For example, different sub-preconditioners can be readily tested since the block-preconditioner can be written so that it is independent of the inner workings of the sub-preconditioner. Similarly, the GMRES solver is independent on the inner workings of the block-preconditioner. This inherent modularity is reflected in the object-oriented design of our code.

4.1.6 Sub-Preconditioners

In the next few sections, we consider particular forms for the sub-preconditioners P_A and P_S. Note that the use of either the Schur-reduction technique or the block-preconditioner requires two sub-preconditioners, and so the following comments are generally applicable. First, however, we note that good preconditioners for a Schur-complement reduction approach may not necessarily be as effective when working with a block-preconditioned approach, and vice versa. To illustrate this idea, we observe that in the Schur-reduction technique the key bottleneck is the repeated solves with A. As a result, significant performance increases can be observed by developing fast preconditioners for A. In general, a preconditioner that is cheap to build and apply may be more effective than a more expensive one, even if the cheaper one requires a few more iterations. Ultimately it is only the total wall time that matters. In the block-preconditioned approach, however, each iteration is more expensive since we are working with the entire $(n_u + n_p) \times (n_u + n_p)$ system at once, and have to apply a more complicated, block-structured preconditioner. In this case, reducing the number of iterations has a bigger payoff, and so a more expensive but better A^{-1} approximation may be favored. Nevertheless, since both methods are based on similar forms, good P_A and P_S preconditioners for

one should continue to be good for the other. What is *best* for one, however, may not be best for the other.

Sparse Schur Approximations

The first question to be addressed is how to precondition the exact Schur-complement $S = C - B_2 A^{-1} B_1$. While C is sparse, the presence of the inverse stiffness matrix A^{-1} in the second term leads to a dense S which is awkward to deal with. At least two options are available to avoid working with dense matrices.

The first is to approximate the action of S^{-1} through its own linear solver, exactly as in the Schur-reduction technique. In this case, however, only an approximate solve is needed and a large tolerance for convergence can be used. Cheaper approximations of the component matrices could also be used, in an attempt to balance the ease of computing P_S^{-1} with the quality of the approximation.

The second approach is to develop an explicit but sparse approximation to S, and then compute the preconditioner based on this sparse approximation. Again, we consider a few possibilities for how this can be accomplished. Perhaps the simplest approximation, and the one we will focus on in this work, is the approximation

$$\tilde{S} = C - B_2 \operatorname{diag}(A)^{-1} B_1 \tag{4.27}$$

This approximation is easy to compute and has a good sparsity pattern. Another appealing feature is that it accounts for possible assymmetries in B_2 and B_1 (i.e. in the case $B_2 \neq B_1^T$). Of course, the quality of this approximation depends on the diagonal dominance of A.

For the symmetric $B_2 = B_1^T$ case, another possibility to consider is

$$\tilde{S} = C - \alpha M_p, \qquad [M_p]_{ab} = \int_\Omega \phi_a \phi_b \, \mathrm{dV}, \qquad \alpha = \frac{1}{\lambda + 2\mu} \tag{4.28}$$

Here, M_p is the $n_p \times n_p$ pressure mass matrix, i.e., the product of all pressure shape functions $\{\phi_a\}$. The use of the pressure mass matrix to approximate $B_2 A^{-1} B_1$ is motivated by observations in [Verfürth 1984] for the cases of incompressible elasticity and Stokes flow. If we have a linear elastic material model and examine the undrained limit ($k \rightarrow 0$), the Schur complement operator $B_2 A^{-1} B_1$ is the same as that encountered for Stokes flow or incompressible elasticity. In this case, Verfurth demonstrated that the pressure mass matrix is spectrally equivalent to $B_2 A^{-1} B_1$. The current case is somewhat different in that we have nonlinear material models, locally-drained conditions ($k > 0$), and a nonzero C. Nevertheless we have obtained good results using the form (4.28). For low permeability situations, the M_p term tends to dominate, while for high permeability situations the C term tends to dominate. The scalar α is a weighting factor, based on the compressibility of the solid, which is used to ensure that the two components M_p and C of the sparse approximation have appropriate relative magnitudes.

Black-Box Algebraic Preconditioners

Assuming we have explicit (and sparse) representations of A and $\tilde{S}$, the next step is to compute preconditioners P_A and P_S. In our numerical experiments we examine three possibilities, all of which can be considered black-box preconditioners. These preconditioners can be constructed on a purely algebraic basis, without access to the mesh topology, element subroutines, or knowledge of the underlying equations being discretized. In our experiments, we use the same preconditioning strategy for both A and $\tilde{S}$—though with the parameters tuned for each separate matrix. In principle of course one can use differing strategies that are better optimized for each sub-problem. We have found that all three preconditioning strategies, however, work well for both the stiffness and the sparse Schur-complement matrices.

In the first approach, we use a sparse direct solver to factorize A and $\tilde{S}$. The "preconditioners" P_A^{-1} and P_S^{-1} are then essentially exact inverses of A and $\tilde{S}$. Obviously, this approach should minimize the subsequent GMRES iteration count, since the only approximation that has been made is to replace the dense S with the sparse $\tilde{S}$. This approach therefore provides a good means for testing the quality of the sparse Schur complement approximation. These direct factorizations are quite expensive and quickly become memory limited—though less quickly than if we attempted to invert J directly. In a parallel setting, we consider a block-Jacobi approach in which the direct factorization is computed only for the portion of the matrix local to each processor. For the factorizations we use the SUPERLU package [Demmel et al. 1999].

Second, we consider an Incomplete LU preconditioning (ILU). An incomplete LU decomposition of a matrix is a sparse approximation to a true LU decomposition. An appealing feature of ILU is that the quality of the approximation can be controlled by specifying how many non-zero entries are kept in the decomposition. As a result, one can control whether to use a fast and cheap approach, or a more expensive preconditioner with better approximation properties. In this work, we use an ILU(0) preconditioner implemented in the Trilinos-IFPACK package [Sala and Heroux 2005]. In parallel, the ILU preconditioning is computed processor-wise, though with the possibility of some overlap between processors.

Third, we consider Algebraic Multigrid Preconditioning (AMG). For this purpose, we use the AMG preconditioner implemented in the Trilinos-ML package [Gee et al. 2006]. Multigrid methods stem from the recognition that, when solving a linear system, the error has high and low frequency components with respect to the grid spacing. While the high-frequency error can be quickly damped, the low-frequency error tends to persist and slows down the convergence of the solver. In multigrid methods, one instead defines a hierarchy of grids, each coarser than the next, and solves coarsened versions of the problem on this hierarchy. On a coarse grid, the low-frequency errors become high-frequency errors, and can be effectively damped. By using this hierarchy, the error at all frequencies can be removed and swift convergence can be achieved. A key challenge in multigrid methods, however, is defining this grid hierarchy. In *geometric* multigrid, the mesh hierarchy is based on the finite element/difference/volume mesh used to discretize the problem. Because of the difficulty in defining a coarsening scheme on unstructured meshes, however, geometric multigrid has mostly

been limited to structured meshes. *Algebraic* multigrid takes a different approach, constructing the coarser "grids" algebraically using only the entries in the linear system, without any reference to the actual mesh used for the numerical discretization. As a result, algebraic methods can be used on arbitrary mesh topologies. Algebraic multigrid has found the most success not as a solution method in itself, but rather as a method for forming quality preconditioners for use in an iterative solver.

We emphasize the all of the above preconditioning strategies (LU, ILU, and AMG) have shown themselves to be robust and flexible enough to deal with a wide array of model problems. Nevertheless, they are best suited to systems arising from scalar partial differential equations or vector differential equations deriving from a single physics application. Their performance when applied directly to coupled, multiphysics problems can be quite poor. In the methodology proposed here, however, the black-box preconditioners are only applied at the block level, to precondition the stiffness and Schur-complement matrices. As a result, we use these preconditioners to their best advantage, while still endowing the global preconditioning strategy with intuition about the block-structure of the coupled problem

4.1.7 Parallelization Strategies

In the numerical examples to follow, we explore both serial and parallel implementations of the block preconditioner. In serial, the implementation is straightforward. In parallel, however, additional questions arise as to how best to use the block preconditioner to achieve good parallel scalability. In the current work, we have approached the problem as follows: The system matrix J is distributed across processors, and the preconditioners P_A^{-1} and P_S^{-1} are similarly computed in a distributed manner. A global GMRES or BICGSTAB iteration is then used to solve the coupled problem, preconditioned with the distributed block-preconditioner. This approach involves two levels of "decomposition". First, the block approach decomposes the expensive coupled problem into an easier one of computing good preconditioners P_A^{-1} and P_S^{-1}. Second, the expensive problems of computing P_A^{-1} and P_S^{-1} are made easier by distributing them across many processors. Thus, loosely speaking, we first do a block decomposition followed by a domain decomposition.

While this is a natural way to approach the issue of parallelization, we can consider other alternatives. In particular, consider the reverse situation: first decompose by domain, and only then apply the block decomposition. For example, we could use an additive-Schwarz type domain decomposition preconditioner for the global preconditioner. Such preconditioners are based on smaller, approximate inverses built on each processor. To approximate the local Jacobian inverse (which remains coupled) we could then use processor-local block u/p preconditioners.

In this work, we explore the first option. It is nevertheless interesting to consider the second alternative, as we can envision certain situations where processor-local variants of the u/p preconditioner could prove useful.

4.1.8 Extension to $u/p/w$ Discretizations

We can consider a similar block preconditioning methodology for linear systems arising from $u/p/w$ discretizations (or many other three-field formulations). These systems have a 3×3 block-tridiagonal structure:

$$\begin{bmatrix} A & B_1 & \\ B_2 & C & D_1 \\ & D_2 & E \end{bmatrix}_k \begin{bmatrix} \Delta u \\ \Delta p \\ \Delta w \end{bmatrix} = - \begin{bmatrix} R_{\text{mom.}} \\ R_{\text{mass}} \\ R_{\text{darcy}} \end{bmatrix} \tag{4.29}$$

The Jacobian J can be decomposed as

$$\begin{bmatrix} I & & \\ B_2 & I & D_1 \\ & & I \end{bmatrix} \begin{bmatrix} A & & \\ & S & \\ & & E \end{bmatrix} \begin{bmatrix} A & B_1 & \\ & C & \\ & D_2 & E \end{bmatrix} \tag{4.30}$$

where the Schur-complement S now has the form,

$$S = C - B_2 A^{-1} B_1 - D_1 E^{-1} D_2 \tag{4.31}$$

Note that a block LDU decomposition can be recovered from eqn. (4.30) by a suitable rearrangement of rows and columns. We have retained the same $u/p/v$ ordering here merely for convenience of presentation. We propose the following form for the left and right preconditioners,

$$\tilde{P}_L = \begin{bmatrix} A & & \\ B_2 & \tilde{S} & D_1 \\ & & E \end{bmatrix} \quad \text{and} \quad \tilde{P}_R = \begin{bmatrix} A & B_1 & \\ & \tilde{S} & \\ & D_2 & E \end{bmatrix} \tag{4.32}$$

where we have approximated the dense Schur complement S with a sparse $\tilde{S}$ defined as

$$\tilde{S} = C - B_2 \operatorname{diag}(A)^{-1} B_1 - D_1 \operatorname{diag}(E)^{-1} D_2 \tag{4.33}$$

We then use a suitable sub-preconditioning strategy to compute P_A^{-1}, P_S^{-1}, and P_E^{-1} as approximate inverses for A, $\tilde{S}$, and E respectively. To represent the action of $\tilde{P}_L^{-1}$ and $\tilde{P}_R^{-1}$ consider the following four block matrices:

$$M_1 = \begin{bmatrix} I & & \\ & P_S^{-1} & \\ & & I \end{bmatrix} \quad M_2^L = \begin{bmatrix} I & & \\ B_2 & I & D_1 \\ & & I \end{bmatrix} \quad M_2^R = \begin{bmatrix} I & B_1 & \\ & I & \\ & D_2 & I \end{bmatrix} \quad M_3 = \begin{bmatrix} P_A^{-1} & & \\ & I & \\ & & P_E^{-1} \end{bmatrix}$$

$$\tag{4.34}$$

The left and right-acting inverses can then be written as

$$\tilde{P}_L^{-1} = M_1 M_2^L M_3 \quad \text{and} \quad \tilde{P}_R^{-1} = M_3 M_2^R M_1 \qquad (4.35)$$

which, in analogy to the 2×2 cases, furnish a ready way to compute the necessary matrix-vector multiply operations. In the case of exact sub-preconditioners (i.e. using exact inverses A^{-1}, S^{-1}, and E^{-1}) one can again show that the preconditioned system matrix has only a single distinct eigenvalue $\lambda = 1$. This block 3×3 preconditioning strategy therefore provides an effective framework for tackling the ill-conditioning inherent in J. To some extent, in the engineering community there is hesitation to adopt a three-field formulation in many problems due to the difficulty in developing a good solver methodology. We believe that the method proposed here is an effective step to removing this hurdle.

At this point, we end our discussion of *linear* solver methods, and now examine how the proposed iterative strategies can be embedded in a *nonlinear* solver framework—since ultimately we are interested in nonlinear behavior.

4.2   Nonlinear Solvers

Using the block-preconditioning approach we have seen that we can efficiently solve the linear systems required to determine the Newton update Δx_k. These systems are of the form

$$J(x_k)\Delta x_k = -R(x_k) \qquad (4.36)$$

A key feature of using iterative solvers is that these linear systems are only solved to a user specified tolerance. That is, the linear solver finds Δx_k such that

$$\|R(x_k) + J(x_k)\Delta x_k\| \leq \eta \|R(x_k)\| \qquad (4.37)$$

where η is the desired tolerance.

An important point to recognize is that the update step Δx_k is determined as the solution of a linear approximation to the true, nonlinear residual function R. In a neighborhood of the nonlinear solution $\overset{*}{x}$ this linear approximation is a good one and the method shows quadratic convergence. Far from the solution, however, the linear approximation may be quite poor. As a result, solving the linear system (4.36) to a strict tolerance does not make much sense. The predicted Newton step using a rough, *inexact* solve may be just as good as an exact solve. Only close to the solution, where there is good agreement between the linear and nonlinear models, is it reasonable to expect that solving (4.36) to high accuracy will lead to a better Newton step.

Given these considerations, a logical approach is to introduce a variable tolerance η_k that changes from step to step. Close to the nonlinear solution, when $\|R(x_k)\|$ is small in some sense, then η_k is chosen to be small as well. Far from the solution, when $\|R(x_k)\|$ has not yet been reduced much, η_k is kept large, allowing for significant computational savings. If the choice of the sequence $\{\eta_k\}$ is done well, Newton's method will continue to exhibit quadratic convergence in a neighborhood of

the solution $\overset{*}{x}$. Formally, this methodology is known as an Inexact Newton Method:

Input: Initial guess x_0, nonlinear tolerance $\epsilon \in [0, 1)$
k=0;
while $\|R(x_k)\| \geq \epsilon\|R(x_0)\|$ **do**
  Choose linear tolerance $\eta_k \in [0, 1)$;
  Solve for Δx_k such that (4.37) is satisfied;
  $x_{k+1} = x_k + \Delta x_k$;
  $k \leftarrow k + 1$;
end

Algorithm 1: Inexact Newton Method

The difficult part is to develop a methodology for choosing η_k that leads to significant computational savings but still preserves the quadratic convergence behavior of an Exact Newton Method. In our code we choose the forcing sequence $\{\eta_k\}$ using a method proposed in [Eisenstat and Walker 1996]. In the numerical examples we demonstrate that an Inexact Newton Method maintains quadratic convergence but provides significant performance improvement over an Exact Newton Method.

As a final note, we have modified the standard Algorithm 1 to include a backtracking linesearch procedure for additional robustness. The resulting approach is typically more stable when the initial guess x_0 is far from the solution.

4.3 Numerical Examples

We now provide three numerical examples to test the proposed preconditioning approach. The first examines a linear elastic footing, and is used to examine the behavior of the preconditioner upon mesh refinement. The second example is similar to the first, except the computations are performed in parallel on a distributed-memory cluster. This example is used to test the parallel scaling properties of the approach. Finally, we present another serial example, but this time involving a sophisticated elastoplastic material model. This example is meant to test the behavior of the preconditioning approach with respect to more complicated constitutive behavior. It also provides an opportunity to compare exact and inexact variants of Newton's method.

4.3.1 Footing Example

We first test the performance of the block preconditioning approach in a serial (single-processor) environment. In particular, we wish to examine the behavior of the preconditioner with respect to mesh refinement. Typically, as the mesh is refined the discrete problem becomes increasingly ill-conditioned. Ideally, however, a preconditioner should exhibit mesh independent scaling—that is, the number of solver iterations to convergence for the preconditioned system remains constant as $h \rightarrow 0$. Without this mesh independence it is difficult to construct a truly scalable algorithm. Of course, the total solution time is expected to grow with the problem size even if the number of iterations is constant, since the required matrix-vector operations must be performed on increasingly
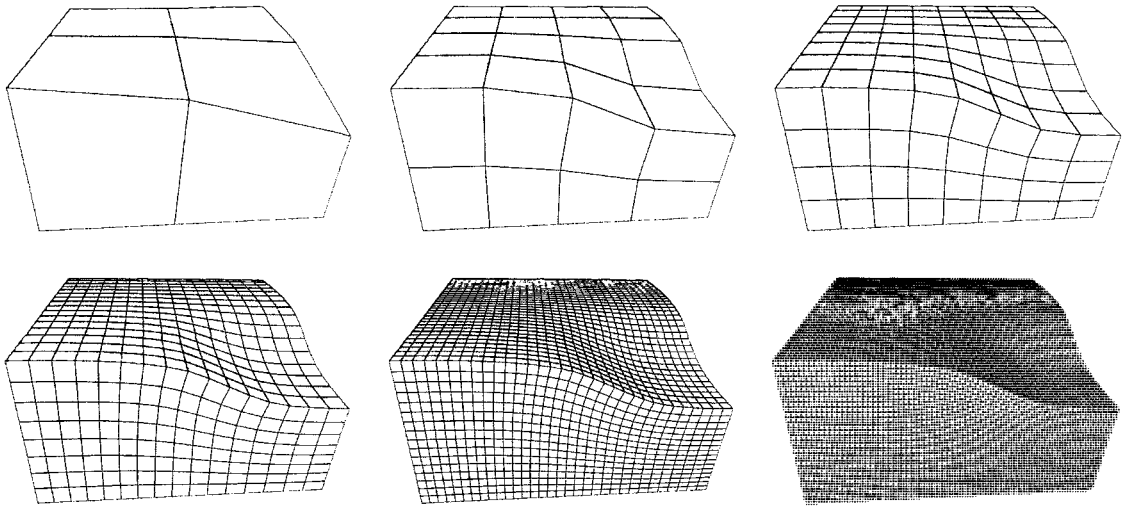
Figure 4.4: Geometry for the footing example, showing the mesh at six refinement levels.

large systems. In later sections we will examine how parallel processing can be used to reduce the wall-time cost of each iteration and work towards a constant-in-time algorithm (the ideal goal).

For this test problem, we examine an elastic, saturated soil subject to a surface traction representing a square footing load. In later examples we will examine more sophisticated elastoplastic models and unsaturated soils. We note that the proposed preconditioning methodology, however, is independent of the constitutive assumptions. In fact, the only requirements on the underlying model are that A and S are invertible—otherwise, very little has been said about the basic governing equations. This generality implies that the block preconditioning approach can be applied to a wide array of applications—e.g. Stokes and Navier-Stokes flow, incompressible elasticity, Darcy-flow, constrained optimization, etc. Only at the level of the sub-preconditioners P_A and P_S does it become useful to incorporate additional, physics-based knowledge.

Figure 4.4 shows the example mesh at six progressive refinement levels. The domain dimensions are 10m × 10m × 5m. At each refinement level, the hexahedral elements from the previous level are partitioned into 8 new elements. At the finest level the mesh has 557,700 degrees of freedom (418,275 displacement degrees of freedom and 139,425 pressure degrees of freedom). Figure 4.4 also shows the resulting deformed configuration and excess pressure distribution at the end of the simulation. Note that the displacements have been exaggerated for clarity.

The footing load is applied as a surface traction in one corner of the domain (Figure 4.5). This load has a uniform magnitude w in a 2.5m × 2.5m region, which then tapers linearly to zero in a 5m × 5m region. All boundaries are treated as no-flux, no-normal-displacement boundaries, except for the upper surface. The pressure at this surface is fixed to atmospheric ($p = 0$) and the displacements are unconstrained.

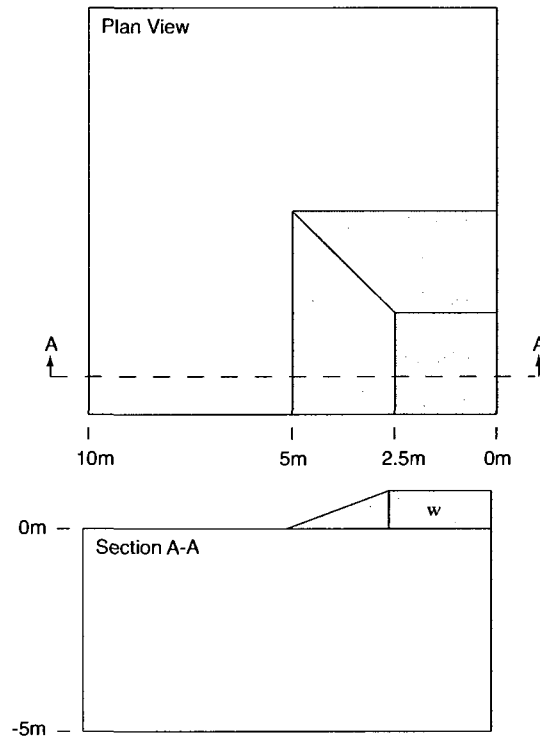At the start of the simulation, a growing surface traction is applied at a rate of 0.1 kPa/s. The

Figure 4.5: Plan and section geometry for the footing example.

remaining simulation parameters are as follows: bulk modulus $K = 5$ MPa, Poisson ratio $\nu = 0.25$, intrinsic permeability $k = 10^{-10}$ m^2, dynamic viscosity of water $\eta = 10^{-6}$ kPa·s, stabilization $\tau = 1.0$, and time step $\Delta t = 1$ s. Because we employ a saturated, linear elastic model the governing equations are linear and therefore only a single linear system must be solved in each time step. Since we have a transient problem, the simulation was run for twenty time steps and all metrics were averaged over these steps.

Table 4.1 contains a summary of the analysis. Four preconditioning strategies were analysed:

- None: No preconditioning.

- BP-LU: Block-preconditioning with direct LU factorizations for sub-preconditioners.

- BP-ILU: Block-preconditioning with ILU sub-preconditioners.

- BP-AMG: Block-preconditioning with AMG sub-preconditioners.

Table 4.1 records two metrics of solver performance: the number of GMRES iterations required for convergence in each step, and the average CPU time to perform each solve. A restarted GMRES(m) method was used with $m = 200$ saved basis vectors. While not presented here, the same study using a BICGSTAB solver produced very similar results. These results are for right-preconditioning, though

Table 4.1: Comparison of preconditioning strategies at several refinement levels for the serial u/p footing problem.

		0	1	2	3	4	5
	Refinement Level	0	1	2	3	4	5
	Elements	4	32	256	2,048	16,384	131,072
	Degrees of Freedom	72	300	1,620	10,404	74,052	557,700
GMRES Iterations	None	45	115	538	*	*	*
	UP-LU	2	2	2	2	2	*
	UP-ILU	5	9	16	30	61	*
	UP-AMG	2	9	7	12	13	15
CPU Time (seconds)	None	0.017	0.051	1.74	*	*	*
	UP-LU	0.004	0.006	0.086	11.1	1350.0	*
	UP-ILU	0.006	0.010	0.063	0.764	11.6	*
	UP-AMG	0.004	0.011	0.039	0.414	3.6	37.0

* Solver failed.

left-preconditioning is again similar. The criterion for convergence was a relative residual reduction tolerance of

$$\frac{\|r\|_2}{\|r_0\|_2} \leq 10^{-8} \tag{4.38}$$

The solve was deemed to have failed if it took more than 1000 iterations or if memory-resources were insufficient. All simulations were run on individual, identical nodes of a distributed-memory cluster. Each node had 2GB of memory.

The results for the unpreconditioned systems clearly illustrate the inherent ill-conditioning of the coupled system, and that this ill-conditioning worsens with mesh refinement. Beyond about 10,000 unknowns the number of iterations grows to be well more than a 1000. Even for relatively few unknowns, the unpreconditioned solver cannot compete with any of the other three approaches.

In terms of iteration counts, BP-LU is the clear winner, as expected. For this case, the only difference between the current preconditioner and an "exact" preconditioner is the introduction of a sparse approximation for S. This sparse approximation appears to be a good one, as the solver converges in two iterations. This is the same performance as would be expected from the "exact" preconditioner. This approach also clearly demonstrates mesh-independent scaling with respect to the iteration count. Of course, iterations counts only tell a piece of the story. This approach quickly becomes memory-limited due to fill-in during the direct factorizations. As a result, the solver fails on the most refined mesh. Also, because the construction of the preconditioner is very expensive, this approach cannot compete with the remaining two in terms of total CPU time.
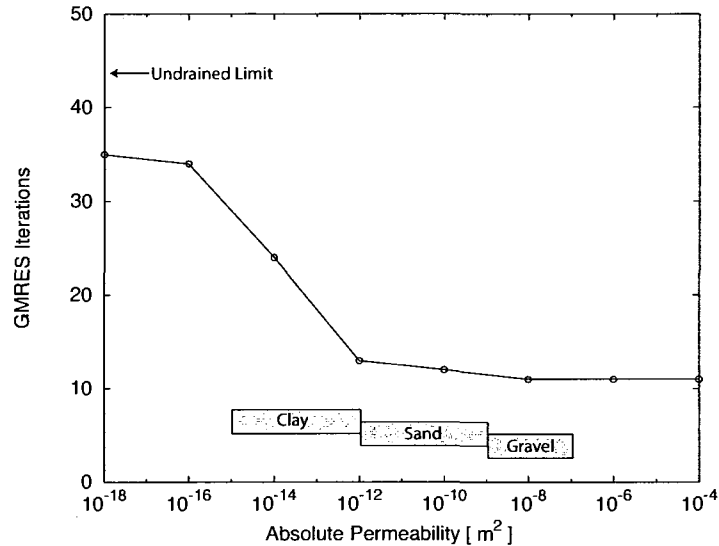
Figure 4.6: Sensitivity of the block-preconditioner to the soil's permeability. Gray bars indicated typical permeability ranges for some common geomaterials.

The remaining two options, BP-ILU and BP-AMG both appear to be the most practical approaches. BP-ILU shows good performance on relatively small problems. Unfortunately, the iteration count grows quickly with mesh refinement and the solver ultimately fails on the most refined mesh. BP-AMG, on the other hand, shows much better refinement behavior—at least close to (if perhaps not exactly) h-independent scaling.

We have found that these performance trends have held across a wide array of problems and configurations, with BP-AMG often outperforming BP-ILU. There are certain situations, however, where the performance of AMG may degrade severely. For example, on distorted or anisotropically refined meshes the AMG preconditioner (which has no knowledge of the mesh topology) may choose a poor coarsening strategy. Our experience is that BP-ILU is often more robust in difficult situations and is a good fall-back approach—even if it cannot compete with BP-AMG in favorable situations. It should also be noted that significant research is currently devoted to improving the robustness of AMG preconditioners. See, for example, [Gee et al. 2009].

Before concluding this example, we make a final observation about the sensitivity of the preconditioner to the permeability of the medium. As the permeability of the medium and the size of the time increment changes, the resulting linear systems will have very different character. A natural question is whether the proposed preconditioning approach is robust across the entire range of behaviors that may be encountered. To examine this point, Figure 4.6 shows the sensitivity of the iteration count as the permeability of the medium changes (the time increment remains fixed). As we can see, as the undrained limit is approached the number of iterations grows—suggesting that the conditioning of the system worsens as the permeability drops. Nevertheless, we see that the number of required iterations ($\sim$35) is bounded and that the preconditioner continues to perform
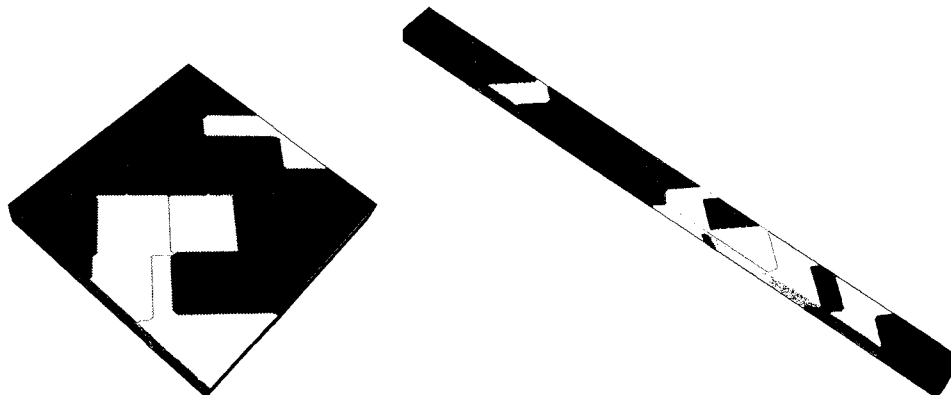
Figure 4.7: Geometric configurations for the weak scaling test, partitioned for 16 processors: (left) configuration one, which is 4m × 4m × 0.5m; (right) configuration two, which is 16m × 1m × 0.5m. Each domain has 1,131,588 degrees of freedom.

well. This robustness with respect to changing permeability regimes is reassuring.

4.3.2 Parallel Scaling

We now test the behavior of the preconditioner in a parallel setting. For this example, we perform a weak scaling test: the number of elements in the domain grows with the number of processors such that the number of elements per processor remains approximately constant. We have also examined two geometric configurations for the test (Figure 4.7). Given the simple geometry of the two domains, it is easy to define a partitioning that leads to an exact load balancing of the number of elements per processor. Since we are more interested in the performance with respect to unstructured computations, however, we have used the METIS partitioning package to allocate the elements to each processor. The resulting partitioning is much more irregular.

We test the performance of the preconditioned solver on 1, 4, 9, and 16 processors. The single processor version contains 74,052 degrees of freedom, while the 16 processor version has 1,131,588 degrees of freedom. Instead of the footing load, we have defined a spatially periodic footing traction that grows with time as

$$w = w_o t \quad \text{with} \quad w_o = 0.1 \sin(\pi x) \sin(\pi y) \tag{4.39}$$

Units are kPa. The remaining simulation parameters are the same as for the previous serial test. In this test we used BICGSTAB as the solver, as we found runtimes to be slightly faster than with GMRES.

The results of the weak scaling tests are presented in Table 4.2. We record both the number of iterations to convergence, as well as the total wall time. We should note that these simulations were run on a heavily-trafficked cluster, and the total wall time for two identical runs could vary quite a bit—often on the order of 20 or 30% and occasionally as high as 100%. We have attempted

Table 4.2: Parallel scaling performance of the BP-AMG preconditioner.

		1	4	9	16
	Processors	1	4	9	16
	Elements	16,384	65,536	147,459	262,144
	Degrees of Freedom	74,052	287,300	639,812	1,131,588
Configuration One	Iterations	9	9	9	9
	Wall Time (s)	4.41	8.30	8.81	10.69
Configuration Two	Iterations	9	9	10	10
	Wall Time (s)	4.41	8.38	10.66	11.24

to minimize this effect by averaging over several runs. We believe that the resulting run times are representative—they are neither the best nor the worst we encountered. This sensitivity, however, is a clear indication that the solution process is communication intensive and that we suffer large communication-overhead penalties.

Examining the results of Table 4.2, the scaling properties of the preconditioner with respect to the number of iterations is excellent. The number of iterations remains essentially constant regardless of configuration or number of processors. The story with respect to the actual wall time is not so good however. The wall time efficiencies for the two 16-processor runs (defined as T_1/T_{16}) are 41% and 39%, respectively. Our tests were limited to 16 processors, so we could not investigate the scaling behavior beyond this point. Hopefully with additional work we can improve this behavior. There is, at least, a tapering of the degradation as one goes to larger numbers of processors. That is, the efficiency penalty of going from 1 processor to 4 is worse than going from 4 to 16. Also, the ability to solve coupled problems with 1.1 million degrees of freedom in $\sim$11 seconds per iteration may be entirely sufficient for many applications.

4.3.3 Plane-Strain Compression Example

We now explore the performance of the preconditioning technique with respect to more sophisticated material models. This example considers the plane-strain compression of a 50mm × 100mm saturated sand specimen. The geometry and boundary conditions are indicated in Figure 4.8. Since this is a simple two-dimensional test, the simulation was run on a single processor.

A critical state plasticity model suitable for dense sands is used for the constitutive behavior of the specimen. The formulation of this model is the subject of [Borja and Andrade 2006] and so we skip over a detailed presentation here. A key feature of the model, however, is that the strength of the material is dependent on the density of the sand with respect to its critical state density. In this example, we have introduced a high void ratio zone in the upper left corner. This zone is inherently
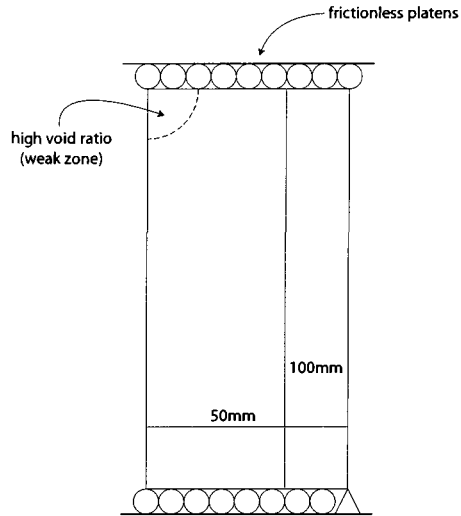
Figure 4.8: Geometry and boundary conditions for plane-strain compression of a sand specimen.
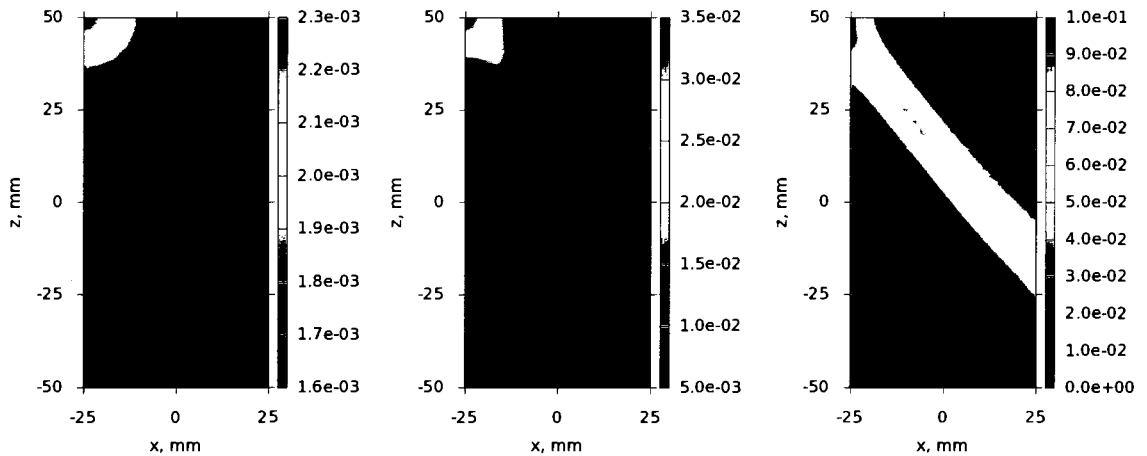

Figure 4.9: Snapshots of cumulative plastic strain—at 25, 50, and 75 s—for the plane-strain compression example.
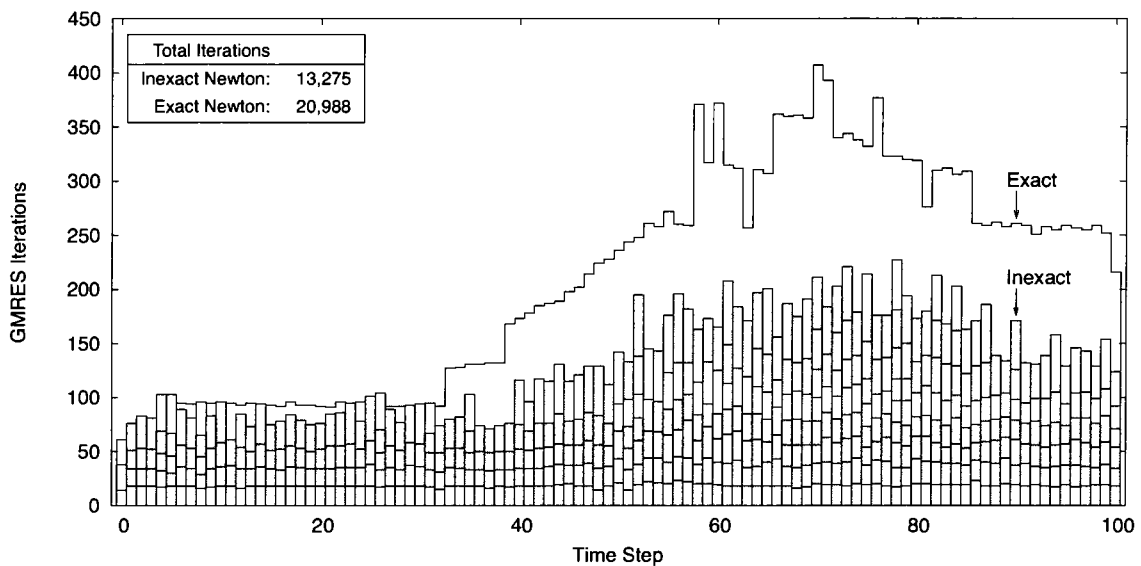
Figure 4.10: Comparison of iteration count using exact and inexact versions of Newton's method on the plane-strain compression problem.

weaker than the rest of the specimen, and upon loading undergoes significant plastic volumetric compaction (Figure 4.9). As the sample is compressed further, however, this weak spot triggers shear localization along a band within the specimen. The resulting deformations and fluid pressures are highly nonlinear and provide a rigorous test of both the linear preconditioning technique and the nonlinear outer Newton iteration.

We recall that at each time step, a certain number of Newton iterations (say, m) must be performed to drive the nonlinear residual to zero. To determine the Newton update within each iteration, however, a certain number of iterations of a GMRES solver (say, k) are required. Figure 4.10 provides a detailed account of the nonlinear solution procedure using an inexact Newton method. At each time step, the vertical bar is divided into m sub-increments, with the height of each sub-bar equal to the solver iterations k required to determine the Newton update. The total height of the bar therefore represents the total number of linear GMRES iterations required for each nonlinear solve. The key idea behind an inexact Newton method is that using exact solves in early iterations will lead to an increase in k but no decrease in m. Far from the solution the linearized approximation is a poor representation of the true residual function behavior, and so an exact solve may be no better than an inexact solve. Therefore, it only makes sense to increase the linear solver tolerance as one approaches the nonlinear solution, where the linear approximation is increasingly meaningful. Significant savings can then be achieved by using inexact solves in the early iterations far away from the nonlinear solution. As a baseline for comparison, Figure 4.10 indicates the total number of GMRES iterations required in each time step using an exact (constant-tolerance) Newton's method. The exact method requires 20,988 total GMRES iterations, while the inexact variant shows the same quadratic convergence behavior but only requires 13,275 total iterations—a computational savings of

37%. The savings are most significant at later times with the onset of the shear localization. At this point the problem becomes more nonlinear and more Newton updates are required before the scheme converges to a solution. Examining the figure in more detail, we see that the lower bars (representing the first few Newton updates) are indeed shorter than the upper bars (representing Newton updates close to the nonlinear solution). Furthermore, we see that even for the nonlinear material model the preconditioner continues to perform well, keeping the number of GMRES iterations k within each Newton iteration relatively small—typically between 10 and 50.

4.4 Conclusion

This chapter has focused on two key ideas: (1) developing efficient preconditioning techniques for block-structured linear systems, and (2) using inexact Newton methods to improve the performance of the nonlinear solver.

We have proposed a block-structured preconditioning approach which works for both two-field (u/p) and three-field $(u/p/w)$ formulations. The approach avoids the inner-solver bottleneck associated with the more classical Schur-reduction method. Furthermore, we have seen that through block-preconditioning we can leverage existing black-box preconditioners to their best advantage, while still endowing the global preconditioning strategy with knowledge of the block-structured nature of our problem. The resulting preconditioner leads to impressive results. On a single-processor, a three-dimensional, fully-coupled problem with more than 0.5 million degrees of freedom averaged 37 seconds per solve. In parallel, a problem with 1.1 million degrees of freedom took 11 seconds per solve.

We have also seen that iterative approaches provide additional savings in the context of the nonlinear iteration, because we are able to implement an inexact Newton methodology. For the nonlinear test problem, the inexact Newton method took 37% less work than its exact counterpart.

As far as future work is concerned, our first priority is to address the parallel efficiency of our solver—as we believe better scaling can be achieved with additional optimization. Long term, we are also interested in *nonlinear* preconditioning strategies. Just as a linear preconditioner seeks to reduce the number of solver iterations in a Krylov method, a nonlinear preconditioner seeks to reduce the number of Newton steps needed to drive the nonlinear residual to zero. Given the stiff nature of the coupled problems under consideration, any improvement in the nonlinear convergence behavior would be of tremendous computational advantage.

5 Hydrologically-Driven Slope Instability

In this chapter, we explore a specific application of the coupled formulation: modeling rainfall-driven slope failure.

5.1 Introduction

Hydrologically-driven slope instability threatens lives and property worldwide. Examples of large-scale slope failure all over the world are numerous. Landslides in the San Francisco Bay Area during the January 1982 storm resulted in 24 fatalities and millions of dollars in property damage [Brown et al. 1984; Smith and Hart 1982]. In January 4, 1982, a 0.22 m storm superimposed on approximately 0.6 m of pre-storm seasonal rainfall triggered thousands of landslides in the central Coast Ranges of California. In Mameyes, Puerto Rico, 1985, rainfall intensity as high as 0.56 m within a 24-hour period, with rates as high as 70 mm/hr, triggered debris flows and resulted in 129 deaths [Jibson 1992]. In Rio Limón, Venezuela, 1987, rainfall as much as 0.174 m in less than five hours triggered numerous shallow landslides and debris flows resulting in 210 deaths [Schuster et al. 2002]. In Antofagasta, Chile, 1991, rainfall rates as great as 60 mm/hr during a three-hour period triggered landslides that resulted in 101 deaths [Van Sint Jan and Talloni 1993]. In Vargas, Venezuela, 1999, a heavy rainfall exceeding 0.9 m over a three-day period, with daily values greater than the 1,000 year return period [Martinez 2000], triggered thousands of landslides that, along with severe flooding, resulted in an estimated 30,000 deaths [USAID 2000]. And in 2006, a heavy rainfall in Guinsaugon, Philippines, triggered massive landslides burying an elementary school that had 246 students and 7 teachers [Lagmay et al. 2006].

Despite decades of extensive slope stability model development, the fundamental controls connecting the hydrologic and geotechnical processes triggering slope failure are still not well quantified. This is evident from the La Conchita landslide of January 11, 2005 in southern California that occurred without warning. This lack of understanding is a direct result of the simplified physics in current models, with the omission of the effect of partial saturation from slope stability calculations. It is known that increasing the degree of saturation decreases the capillary pressure, which in turn weakens the slope. Moreover, fluid flow generates frictional drag on the soil matrix, thus increasing the load that the soil is expected to carry. Despite the expected significant impact, such interplay between the increase in saturation accompanied by fluid flow, and increase in potential for slope failure is yet to be fully quantified.

In this paper, we develop and test a 3D physics-based slope deformation model that couples solid deformation with fluid flow in variably-saturated soils. We also assess the potential of the coupled model to realistically quantify stresses and deformation necessary to trigger slope failure. We emphasize the continuum nature of the modeling approach in that it does not quantify the factor of safety of the slope *per se*, unlike methods based on limit equilibrium concepts. Instead, our model predicts the spatial and temporal variations of internal stresses, pore water pressure, degree of saturation, and deformation within the slope, as functions of the spatially and temporally varying rainfall intensity. Thus, instead of having one scalar quantity called factor of safety, we assess the failure potential of a given slope based on the stresses and deformation arising from a prescribed forcing function (i.e., rainfall intensity). Stability criteria based on deformation bands are used in this paper to assess the potential for slope failure.

To assess whether the proposed model can successfully integrate important variables in a physics-based characterization of the field conditions and still obtain a realistic description of slope failure, we test the approach with comprehensive and exhaustive data from the Coos Bay experimental catchment (CB1) [Ebel et al. 2007a,b]. The highly instrumented site was originally chosen as a hillslope-scale "laboratory" for conducting sprinkling experiments aimed at developing and testing hydrologic response models. Experimental and field data generated from the site are plentiful, albeit most of them pertain to hydrologic response data [Anderson et al. 2002, 1997a,b; Montgomery and Dietrich 2002; Montgomery et al. 2002, 1997; Torres et al. 1998]. However, the highly instrumented CB1 slope failed as a large debris flow in November 1996, raising some interesting questions related to the geotechnical aspects of the site. Given what we know about the topography, hydrologic constraints, and the geotechnical outcome at CB1, this case study seems ideal to test the proposed continuum slope model. However, it must be noted that despite much effort to constrain the site with extensive field instrumentation and investigation, much uncertainty remains, particularly with respect to the geotechnical and hydrologic boundary conditions appropriate for the site. Therefore, to underscore what we still do not know about CB1, we emphasize that the analysis reported in this paper pertains only to a slope similar to CB1, and not to the slope at CB1 itself.

The slope at CB1 is steep, on the order of 43°. Any critical state soil mechanics model would predict an initial stress condition within the slope that lies on the "dilatant side" of the critical state line because of the high shear stress to effective-mean-normal stress ratio generated by the steep slope condition [Moriguchi et al. 2009]. It would thus be reasonable to characterize the constitutive response of the soil skeleton with a Mohr-Coulomb plasticity or similar models, since the compression cap is unlikely to be activated particularly when the effective mean normal stress decreases further as a result of the loss of suction with increased saturation. A simpler plasticity model such as the Mohr-Coulomb model also requires fewer material parameters than any of the more sophisticated constitutive models available in the literature (see [Young et al. 2009]). For the soil at CB1, which is mostly colluvium, laboratory-determined values of friction angle are available [Schmidt 1994, 1999]. Furthermore, representative values of lateral root cohesion in clearcut forests in the Oregon Coast Range have been estimated in [Schmidt et al. 2001]. Finally, the thickness of the colluvium at CB1 is known to be much smaller than the plan area of the mobilized slide, thus making a continuum

analysis approach appropriate for this problem.

5.2 Continuum slope model

For a solid-water-air mixture the constitutive stress tensor $\boldsymbol{\sigma}'$ (positive for tension) may be defined as

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + \Big(1 - \frac{K}{K_s}\Big)\bar{p}\mathbf{1}\,, \qquad \bar{p} = \psi^w p_w + (1 - \psi^w)p_a\,, \tag{5.1}$$

where $\boldsymbol{\sigma}$ is the total Cauchy stress tensor, ψ^w is the degree of saturation, K and K_s are the elastic bulk moduli for the solid matrix and solid constituent, respectively, p_w and p_a are the intrinsic pore water and pore air pressures, respectively, and $\mathbf{1}$ is the Kronecker delta tensor. Borja [Borja 2006c; Borja and Koliji 2009] demonstrated this expression for the constitutive stress tensor to be energy-conjugate to the solid rate of deformation. Under special cases, $\boldsymbol{\sigma}'$ reduces to Bishop's stress when $K/K_s = 0$ (incompressible solid grains) and $\psi^w = \chi$; to the Skempton [Skempton 1961] and Nur-Byerlee [Nur and Byerlee 1971] stress when $\psi^w = 1$; and to the Terzaghi [Terzaghi 1943] stress when $K/K_s = 0$ and $\psi^w = 1$. The Bishop parameter χ can be determined experimentally, and the substitution of degree of saturation ψ^w in lieu of this parameter is a simplified approximation derived from the volume averaging over a representative elementary volume (REV). For purposes of constitutive modeling of the soil layer, we will use the Bishop stress ($K/K_s = 0$), with $\chi = \psi^w$.

Ignoring inertia forces and setting p_a equal to the atmospheric pressure (i.e., zero) for near-surface condition, balance of linear momentum in the subsurface is given by the equation

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} + \bar{\rho}\boldsymbol{g} = \mathbf{0}\,, \qquad \bar{\rho} = \phi^s\rho_s + \phi^w\rho_w\,, \tag{5.2}$$

where ϕ^s and ϕ^w are the solid and water volume fractions in the REV, respectively; ρ_s and ρ_w are the intrinsic solid and water mass densities (i.e., mass of constituent α per unit volume of constituent α), respectively; $\bar{\rho}$ is the total mass density of the mixture (ignoring the mass of air phase); and $\boldsymbol{g}$ is the gravity acceleration vector. Assuming that water is incompressible, balance of water mass in the subsurface takes the form [Borja 2004, 2006c]

$$\frac{\mathrm{d}\phi^w}{\mathrm{d}t} + \phi^w\boldsymbol{\nabla} \cdot \boldsymbol{v} + \boldsymbol{\nabla} \cdot \boldsymbol{q}^w = \pm\theta^b \pm \theta^e\,, \tag{5.3}$$

where $\boldsymbol{v}$ is the velocity of the solid matrix, $\boldsymbol{q}^w = \phi^w(\boldsymbol{v}^w - \boldsymbol{v})$ is the relative discharge velocity, θ^b is the specified rate source/sink, θ^e is the rate of water exchange with the surface continuum, and $\mathrm{d}(\cdot)/\mathrm{d}t$ denotes a material time derivative following the solid motion. The relative discharge velocity is given by the constitutive equation

$$\boldsymbol{q}^w = k_{rw}\boldsymbol{K} \cdot \boldsymbol{\nabla}\Big(\frac{p_w}{\rho_w\mathrm{g}} + z\Big)\,, \tag{5.4}$$

where $\boldsymbol{K}$ is the hydraulic conductivity of the porous medium at complete saturation, k_{rw} is the

relative permeability that varies with degree of saturation ψ^w, g is the gravity acceleration constant, and z is the vertical coordinate.

The two independent variables of the formulation are the solid velocity v and pore water pressure p_w satisfying the two balance laws (5.2) and (5.3), along with the appropriate boundary and initial conditions. The solution is fully coupled in the sense that the independent variables are determined simultaneously. However, the governing equations must be closed by introducing a pair of constitutive laws determined from continuum principles of thermodynamics [Borja 2004, 2006c; Borja and Koliji 2009; Houlsby 1979, 1997; Jha and Juanes 2007; Khalili et al. 2004; Nuth and Laloui 2008]: (a) a constitutive law relating the Cauchy effective stress rate tensor $\dot{\sigma}'$ with the strain rate tensor $\nabla^s v$ for the solid matrix; and (b) a constitutive law relating the suction stress $s = p_a - p_w$ with degree of saturation ψ^w. The second constitutive law is determined from the imbibition portion of the soil-water retention curve for a given soil [Birle et al. 2008; van Genuchten 1980].

Continuum principles of thermodynamics suggests an elasto-plastic constitutive framework for the soil skeleton given by a yield function of the form [Borja 2004, 2006c; Borja and Koliji 2009; Houlsby 1979, 1997]

$$F(\sigma', s, \bar{p}_c) = 0, \tag{5.5}$$

where $\bar{p}_c$ is a stress-like plastic internal variable that determines the size of the compression cap. The suction stress has a positive value for negative pore water pressure (e.g., above the water table), and approaches zero as the soil approaches perfect saturation. However, as noted earlier, the compression cap is not likely to be activated for steep hillside slope conditions, and thus, this explicit dependence of the yield surface on the suction stress may be ignored when using the Mohr-Coulomb or similar yield criteria. We emphasize, however, that the yield function still depends on the suction stress through the Cauchy effective stress tensor σ'.

Mixed finite element (FE) equations may be readily developed from the variational forms of (5.2) and (5.3). The independent variables in the present case are the nodal solid displacement vector d and nodal pore water pressure vector p. The coupled FE equations take the form

$$\boldsymbol{F}_{INT}(\boldsymbol{d}) + \boldsymbol{G}\boldsymbol{p} = \boldsymbol{F}_{EXT}, \tag{5.6}$$

$$\boldsymbol{G}^{\mathsf{T}}\dot{\boldsymbol{d}} + \boldsymbol{M}\dot{\boldsymbol{p}} + \boldsymbol{\Phi}\boldsymbol{p} = \boldsymbol{G}_{EXT}, \tag{5.7}$$

where $\boldsymbol{F}_{INT}(\boldsymbol{d})$ is the internal nodal force vector arising from the effective stress σ', $\boldsymbol{G}$ and $\boldsymbol{G}^{\mathsf{T}}$ are the discrete gradient and discrete divergence operators, respectively, which depend on p through the degree of saturation in the unsaturated regime, $\boldsymbol{M}$ is a coupling matrix that vanishes at full saturation, $\boldsymbol{\Phi}$ is an effective conductivity matrix that also depends on the degree of saturation, and $\boldsymbol{F}_{EXT}$ and $\boldsymbol{G}_{EXT}$ are prescribed vectors of momentum and fluid supplies. Note that the degree of coupling implied by the above equations is more intricate than in the fully saturated case because of the presence of ψ^w in many terms.

The coupled formulation presented above requires mixed finite elements with nodal pressure and displacement degrees of freedom. Low-order mixed finite elements, with equal order interpolation
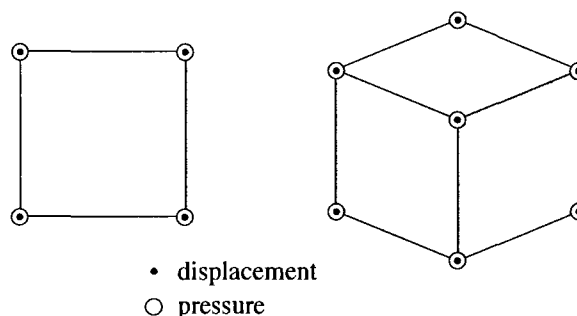
• displacement
○ pressure

Figure 5.1: Stabilized low-order mixed finite elements for coupled solid-deformation/fluid-diffusion: four-node quadrilateral for 2D (left), and eight-node hexahedral for 3D (right). After White and Borja [White and Borja 2008].

for displacement and pressure, would be ideal—particularly in 3D when the number of equations can increase dramatically. Unfortunately, low-order mixed finite elements create a numerical difficulty in that in the limit of full saturation and either low permeability or fast loading rates, the pore fluid can impose near or exact incompressibility on the deformation of the solid matrix and create spurious pore pressure oscillation. To address this difficulty, we use a stabilization scheme proposed by White and Borja [White and Borja 2008] based on the concept of Polynomial Pressure Projections (PPP), which was proposed in [Bochev and Dohrmann 2006; Bochev et al. 2006; Dohrmann and Bochev 2004] to stabilize the Stokes problem. The primary motivation for using stabilization is computational efficiency: with this technique, it would now be possible to use low-order finite elements, such as those shown in Fig. 5.1, without the undesirable pore pressure oscillation that otherwise afflicts the solution without stabilization.

5.3   Hydrologic and geotechnical conditions at CB1

The CB1 experimental catchment [Anderson et al. 2002, 1997a,b; Montgomery and Dietrich 2002; Montgomery et al. 2002, 1997; Torres et al. 1998], clearcut in 1987, is located along Mettman Ridge approximately 15 km north of Coos Bay in the Oregon Coast Range (Fig. 5.2). CB1 is a 51 m long (860 m^2) unchanneled valley, with a north-facing aspect, that has an average slope of 43° (see Fig. 5.3). Three sprinkling experiments were conducted at CB1: experiment #1 conducted in May 1990 at 1.5 mm/hr for 6 days; experiment #2 conducted in May 1990 at 3.0 mm/hr for 4 days; and experiment #3 conducted in May 1992 at 1.7 mm/hr for 7 days. The instrumentation at CB1 used to characterize the spatial and temporal variability in near-surface hydrologic response for the three experiments included an exhaustive grid of rain gauges, piezometers, tensiometers, TDR wave guide pairs (for estimating soil-water content), lysimeters, meteorological sensors (on a tower), atmometers, and weirs. Continuous measurements from rainfall, discharge, and total head (from selected piezometers) are available from 1990 through 1996. In November 1996 the CB1 slope
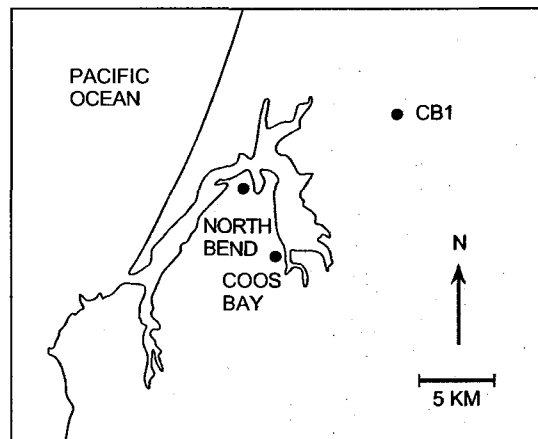
Figure 5.2: Location map for CB1 experimental catchment near Coos Bay in the Oregon Coast Range, USA. Reproduced from Ebel et al. [Ebel et al. 2007a].

failed as a large debris flow. The instrumentation at CB1 provides one of the most comprehensive hydrological response data sets in existence for a steep, deforested catchment that has experienced slope failure.

The sediment at CB1 is colluvium, a surficial sediment derived from weathered or fresh bedrock, and the soil has no input from aeolian transport. The parent rock for the colluvium is an Eocene turbidite sandstone from the Tyee and Flournoy formations [Schmidt et al. 2001]. The soils are well mixed, nonplastic (plasticity index of 0), gravelly sands [Schmidt 1999]. The geometry and thickness of the colluvium are well defined from soil borings. Saturated hydraulic conductivity were determined from slug tests, soil-water content and porosity from TDR measurements, and hysteretic capillary pressure relationships were established from experiments. Discharge chemistry data suggest that runoff generation occurs primarily from water stored in small, poorly connected pores and fractures in the bedrock and saprolite connecting with larger macropores during storms [Montgomery 1991]. Tracer data (bromide and isotopically-tagged water) suggest that the two most important flow paths at CB1 are rapid saturated flow through the shallow, fractured bedrock and vertical percolation in the vadose zone [Anderson 1995]. Piezometer records show that subsurface storm flow in the shallow, fractured-rock zone exerts the most significant control on pore pressure development in the CB1 colluvium [Montgomery 1991]. Tensiometer data indicate that the flux of water through the unsaturated zone provides an additional control on pore pressure development and cannot be ignored in slope stability assessment models [Torres 1997].

Low confining stress triaxial shear tests demonstrate that the colluvium at CB1 is cohesionless (consisting of a sandy matrix with a friction angle of 33°) [Schmidt 1994, 1999]. Lateral root cohesion in clearcut forests in the Oregon Coast Range was estimated to be uniformly less than 10 kPa [Schmidt et al. 2001], and may have explained how such a steep slope at CB1 could be sustained by the sediment (apart from the capillary tension that develops in the vadose zone). Low
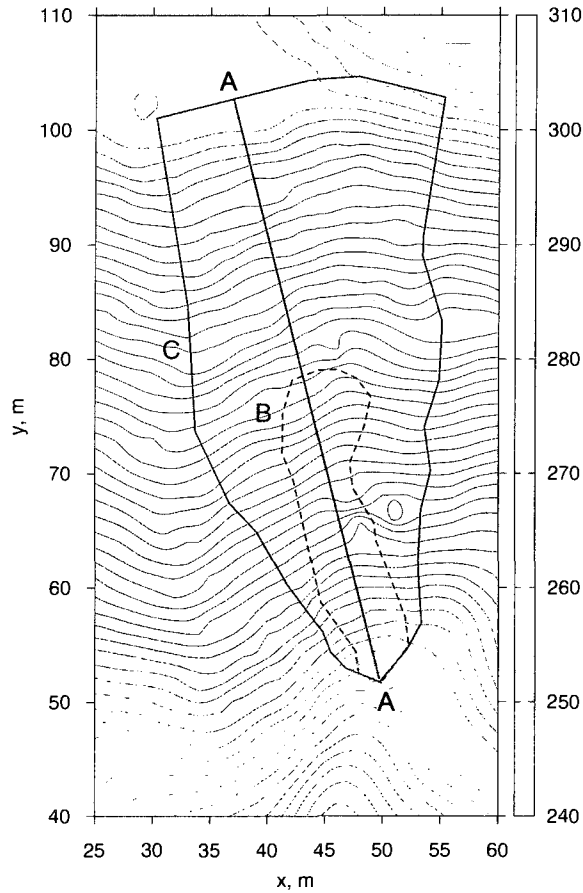
Figure 5.3: Topographic map for CB1 experimental catchment. Two-dimensional plane strain condition is assumed along section A-A for the coupled solid deformation-fluid flow analysis. Dashed curve B delineates the extent of debris flow zone at CB1 from event of November 1996. Larger region C defines the boundary of the Ebel et al. [Ebel et al. 2007b] 3D hydrologic FE model. Color bar is elevation in meters.

confining stress triaxial tests gave an internal friction angle of 40° and zero cohesion [Schmidt et al. 2001]. This is in agreement with other measurements near the site that gave internal friction angles ranging from 35-44° [Burroughs et al. 1985; Schroeder and Alto 1983; Wu et al. 1988; Yee and Harr 1977] and zero cohesion [Yee and Harr 1977]. The stress-strain behavior, observed in low confining stress triaxial strength testing, was approximately linear for low applied stresses with a distinct transition to a non-linear material behavior at axial strains greater than 1-2% [Schmidt 1999]. The saturated density of the soil is about 1600 kg/cu.m. [Schmidt et al. 2001] and the bulk density at soil water contents of 20-30% is around 1200 kg/cu.m. [Schmidt 1999]. The Van Genuchten [Houston et al. 1999; van Genuchten 1980] parameters for the CB1 soil are discussed in the next section.

5.4 Hydro-mechanical model

A challenging aspect of the numerical simulation of a hillslope-scale problem is determining the extent of the spatial and temporal discretizations, and what initial and boundary conditions to impose on the problem. Ebel et al. [2007b] conducted full 3D fluid flow simulations based on Richards equation of variably-saturated flow in the subsurface [Richards 1931] using a comprehensive FE code InHM (Integrated Hydrology Model) [VanderKwaak 1999]. They simulated flow over a portion of the sediment and fractured bedrock delineated by vertical boundaries on which they believed they could reasonably impose a no-flow boundary condition. The spatial extent of their domain description is shown in Fig. 5.3.

For our purposes, the deformation in the bedrock would be too small to be of concern in the coupled solid-deformation/fluid-flow analysis, and thus we only represent the sediment domain in our simulations. However, there are two important factors that make the present simulations significantly more demanding than the fluid flow simulations conducted by Ebel et al. [2007b]: (a) each node in the FE mesh is now composed of 3 degrees of freedom (DOFs) in 2D (two displacement and one pressure), and 4 DOFs in 3D (three displacement and one pressure); thus the number of equations increases very quickly; and (b) the water-retention curve typically exhibits a steep slope near the wetting front, necessitating a high-resolution mesh everywhere in the unsaturated zone (since the wetting front propagates through the unsaturated zone).

Considering the difficulty with modeling the exact CB1 conditions due to the uncertainties in defining what constitutes a 'sufficient' 3D representation of CB1 and what boundary conditions to impose on this model, we believe that it would be more enlightening to consider a much simpler 2D plane strain representation for now, where the boundary conditions can be constrained more easily. To this end, we selected a cross-section passing through the steepest portion of the CB1 catchment, shown as section A-A in Fig. 5.3, and constructed the FE mesh shown in Fig. 5.4. This 2D representation of the slope at CB1 is conservative in the sense that out-of-plane strengths from 3D effect emanating from lateral root cohesion, friction, etc. has been ignored. The mesh shown in Fig. 5.4 has 20,000 nodes and 18,981 (stabilized) quadrilateral mixed finite elements, resulting in a total of 60,000 DOFs.

Figures 5.5 and 5.6 show the suction/saturation and saturation/relative permeability relationships used in the simulations. The degree of saturation ψ^w is determined from the van Genuchten [1980] model as

$$\psi^w(s) = \psi_1 + (\psi_2 - \psi_1)\left[1 + \left(\frac{s}{s_a}\right)^n\right]^{-m}.$$
(5.8)

The model contains four parameters: ψ_1 is the residual water saturation, ψ_2 is the maximum water saturation, s_a is a scaling pressure, and n and m are empirical constants defining the shape of the saturation curve. The constants n and m are not independent, but are rather related to one another
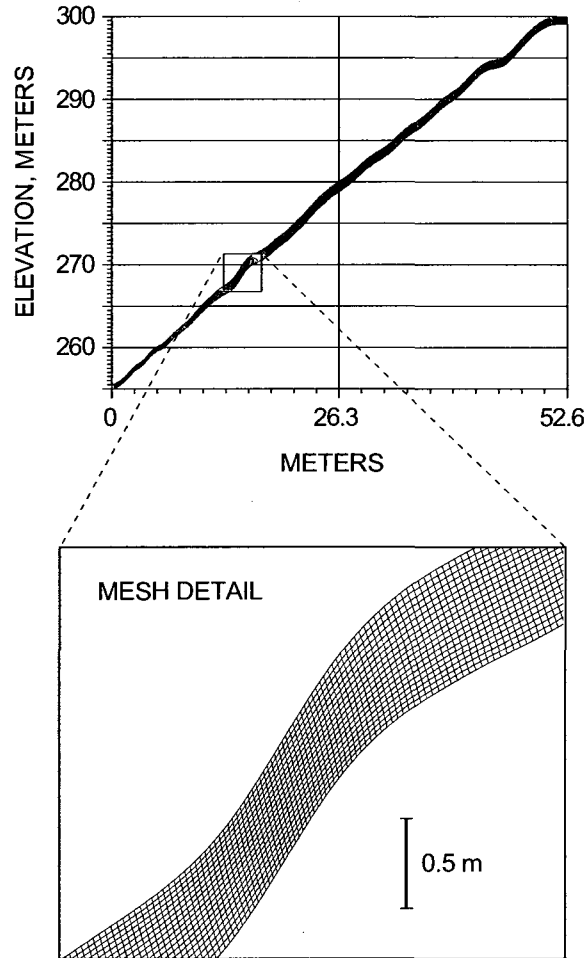
Figure 5.4: Finite element mesh for problem simulations (cf. section A-A in Fig. 5.3).

as $m = (n - 1)/n$. The water phase relative permeability is similarly defined as

$$k_{rw}(\theta) = \theta^{1/2} \left[1 - \left(1 - \theta^{1/m}\right)^m\right]^2, \qquad \theta = \frac{\psi^w - \psi_1}{\psi_2 - \psi_1} \tag{5.9}$$

The parameter values used in this work are given in Table 5.1. In order to calibrate the model, we have used the *in situ* retention curves as measured by Torres et al. [Torres et al. 1998] for the CB1 site. We ignored hysteretic effects, and have only used the wetting measurements for calibration. The chosen values are nearly the same as those used by Ebel et al. [2007b] in their hydrologic model except for a slight adjustment in the scaling pressure. The steep geometry of the saturation and relative permeability relationships proved to be numerically challenging on the relatively coarse mesh, and a slight flattening of the curve improved the robustness of the model.
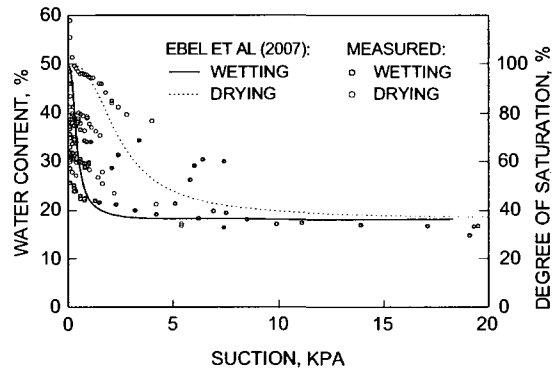
Figure 5.5: Soil-water retention curve data [Torres et al. 1998] and van Genuchten curves used by Ebel et al. [2007b] for hydrologic simulations. Our saturation-suction curve is very close to the Ebel et al. [2007b] wetting curve.
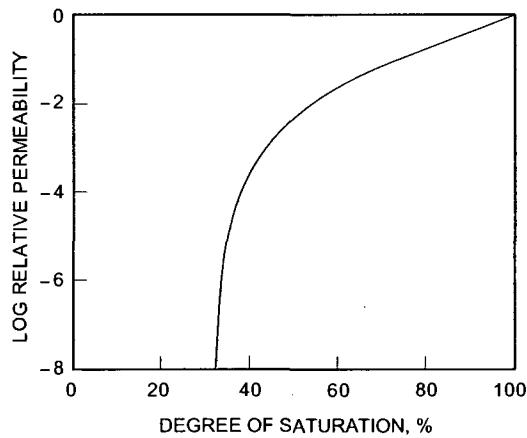

Figure 5.6: Variation of relative permeability with degree of saturation.

Table 5.1: Material parameters.

Intrinsic permeability	k	3.4×10^{-11}	m^2
Fluid viscosity	μ	1.0×10^{-6}	kPa·s
Residual saturation	ψ_1	0.32	
Maximum saturation	ψ_2	1.00	
Shape constant	n	3.00	
Scaling pressure	s_a	0.40	kPa
Bulk Modulus	K	50	MPa
Poisson Ratio	ν	0.25	
Cohesion	c	0–10	kPa
Friction angle	ϕ	33–40	deg.
Dilatancy angle	ψ	≈ 25	deg.
Porosity	ϕ^s	0.50	
Fluid density	ρ_w	1.0	Mg/m^3
Solid density	ρ_s	2.2	Mg/m^3

The elastic parameters K and ν are typical for sand/gravel mixture subjected to a confining pressure comparable to those prevailing in the CB1 sediment [Lambe and Whitman 1969]. The elastic parameters influence the displacements of the sediment but have little effect on the mechanism of failure, which is determined largely by the plasticity model. The parameters of the Mohr-Coulomb plasticity model are the cohesion c, friction angle ϕ, and dilatancy angle ψ, and their range of values is also summarized in Table 5.1.

5.5 Results

In general, failure mechanisms generated by the 2D slope model are complex and very much dependent on the imposed flow boundary conditions. Furthermore, the timing and location of initial failure are dependent on the intensity of rainfall and local flow conditions, such as the presence of a source or sink. In this section we summarize the results of numerical simulations on the 2D slope model. The section is divided into two parts. In the first part we show the impact of flow boundary conditions on the resulting deformation and failure patterns. In the second part we conduct parametric studies to investigate the sensitivity of the calculated deformation and failure patterns to variation in material parameters.

5.5.1 Impact of flow boundary conditions

We consider three fluid flow boundary conditions—Case A: a zero fluid pressure is prescribed on the slope surface; Case B: rainfall infiltration is prescribed on the slope surface in the form of fluid flux; and Case C: fluid infiltration into the sediment is prescribed on the interface between the sediment and fractured bedrock. The material parameters are summarized in Table 5.1. For the present simulations, baseline values of $c = 4$ kPa for cohesion, $\phi = 40°$ for friction angle, and $\psi = 25°$ for dilatancy angle were used.

In all the simulations the sediment was assumed fixed to the bedrock, and any relative sliding between the sediment and bedrock may only take place in the form of plastic deformation on the sediment. The slope is very thin compared to its length, so the displacement boundary conditions on the top and bottom ends of the slope are expected to play a very minor role on the calculated mechanical responses. Thus, in the present simulations we simply assumed the top and bottom ends of the slope to be fixed to the support. However, initial conditions do play a significant role on the hydro-mechanical responses, particularly the initial value of negative pore water pressure within the sediment. Based on the hydrologic simulations of Ebel et al. [2007b], we assumed an initial pore water pressure of -1.5 kPa throughout the sediment. Gravity load was turned on along with this initial pore water pressure to obtain the initial effective stresses, after which the nodal displacements were reset to zero. A standard 2×2 Gauss integration rule was used for the stabilized quadrilateral mixed elements, and a backward implicit scheme was used for time integration.

For Case A, we assumed the pore water pressure on the slope surface to ramp up linearly from -1.5 kPa at $t = 0$, to 0 kPa at $t = 3$ hr (Dirichlet boundary condition), after which the pressure was held fixed at 0 kPa. All other fluid flow boundaries were assumed to be impermeable (Neumann boundary condition). This represents a rainfall heavy enough to saturate the slope surface within a period of three hours and keep it saturated thereafter. Figure 5.7 summarizes the calculated hydro-mechanical responses at the inception of primary and secondary failure mechanisms within the slope. The general mechanism is as follows. As the sediment becomes saturated, local plastic zones develop on the base and stretch the slope segments unevenly. Eventually, plastic zones curve upwards due to local stretching of the slope segments and emerge on the ground surface approximately at coordinates $x = 43$ m and $x = 50$ m. The resulting slope failure mechanism is similar to multiple slide block described by Varnes [Varnes 1978]. Drained bifurcation analyses based on the Rudnicki and Rice [Rudnicki and Rice 1975] procedure allowed the calculation of localization function [Andrade and Borja 2006, 2007; Borja 2006a,b,c; Runesson et al. 1991], showing negative values. This suggests that a shear band-type bifurcation is likely to occur where the plastic deformation concentrates. On a related note Borja [2004] demonstrated that drained shear-band bifurcation is likely to be more critical than undrained bifurcation in partially saturated dilatant frictional materials. At the inception of initial and secondary failure mechanisms, the critical slope segment is nearly saturated, and the pore water pressure is approximately hydrostatic relative to the slope surface.

For Case B, we specified a seepage infiltration of 15 mm/hr on the slope surface (Neumann boundary condition) while maintaining the same flow boundary conditions on the other faces as
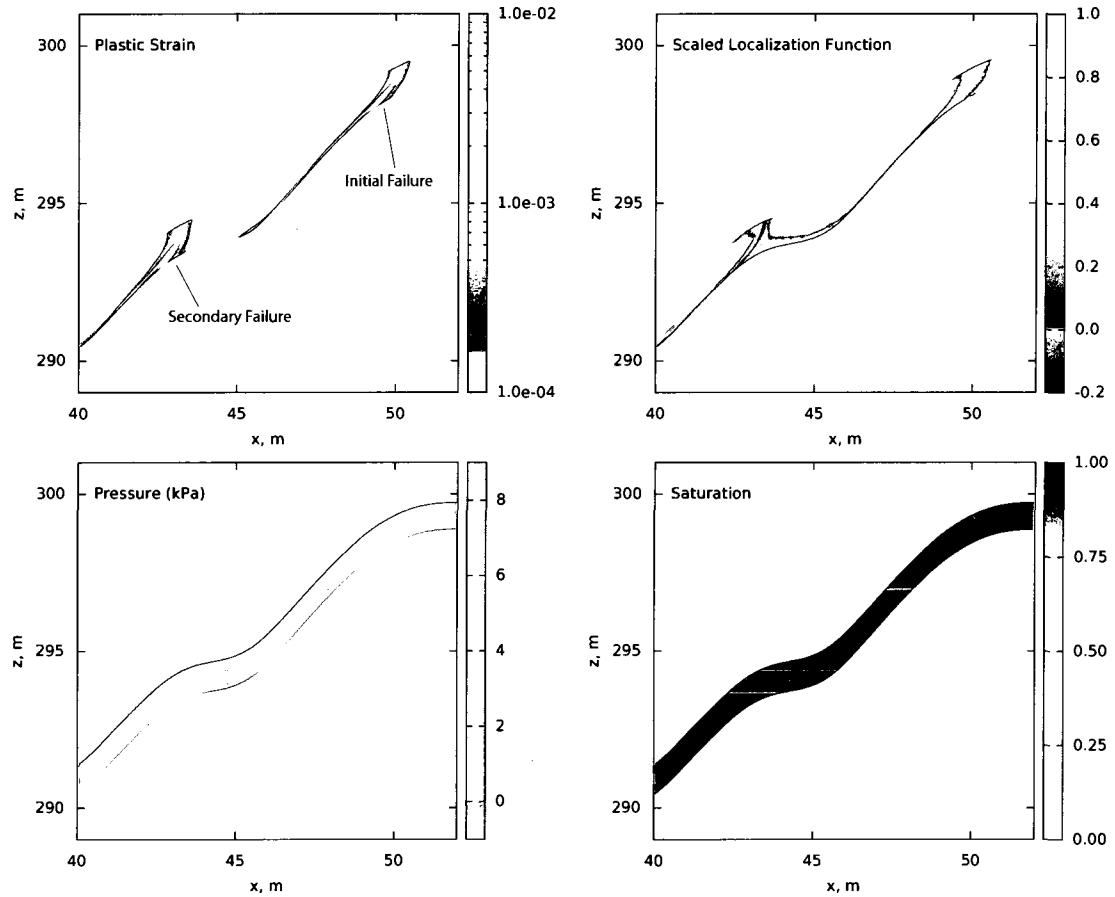
Figure 5.7: Multiple slide block mechanism in a steep hillside slope subjected to rising pore water pressure boundary condition from -1.5 kPa up to 0 kPa in 3 hr on slope surface.

in Case A (an initial all-Neumann set of flow boundary conditions). This eventually caused the bottom end of the slope to become saturated and the pore water pressure there to rise indefinitely. Since the pore water pressure cannot exceed zero on the slope surface, the boundary condition switched from a seepage type (Neumann) to a pressure-type (Dirichlet) on the slope surface where the pore water pressure reached zero. Figure 5.8 shows the calculated hydro-mechanical responses. The failure mechanism is similar to a multiple block type as in Case A, but note that the new boundary conditions have mobilized failure within a different segment of the slope (at $x = 16$ m and $x = 25$ m). The localization function also shows a propensity of the sediment to develop shear strain localization where plastic deformation concentrates. Furthermore, even though the sediment is saturated all throughout the depth near the bottom end of the slope, initial failure still developed upslope, at $x = 25$ m where the sediment is still partially saturated. This suggests that the sediment does not have to be completely saturated to experience failure by plastic deformation and shear strain localization.

For Case C, we also assumed an initial all-Neumann boundary condition with zero flux on all
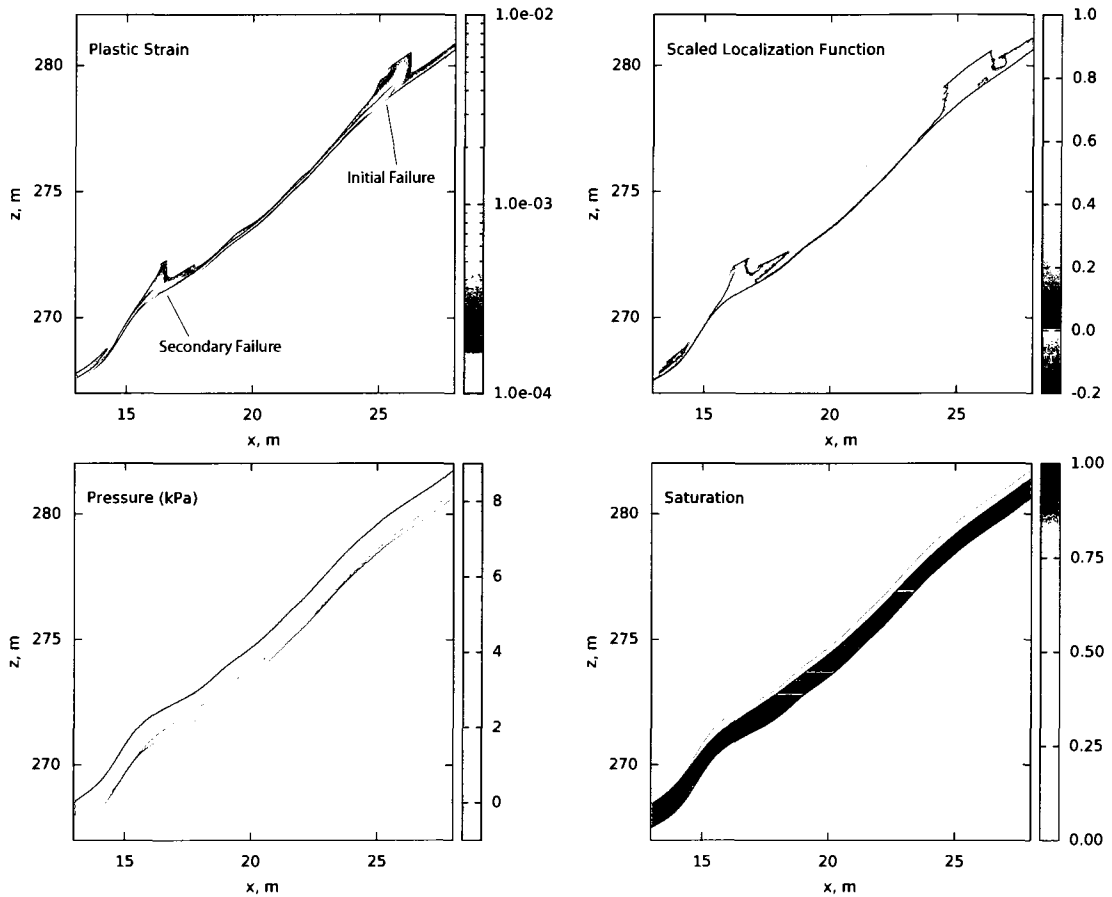
Figure 5.8: Multiple slide block mechanism in a steep hillside slope subjected to rainfall infiltration of 15 mm/hr on slope surface.

surface boundaries except over a finite segment between $x = 12$ m to $x = 18$ m on the sediment-bedrock interface. Thus, this segment of the fractured bedrock was assumed to be a fluid source. There has been much speculation about the fluid flux from the fractured bedrock at CB1. Whereas its importance on the hydrologic aspects has been elucidated from previous fluid flow simulations [Ebel et al. 2007b], its impact on the sediment mechanical responses is poorly understood. To isolate the effect of fluid infiltration from the fractured bedrock, we assumed the top surface to be a no-flux boundary unless the pressure plume reaches it, at which point it switches to a $p = 0$ face. Figure 5.9 suggests that with only fluid influx from the fractured bedrock it takes a large amount of seepage to build up enough pressure to fail the slope. Note that the pressure plume extends all the way to the surface, implying that upward seepage would be observed at the surface. It thus appears that upward seepage is unlikely to be the primary failure mechanism for a slope similar to CB1. However, upward seepage could certainly aggravate a slope that has already been weakened by other infiltration.
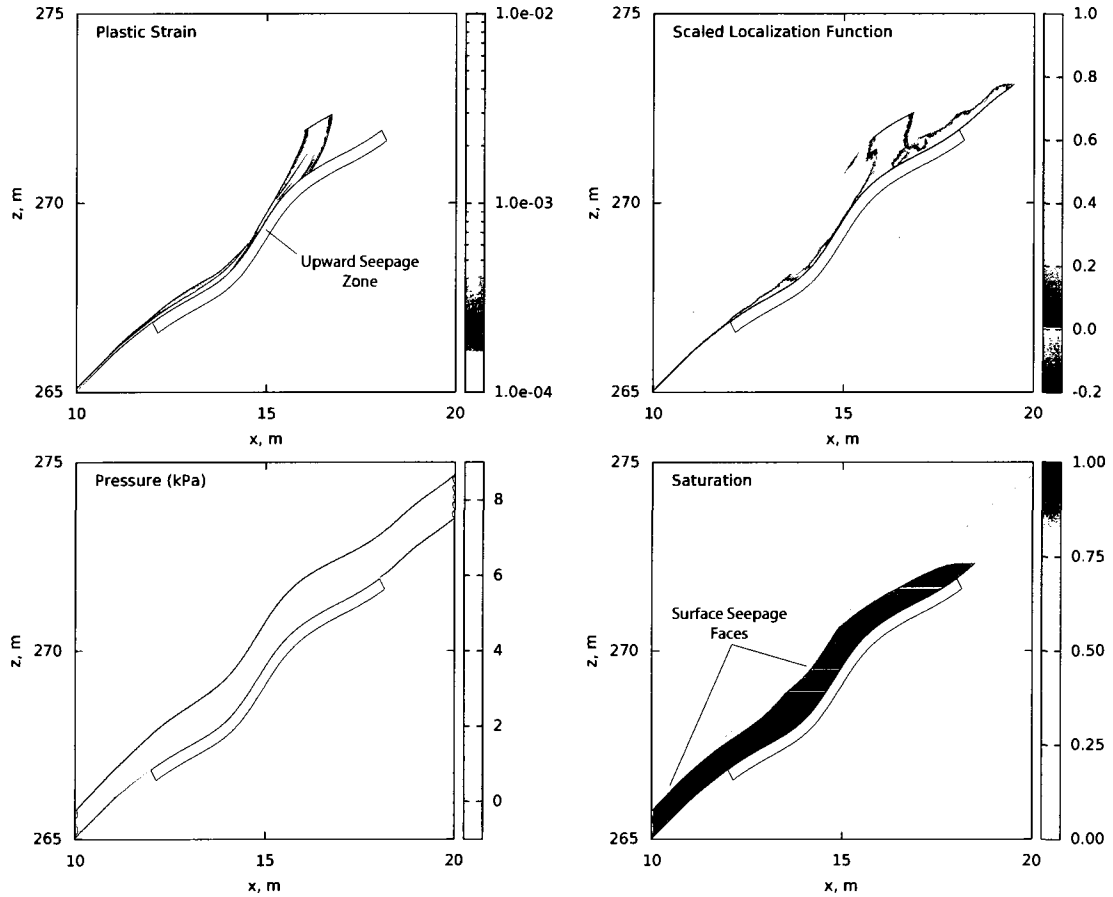
Figure 5.9: Failure mechanism in a steep hillside slope subjected to upward seepage on the base of the slope.

5.5.2 Impact of strength parameters variation on failure mechanism

The comprehensive hydrologic FE simulations conducted by Ebel et al. [2007b] highlighted the importance of being able to constrain the parameters of the fluid flow model in general, and the soil-water retention curve in particular, for a realistic prediction of the field responses at CB1. In this section, we address the implications of uncertain strength parameters on the predicted failure mechanism and the timing of such mechanism for a steep hillside slope similar to CB1. Fortunately, the perfectly plastic Mohr-Coulomb model has only three strength parameters: cohesion c, friction angle ϕ, and dilatancy angle ψ, and thus the parametric studies can be conducted with relative ease. Any unknown hardening or softening responses may be viewed as part of the uncertainties in these strength parameters.

Figure 5.10 shows the impact of an uncertain cohesion on the failure mechanism for the 2D slope model. The friction and dilatancy angles were fixed at 40° and 25°, respectively, and the loading condition was defined by Case B (prescribed seepage on the slope face). We recall that at the CB1
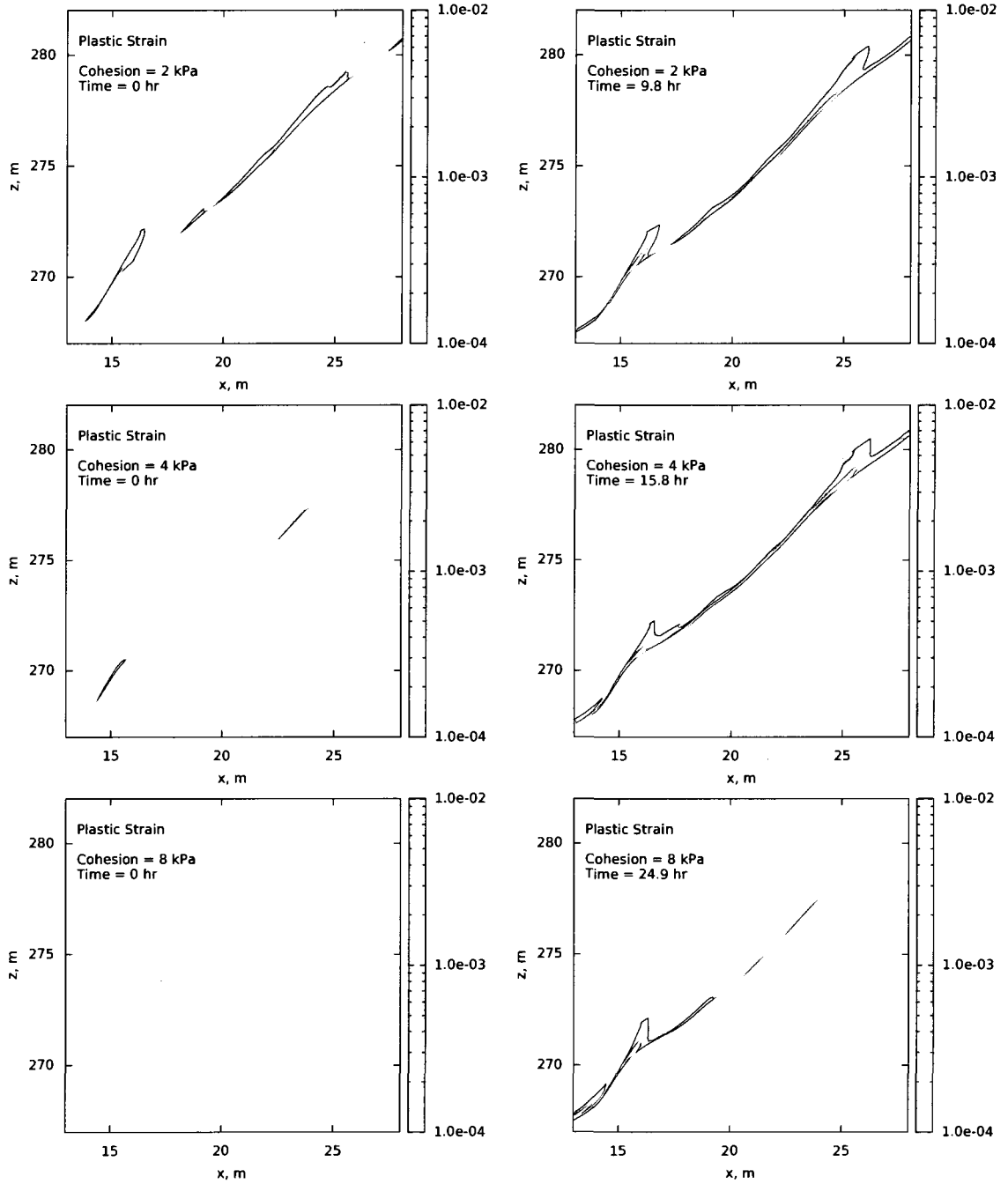
Figure 5.10: Variation of failure mechanism and timing of failure with cohesion.
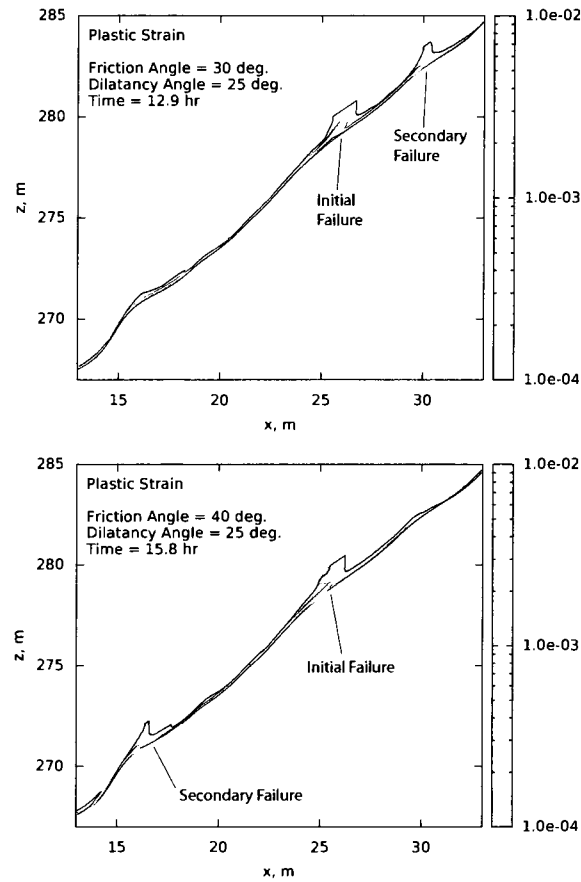
Figure 5.11: Variation of failure mechanism and timing of failure with friction angle.

site the parameter c was derived mainly from lateral root cohesion, which could vary in space and time. For example, as the sediment is stretched the root strength could be mobilized, resulting in an increase in c, but it could also subsequently decrease when the roots are pulled from the host sediment. The figure shows the failure mechanisms and timing of failure at cohesion of 2, 4, and 8 kPa. At 2 kPa the slope failed for the initial configuration, just under its own self-weight. The observed trends are what one would expect: larger cohesion delays the timing of failure. However, at $c = 8$ kPa only a single plastic zone emerged on the slope face, not multiple ones.

Figure 5.11 shows the impact of an uncertain friction angle ϕ on the failure mechanism for the 2D slope model. In the simulations c was fixed at 4 kPa, and the loading condition was again defined by Case B. For the two values of friction angle tested ($\phi = 30°$ and $40°$) the initial plastic zone emerged on the slope face at approximately the same location (at $x = 25$ m). However, for $\phi = 30°$ the secondary failure zone developed at an earlier time and above the initial failure zone (at $x = 30$ m), whereas for $\phi = 40°$ the secondary plastic zone emerged on the slope face at a later time and below the initial failure zone (at $x = 17$ m).
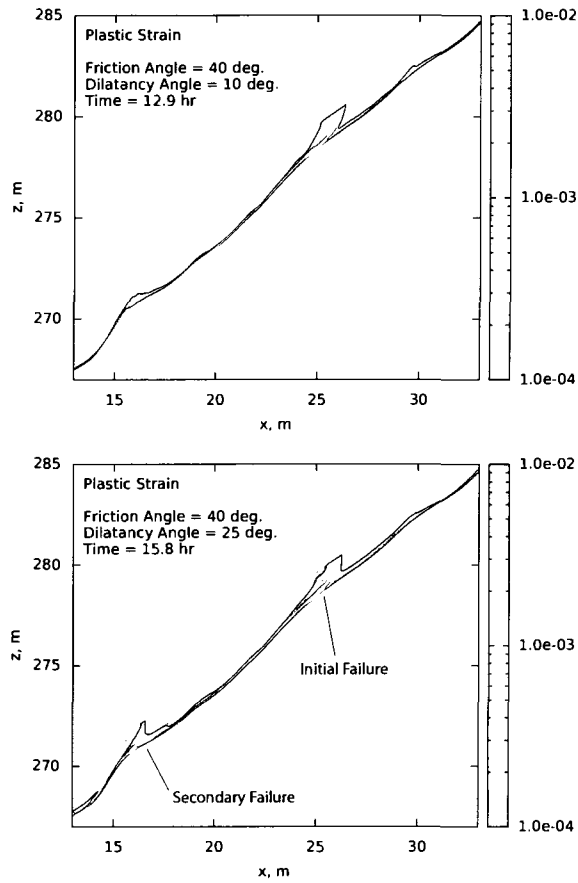
Figure 5.12: Variation of failure mechanism and timing of failure with dilatancy angle.

Finally, we consider the effect of the dilatancy angle on the predicted failure kinematics. Figure 5.12 compares the plastic strain distribution in the slope using two dilatancy angles, 10° and 25°, while the friction angle remains fixed at the baseline value of 40°. In both cases, a failure surface extending to slope surface is observed at approximately $x = 25$ m, suggesting that this failure mode is insensitive to dilatancy. In the low dilatancy angle simulation, however, the secondary failure in the $x = 16$ m region is suppressed. There is some indication that with additional deformation a secondary failure may initiate at this point, but the timing of the two would not be as closely linked as in the high dilatancy angle case. Close examination of the plastic strain contours in both simulations also indicates that the plastic zone in the low dilatancy case is less "diffuse" than in the high dilatancy case.

5.6   Summary and conclusions

We have presented a continuum physics-based framework for analysis of coupled solid deformation-fluid flow processes in partially saturated earthen slopes. The formulation for constitutive modeling of the unsaturated sediment is based on the use of an effective stress measure that represents the sole stress state variable responsible for deforming the solid skeleton, and on direct use of suction in the constitutive relations [Borja 2004, 2006c; Borja and Koliji 2009; Koliji et al. 2008]. Other alternative approaches exist in which different pairs of stress state variables are chosen instead of a single effective stress, such as the net stress and matric suction [Alonso et al. 1990; Fredlund and Morgenstern 1977; Gallipoli et al. 2003; Vaunat et al. 2000; Wheeler et al. 2002]. However, the effective stress used in this work appears to have an advantage in terms of FE implementation since the coupled formulation follows exactly the same lines as those developed for the fully saturated case. This implies that stabilized low-order mixed FE elements can be used just as well for the unsaturated case.

We have used the coupled continuum FE model to analyze the deformation and stability of a steep hillside slope similar to the CB1 site. Despite the simplified 2D plane strain representation adopted by the model, the predicted failure mechanisms are complex and nowhere near what one would normally obtain from an infinite slope assumption. For a steep hillside slope similar to CB1, it appears that failure mechanism would likely initiate in the form multiple failure blocks, which could transform into a debris flow similar to what was observed at the CB1 site during the November 1996 event. An obvious advantage of the proposed continuum modeling approach lies in its ability to integrate important hydro-mechanical processes responsible for triggering rainfall-induced movement of earthen slopes, including increased saturation, fluid flow, and inelastic solid deformation. Work is currently underway to incorporate 3D effects into the numerical model.

6 Conclusion

As we have seen, in many applications it is necessary to strongly couple the hydrologic and geome-chanical responses in order to accurately predict system behavior. Numerical simulations provide a natural tool for these investigations, and provide a powerful complement to theory, experiment, and field observation. Unfortunately, there has been hesitancy to adopt fully-coupled formulations in practice. Much of this hesitancy stems from the expense of performing these simulations. Even with today's growing computer power, the numerical challenges associated with these problems are sufficiently complex that they defy "brute force" solution methods. Rather, we need scalable, effi-cient approaches that use modern computing platforms to their best advantage. Chapters 3 and 4 of this thesis have been aimed at precisely this issue.

First, we examined the inherent stability restrictions that are present in mixed finite element formulations. Standard, stable discretizations that satisfy these restrictions can be very expensive, leading to algebraic problems with many more unknowns than is likely justified by the convergence behavior of the approximation. To deal with this fact, we have proposed a stabilized finite element method that circumvents these restrictions and allows for the successful use of equal-order linear interpolation for all field variables. For very large, three-dimensional problems, the resulting system matrices will have approximately one-sixth as many degrees of freedom as their stable counterparts. The stabilized scheme also has several appealing computational properties that set it apart from other approaches: it only requires standard shape function information, is entirely element local, preserves the original symmetry and sparsity patterns of the system matrix, and is computationally cheap to apply. Through several numerical examples we have demonstrated that the method shows good convergence properties while being significantly more tractable than standard, unequal-order approaches.

Even with stabilization, however, coupled formulations lead to algebraic systems that can be on the order of hundreds of thousands to millions of degrees of freedom. Furthermore, these problems are nonlinear, stiff, and extremely ill-conditioned. We have demonstrated that despite these diffi-culties efficient solution procedures can be designed. For the linear solver, we use memory-efficient Krylov subspace methods like GMRES or BICGSTAB. To deal with ill-conditioning, we have proposed a block-structured preconditioning approach. This block approach provides a middle ground be-tween "physics-based" and "algebraic" preconditioners. Algebraic preconditioners work very well for single-physics applications, and many linear algebra packages provide high-performance imple-mentations. Unfortunately, since they have no knowledge of the coupled nature of multiphysics

problems, they typically perform poorly in these situations. The block-preconditioning approach directly addresses this issue, by building the preconditioner through block-decompositions of the system matrices. In doing so, we endow the preconditioner with physics-based knowledge about the coupled system, but also decompose the problem into simpler pieces for which algebraic preconditioners work extremely well. When combined with Incomplete LU (ILU) or Algebraic Multigrid (AMG) sub-preconditioners, we have demonstrated fast, scalable performance of both serial and parallel processors. With additional work, we also believe we can minimize the communication overhead in a parallel setting, and achieve better parallel efficiency. We have also examined how the iterative solver can provide additional computational savings in the context of inexact Newton methods, automatically tuning the amount of work performed based on the expected increase in accuracy it will provide.

As far as future numerical work is concerned, we have only indirectly addressed the inherent stiffness in the nonlinear problems we have been solving. These nonlinearities can lead to many Newton updates before quadratic convergence is observed. Furthermore, several robustness checks are required to ensure divergence does not take place. An interesting problem to consider is that of "nonlinear preconditioning." Just as a linear preconditioner reduces the number of linear solver iterations, nonlinear preconditioning seeks to reduce the number of nonlinear Newton iterations. By reducing and balancing the nonlinearities in the problem, significant speed and robustness improvements can be obtained. As another avenue of research with respect to stabilization, we note other physical mechanisms besides divergence constraints can lead to oscillatory solutions. These can include oscillations around the self-sharpening wetting fronts that may appear in modeling unsaturated fluid flow. We would like to explore the application of advective stabilization techniques to these numerical difficulties.

In the last chapter, we have used our physics-based model to examine a critical application: hydrologically-driven landslides. As a motivating problem, we have used the 1996 landslide at the CB1 experimental catchment. In order to limit the number of uncertainties in the model, however, we have implemented a highly-simplified, two-dimensional approximation. We have explored the sensitivity of the failure kinematics to changes in the material parameters, which are not as well constrained as the hydrologic parameters at the CB1 site. Even the simple model indicates that the failure kinematics are quite complex, involving a sliding block mechanism in which multiple failure surfaces interact.

Of course, the work presented in this dissertation is by no means the end of the story. With the knowledge obtained about material uncertainties from the 2D analyses, we intend to examine the effect of the 3D geometry on the expected failure mechanisms. Also, the hydrologic modeling we have used to this point has been simplified—we assume, for example, that the rainfall can be modeled as a constant flux, rather than specifying the actual time variation as measured during the failure storm. The boundary and initial conditions were also simplified. At the actual CB1 site, hydrologic mechanisms associated with sapprolite layers and the fractured bedrock are complex. These additional features may be essential to modeling the flow and pressure distribution in a realistic way. Fortunately, sprinkling experiments and other field observations provide a wealth of

data to test a more sophisticated hydrologic model.

The basic premise of this dissertation is that fully-coupled simulations should become a standard analysis technique, not one that is put off as too complex or expensive to provide meaningful insight. Our hope is that the work presented here moves us closer to this goal.

Bibliography

Alonso, E., Gens, A., and Josa, A. (1990). A constitutive model for partially saturated soils. *Géotechnique*, 40:405–430.

Anderson, S. (1995). *Flow paths, solute sources, weathering, and denudation rates: The chemical geomorphology of a small catchment.* PhD thesis, University of California, Berkeley.

Anderson, S., Dietrich, W., and Brimhall, G. (2002). Weathering profiles, mass balance analysis, and rates of solute loss: Linkages between weathering and erosion in a small, steep catchment. *GSA Bulletin*, 114:1143–1158.

Anderson, S., Dietrich, W., Montgomery, D., Torres, R., Conrad, M., and Loague, K. (1997a). Subsurface flowpaths in a steep, unchanneled catchment. *Water Resources Research*, 33:2637–2653.

Anderson, S., Dietrich, W., Montgomery, D., Torres, R., and Loague, K. (1997b). Concentration discharge relationships in runoff from a steep, unchanneled catchment. *Water Resources Research*, 33:211–225.

Andrade, J. and Borja, R. (2006). Capturing strain localization in dense sands with random density. *International Journal for Numerical Methods in Engineering*, 67:1531–1564.

Andrade, J. and Borja, R. (2007). Modeling deformation banding in dense and loose fluid-saturated sands. *Finite Elements in Analysis and Design*, 43:361–383.

Arnold, D. (1990). Mixed finite element methods for elliptic problems. *Comp. Meth. Appl. Mech. Eng.*, 82:281–300.

Arnold, D., Brezzi, F., and Fortin, M. (1984). A stable finite element for the stokes equations. *Calcolo*, 21(4):337–344.

Bangerth, W., Hartmann, R., and Kanschat, G. (2007). deal.II — a general purpose object oriented finite element library. *ACM Transactions on Mathematical Software*, 33(4):24. Article 24, 27 pages.

Becker, R. and Braack, M. (2001). A finite element pressure gradient stabilization for the Stokes equations based on local projections. *Calcolo*, 38:173–199.

Benzi, M., Golub, G., and Liesen, J. (2005). Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137.

Birle, E., Heyer, D., and Vogt, N. (2008). Influence of the initial water content and dry density on the soil-water retention curve and the shrinkage behavior of a compacted clay. *Acta Geotechnica*, 3:191–200.

Bochev, P. and Dohrmann, C. (2006). A computational study of stabilized, low-order C^0 finite element approximations of Darcy equations. *Computational Mechanics*, 38:323–333.

Bochev, P., Dohrmann, C., and Gunzburger, M. (2006). Stabilization of low-order mixed finite elements for the Stokes Equations. *SIAM J. Numer. Anal.*, 44(1):82–101.

Borja, R. (2004). Cam-Clay plasticity. Part V: A mathematical framework for three-phase deformation and strain localization analyses of partially saturated porous media. *Comp. Meth. Appl. Mech. Eng.*, 193(48-51):5301–5338.

Borja, R. (2006a). Condition for liquefaction instability in fluid-saturated granular soils. *Acta Geotechnica*, 1(4):211–224.

Borja, R. (2006b). Conditions for instabilities in collapsible solids including volume implosion and compaction banding. *Acta Geotechnica*, 1(2):107–122.

Borja, R. (2006c). On the mechanical energy and effective stress in saturated and unsaturated porous continua. *International Journal of Solids and Structures*, 43(6):1764–1786.

Borja, R. and Andrade, J. (2006). Critical state plasticity. Part VI: Meso-scale finite element simulation of strain localization in discrete granular materials. *Computer Methods in Applied Mechanics and Engineering*, 195(37-40):5115–5140.

Borja, R. and Koliji, A. (2009). On the effective stress in unsaturated porous continua with double porosity. *Journal of the Mechanics and Physics of Solids*, doi:10.1016/j.jmps.2009.04.014.

Bramble, J. and Pasciak, J. (1988). A preconditioning technique for indefinite systems resulting from mixed approximation of elliptic problems. *Math. Comput.*, 50(181):1–17.

Brezzi, F. (1990). A discourse on the stability conditions for mixed finite element formulations. *Comp. Meth. Appl. Mech. Eng.*, 82(1-3):27–57.

Brezzi, F. and Pitkäranta, J. (1984). On the stabilization of finite element approximations of the Stokes problem. *Efficient Solution of Elliptic Systems*, pages 11–19.

Brown, III, W., Sitar, N., Saarinen, T., and Blair, M. (1984). *Debris flows, landslides, and floods in the San Francisco Bay region, January 1982. Overview of and summary of a conference held at Stanford University, August 23–26, 1982.* National Research Council and USGS, Washington D.C.

Burman, E. (2007). Pressure projection stabilizations for Galerkin approximations of Stokes' and Darcy's problem. *Num. Meth. Partial Differential Equations*, 24(1):127–143.

Burroughs, Jr., E., Hammond, C. J., and Booth, G. D. (1985). Relative stability estimation for potential debris avalanche sites using field data. In *Proceedings of the International Symposium on Erosion, Debris Flow and Disaster Prevention*, pages 335–339, Tsukuba, Japan.

Cryer, C. (1963). A comparison of the three-dimensional consolidation theories of Biot and Terzaghi. *The Quarterly Journal of Mechanics and Applied Mathematics*, 16(4):401–412.

Cuthill, E. and McKee, J. (1969). Reducing the bandwidth of sparse symmetric matrices. In *Proceedings of the 1969 24th national conference*, pages 157–172. ACM Press New York, NY, USA.

Demmel, J. W., Eisenstat, S. C., Gilbert, J. R., Li, X. S., and Liu, J. W. H. (1999). A supernodal approach to sparse partial pivoting. *SIAM J. Matrix Analysis and Applications*, 20(3):720–755.

Dohrmann, C. and Bochev, P. (2004). A stabilized finite element method for the Stokes problem based on polynomial pressure projections. *Int. J. Numer. Meth. Fluids*, 46:183–201.

Douglas, J. and Wang, J. (1989). An Absolutely Stabilized Finite Element Method for the Stokes Problem. *Mathematics of Computation*, 52(186):495–508.

Ebel, B., Loague, K., Dietrich, W., Montgomery, D., Torres, R., Anderson, S., and Giambelluca, T. (2007a). Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 1. Sprinkling experiments. *American Journal of Science*, 307:678–708.

Ebel, B., Loague, K., Vanderkwaak, J., Dietrich, W., Montgomery, D., Torres, R., and Anderson, S. (2007b). Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 2. Physics-based simulations. *American Journal of Science*, 307:709–748.

Eisenstat, S. and Walker, H. (1996). Choosing the forcing terms in an inexact Newton method. *SIAM J. Sci. Comput.*

Elman, H., Howle, V., Shadid, J., Shuttleworth, R., and Tuminaro, R. (2008). A taxonomy and comparison of parallel block multi-level preconditioners for the incompressible Navier–Stokes equations. *Journal of Computational Physics*, 227(3):1790–1808.

Fredlund, D. and Morgenstern, N. (1977). Stress state variables for unsaturated soils. *Journal of the Geotechnical Engineering Division, ASCE*, 103:354–358.

Fredrich, J., DiGiovanni, A., and Noble, D. (2006). Predicting macroscopic transport properties using microscopic image data. *Journal of Geophysical Research*, 111(B03201).

Gallipoli, D., Gens, A., Sharma, R., and Vaunat, J. (2003). An elasto-plastic model for unsaturated soil incorporating the effects of suction and degree of saturation on mechanical behaviour. *Géotechnique*, 53(1):123–136.

Gawin, D., Baggio, P., and Schrefler, B. (1995). Coupled heat, water and gas flow in deformable porous media. *Int. J. Num. Meth. Fluids*, 20:969–987.

Gee, M., Hu, J., and Tuminaro, R. (2009). A new smoothed aggregation multigrid method for anisotropic problems. *Numerical Linear Algebra with Applications*, 16(1).

Gee, M., Siefert, C., Hu, J., Tuminaro, R., and Sala, M. (2006). ML 5.0 smoothed aggregation user's guide. Technical Report SAND2006-2649, Sandia National Laboratories.

George, A. and Liu, J. (1989). The Evolution of the Minimum Degree Ordering Algorithm. *SIAM Review*, 31:1.

Gibbs, J. (1898). Fourier series. *Nature*, 200.

Hardebeck, J. L. and Hauksson, E. (1999). Role of Fluids in Faulting Inferred from Stress Field Signatures. *Science*, 285(5425):236–239.

Heroux, M., Bartlett, R., Howle, V., Hoekstra, R., Hu, J., Kolda, T., Lehoucq, R., Long, K., Pawlowski, R., Phipps, E., et al. (2005). An overview of the Trilinos project. *ACM Transactions on Mathematical Software*, 31(3):397–423.

Houlsby, G. (1979). The work input to a granular material. *Géotechnique*, 29:354–358.

Houlsby, G. (1997). The work input to an unsaturated granular material. *Géotechnique*, 47:193–196.

Houston, S., Houston, W., Zapata, C., Manepally, C., and C., L. (1999). Influence of compressibility on use and interpretation of soil water characteristic curve. In *Proceedings of the XI Pan-American Conference on Soil Mechanics and Geotechnical Engineering*, pages 947–954, Foz do Iguassu, Brazil.

Hughes, T., Feijóo, G., Mazzei, L., and Quincy, J. (1998). The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24.

Hughes, T., Franca, L., and Balestra, M. (1986). 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. *Comp. Meth. Appl. Mech. Eng.*, 59(1):85–99.

Iverson, R. M., Reid, M. E., and LaHusen, R. G. (1997). Debris-flow mobilization from landslides. *Annual Review of Earth and Planetary Sciences*, 25(1):85–138.

Jha, B. and Juanes, R. (2007). A locally conservative finite element framework for the simulation of coupled flow and reservoir geomechanics. *Acta Geotechnica*, 2(3):139–153.

Jibson, R. (1992). The Mameyes, Puerto Rico, landslide disaster of October 7, 1985. In Slosson, J., Keene, A., and Johnson, J., editors, *Landslides/Landslide Mitigation*, Reviews in Engineering Geology, pages 37–54. Geological Society of America, Boulder, Colorado.

Jibson, R. (2005). Landslide hazards at La Conchita, California. Technical report, United States Geological Survey.

Khalili, N., Geiser, F., and Blight, G. (2004). Effective stress in unsaturated soils: review with new evidence. *International Journal of Geomechanics*, 4:115–126.

King, I. (1970). An automatic reordering scheme for simultaneous equations derived from network systems. *Int. J. Numer. Meth. Eng*, 2:479–509.

Knoll, D. and Keyes, D. (2004). Jacobian-free Newton–Krylov methods: a survey of approaches and applications. *Journal of Computational Physics*, 193(2):357–397.

Koliji, A., Vulliet, L., and Laloui, L. (2008). New basis for the constitutive modelling of aggregated soils. *Acta Geotechnica*, 3:61–69.

Lagmay, A., Ong, J., Fernandez, D., Lapus, M., Rodolfo, R., Tengonciang, A., Soria, J., Baliatan, E., Quimba, Z., Uichanco, C., Paguican, E., Remedio, A. C., Lorenzo, G., Validivia, W., and Avila, F. (2006). Scientists investigate recent Philippine landslide. *Eos Transactions American Geophysical Union*, 87(12):121–128.

Lambe, T. and Whitman, R. (1969). *Soil mechanics*. John Wiley & Sons, New York.

Li, X., Han, X., and Pastor, M. (2003). An iterative stabilized fractional step algorithm for finite element analysis in saturated soil dynamics. *Comp. Meth. Appl. Mech. Eng.*, 192(35-36):3845–3859.

Liakopoulos, A. (1964). *Transient Flow Through Unsaturated Porous Media*. PhD thesis, University of California, Berkeley.

Martinez, E. (2000). Evento Meteorologico sobre el Litoral Central en Diciembre 1999. *Informe inedito*.

Masud, A. and Hughes, T. (2002). A stabilized mixed finite element method for Darcy flow. *Comp. Meth. Appl. Mech. Eng.*, 191(39-40):4341–4370.

Masud, A. and Xia, K. (2005). A Stabilized Mixed Finite Element Method for Nearly Incompressible Elasticity. *Journal of Applied Mechanics*, 72:711.

Masud, A. and Xia, K. (2006). A variational multiscale method for inelasticity: Application to superelasticity in shape memory alloys. *Computer Methods in Applied Mechanics and Engineering*, 195(33-36):4512–4531.

Ming, H. and Li, X. (2003). Fully coupled analysis of failure and remediation of lower San Fernando Dam. *Journal of geotechnical and geoenvironmental engineering*, 129(4):336–349.

Montgomery, D. (1991). *Channel initiation and landscape evolution*. PhD thesis, University of California, Berkeley.

Montgomery, D. and Dietrich, W. (2002). Runoff generation in a steep, soil-mantled landscape. *Water Resources Research*, 38(9):1168.

Montgomery, D., Dietrich, W., and Heffner, J. (2002). Piezometric response in shallow bedrock at CB1: Implication for runoff generation and landsliding. *Water Resources Research*, 38(12):1274.

Montgomery, D., Dietrich, W., Torres, R., Anderson, S., Heffner, J., and Loague, K. (1997). Hydrologic response of a steep, unchanneled valley to natural and applied rainfall. *Water Resources Research*, 33:91–109.

Moriguchi, S., Borja, R., Yashima, A., and Sawada, K. (2009). Estimating the impact force generated by granular flow on a rigid obstruction. *Acta Geotechnica*, 4:57–71.

Murad, M. and Loula, A. (1994). On stability and convergence of finite element approximations of Biot's consolidation problem. *Int. J. Numer. Meth. Eng.*, 37:645–667.

Nur, A. and Byerlee, J. (1971). An exact effective stress law for elastic deformation of rock with fluids. *Journal of Geophysical Research*, 76:6414–6419.

Nuth, M. and Laloui, L. (2008). Effective stress concept in unsaturated soils: Clarification and validation of a unified framework. *International Journal for Numerical and Analytical Methods in Geomechanics*, In press.

Oñate, E., Rojek, J., Taylor, R., and Zienkiewicz, O. (2004). Finite calculus formulation for incompressible solids using linear triangles and tetrahedra. *Int. J. Numer. Meth. Eng*, 59:1473–1500.

Pastor, M., Li, T., Liu, X., Zienkiewicz, O., and Quecedo, M. (2000). A fractional step algorithm allowing equal order of interpolation for coupled analysis of saturated soil problems. *Mechanics of Cohesive-Frictional Materials*, 5(7):511–534.

Pastor, M., Zienkiewicz, O., and Li, T. (1999). Stabilized finite elements with equal order of interpolation for soil dynamics problems. *Archives of Computational Methods in Engineering*, 6(1):3–33.

Richards, L. (1931). Capillary conduction of liquids in porous mediums. *Physics*, 1:318–333.

Romero, I. and Bischoff, M. (2007). Incompatible Bubbles: A non-conforming finite element formulation for linear elasticity. *Computer Methods in Applied Mechanics and Engineering*, 196(9-12):1662–1672.

Roscoe, K. and Burland, J. (1968). On the generalized stress–strain behaviour of 'wet' clay. *Engineering Plasticity*, 3:539–609.

Roscoe, K., Schofield, A., and Wroth, C. (1958). On the yielding of soils. *Geotechnique*, 8(1):22–53.

Rudnicki, J. and Rice, J. (1975). Conditions for the localization of deformation in pressure-sensitive dilatant materials. *Journal of the Mechanics and Physics of Solids*, 23:371–394.

Runesson, K., Ottosen, N., and Perić, D. (1991). Discontinuous bifurcations of elastic-plastic solutions at plane stress and plane strain. *International Journal of Plasticity*, 7:99–121.

Saad, Y. (2003). *Iterative Methods for Sparse Linear Systems*. Society for Industrial Mathematics, Philadelphia, 2nd edition.

Sala, M. and Heroux, M. (2005). Robust algebraic preconditioners with IFPACK 3.0. Technical Report SAND-0662, Sandia National Laboratories.

Schmidt, K. (1994). Mountain scale strength properties, deep-seated landsliding and relief limits. Master's thesis, University of Washington, Seattle.

Schmidt, K. (1999). *Root strength, colluvial soil depth, and colluvial transport on landslide-prone hillslopes*. PhD thesis, University of Washington, Seattle.

Schmidt, K., Roering, J., Stock, J., Dietrich, W., Montgomery, D., and Schaub, T. (2001). The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range. *Canadian Geotechnical Journal*, 38:995–1024.

Schofield, A. and Wroth, P. (1968). *Critical state soil mechanics*. McGraw-Hill, New York.

Schrefler, B. and Gawin, D. (1996). The effective stress principle: incremental or finite form? *International Journal for Numerical and Analytical Methods in Geomechanics*, 20(11):785–814.

Schroeder, W. L. and Alto, J. V. (1983). Soil properties for slope stability analysis; Oregon and Washington coastal mountains. *Forest Science*, 29:823–833.

Schuster, R., Salcedo, D., and Valenzuela, L. (2002). Overview of catastrophic landslides of South America in the Twentieth Century. In Evans, S. and DeGraff, J., editors, *Catastrophic Landslides: Effects, Occurrence, and Mechanisms*, volume 15 of *Reviews in Engineering Geology*, pages 1–33. Geological Society of America, Boulder, Colorado.

Seed, H., (US, N. S. F., of Water, D., Power, and California (1973). *Analysis of the slides in the San Fernando dams during the earthquake of Feb. 9, 1971*. College of Engineering, University of California.

Segall, P. and Rice, J. (1995). Dilatancy, compaction, and slip instability of a fluid-infiltrated fault. *J. Geophys. Res.*, 100(B11):22155–22172.

Skempton, A. (1961). Effective stress in soils, concrete and rocks. In *Pore Pressure and Suction in Soils*, pages 4–16. Butterworths, London.

Sleep, N. and Blanpied, M. (1992). Creep, compaction and the weak rheology of major faults. *Nature*, 359(6397):687–692.

Smith, T. and Hart, E. (1982). Landslides and related storm damage, January 1982, San Francisco Bay region. *California Geology*, 35:139–152.

Terzaghi, K. (1943). *Theoretical Soil Mechanics.* John Wiley & Sons, New York.

Terzaghi, K. and Peck, R. (1967). *Soil mechanics in engineering practice.* Wiley, New York.

Torres, R. (1997). *Unsaturated zone processes and the hydrologic response of a steep, unchanneled valley.* PhD thesis, University of California, Berkeley.

Torres, R., Dietrich, W., Montgomery, D., Anderson, S., and Loague, K. (1998). Unsaturated zone processes and the hydrologic response of a steep, unchanneled catchment. *Water Resources Research*, 34:1865–1879.

Truty, A. (2001). A Galerkin/least-squares finite element formulation for consolidation. *Int. J. Numer. Meth. Eng.*, 52:763–786.

Truty, A. and Zimmermann, T. (2006). Stabilized mixed finite element formulations for materially nonlinear partially saturated two-phase media. *Comp. Meth. Appl. Mech. Eng.*, 195:1517–1546.

USAID (2000). Venezuela factsheet. Technical report, USAID Office of Foreign Disaster Assistance.

van Genuchten, M. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J*, 44(5):892–898.

Van Sint Jan, M. and Talloni, P. (1993). Flujo de sedimentos del 18 de Juniode 1991 en Antofagosta: La Serena, Chile. In *Tercer Congreso Chileno de Ingenieria Geotecnia*, volume 1, pages 247–265.

VanderKwaak, J. (1999). *Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems.* PhD thesis, University of Waterloo, Ontario, Canada.

Varnes, D. (1978). Slope movement types and processes. In Schuster, R. and Krizek, R., editors, *Landslides Analysis and Control*, Special Report 176. National Academy of Sciences, Washington D.C.

Vaunat, J., Cante, J., Ledesma, A., and Gens, A. (2000). A stress point algorithm for an elastoplastic model in unsaturated soils. *International Journal of Plasticity*, 16:121–141.

Verfürth, R. (1984). Error estimates for a mixed finite element approximation of the Stokes equations. *RAIRO. Analyse numérique*, 18(2):175–182.

Wan, J. (2002). *Stabilized finite element methods for coupled geomechanics and multiphase flow.* PhD thesis, Stanford University.

Wheeler, S., Gallipoli, D., and Karstunen, M. (2002). Comments on the use of Barcelona Basic Model for unsaturated soils. *International Journal for Numerical and Analytical Methods in Geomechanics*, 26:1561–1571.

White, J. and Borja, R. (2008). Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients. *Comp. Meth. Appl. Mech. Eng.*, 197(49–50):4353–4366.

White, J., Borja, R., and Fredrich, J. (2006). Calculating the effective permeability of sandstone with multiscale lattice Boltzmann/finite element simulations. *Acta Geotechnica*, 1(4):195–209.

White, J. and Fredrich, J. (2007). An object-oriented computational framework for rock properties prediction. Technical report, Exploration and Production Technology, BP America, Inc.

Wu, T., Beal, P., and Lan, C. (1988). In-situ shear test of soil-root systems. *Journal of Geotechnical Engineering, ASCE*, 114:1376–1394.

Yee, C. and Harr, D. (1977). Influence of soil aggregation on slope stability in the Oregon Coast Range. *Environmental Geology*, 1:367–377.

Young, Y., White, J., Xiao, H., and Borja, R. (2009). Tsunami-induced liquefaction failure of coastal slopes. *Acta Geotechnica*, 4:17–34.