T H E U N I V E R S I T Y O F T U L S A

THE GRADUATE SCHOOL

ON SUFFICIENT CONDITIONS FOR THE CONVERGENCE OF DAMPED NEWTON METHODS FOR THE IMPLICIT SIMULATION OF TRANSPORT

by

Shahriyar Alkhasli

A thesis submitted in partial fulfillment of

the requirements for the degree of Master of Science

in the Discipline of Petroleum engineering

The Graduate School

The University of Tulsa

2015

T H E U N I V E R S I T Y O F T U L S A

THE GRADUATE SCHOOL

ON SUFFICIENT CONDITIONS FOR THE CONVERGENCE OF DAMPED NEWTON METHODS FOR THE IMPLICIT SIMULATION OF TRANSPORT

by

Shahriyar Alkhasli

A THESIS

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Thesis Committee

_____, Advisor
Rami M. Younis

_____
Randy Hazlett

_____
Richard Redner

_____
Mohammad Shahvali

ABSTRACT

Shahriyar Alkhasli (Master of Science in Petroleum engineering)

On sufficient conditions for the convergence of damped Newton methods for the implicit simulation of transport

Directed by Rami M. Younis

61 pp., Chapter 4: Discussion

(435 words)

Fluid flow in porous media is represented by nonlinear flow and transport equations, which create difficulties in obtaining a solution for the system of equations. Analysis for implicit treatment of the transport equation shows that Newton's method fails to converge in most of the cases with an S-shaped flux function for relatively large time step sizes. Thus, current time step size is chopped, and previous iterations are wasted. These circumstances increase total computational time for simulations, which serves as motivation for development of safeguarding approaches for solution techniques used in reservoir simulation problems. So far, the most popular method, which is found in commercial simulators, is the Eclipse Appleyard method, a heuristic nonlinear solver which is unpredictable and not robust. No classical proof of convergence is known for the recently-developed trust region based solver.

The new safeguarding approaches proposed in this thesis are based on the contraction mapping theorem, proven by Banach in 1922, provides sufficient conditions for guaranteed convergence of Newton's method. Standard Newton's method does not satisfies this condition all the time; hence, a damping based method was formulated. The method can be applied for any flux function shape, even for shapes more complicated than S-shaped. Robustness and accuracy are independent of initial guess values, and convergence is at least linear.

The first part of this work describes the so-called fixed damping based Newton's method, where a damping factor determines the contraction factor values for all possible saturations below unity. This value remains the same during all iterations, which leads to unnecessary chops in neighborhood close of the solution and causes a high number of nonlinear iterations.

Thus, a dynamic damping based method is developed, which has the same level of robustness and accuracy as fixed damping based method. Moreover, efficiency of the method is improved due to the strategy constructed, where chops are applied only for saturation values that are not in the contraction region. Dynamic damping based method has been compared to state-of-the-art nonlinear solvers. Our method demonstrates the following features: better robustness, the same accuracy, and problem-dependent efficiency. Empirical results indicate that: the new method can be faster, the same as, or slower than the other methods.

Numerical schemes for both fixed and dynamic methods have been generated for one dimensional, two cell cases, where each grid cell is treated separately, using the same strategy as for single cell problems. Performance and adaptive damping of the Newton update shows identical results with single cell problems for a variety of cases.

ACKNOWLEDGEMENTS

My journey at The University of Tulsa has been very interesting and challenging. During my course and research studies, my advisor, Dr. Rami M. Younis, has taught me to never give up and to keep moving forward. He became for me spiritual mentor, and I am very thankful for everything that he did for me.

I extend special thanks to FuRSST team for their valuable discussions and support for the past two years.

I would also like to thank Dr. Randy Hazlett for his guiding throughout my studies and my other committee members, Dr. Richard Redner and Dr. Mohammad Shahvali, for their valuable input and feedback.

Although my parents, Babak Alkhasov and Nargiz Alkhasova, could not be with me during these two years, I always felt their support and motivation which helped me to step forward. On the other hand, there were a lot of friends who stayed with me in tough and relaxed situations, Rahman Mustafayev, Soham Sheth, Giorgiy Lutidze, Walter Ernesto Poquioma Escorcia, Cintia Goncalves Machado, and others too numerous to list. Thank you very much for all moments we shared together, I hope to keep our friendship forever.

I also would like to thank the administrative staff, Loreta Watkins and Donna Trankley: without them this work would not have been possible.

And finally, I would like to thank The University of Tulsa's Board for believing in me, providing with the opportunity to be a teaching assistant, and funding my study.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Reservoir simulations numerically model the dynamics of petroleum fluid flow within the subsurface given a prior characterization of the highly uncertain physical system. Simulation is often the primary workhorse behind data assimilation, uncertainty quantification, and optimization components of modern reservoir engineering workflows [11]. In light of their practical role, industry-grade simulators are required to be:

- Flexible and scalable enough to accommodate sufficient physical complexity and detail,
- Computationally efficient enough to expedite practical decision-making timelines, and,
- Robust enough to be relied upon within automated, iterative driving processes such as optimization.

Multiphase, multicomponent flows through subsurface porous media couple several physical phenomena with vastly differing characteristic scales. The fastest processes, such as phase equilibria, occur almost instantaneously and are modeled by nonlinear algebraic constraints. Mass conservation laws govern the transport of chemical species that propagate through an underlying flow velocity field [21]. In the limit of low capillary numbers, the transport equations are purely advective, and they give rise to an evolution with a finite domain-of-dependence. In the other limit, the transport problem is advective-diffusive. Moreover, the underlying flow velocity field is itself also transient, and it evolves with parabolic character. In the limit of no total compressibility, the flow field reaches instantaneous equilibrium, and is governed by an elliptic equation. Constitutive relations such as the multiphase extension to Darcy's law couple the variables across governing equations in a strongly nonlinear manner [23]. Additional sources of complexity include the heterogeneity of the underlying

porous media, body forces, and the presence of wells that are typically operated using busy, and often discrete, control schedules. This leads to a coupled system of governing Partial Differential Algebraic Equations (PDAEs) with mixed parabolic-hyperbolic character and stiff nonlinear algebraic constraints.

In reservoir simulation, the governing equations are approximated using spatial and temporal discretization. In space, a rich collection of discretization alternatives may be applied, including finite-volume and finite-element methods (see for example, [4, 6, 12]). In time, low-order implicit methods are a main staple of reservoir simulation. In implicit methods, all variables are treated at the new time-level leading to unconditional stability with respect to time-step size in the sense of discrete approximations. This robustness comes at the computational expense of requiring that a large, coupled, nonlinear system of algebraic equations be solved at each time-step; an often cited limitation of implicit methods, particularly on massively parallel architectures (e.g. [15, 18]).

Modern simulators rely on a fixed-point iteration, such as a variant of Newton's method in order to solve these problems (see for example [29, 4, 26, 8]). For general problems, Newton's method is not guaranteed to converge, and it is known to be sensitive to the initial guess that must be supplied. In [29], the nonlinear system is viewed as a homotopy with time-step size as the natural parameter. A numerical continuation was proposed to solve the system for a target time step size. In most reservoir simulators, however, the initial guess to the iteration is the old state, amounting to a zero-order continuation prediction step. For small time-step sizes, this may be a good approximation to the new state, and is, therefore, thought to be a good starting point for the Newton iteration. For larger time-steps, however, this is less likely to be the case, and the iteration may converge too slowly, or even diverge. Given a target time-step size, the potential for slow convergence or divergence is difficult to predict. Generally, it may depend on the local speed of transients as well as on the structure of the underlying nonlinearity of the specific problem at hand. This uncertainty in the efficacy of the nonlinear solver leads modern simulator developers to adopt the computationally wasteful try-adapt-try-again strategy for time-stepping. A

time-step size is chosen based on local accuracy considerations, and a solution attempt is made. If the solution attempt is deemed to be a failure, heuristics are then applied to cut the time-step and to try again.

Globalization or safeguarding strategies are universally applied in order to improve the robustness of Newton-like methods for varying time-step size. There are several classic approaches to safe-guarding Newton's method, such as the line-search and trust-region algorithms described in [26] and [8], for example. In these methods, the Newton direction is scaled by a constant factor in order to locally and approximately minimize the residual norm along the Newton direction or to maintain iterates within a region of trust. While substantial theory and empirical evidence support the effectiveness of such globalization strategies, modern reservoir simulators seldom rely on them. In commercial reservoir simulators, heuristic strategies are devised in order to safeguard Newton's method. Such strategies aim to reduce the computational cost of backtracking or trust region adaptation while retaining a similar globalization effect. In particular, since the evaluation of the residual itself can be computationally demanding, backtracking and other iterative step length selection algorithms can become computationally prohibitive. By specializing to the precise underlying functional form of the nonlinear residual for the problem at hand, engineers have sought strategies to select step lengths at a constant computational cost. Empirically, these safeguarding strategies have proven their worth [29, 16, 27], providing improved robustness for a wide range of subproblems at negligible computational cost. Serious shortcomings of reported heuristic safeguarding strategies to date are twofold:

1. They require strict assumptions about the precise functional form of the constitutive models. Outside of this range of applicability, these methods may fail in practice. Generalizing these methods to growing complexity requires new analysis.

2. They are not proven to guarantee convergence for general multidimensional problems even within the class of special physics that they are designed for.

One goal of this thesis is to devise a general safeguarding strategy for reservoir simulation that requires no more than the ability to evaluate an element of the residual and its scalar first and second derivatives. Moreover, this work will show that the safeguarding strategy is sufficient to guarantee at least linear convergence. While we limit our attention to immiscible incompressible two phase flow, we do not assume any specific functional form for any constitutive relation.

In this chapter, the canonical form of the equations is presented, followed by a review of the state-of-the-art in heuristic safeguarding strategies, and finally, a formal statement of the objectives.

1.1 Immiscible, incompressible two-phase flow

1.1.1 The governing equations

We consider the flow of two phases (a wetting phase w, and a non-wetting phase n), where the independent state variables are the two phase saturations, S_α, $\alpha \in \{w, n\}$. The phase saturations are constrained by,

$$S_w + S_n = 1 \tag{1.1.1}$$

and

$$S_\alpha \in [0, 1]\,, \tag{1.1.2}$$

leading to one independent variable, $0 \leq S_w \leq 1$. Assuming compatible auxiliary conditions, the conservation equation for incompressible immiscible flow is a scalar nonlinear conservation law,

$$\begin{cases} \phi \frac{\partial S_w}{\partial t} + \nabla \cdot u_w = 0 & x \in \Omega, t \geq 0 \\ S_w\left(x, 0\right) = S_0\left(x\right) & x \in \Omega \\ S_w\left(x, t\right) = S_{inj}\left(x\right) & x \in \partial\Omega, t \geq 0, \end{cases} \tag{1.1.3}$$

where ϕ represents porosity, and u_w is the phase velocity that is nonlinearly dependent on

saturation and the underlying flow field. The phase velocity is represented by Muskat's extension to Darcy's law,

$$u_w = -\frac{kk_{rw}}{\mu_w}\left(\nabla p_w + \rho_w g \nabla h\right), \tag{1.1.4}$$

where k is the heterogeneous permeability field, k_{rw} is the nonlinear relative permeability, μ_w is viscosity, p_w is the phase pressure, and ρ_w is the phase density. In one dimension, equation 1.1.4 becomes

$$u_w = -\frac{kk_{rw}}{\mu_w}\left(\frac{\partial p_w}{\partial x} + \rho_w g\ \sin\alpha\right), \tag{1.1.5}$$

where α is the dip angle with respect to gravity. The assumption of incompressible flow leads to the requirement that the total velocity field $u_t = u_w + u_n$ be divergence-free. This homogeneous total velocity field can therefore be used to eliminate the pressure dependence within the conservation equation. In particular, with a fixed total velocity field, the fractional flow of water is

$$f_w = \frac{u_w}{u_t} \tag{1.1.6a}$$

$$= \frac{1 + \frac{kk_{ro}}{u_t\mu_o}\left(\frac{\partial p_c}{\partial x} - g\Delta\rho\ \sin\alpha\right)}{1 + \frac{\mu_w}{\mu_o}\frac{k_{ro}}{k_{rw}}}, \tag{1.1.6b}$$

where $p_c = p_n - p_w$ is the capillary pressure. In equations 1.1.6, the relative permeability and capillary pressures are generally nonlinear functions of saturation. This leads to the hyperbolic conservation law form with a general nonlinear flux function,

$$\begin{cases} \frac{\partial}{\partial t}S_w + u_t\frac{\partial}{\partial x}f\left(S_w\right) = 0 & x \in [0,1], t \geq 0 \\ S_w\left(x,0\right) = S_0\left(x\right) & x \in [0,1] \\ S_w\left(0,t\right) = S_{inj}\left(x\right) & t \geq 0. \end{cases} \tag{1.1.7}$$

The nonlinear flux function.

In the limiting case of horizontal displacements and neglecting capillary forces, the

fractional flow curve is simply,

$$f_w = \frac{1}{1 + \frac{\mu_w}{\mu_o}\frac{k_{ro}}{k_{rw}}}. \tag{1.1.8}$$

Additional nonlinearity in this form of the flux function arises due to the functional dependence of relative permeability on the water saturation. One of the widely used models for relative permeability is the Corey relation [7],

$$k_{rw} = k_{rw0}\left(\frac{S - S_{wr}}{1 - S_{wr} - S_{or}}\right)^{n_w}, \text{ and,} \tag{1.1.9a}$$

$$k_{ro} = k_{ro0}\left(\frac{1 - S - S_{or}}{1 - S_{wr} - S_{or}}\right)^{n_o}, \tag{1.1.9b}$$

where n_w and n_o are empirical exponents, S_{wr} and S_{or} are the saturation end-points, and k_{rw0} and k_{ro0} are the end-point permeabilities. The choice of quadratic exponents, $n_w = n_o = 2$ is widely used. Figure 1.1 presents model relative permeability curves using the Corey model.

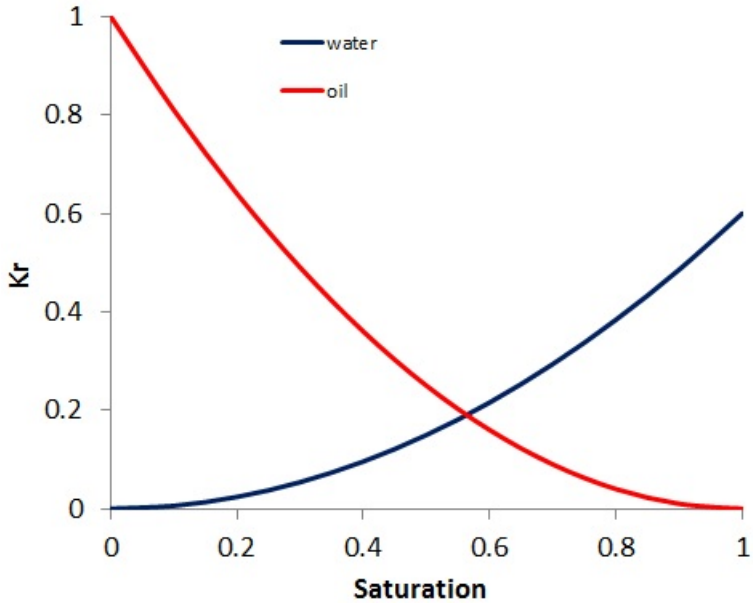

Figure 1.1: Corey's model example: $n_w = n_o = 2, k_{rw0} = 0.6, k_{ro0} = 1, S_{or} = S_{wr} = 0$.

There are other relative permeability models such as the Brooks-Corey model [5],

$$k_{rw} = S_w^{3+2/\lambda}, \text{ and,}$$

$$k_{ro} = (1 - S_w)^2\left(1 - S_w^{1+2/\lambda}\right),$$

where λ is an indicator of pore size distribution. Generally, relative permeability curves may be modeled using curve fitting through experimental data, and the functional dependence may have rich nonlinear character.

Substitution of the quadratic Corey relative permeability into the fractional flow function leads to the so-called *S-shaped* flux function,

$$f(S_w) = \frac{S_w^2}{S_w^2 + M\left(1 - S_w\right)^2}, \tag{1.1.11}$$

where $M = \frac{\mu_w}{\mu_o}$ is the viscosity ratio. Figure 1.2 illustrates the effects of various viscosity ratios on the S-shaped fractional flow curve in the absence of capillary and buoyancy forces, where irreducible saturations are equal to zero.

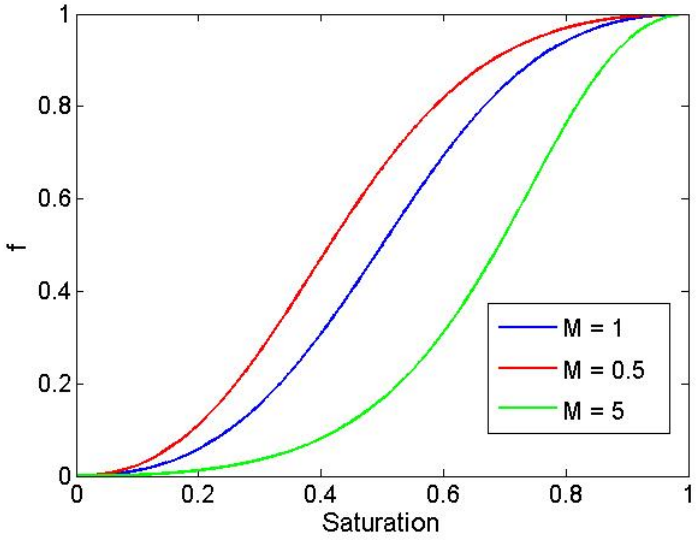

Figure 1.2: S-shaped flux function without capillary and buoyancy forces: M = 0.5, 1, 5

1.1.2 The discrete approximation

We apply a backward Euler discretization in time, and a first-order upwind discretization in space. On a uniform mesh with N cells, the numerical is saturation unknown at the n^{th} timestep and the i^{th} cell is denoted as S_i^n. The numerical scheme is written as

$$S_i^{n+1} - S_i^n + \frac{\Delta t}{\Delta x}\left[F\left(S_i^{n+1}, S_{i+1}^{n+1}\right) - F\left(S_{i-1}^{n+1}, S_i^{n+1}\right)\right] = 0,\ i = 1, \ldots, N \tag{1.1.12}$$

In Equation 1.1.12, the timestep size is denoted as Δt, the mesh spacing as $\Delta x = \frac{1}{N}$, and the numerical flux as F. Neglecting gravitational effects and assuming a positive total velocity, we apply a Dirichlet boundary condition on the left edge of the domain, and a second order treatment of a free boundary condition on the right. These are numerically prescribed as the conditions,

$$S_0^n = S_{inj}, \text{ and,} \tag{1.1.13}$$

$$S_{N+1}^n = 2S_N^n - S_{N-1}^n. \tag{1.1.14}$$

For general fractional flow functions it is necessary to apply an entropy satisfying upwind condition for the numerical flux. This condition corresponds to the analytical solution of cell-face Riemann problems. The condition applied is described by Equation 1.1.15 below.

$$F(a,b) = \begin{cases} \min\limits_{a \le s \le b} f_w(s) & a \le b \\ \max\limits_{b \le s \le a} f_w(s) & \text{otherwise.} \end{cases} \tag{1.1.15}$$

Note that for fractional flows with sonic points, the numerical flux at a cell interface may be independent of both the left and right cell saturations when it is evaluated at the sonic point. On the other hand, monotonically increasing or decreasing fractional flows lead to the simplified left- and right-looking upwind schemes respectively.

1.1.3 The nonlinear implicit time stepping problem

Equations 1.1.12 along with the auxiliary conditions 1.1.13 form a coupled system of nonlinear algebraic equations for the new saturation approximation, S_i^{n+1}, $i = 1, \ldots, N$, over the target time step, Δt, given the initial state, S^n. This system of equations is referred to

as the nonlinear residual system, and it has the general form,

$$R\left(S^{n+1}; S^n, \Delta t\right) = 0.$$

An implicit time-step requires the solution of this nonlinear residual system. A fixed-point iteration is typically applied [1], and using the iteration index ν, it is written as,

$$\left[S^{n+1}\right]^{\nu+1} = G\left(\left[S^{n+1}\right]^{\nu}; S^n, \Delta t\right), \qquad \nu = 0, 1, 2... \tag{1.1.16}$$

where G is a nonlinear operator. If this operator is,

$$G\left(\left[S^{n+1}\right]^{\nu}; S^n, \Delta t\right) = \left[S^{n+1}\right]^{\nu} - R'\left(\left[S^{n+1}\right]^{\nu}; S^n, \Delta t\right)^{-1} R\left(\left[S^{n+1}\right]^{\nu}; S^n, \Delta t\right),$$

where R' is the Jacobian matrix of the residual, then the fixed-point iteration is Newton's method. In order to start this iterative process, an initial guess is required. This is typically taken as the initial state, i.e.,

$$\left[S^{n+1}\right]^{0} = S^n.$$

In general, the time stepping solutions can exhibit complex structure including approximations to weak solutions with local discontinuity and counter-current traveling waves as depicted in Figure 1.3.

1.1.4 Mathematical model problems under consideration

In this thesis, the focus is on two model problems of the general form in Equation 1.1.12. The first problem is a single cell problem ($N = 1$), and it is selected to provide a scalar residual that incorporates the nonlinearity of the flux function. The second problem is one-dimensional.

In both problems, we neglect the effects of capillarity and gravity. Subsequently, the continuous initial boundary value problem is hyperbolic, and the nonlinear flux function

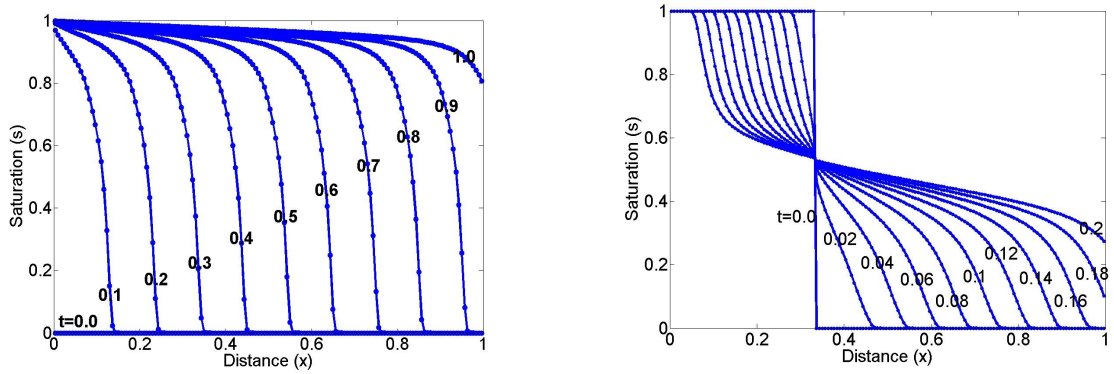

Figure 1.3: Saturation solution profiles for a 1-dimensional with (a) cocurrent flow, and (b) counter-current gravity effects.

is monotonic and contains no sonic points. Mathematically, we will restrict our attention to certain flux functions, $f : \mathbb{R} \to \mathbb{R}$, with assumptions on their restriction to the set $I = [0, 1] \subset \mathbb{R}$, denoted $f|_I$. These assumptions lead to Definition 1.1.1 that introduces admissible flux functions.

Definition 1.1.1. A function $f : \mathbb{R} \to \mathbb{R}$ is called an *admissible flux function* provided that:

- it is continuously differentiable; $f \in C^1(\mathbb{R})$,
- the restriction is an injection; $f|_I : I \to I$,
- the first derivative is nonnegative and bounded above; $f'|_I : I \to [0, f'_{max}]$, where $f'_{max} < \infty$, and,
- the second derivative is bounded; $f''|_I : I \to [f''_{min}, f''_{max}]$, where $f''_{min} > -\infty$ and $f''_{max} < +\infty$.

Model problem 1.

A limiting case is the single-cell problem that is often studied in order to motivate heuristic reservoir simulation safeguarding methods. This model problem leads to a scalar nonlinear residual that must be solved at each time-step. The independent variable, $S \in I = [0, 1]$, is the saturation in the cell after a time step, $\Delta t \geq 0$. The initial saturation before the time

step is fixed, $S_{init} \in I$, and a constant injection saturation, $S_{inj} \in I$, is applied. We assume an admissible flux function holds. The nonlinear residual problem is,

$$R\left(S; S_{init}, \Delta t\right) = S - S_{init} + \Delta t \left[f\left(S\right) - f\left(S_{inj}\right)\right] = 0. \tag{1.1.17}$$

Since the flux function is assumed to be admissible, the derivative of the residual is well-defined,

$$R'\left(S; S_{init}, \Delta t\right) = 1 + \Delta t f'\left(S\right), \tag{1.1.18}$$

and is strictly positive. This implies that the Newton direction is well-defined, and the iteration is,

$$S^{\nu+1} = G\left(S^{\nu}\right) = S^{\nu} - \frac{R\left(S^{\nu}; S_{init}, \Delta t\right)}{R'\left(S^{\nu}; S_{init}, \Delta t\right)}. \tag{1.1.19}$$

Remark. Since the derivative of the residual is nonzero for admissible flux functions, an iterate of the fixed point iteration is a stationary point if and only if it is a solution of the residual equation.

Even in the simplest of cases such as problems with quadratic relative permeability, the exact standard Newton's method may diverge. The residual curve and iterates for such a case are illustrated in Figure 1.4.

Model problem 2.

The one-dimensional problem models the propagation of saturation waves under a constant total velocity field. With an admissible flux function, cocurrent flow takes place, and rarefaction and shock waves travel from left to right. At each time-step, the nonlinear system of discrete residual equations must be solved in order to obtain the independent variable $S \in I^N = [0,1]^N$. Given a fixed injection saturation, $S_{inj} \in I$, and an initial state, $S_{init} \in I^N$,

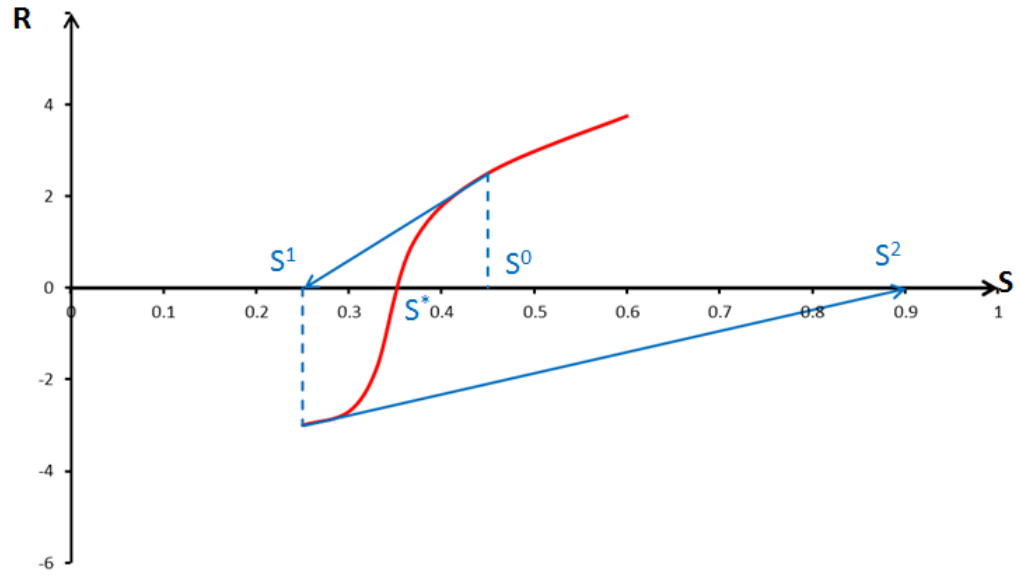

Figure 1.4: An example where Newton's method diverges for a non-monotonic residual.

the residual system is defined as,

$$\begin{aligned} R_1\left(S; S_{init}, \Delta t\right) &= S_1 - S_{1,init} + \frac{\Delta t}{\Delta x}\left[f\left(S_1\right) - f\left(S_{inj}\right)\right], \\ R_i\left(S; S_{init}, \Delta t\right) &= S_i - S_{i,init} + \frac{\Delta t}{\Delta x}\left[f\left(S_i\right) - f\left(S_{i-1}\right)\right], \quad i = 2, \ldots, N. \end{aligned} \tag{1.1.20}$$

Clearly, the system can be solved sequentially by zero-finding a sequence of scalar problems in the order $R_i = 0, i = 1, \ldots, N$. Alternately, a Newton method may be applied to this nonlinear lower triangular system. The Jacobian matrix is well-defined for admissible flux functions, and it is lower bidiagonal,

$$R_{ij}{}' = \begin{cases} 1 + c_i & i = j \\ -c_j & j = i - 1, i = 2, \ldots, N \\ 0 & \text{otherwise,} \end{cases} \tag{1.1.21}$$

where the local wave speeds are defined as,

$$c_i = \frac{\Delta t}{\Delta x} f'\left(S_i\right) \tag{1.1.22}$$

Since the diagonal entries of the Jacobian matrix are nonzero for admissible flux functions, the Jacobian is invertible for any $S \in \mathbb{R}^N$, and the Newton update, $\delta = {R'}^{-1}R$, is well defined,

$$\begin{aligned} \delta_1 \left(S; S_{init}, \Delta t\right) &= \frac{R_1}{1 + c_1}, \\ \delta_i \left(S; S_{init}, \Delta t\right) &= \frac{R_i + c_{i-1}\delta_{i-1}}{1 + c_i}, \quad i = 2, \ldots, N. \end{aligned} \tag{1.1.23}$$

Subsequently the Newton iteration is,

$$S^{\nu+1} = G\left(S^{\nu}\right) = S^{\nu} - \delta\left(S^{\nu}; S_{init}, \Delta t\right). \tag{1.1.24}$$

Remark. Since the Jacobian matrix is invertible, an iterate of the fixed point iteration is a stationary point if and only if it is a solution of the residual system.

1.2 State-of-the-art in reservoir simulation safeguarding

Given the possibility of divergence of Newton's method for general problems, globalization or safeguarding strategies are important in order to avoid time-step cuts. Safeguarding strategies can be viewed as different ways to specify a diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}\left(\lambda_{\mathbf{1}}, \ldots, \lambda_{\mathbf{N}}\right)$ in a safeguarded Newton iteration,

$$S_{\nu+1} = S_{\nu} - \mathbf{\Lambda}_{\nu} R'\left(S_{\nu}\right)^{-1} R\left(S_{\nu}\right). \tag{1.2.1}$$

The standard Newton's method selects all diagonal weights to be one, and subsequently, in cases where the underlying Newton flow changes too rapidly, the iteration may not converge.

There are several classic approaches to safe-guarding Newton's method, such as the line-search and trust-region algorithms described in [26] and [8], for example. In line-search methods, all entries of the diagonal are identical (i.e. $\mathbf{\Lambda}_{\nu} = \lambda_{\nu} I$), implying that the Newton direction is simply scaled by a constant damping factor, λ. In these methods, the choice of this damping factor is dictated by the rate of change in the residual norm along the Newton direction or within a neighborhood about the current iterate. Algorithms for the choice of

λ include backtracking and trust-region adaptation which require the residual vector to be evaluated possibly several times. In reservoir simulation, the residual evaluation can involve computationally demanding thermodynamic flash calculations, making the computational cost of evaluating the residual non-negligible.

In commercial reservoir simulators, heuristic strategies have been devised in order to safeguard Newton's method. Such strategies select the diagonal scaling entries on a cell-by-cell basis, using physical arguments. In practice, these heuristics are known to improve convergence dramatically (larger time step sizes converge from the same initial state at a faster rate than classical methods). There is a common underlying hypothesis shared by the most commonly employed heuristics. The hypothesis may be motivated by the shape of common fractional flow curves (e.g. Figure 1.2). The hypothesis is that the nonlinearity of the residual is in some sense dictated by the structure of the flux function in each cell. That is, suppose that a nonlinear Gauss-Seidel iteration is applied to the residual system as a whole. Then for each Gauss-Seidel iteration, the residual in each cell must be solved sequentially. To solve each of these single cell nonlinear residuals, we may apply a scalar Newton process, obtaining the saturation in the current single cell, assuming that all other cell saturations are known. These scalar Newton iterations need to be safe-guarded by some scalar damping protocol ([19]). If this protocol assures cell-wise convergence of a Newton iteration, then the outer Gauss-Seidel iteration at least has a hope of converging. The hypothesis is that if the same **scalar** damping protocols are also applied on saturation variables, cell-by-cell, in the context of a **system** Newton updates, then that too is more likely to converge. Inspecting Figure 1.2, we can deduce that saturations around end-points and inflection points produce particularly sensitive Newton directions. The hypothesis is to apply a cell-by-cell (diagonal) damping factors to limit large changes around such physically sensitive boundaries.

The following is a review of heuristic local damping protocols for Newton methods in reservoir simulation:

1. The Modified Eclipse Appleyard (MEA) safeguarding heuristic limits the Newton update in every grid cell to a maximum magnitude of 0.2 ([29, 13, 24]). This ad hoc choice

of λ_i^ν is the result of extensive empirical tuning that was conducted by the developers of the *Eclipse* reservoir simulator [13]. It was observed that since saturation can vary between 0 and 1 that limiting any Newton update to 0.2 yields improved convergence for a vast body of simulations conducted using the popular commercial code. While effective for the typical case, it is easy to find examples where the heuristic fails. Moreover, for growing complexity, such as compositional models, the heuristic is not useful. On the other hand, the computational cost of calculating the local damping factors is negligible.

2. An improved heuristic was proposed by specializing to problems with an S-shaped flux function [16]. The authors propose a cell-based choice of λ_i^ν such that Newton updates that cross the inflection point of the flux function are damped to a small neighborhood about the inflection point. Through empirical testing, this strategy proved to perform very well for problems with no gravity or capillarity effects. The strategy was empirically shown to be more robust than the MEA strategy, but it is specific to the S-shaped flux function and requires knowledge of the inflection point location.

3. More recently, a trust-region based strategy was proposed to account for buoyancy and capillary forces [27]. The proposed scheme accommodates sonic points and two inflection points as in Figure 1.5. It is a direct extension of [16] to flux functions with two inflection points and one sonic point (maximizer or minimizer).

4. The trust region idea was later extended to analyze jumps across inflection manifolds in the discrete flux function (Equation 1.1.15). The discrete flux function incorporates the upwinding condition to evaluate the fractional flow at either saturation on either side of a grid cell face. While this heuristic improves robustness and efficiency even further, it requires a multidimensional analysis of the structure of the discrete flux

function in a pre-processing step. It is not possible to automate this to general flux functions with arbitrary structure.

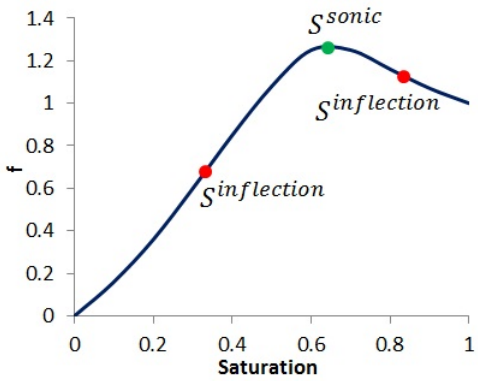

Figure 1.5: Fractional flow curve with one sonic point and two inflection points.

A common denominator amongst these heuristic local damping safeguards is that they alter the update direction as well as its magnitude. They do so by scaling the update element by element such that some local measure of fitness is satisfied. When they work, empirical evidence shows that they are extremely efficient. On the other hand, they are crafted for very specific flux functions with limited complexity. Another observation is that while some safeguards may be proven theoretically to guarantee convergence of Newton's method for the scalar subproblem, there is no theoretical proof that they should improve convergence on Newton's method for the system and counter-examples are abundant.

In practice, the flux function may be any general nonlinear function and it may have considerable structure (e.g. Figure 1.6). There is value in simple safeguarding strategies that rely on less information about the flux function and that simultaneously can be shown to guarantee certain convergence properties.

1.3 Objectives and outline

The objectives of this thesis are to devise simple damping strategies that will guarantee at least theoretical linear convergence. The damping strategies will not assume any specific functional form for the flux function beyond admissibility. The key idea is to analyze the Fréchet derivative of the fixed point mapping and to select a damping factor such that

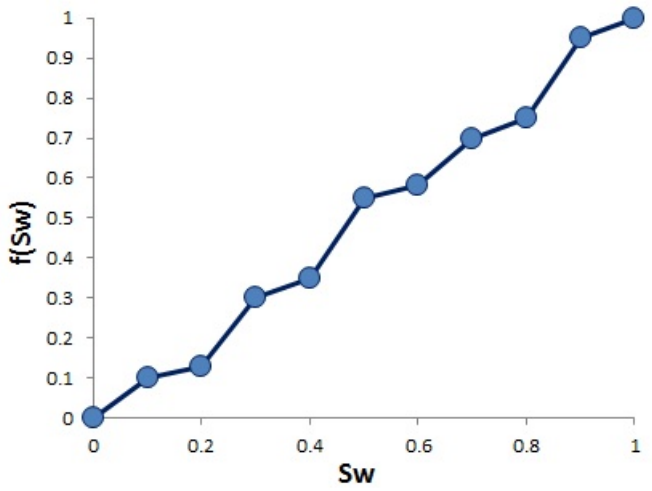

Figure 1.6: Possible fractional flow curve with nine inflection points.

the mapping is contractive and into. The calculation of the damping factors will require negligible constant computational cost.

In Chapter 2, a fixed damping strategy is devised in order to guarantee at least linear convergence for general problems. The damping is fixed because a constant $\mathbf{\Lambda}$ is predetermined and is used for all subsequent iterations. This is generalized to systems of equations in the infinity induced operator norm. Chapter 3 approaches the development of a dynamic damping strategy that weakens the conditions developed in Chapter 2. The result is a less conservative safeguard that varies with iteration number, $\mathbf{\Lambda}^{\nu}$. Finally, Chapter 4 discusses the potential for future work and implications of this thesis on modern implicit reservoir simulation.

CHAPTER 2

FIXED DAMPING STRATEGY

The residual equations for a particular time step (see for example, Equations 1.1.17 and 1.1.20) are the discrete approximation to an underlying Initial Boundary Value Problem (IBVP). An assumption of well-posedness of the underlying IBVP guarantees the existence and uniqueness of solutions to the IBVP. The implicit discretization is an accurate approximation in space and time, and this fact can be used to show the existence of a solution to the discrete residual equation in the asymptotic sense of a mesh refinement path (see for example, [28, 10]). Uniqueness of the discrete solution is not generally guaranteed, however. In [20], both existence and uniqueness are demonstrated specifically for problems with phase-based upwinding and monotone flux functions. For general flux functions, it is easy to construct problems where the discrete residual has multiple solutions in $\mathbb{R}^N$, and in this work, we consider any such solution as admissible since the corresponding discrete approximation is mass conservative. We will assume existence of solutions to both model problems under consideration, and a secondary outcome of this chapter is a demonstration that the solution is unique within the hypercube $I = [0,1]^N$.

The focus of this chapter is to devise a fixed damping strategy for Newton's method. In particular, a single diagonal damping factor, $\lambda\left(S_{init}, \Delta t\right)$, is to be determined for each time step such that the iteration,

$$S^1 = S_{init}, \text{ and,} \tag{2.0.1}$$

$$S^{\nu+1} = S^\nu - \lambda\left(S_{init}, \Delta t\right) R'\left(S^\nu; S_{init}, \Delta t\right)^{-1} R\left(S^\nu; S_{init}, \Delta t\right), \nu = 1, 2, \ldots \tag{2.0.2}$$

converges to a solution, S^*. Moreover, the convergence rate will be at least linear. Since

$\lambda(S^0, \Delta t)$ is fixed and does not vary with ν, we refer to this as a fixed damping strategy. The basic idea is to select λ such that the conditions of a Contraction Mapping Theorem are met. We proceed by reviewing established global results on the convergence of fixed point iterations, and then, these results are applied to model problems 1 and 2.

2.1 Preliminary results

We consider the general fixed point iteration,

$$S^{\nu+1} = G(S^{\nu}), \quad \nu = 1, 2, \ldots, \tag{2.1.1}$$

where the map $G : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$ is a general generating function.

Theorem 2.1.1 (Contraction-Mapping Theorem [17, 25])**.** *Suppose that $G : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$ maps a closed set $D_0 \subset D$ into itself, and that there exists an $0 < \alpha < 1$ such that*

$$\|G(x) - G(y)\| < \alpha \|x - y\|, \quad \forall x, y \in D_0.$$

Then for any $S^0 \in D_0$, the sequence 2.1.1 converges to the unique fixed point S^* of G in D_0, and the following error estimate holds,

$$\|S^* - S^{\nu}\| \leq \frac{\alpha^{\nu}}{1-\alpha} \left\|S^1 - S^0\right\|, \quad \nu = 1, 2, \ldots.$$

Proof. Suppose that $x^*, y^* \in D_0$ are any two fixed points, then,

$$\|x^* - y^*\| \leq \|G(x^*) - G(y^*)\| \leq \alpha \|x^* - y^*\|, \tag{2.1.2}$$

implying that $x^* = y^*$ since, by assumption, $0 < \alpha < 1$. Next, consider any $x^0 \in D_0$. The subsequent sequence $\{x^{\nu}\}$ of iteration 2.1.1 is well-defined and is contained in D_0 since by assumption, $G(D_0) \subseteq D_0$. We will show that any such sequence is a Cauchy sequence, and by the completeness of D_0 which is a closed subset of $\mathbb{R}^N$, the sequence must converge to a

point within D_0. Since $x^{\nu+1} = G\left(x^{\nu}\right) = G^2 x^{\nu-1}$ for any $\nu \geq 1$, we have that,

$$\left\|x^{\nu+1} - x^{\nu}\right\| = \left\|G^{\nu}\left(x^1\right) - G^{\nu}\left(x^0\right)\right\| \leq \alpha^{\nu}\left\|x^1 - x^0\right\|, \tag{2.1.3}$$

and subsequently, using the triangle inequality, and for any $n \geq m \geq 0$, we have,

$$\begin{aligned}
\left\|x^n - x^m\right\| &= \left\|G^n\left(x^0\right) - G^m\left(x^0\right)\right\| \\
&\leq \alpha^m \left\|G^{n-m}\left(x^0\right) - x^0\right\| \\
&\leq \alpha^m \left(\left\|G^{n-m}\left(x^0\right) - G^{n-m-1}\left(x^0\right)\right\| + \ldots + \left\|G\left(x^0\right) - x^0\right\|\right) \\
&\leq \alpha^m \left(\sum_{\nu=0}^{n-m-1} \alpha^{\nu}\right) \left\|x^1 - x^0\right\| \\
&\leq \alpha^m \left(\sum_{\nu=0}^{\infty} \alpha^{\nu}\right) \left\|x^1 - x^0\right\|
\end{aligned}$$

Since $0 < \alpha < 1$, the power series converges, and we arrive at the inequality,

$$\left\|x^n - x^m\right\| \leq \frac{\alpha^m}{1-\alpha}\left\|x^1 - x^0\right\|, \quad n \geq m \geq 0, \tag{2.1.4}$$

which directly shows that the sequence is a Cauchy sequence, and that it converges to a point $x^* \in D_0$.

By continuity, x^* must also be a fixed point since,

$$G\left(x^*\right) = G\left(\lim_{\nu\to\infty} x^{\nu}\right) = \lim_{\nu\to\infty} G\left(x^{\nu}\right) = \lim_{\nu\to\infty} x^{\nu+1} = x^*.$$

Finally, the error estimate follows directly from Equation 2.1.4 with $n \to \infty$. □

Remark. The contraction constant α in Theorem 2.1.1 is clearly norm-dependent. A mapping may satisfy the assumptions of the theorem in one norm but may not in another.

A corollary to Theorem 2.1.1 provides a characterization for the number of iterations that are required in order to reduce the error between iterates and the fixed point to a certain tolerance. The result is a direct application of the *a priori* error estimate that is provided

by the contraction-mapping theorem.

Corollary 2.1.1.1. *Consider iteration 2.1.1, and suppose that G satisfies the assumptions of Theorem 2.1.1 with the unique fixed-point S^*. Fix a convergence tolerance $\epsilon > 0$. Then for any nonstationary starting point, S^0, such that $\|S^1 - S^0\| > \epsilon > 0$, the error satisfies,*

$$\|S^\nu - S^*\| \leq \epsilon,$$

provided that,

$$\nu \geq \left\lceil \frac{1}{\log(\alpha)} \left(\log\left(\frac{\epsilon}{\|S^1 - S^0\|} \right) + \log(1-\alpha) \right) \right\rceil.$$

Proof. By the error estimate provided by Theorem 2.1.1, the error at iteration $m > 1$ is,

$$\|e^m\| = \|S^m - S^*\| \leq \frac{\alpha^m}{1-\alpha} \left\| S^1 - S^0 \right\|,$$

and since $0 < \alpha < 1$, the error is strictly decreasing with iteration number; i.e., $\|e^{m+1}\| < \|e^m\|$. Let,

$$\begin{aligned} n &= \frac{1}{\log(\alpha)} \left(\log\left(\frac{\epsilon}{\|S^1 - S^0\|} \right) + \log(1-\alpha) \right) \\ &= \log_\alpha \left(\frac{\epsilon(1-\alpha)}{\|S^1 - S^0\|} \right) \end{aligned}$$

and $n* = \lceil n \rceil \geq n$. Then,

$$\|e^{n*}\| \leq \epsilon$$

and for each $\nu \geq n^*$, we have that $\|e^\nu\| \leq \epsilon$. □

The Taylor remainder and mean value theorems supply a characterization for the contraction constant in Theorem 2.1.1 in terms of the operator norm of the Fréchet derivative. This is a useful result to demonstrate that a fixed point mapping is a contraction.

Theorem 2.1.2 ([25])**.** *Suppose that $G : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$ is a continuously differentiable*

map, i.e. $G \in C^1(D)$*, and that* D *is a convex set. Then if in any p-norm,*

$$\sup_{x \in D} \|G'(x)\|_p = \alpha < 1,$$

holds, then for every $x, y \in D$*,*

$$\|G(x) - G(y)\|_p \leq \alpha \|x - y\|_p$$

Proof. Let $x, y \in D$, and consider the line $z = x + t(y - x)$, for $t \in [0, 1]$. By the convexity of D, we must have that $z \in D$, and subsequently, we can define the map, $H(t) = G(x + t(y - x))$ for $t \in [0, 1]$. Since G is differentiable, and using the chain rule, we have that,

$$H'(t) = G'(x + t(y - x))(y - x),$$

where G' is the Jacobian matrix. Moreover, by the fundamental theorem of calculus, we have,

$$G(y) - G(x) = H(1) - H(0) = \int_0^1 H'(t)\, dt = \int_0^1 G'(x + t(y - x))(y - x).$$

By the triangle inequality, we have that,

$$\begin{aligned}
\|G(y) - G(x)\|_p &= \left\| \int_0^1 G'(x + t(y - x))(y - x) dt \right\|_p \\
&\leq \int_0^1 \|G'(x + t(y - x))(y - x)\|_p dt \\
&\leq \int_0^1 \|G'(x + t(y - x))\|_p \, dt \, \|(y - x)\|_p \\
&\leq \alpha \|(y - x)\|_p
\end{aligned}$$

□

Remark. The result of Theorem 2.1.2 carries through to the scalar univariate case directly.

Scalar problems will require the supremum of the absolute value of the scalar derivative as the contraction constant.

We now consider the nonlinear algebraic problem

$$R(S) = 0, \tag{2.1.5}$$

where $R : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$ is continuously differentiable. Moreover, we assume that for $S \in D$ the Frèchet derivative $R'(S)$ is invertible. Given $S^0 \in D$, the modified (damped) Newton iteration is,

$$S^{\nu+1} = S^\nu - \Lambda\delta(S^\nu), \nu = 1, 2, \ldots, \tag{2.1.6}$$

where

$$\delta(S) = R'(S)^{-1} R(S),$$

is the Newton update direction, and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_N)$ is a fixed diagonal weighting. Iteration 2.1.6 can be regarded as a family of fixed point iterations parameterized by Λ,

$$S^{\nu+1} = G_\Lambda(S^\nu) = S^\nu - \Lambda\delta(S^\nu), \quad \nu = 1, 2, \ldots. \tag{2.1.7}$$

Next, we will apply the contraction-mapping theorem and its corollaries to the damped Newton iteration. We first provide a characterization of fixed points in terms of solutions to the nonlinear system.

Lemma 2.1.3. *Let $R : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$ be continuously differentiable with Frèchet derivative $R'(S)$ that is invertible for $S \in D$. Suppose that $G(D_0) \subseteq D_0$ and $D_0 \subseteq D$. Then iteration 2.1.7 is well-defined, and $S^* \in D_0$ is a stationary point of the iteration if and only if it is a solution to problem 2.1.5.*

Proof. Suppose that $S^* \in D$ is a solution to problem 2.1.5. Then,

$$G_\Lambda(S^*) = S^* - \Lambda R'(S^*)^{-1} R(S^*) = S^*,$$

implying that it is a fixed point. Next, suppose that $G_\Lambda(S^*) = S^*$, then,

$$0 = S^* - G_\Lambda(S^*) = \Lambda R'(S^*)^{-1} R(S^*).$$

Since Λ and $R'(S^*)$ are invertible, we must have that $R(S^*) = 0$. □

Theorem 2.1.4. *Let $D_0 \subset \mathbb{R}^N$ be closed and convex. Suppose that $R \in C^1(D_0)$ and that R' is invertible and continuously differentiable at each $S \in D_0$. If there is a $\Lambda = diag(\lambda_1, \ldots, \lambda_N)$ with $0 < \lambda_i < 1$ for $i = 1, \ldots, N$ such that for any $S^0 \in D_0$,*

$$S^0 - \Lambda R'\left(S^0\right)^{-1} R\left(S^0\right) = S^0 - \Lambda\delta\left(S^0\right) \in D_0,$$

and,

$$0 < \alpha = \sup_{S \in D_0} \|I - \Lambda\delta'(S)\|_p < 1,$$

then,

(a) There is a unique zero $S^* \in D_0$ such that $R(S^*) = 0$,

(b) The damped Newton iteration (2.1.6) converges in the p-norm to S^* from any starting point in D_0, and,

(c) The following a priori error estimate holds for $\nu > 1$ and any nonstationary $S^0 \in D_0$,

$$\|S^* - S^\nu\|_p \leq \frac{\alpha^\nu}{1-\alpha} \left\|S^1 - S^0\right\|_p.$$

Proof. By direct application of Theorem 2.1.2, the fixed point iteration is into and contractive. This leads to an application of Theorem 2.1.1. Finally, Lemma 2.1.3 provides the link between the fixed point and the solution of the algebraic system. □

The objective of this chapter will be to show that it is possible to select Λ such that the conditions of Theorem 2.1.4 are fulfilled. In particular, for model problems 1 and 2, we first define a closed and convex set in $\mathbb{R}^n$ that is physically meaningful for saturation. Then,

we will select, or show that it is possible to select, Λ such that the mapping is both into and contractive. There may be several choices for Λ that achieve this, and the implications on optimality and convergence speed are discussed next.

2.1.1 Optimal choice for Λ

Before we move on to the applications to model problems 1 and 2, one matter should be addressed. There may be many Λ that could satisfy the conditions of Theorem 2.1.4; is there however an optimal choice? In order to answer this question, we must first define the objective or measure of optimality.

One measure for example, may be to attempt to minimize the convergence rate ratio,

$$\frac{e^{\nu}}{e^{\nu-1}} = \frac{\|S^{\nu} - S^{*}\|}{\|S^{\nu-1} - S^{*}\|} \leq \alpha(\Lambda) < 1,$$

by minimizing the upper-bound contraction constant, α. While minimizing $\alpha(\Lambda)$ may be computationally tractable, the utility of this objective is rather limited in the sense that it is not sufficient to guarantee that the smallest convergence ratio will actually be attained.

Another measure may be to compare the errors e_1^{ν} and e_2^{ν} at the ν-th iterate corresponding to two contractive iterations each with Λ_1 and Λ_2 respectively, that are both initiated from the same starting point, S^0. Moreover, we can assume that both choices satisfy the conditions of Theorem 2.1.4 and that $\Lambda_1 \neq \Lambda_2$. If a sufficient condition is enforced to ensure that for arbitrary ν, $e_1^{\nu} \leq e_2^{\nu}$, then iteration 1 is more optimal. It is difficult to provide a means to enforce such a condition however without further assumptions on the particular two sequences of Newton updates, δ_1^k and δ_2^k for $k = 0, \ldots, \nu$. Suppose that $M = \max_{k=0,\ldots,\nu} \left\|\delta_{1,2}^k\right\|$ and $m = \min_{k=0,\ldots,\nu} \left\|\delta_{1,2}^k\right\|$. Then the reverse triangle inequality provides an error lower-bound;

$$
\begin{aligned}
e_2^\nu &\geq \left\| S_1^\nu - S_2^\nu \right\| - e_1^\nu \\
&\geq \left\| S_1^\nu - S^0 \right\| - \left\| S_2^\nu - S^0 \right\| - e_1^\nu \\
&\geq \nu \left(\left\| \Lambda_1 \right\| m - \left\| \Lambda_2 \right\| M \right) - e_1^\nu.
\end{aligned}
$$

But by the contractive properties, an upper bound on the error can be derived as,

$$
e_1^\nu \leq \frac{\alpha_1^\nu}{1-\alpha_1} \left\| \Lambda_1 \right\| M,
$$

and subsequently, if,

$$
\frac{\alpha_1^\nu}{1-\alpha_1} \frac{\left\| \Lambda_1 \right\| M}{\left\| \Lambda_1 \right\| m - \left\| \Lambda_2 \right\| M} \geq \frac{\nu}{2},
$$

then $e_1^\nu \leq e_2^\nu$. While such a condition is of very limited use in practice, it motivates that the better choice of Λ is the one that maximizes $\|\Lambda\|$ such that $\alpha < 1$.

2.2 Single cell problem

In Section 1.1.4, the notion of an admissible flux function, f, was introduced (Definition 1.1.1), and model problems 1 and 2 were described. Model problem 1 is based on the implicit discretization of the governing equation for incompressible two-phase flow into and out of a single block in the absence of gravity. The independent variable is saturation, S, which is physically meaningful in the interval $I = [0, 1]$. The task of solving a time-step for the saturation amounts to determining a zero of the scalar problem,

$$
R(S) = S - S_{init} + \Delta t \left[f(S) - f(S_{inj}) \right] = 0, \tag{2.2.1}
$$

where $S_{init}, S_{inj} \in I$, and $\Delta t \geq 0$ are fixed parameters. The objective is to design a fixed

damping iteration of the form,

$$S^{\nu} = G\left(S^{\nu-1}\right) = S^{\nu-1} - \lambda \frac{R\left(S^{\nu-1}\right)}{R'\left(S^{\nu-1}\right)},$$

where the damping factor is bounded as $0 < \lambda \leq 1$, and the iteration satisfies the assumptions of Theorem 2.1.4. The theorem requires essentially three conditions:

1. Differentiability of the residual, existence of the inverse of the derivative, and differentiability of the inverse of the derivative. We have addressed these issues by introducing Definition 1.1.1 for admissible flux functions and we discussed how it takes care of these matters in Chapter 1.

2. That the Newton iterates remain within a closed convex set. In the remainder of this section, the set of concern is $I = [0, 1]$, and we will demonstrate additional conditions in terms of the flux function itself so that the Newton map is into, and $G(I) \subseteq I$.

3. Finally, it is required that the map be a contraction. We will show that under certain conditions on the flux function, it is possible to select a damping factor λ such that this is the case.

2.2.1 Conditions for the into-property of the Newton map

We introduce Lemma 2.2.1 that is applicable to any scalar fixed point iteration in the set $I \subset \mathbb{R}$. Essentially, if the map is contractive and maps the boundary of I into itself, then the map is into I.

Lemma 2.2.1. *Suppose that $G : \mathbb{R} \to \mathbb{R}$ and that the restriction $G\|I$ is continuously differentiable. If,*

1. *For all $S \in I$, $|G'(S)| \leq \alpha \leq 1$*

2. *There exists a $b_1 \in \mathbb{R}$ such that $\alpha |b_1| + |G(b_1)| \leq 1 - \alpha$*

3. *There exists a $b_2 \in \mathbb{R}$ such that $\alpha |1 - b_2| + |G(b_2) - 1| \leq 1 - \alpha$*

then $G(I) \subseteq I$.

Proof. Let $x \in [0,1]$, then by the triangle inequality and assumptions 1 and 2 of the claim,

$$\begin{aligned}
|G(x)| &= |G(x) - 0| \\
&\leq |G(x) - G(b_1)| + |G(b_1) - 0| \\
&\leq \alpha |x - b_1| + |G(b_1)| \\
&\leq \alpha |x - 0| + \alpha |b_1 - 0| + |G(b_1)| \\
&\leq \alpha (1 + |b_1|) + |G(b_1)| \\
&\leq 1
\end{aligned}$$

On the other hand, by the triangle inequality and assumptions 1 and 3, we have,

$$\begin{aligned}
|G(x) - 1| &\leq |G(x) - G(b_2)| + |G(b_2) - 1| \\
&\leq \alpha |x - b_2| + |G(b_2) - 1| \\
&\leq \alpha |x - 1| + \alpha |b_2 - 1| + |G(b_2) - 1| \\
&\leq \alpha (1 + |b_2 - 1|) + |G(b_2) - 1| \\
&\leq 1
\end{aligned}$$

Since $|G(x)| \leq 1$ and $|G(x) - 1| \leq 1$, we must have that $0 \leq G(x) \leq 1$. □

Remark. The conditions of Lemma 2.2.1 may be specialized into conditions concerning the flux function itself as well as the contraction constant. Since the conditions do not necessarily require that they hold for points within I, the extension of the flux function from I to $\mathbb{R}$ may be designed to satisfy the conditions given α itself.

2.2.2 Conditions for the existence of a suitable damping factor

The contractive property is key to the design of an iteration that satisfies the conditions of Theorem 2.1.4. The objective for given $S_{init}, S_{inj} \in I$ and $\Delta t \geq 0$ is to select a

damping parameter $0 < \lambda \leq 1$ such that the iteration is contractive. Before addressing the choice of λ, the conditions for the existence of such a parameter must be studied. According to Theorem 2.1.2, and considering the convex domain I, it is sufficient to show that λ can be chosen such that,

$$\sup_{S \in I} |G'(S)| = \alpha < 1$$

In the particular case of model problem 1, the admissibility of the flux function assures the differentiability of the iteration mapping. Subsequently, we can write,

$$\begin{aligned} G'(S) &= 1 - \lambda \delta'(S) \\ &= 1 - \lambda \left[1 - \frac{R(S) R''(S)}{(R'(S))^2} \right] \\ &= 1 - \lambda \left[1 - \frac{\Delta t (S - S_{init} + \Delta t (f(S) - f(S_{inj}))) f''(S)}{(1 + \Delta t f'(S))^2} \right] \end{aligned}$$

The objective then is to find the largest $\lambda > 0$ such that

$$\sup_{S \in I} |1 - \lambda [1 - \gamma(S)]| = \alpha < 1,$$

where,

$$\gamma(S) = \frac{R(S) R''(S)}{(R'(S))^2}.$$

The particular form of the derivative of the iteration mapping makes it amenable to further analysis. The following results lead to a characterisation for the existence and choice of λ for model problem 1.

Lemma 2.2.2. *Let $\lambda > 0$ and suppose that $\gamma : I \to \mathbb{R}$ is continuous and bounded above and below as,*

$$\gamma_{max} = \sup_{S \in I} \gamma(S),$$

and,

$$\gamma_{min} = \inf_{S \in I} \gamma(S).$$

Then,

$$\sup_{S\in I}|1-\lambda\left[1-\gamma\left(S\right)\right]|=\max\left\{\left|1-\lambda\left[1-\gamma_{max}\right]\right|,\left|1-\lambda\left[1-\gamma_{min}\right]\right|\right\}.$$

Proof. By the assumptions of the lemma, $\gamma : I \to J = [\gamma_{min}, \gamma_{max}]$, and subsequently,

$$\sup_{S\in I}|1-\lambda\left[1-\gamma\left(S\right)\right]|=\sup_{\beta\in J}\left|1-\lambda\left[1-\beta\right]\right|.$$

By definition,

$$|1-\lambda\left(1-\beta\right)|=\begin{cases}-1+\lambda\left(1-\beta\right) & \beta<1-\frac{1}{\lambda}\\ \\ 1-\lambda\left(1-\beta\right) & \beta\geq1-\frac{1}{\lambda}\end{cases},$$

First, suppose that $1-\frac{1}{\lambda}<\gamma_{min}$, then clearly,

$$\begin{aligned}\sup_{\beta\in J}|1-\lambda\left(1-\beta\right)|&=\sup_{\beta\in J}1-\lambda\left(1-\beta\right)\\&=1-\lambda\left(1-\gamma_{max}\right)\end{aligned}$$

Next, suppose that $\gamma_{max}>1-\frac{1}{\lambda}$, then clearly,

$$\begin{aligned}\sup_{\beta\in J}|1-\lambda\left(1-\beta\right)|&=\sup_{\beta\in J}-1+\lambda\left(1-\beta\right)\\&=-1+\lambda\left(1-\gamma_{min}\right)\end{aligned}$$

Finally, if $\gamma_{max}\geq1-\frac{1}{\lambda}\geq\gamma_{min}$, we can select $J_1=\left[\gamma_{min},1-\frac{1}{\lambda}\right]$ and $J_2=\left[1-\frac{1}{\lambda},\gamma_{max}\right]$, so that $J=J_1\cup J_2$. Moreover, we have that,

$$\sup_{\beta\in J}|1-\lambda\left[1-\beta\right]|=\max\left\{\sup_{\beta\in J_1}|1-\lambda\left[1-\beta\right]|,\sup_{\beta\in J_2}|1-\lambda\left[1-\beta\right]|\right\}.$$

and, since $\lambda>0$,

$$\sup_{\beta\in J_1}|1-\lambda\left[1-\beta\right]|=|1-\lambda\left[1-\gamma_{min}\right]|,$$

and,

$$\sup_{\beta \in J_2} |1 - \lambda [1 - \beta]| = |1 - \lambda [1 - \gamma_{max}]| .$$

□

Lemma 2.2.3. *There exists a $\lambda > 0$ such that,*

$$|1 - \lambda (1 - \beta)| < 1,$$

if and only if

$$\beta < 1.$$

Proof. Suppose that $\beta < 0$, then,

$$\begin{aligned} |1 - \lambda (1 - \beta)| &= |1 - \lambda (1 + |\beta|)| \\ &< 1 \end{aligned}$$

for any $\frac{2}{1+|\beta|} > \lambda > 0$. On the other hand, if $1 > \beta \geq 0$, then $\lambda (1 - \beta) > 0$ and,

$$|1 - \lambda (1 - \beta)| < 1,$$

for any $\frac{2}{1+\beta} > \lambda > 0$, In the converse direction, suppose that there exists a $\lambda > 0$ such that,

$$1 > |1 - \lambda (1 - \beta)|$$

Then, after squaring the two sides and rearranging, we must have that

$$0 > (1 - \beta) [\lambda (1 - \beta) - 2] ,$$

which can only happen if either $\beta > 1$ and $\beta < 1 - \frac{2}{\lambda} < 1$ which is a contradiction, or if $1 > \beta > 1 - \frac{2}{\lambda}$. Hence, we must have that $\beta < 1$. □

Theorem 2.2.4. *Suppose that $f : I \to I$ is an admissible flux function. There exists a $\lambda > 0$ such that,*

$$\sup_{S \in I} |1 - \lambda [1 - \gamma (S)]| = \alpha < 1,$$

if and only if

$$\gamma_{max} = \frac{\Delta t \left(1 - S_{init} + \Delta t \left(f_{max} - f \left(S_{inj}\right)\right)\right) f''_{max}}{\left(1 + \Delta t f'_{min}\right)^2} < 1.$$

Proof. Combining the two lemmas, we must have that $\gamma_{min} \leq \gamma_{max} < 1$ if and only if $\alpha < 1$. □

2.2.3 Computational Examples

This section is aimed at analysis and comparison of standard, global damping based and Eclipse Appleyard Newton's methods for different cases according to parameters selected. The new method, called *global damping based* Newton's method, due to the fact that the damping factor calculated for the first iteration will be used throughout all iterations. It is important to visualize and analyze first and second order derivative curves for S-shaped flux function, which helps to ensure that derivatives of the residual equation are continuous as well. Analytical experssion of the first and second derivatives of flux functions are given by equations 2.2.2 and depicted in Figure 1.2 are shown in Figure 2.1. State-of-the-art methods rely their analysis on heuristic observations according to S-shaped curve and its derivatives. For example, the trust region method applies chops of saturation values to the inflection point, where the second order derivative is equal to zero.

$$\frac{\partial f}{\partial S} = \frac{2MS(1-S)}{\left(M + MS^2 + S^2 - 2MS\right)^2} \tag{2.2.2a}$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{2M(M - 3MS^2 + 2MS^3 - 3S^2 + 2S^3)}{\left(M + MS^2 + S^2 - 2MS\right)^3} \tag{2.2.2b}$$

It should be noted, that failure of the Newton based solvers is determined when the number of nonlinear iterations reaches the maximum, which we set to one thousand. Such a high number has been chosen in order to capture even very slow, but still convergent

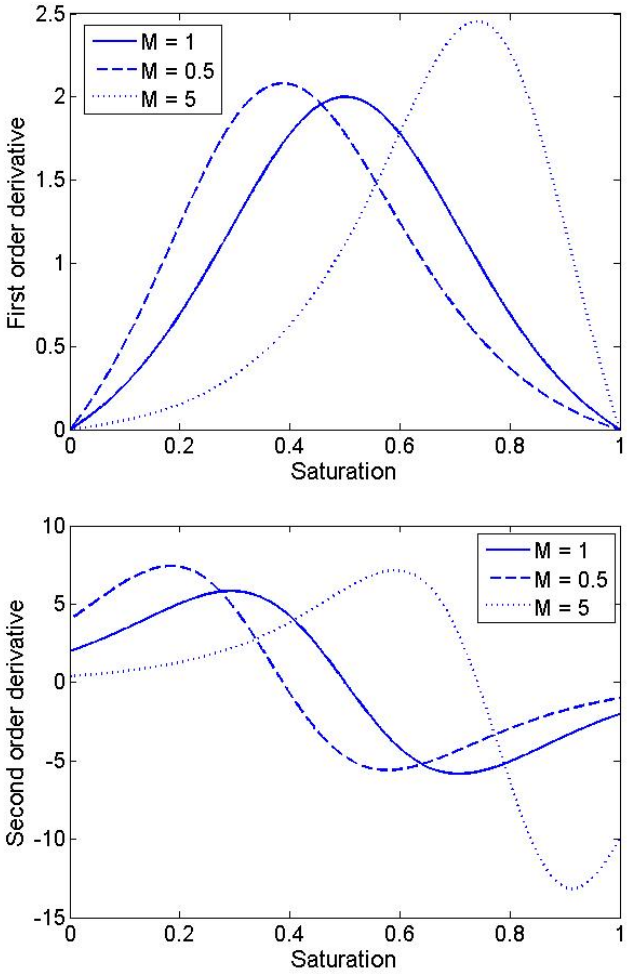

Figure 2.1: First and Second order derivatives of various M numbers.

sequences. The test cases below are based on analysis of errors with variable time step sizes and viscosity ratio M. The first problem statement follows:

Case 1: $S_{in} = 0$, $S_{guess} = 0$, $\Delta t = 1.1$ and the flux function parameter $M = 1$.

As we can from the Figure 2.2, the error, which is defined as the difference of the solution

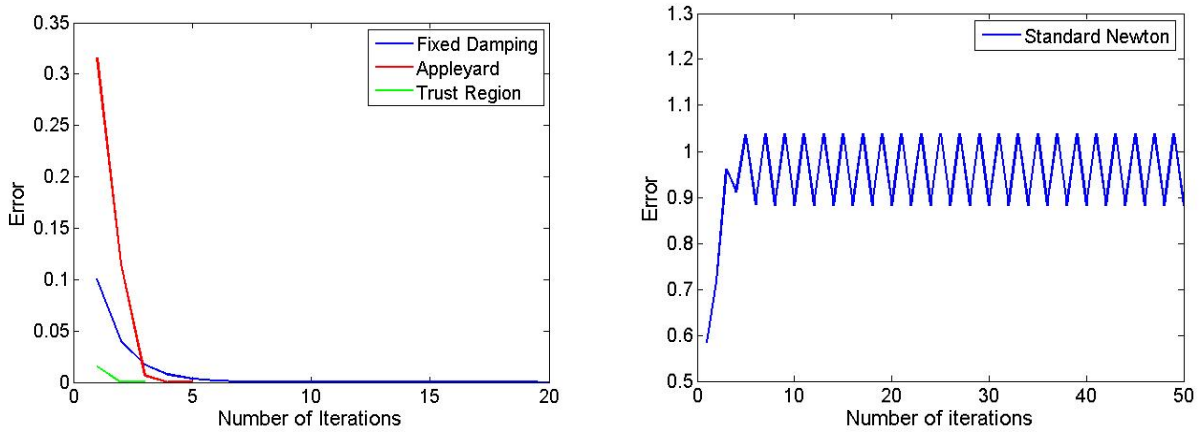

Figure 2.2: Errors for safeguarded and standard Newton's methods for Case 1.

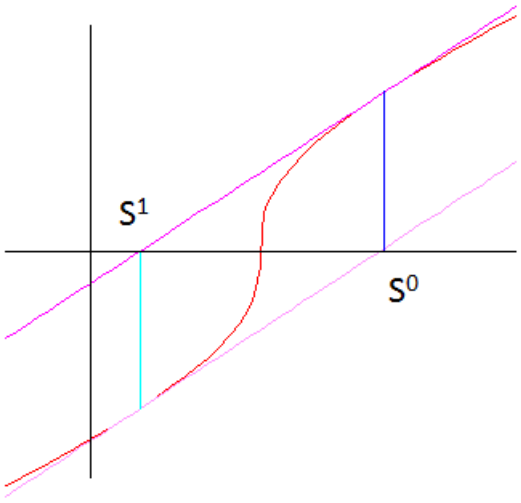

Figure 2.3: Updates in saturation while bouncing.

from the iteration value, is oscillating from the first iterations for standard Newton's method. This may happen for particular cases, as presented in Figure 2.3. The number of nonlinear iterations reached the maximum number for the standard Newton's method, but only the first fifty of them are depicted, since the rest are identical. If we analyze the contraction factor for

the standard method from Figure 2.4, it is shown that all values less than 0.3 and higher than 0.7 are not in the contractive region, which means that if during iterations, saturation value will be equal to any of these values, convergence is not guaranteed by the theorem. However, the theorem considers the sufficient, but not necessary condition, meaning that convergence still can be obtained. Calculated damping factor value for the global method is equal to 0.56, which brings all saturation values to the contraction region, hence convergence is obtained. Comparison with Eclipse Appleyard method indicates that both methods converge to the same solution. However, the number of iterations performed by EA method is less. Fluid

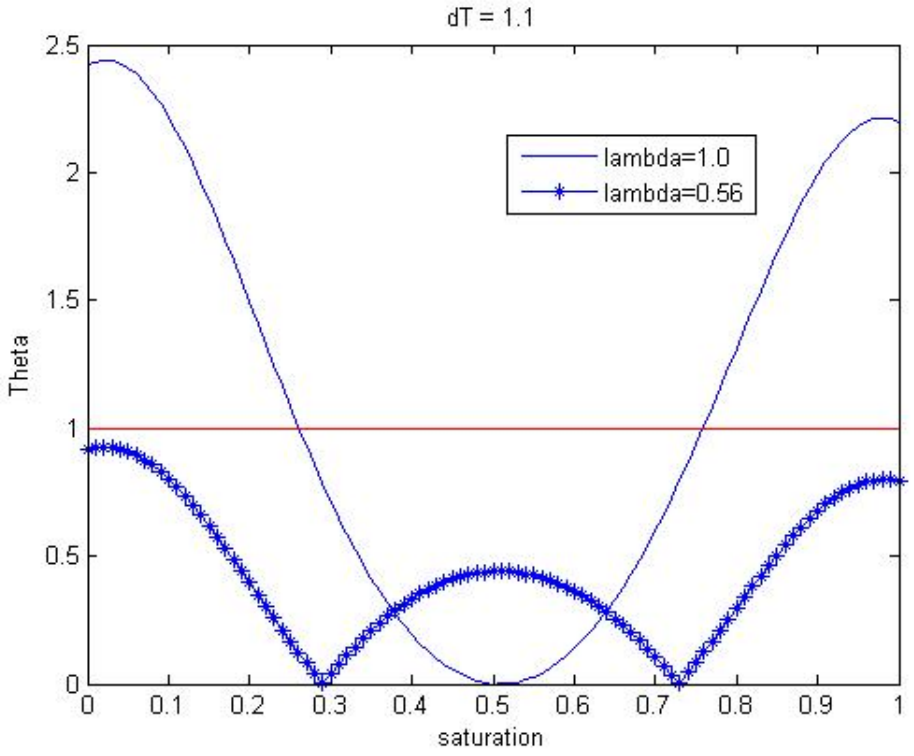

Figure 2.4: Comparison of standard and global damping based Newton methods' contraction factors (Case 1).

properties have always been crucial in recovery processes and flow characterization. So far, oil viscosity, represented in the flux function via M, has been equal to $\mu_{oil} = \mu_{water} = 1cp$. But there are many other types of crude oil worldwide, where viscosity is either less or much higher than water viscosity. As an important concept, lets compare viscosity ratio influence on standard, global damping based and EA Newton methods' performance, where

flux function's derivatives reasonably change their shapes (Figures 2.1).

Case 2: $S_{in} = 0$, $S_{guess} = 0$, $\Delta t = 1.2$ and $M = 0.5$

Number of nonlinear iterations for different method for this case is represented on the

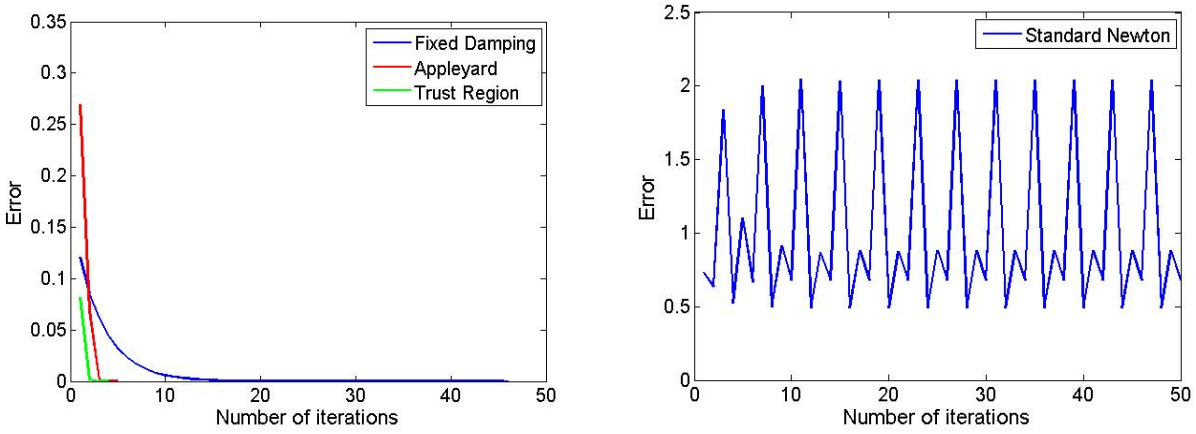

Figure 2.5: Errors for safeguarded and standard Newton's methods for Case 2.

Figure 2.5. As we can see from this figure, oscillation for the standard method's error estimation leads to divergence, while safeguarding methods helps to avoid this issue. Change in viscosity ratio has a noticeable impact on contraction factor number α. If we compare Figures 2.4 and 2.6, curve shapes differ significantly. Lower viscosity ratio M in the problem increases the maximum number of contraction factor value, which makes it difficult to satisfy the condition of the theorem. The standard method failed to converge, while damping based method shows robustness and accuracy in comparison with it. Calculated damping factor value is equal to 0.29, which is a strict chopping strategy.

Case 3: $S_{in} = 0$, $S_{guess} = 0$, $\Delta t = 1.7$ and $M = 5.0$

Relationship between number of nonlinear iterations and error decrease remains the same as in previous examples (Figure 2.7). This test case considers fast changes in the system, where water front moves with constant higher velocity. Behavior of safeguarded methods remains the same: global method confirms robustness and accuracy, while standard method does not (Figure 2.8). Chopping strategy for this case is even more conservative, which leads to small damping factor ($\lambda = 0.1$).

As we can see from the test cases for single cell problems, errors for the global damping

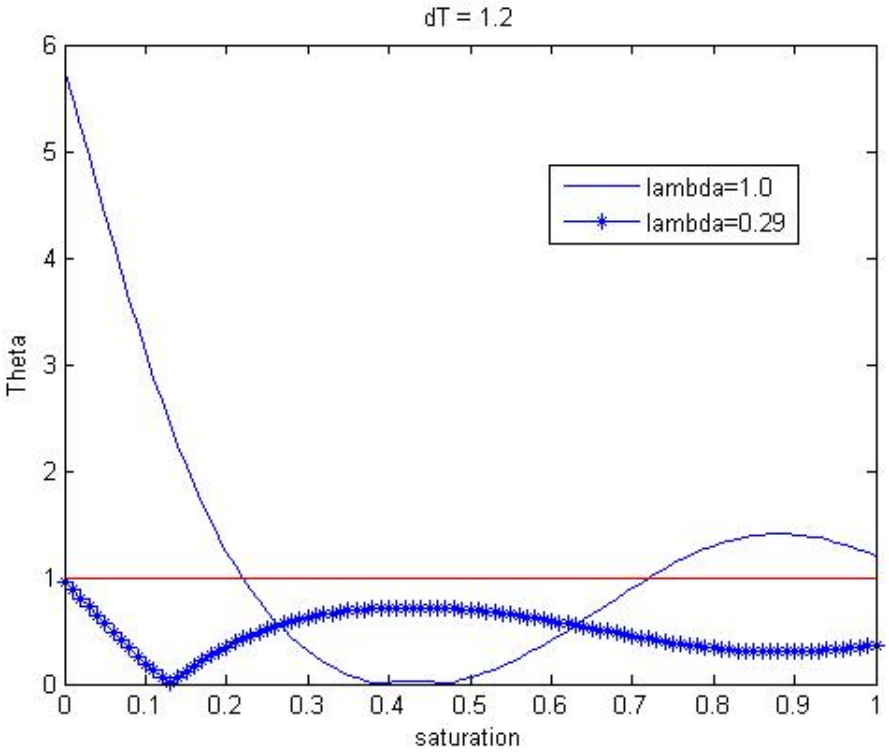

Figure 2.6: Contraction mapping for standard and global damping based methods (Case 2).

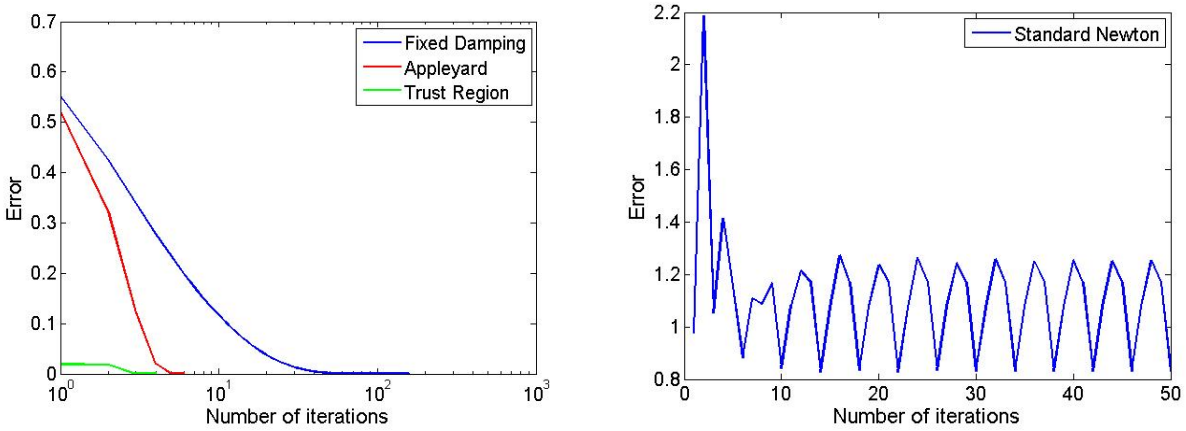

Figure 2.7: Errors for safeguarded and standard Newton's methods for Case 3.

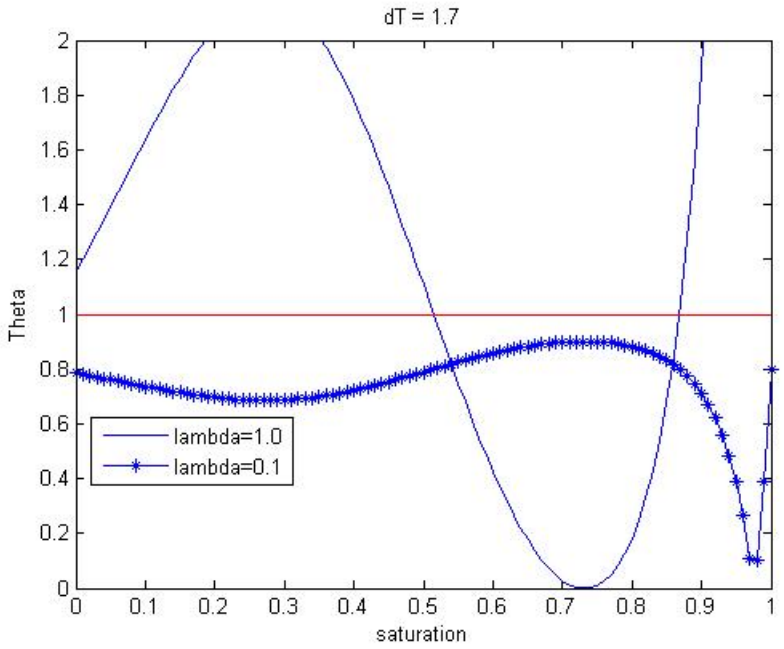

Figure 2.8: Contraction mapping for standard and damping based methods (Case 3).

based method always decrease, which leads to unconditional convergence of the method.

2.3 Multi cell problem

2.3.1 Analytical approach and problem statement

In the previous section the contraction mapping theorem was investigated and applied to a single cell problem, where the damping factor of the Newton update has been analytically calculated in order to satisfy sufficient condition for guaranteed convergence. Some of the heuristic safeguarding nonlinear solvers apply single cell estimations to three dimensional problems, which is unlikely to be a good assumption. This work avoids such assumptions, since analysis is based on theoretical approaches only, which becomes more complicated for a one dimensional problem. Performance of damping based Newton's method for the single cell problem demonstrates valuable features of nonlinear solvers, such as robustness and accuracy. Efficiency of the method highly depends on the problem's complexity. The aim is to keep these features for the one dimensional problem as well by applying contraction mapping theorem. The key difference for one dimensional problems is that the contraction

mapping theorem will be investigated in vector spaces. A numerical scheme for global damping based Newton's method will be developed and implemented in order to compare results with standard and state-of-the-art safeguarded Newton's methods.

Recall that G is a nonlinear operator generating a sequence of iterations using Newton's method for problems investigated in this article. The purpose of the one dimensional approach remains the same, as for a single cell problem: to find such a damping factor λ, which satisfies Equation 2.3.1. For multi cell problem, the equation is simply

$$\|G(x) - G(y)\| \leqslant \alpha \|x - y\|. \tag{2.3.1}$$

Hence, the contraction factor α can be calculated from the following equation:

$$\alpha = \sup_{\mathbf{x} \neq \mathbf{y}} \frac{\|G(x) - G(y)\|}{\|x - y\|} \tag{2.3.2a}$$

$$= \sup_{\mathbf{x} \in [\mathbf{0},\mathbf{1}]} \|G'(x)\|. \tag{2.3.2b}$$

where G' is a matrix. Equation 2.3.2b shows, that we need to estimate the first order derivative of the linear operator G, i.e. calculate the norm of G'. As it was mentioned earlier, there are many different vector space norms, but we use infinity norm. The reason for this is as follows: we estimate contraction factor numbers for each grid cell, which means that different damping factors will be calculated for every cell separately. Hence, it is very important to make sure that all contraction factor numbers are below one. Application of the infinity norm allows us to consider each row, which is not possible, for example, for 1-norm, where the maximum absolute column sum of the matrix considered. Operator G

for one dimensional problem is represents a vector, defined as following:

$$G = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_i \end{bmatrix} - \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_i \end{bmatrix} \begin{bmatrix} \frac{\partial R_1}{\partial S_1} & \frac{\partial R_1}{\partial S_2} & \dots & \frac{\partial R_1}{\partial S_i} \\ \frac{\partial R_2}{\partial S_1} & \frac{\partial R_2}{\partial S_2} & \dots & \frac{\partial R_2}{\partial S_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R_i}{\partial S_1} & \frac{\partial R_i}{\partial S_2} & \dots & \frac{\partial R_i}{\partial S_i} \end{bmatrix}^{-1} \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_i \end{bmatrix} \tag{2.3.3}$$

where the residual vector is represented as

$$\begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_i \end{bmatrix} = \begin{bmatrix} S_1^{n+1} - S_1^n + \frac{\Delta t}{\Delta x}\left[f(S_1) - f(S_{inj})\right] \\ S_2^{n+1} - S_2^n + \frac{\Delta t}{\Delta x}\left[f(S_2) - f(S_1)\right] \\ \vdots \\ S_i^{n+1} - S_i^n + \frac{\Delta t}{\Delta x}\left[f(S_i) - f(S_{i-1})\right] \end{bmatrix}$$

In order to estimate the contraction factor α, we need to evaluate derivatives of the operator with respect to saturation, which leads to creation of matrix from the vector shown in Equation 2.3.3.

$$G' = \begin{bmatrix} \frac{\partial}{\partial S_1}\left[S_1 - \lambda_1 \left(J^{-1}R\right)_1\right] & \frac{\partial}{\partial S_2}\left[S_1 - \lambda_1 \left(J^{-1}R\right)_1\right] & \dots & \frac{\partial}{\partial S_i}\left[S_1 - \lambda_1 \left(J^{-1}R\right)_1\right] \\ \frac{\partial}{\partial S_1}\left[S_2 - \lambda_2 \left(J^{-1}R\right)_2\right] & \frac{\partial}{\partial S_2}\left[S_2 - \lambda_2 \left(J^{-1}R\right)_2\right] & \dots & \frac{\partial}{\partial S_i}\left[S_2 - \lambda_2 \left(J^{-1}R\right)_2\right] \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial S_1}\left[S_i - \lambda_i \left(J^{-1}R\right)_i\right] & \frac{\partial}{\partial S_2}\left[S_i - \lambda_i \left(J^{-1}R\right)_i\right] & \dots & \frac{\partial}{\partial S_i}\left[S_i - \lambda_i \left(J^{-1}R\right)_i\right] \end{bmatrix} \tag{2.3.4}$$

In order to construct the sequence for one dimensional problem with $N > 3$, we introduce following notations with cell number i:

$$a_i = 1 - \frac{c_i}{(1+c_i)^2} R_i - \frac{c_{i-1} c_i' \delta_i}{(1+c_i)^2}, \tag{2.3.5}$$

$$b_i = -\frac{c_{i-1}}{1+c_i} + \frac{c_{i-1}'\delta_i}{1+c_i} + \frac{c_{i-1}}{1+c_i} a_{i-1}. \tag{2.3.6}$$

For calculation of infinity norm of the matrix for general 1D problem, which gives the con-

traction factor α, let us consider example with four grid cells:

$$G' = \begin{bmatrix} 1-\lambda_1 a_1 & 0 & 0 & 0 \\ -\lambda_2 b_1 & 1-\lambda_2 a_2 & 0 & 0 \\ -\lambda_3\left(\frac{c_2}{1+c_3}b_1\right) & -\lambda_3 b_2 & 1-\lambda_3 a_3 & 0 \\ -\lambda_4\left(\frac{c_3}{1+c_4}\frac{c_2}{1+c_3}b_1\right) & -\lambda_4\left(\frac{c_3}{1+c_4}b_2\right) & -\lambda_4 b_3 & 1-\lambda_4 a_4 \end{bmatrix} \tag{2.3.7}$$

The main diagonal can formulate one type of sequence, while the lower bidiagonal formulates another sequence consisting other additional parameters. The remaining part of the lower triangular matrix formulates multiplication sequence included in the sequence of summation. Infinity norm of a matrix, used for calculations, is described in Meyer(2001) [22], as following:

$$\|G'\|_\infty = \max_{1\le j\le i} \sum_{k=1}^{l} |G'_{jk}|, \tag{2.3.8}$$

and particularly for cell number i >3 can be formulated as:

$$\|G'\|_\infty = \sup_{S\in D_0} \left[\lambda_i \left(\sum_{j=1}^{i-2} \left| -b_j \frac{\prod_{k=j+1}^{i-1} c_k}{\prod_{m=j+2}^{i} 1+c_m} \right| + |b_i| \right) + |1-\lambda_i a_i| \right] \tag{2.3.9}$$

which is simply the maximum absolute value of the row sum of the matrix. In other words, the objective is to find such damping factors λ, where the maximum absolute row sum value is less than one, which means that all row sum values should be below unity as well. Hence, the entire system satisfies the contraction mapping principle. In the next section, a two cell problem is implemented and compared with standard and other safeguarding nonlinear solvers.

2.3.2 Two cell example and comparison with state-of-the-art nonlinear solvers

In order to test the one dimensional problem, a two cell case has been generated with water injection on the left boundary of the first cell, as on the Figure 2.9. As in the previous

cases, capillary and buoyancy forces are neglected.

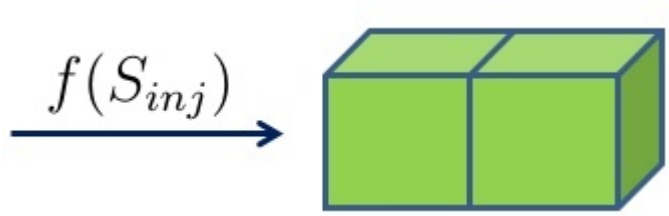

Figure 2.9: One dimensional problem represented by two cells.

$$G = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} - \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 - \frac{\Delta t}{\Delta x} f'(S_1) & 0 \\ -\frac{\Delta t}{\Delta x} f'(S_1) & 1 - \frac{\Delta t}{\Delta x} f'(S_2) \end{bmatrix}^{-1} \begin{bmatrix} S_1^{n+1} - S_1^n + \frac{\Delta t}{\Delta x} \left[f(S_1) - f(S_{inj}) \right] \\ S_2^{n+1} - S_2^n + \frac{\Delta t}{\Delta x} \left[f(S_2) - f(S_1) \right] \end{bmatrix} \tag{2.3.10a}$$

$$= \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} - \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \frac{R_1}{R_1'} \\ \frac{\Delta t}{\Delta x} \frac{f_1' R_1}{R_1' R_2'} + \frac{R_2}{R_2'} \end{bmatrix} \tag{2.3.10b}$$

$$= \begin{bmatrix} S_1 - \lambda_1 \left(\frac{R_1}{R_1'} \right) \\ S_2 - \lambda_2 \left(\frac{\Delta t}{\Delta x} \frac{f_1' R_1}{R_1' R_2'} + \frac{R_2}{R_2'} \right) \end{bmatrix} \tag{2.3.10c}$$

The next step is to differentiate the linear operator with respect to saturation at current iteration:

$$G' = \begin{bmatrix} 1 - \lambda_1 \left(1 - \frac{\Delta t}{\Delta x} \frac{R_1 R_1''}{\left(R_1' \right)^2} \right) & 0 \\ \lambda_2 \left(-\frac{R_1 R_1''}{(R_1')^2 R_2'} \right) & 1 - \lambda_2 \left(1 - \frac{R_2 R_2''}{\left(R_2' \right)^2} - \frac{\Delta t}{\Delta x} \frac{f_1' R_1 R_2''}{R_1' \left(R_2' \right)^2} \right) \end{bmatrix} \tag{2.3.11}$$

After formulating the matrix, in order to find damping factor, which satisfies contraction mapping principle, we need to find infinite norm of G' by applying Equation 2.3.8.

$$\|G'\|_\infty = max \begin{bmatrix} \left| 1 - \lambda_1 \left(1 - \frac{\Delta t}{\Delta x} \frac{R_1 R_1''}{\left(R_1' \right)^2} \right) \right| \\ \left| \lambda_2 \left(-\frac{R_1 R_1''}{(R_1')^2 R_2'} \right) \right| + \left| 1 - \lambda_2 \left(1 - \frac{R_2 R_2''}{\left(R_2' \right)^2} - \frac{\Delta t}{\Delta x} \frac{f_1' R_1 R_2''}{R_1' \left(R_2' \right)^2} \right) \right| \end{bmatrix} < 1 \tag{2.3.12}$$

Each grid cell should be treated separately and all of them should satisfy the contraction

mapping principle. Simple algebraic manipulations, as demonstrated for local damping based single cell problem, has been done for each vector member. Note, that the condition for the first grid cell remains unchanged, except a new term in the governing equation, $\Delta x = \frac{1}{i}$, where i is number of grid cells in the system.

In order to analyze performance and apply the global damping based Newton's solver for two cell problem, the following test case has been generated:

Case 4. $S_1^n = S_2^n = 0$ and $S_1^0 = S_2^0 = 0$ with $\Delta t = 1.0$ and M = 1.0;

In comparison with the single cell problem, for the two cell problem the equation for contraction mapping of the first grid cell changes by factor Δx. For this case it is equal to 0.5. Estimation for the second grid cell becomes even more complicated, due to integration of the first grid cell. For this reason, contraction factor curve becomes a surface, which depends both on S_1 and S_2. The damping factors for this case are equal to: $\lambda_1 = 0.09$; $\lambda_2 = 0.01$. The corresponding contraction mappings for these cells are shown in Figure 2.10, 2.11. As we can see from both figures, any possible saturation value satisfies the contraction mapping principle, hence convergence is guaranteed provided by the theorem. On the other hand, because of very conservative estimations, where damping factors are too small, it takes many iterations to converge to the solution. The main feature of the damping based method to point out is the monotonic decrease of error (or update) during iterations. For standard Newton's method, cycling saturation values still remains, which makes convergence hopeless. Figure 2.12 demonstrates a comparison of these two methods with Eclipse Appleyard, where global damping based and EA converge to the same solution.

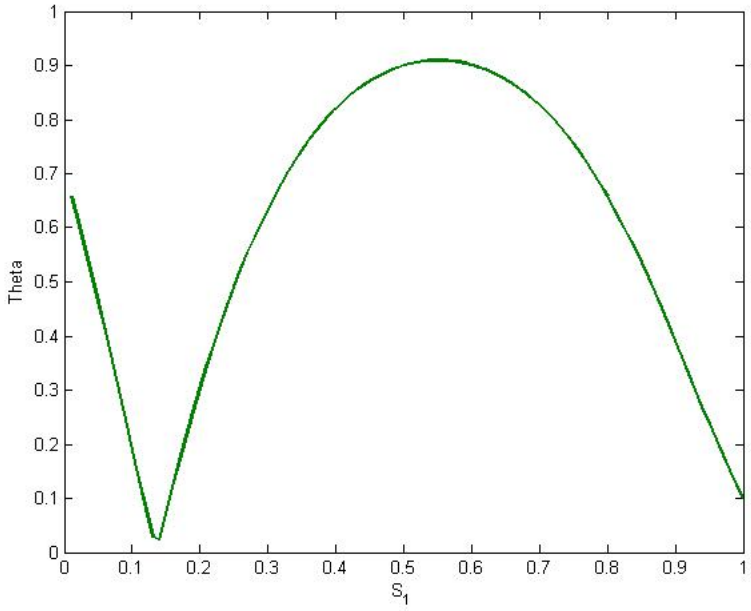

Figure 2.10: Contraction factor curve for first grid cell (Case 4).

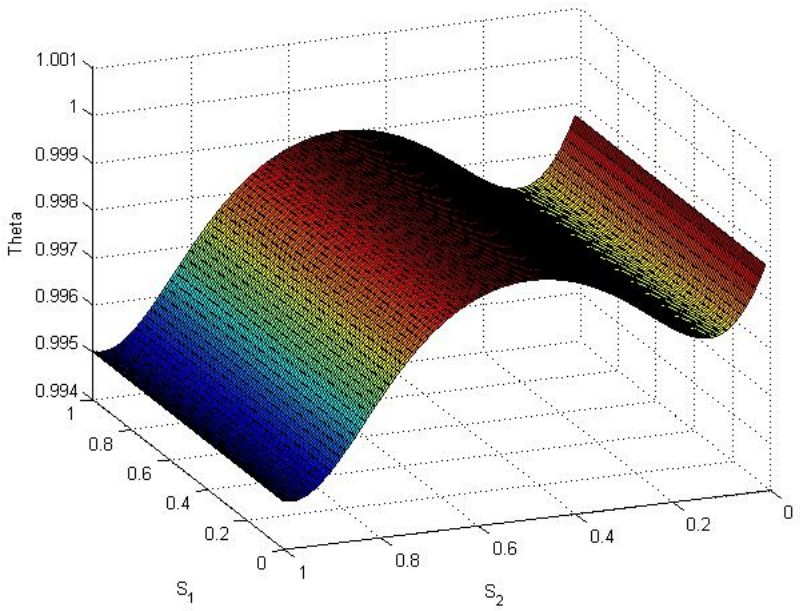

Figure 2.11: Contraction factor surface for second grid cell (Case 4).

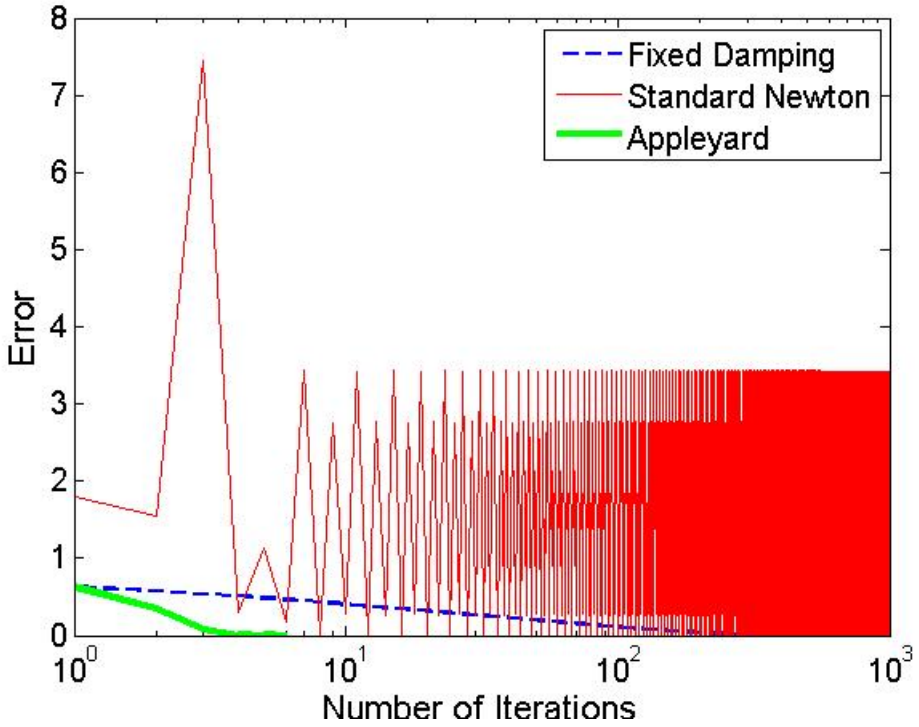

Figure 2.12: Errors for safeguarded and standard methods in Case 4.

CHAPTER 3

DYNAMIC DAMPING BASED NEWTON'S METHOD

3.1 Single cell dynamic damping based Newton's method

3.1.1 Analytical approach and algorithm description

Analysis of examples, described in the previous chapter, shows that fixed damping based Newton's method is good improvement of the standard method, which avoids failures, but still takes relatively high number of iterations to converge. Strategy of the method uses a priori calculated Newton step length throughout all iteration, which is very strict estimate which might not be necessary, especially for values within some neighborhood of the solution. These circumstances became the reason for further improvement of the current approach, resulting in development of a more flexible and adaptive strategy, called *dynamic damping* based Newton's method. It is worth mentioning, that the term *dynamic* related to Newton iterations' length, i.e. for every Newton iteration a new damping factor will be analytically calculated. Like in the previous approach, these values depend on first and second order derivatives of flux function, as well as the time step size. In other words, dynamic damping based approach chops only those saturation values which are not satisfying the contraction condition and hence can not guarantee convergence. Otherwise, if the contraction condition holds, it performs as standard Newton's method with step length size equal to unity. The aim is to reduce the number of nonlinear iterations in comparison with the fixed damping method, without negative influence on solution accuracy and robustness of the method.

To begin with, recall the general contraction factor estimation for single cell problem with $I = [0, 1] \subset \mathbb{R}$:

$$\alpha = \sup_{\mathbf{S}\in\mathbf{I}} \left| 1 + \lambda \left(\frac{R(S)R''(S)}{\left(R'(S)\right)^2} - 1 \right) \right| < 1 \tag{3.1.1}$$

For convenience, we introduce the following notation

$$\gamma = \left(\frac{R(S)R''(S)}{\left(R'(S)\right)^2} - 1 \right) \tag{3.1.2}$$

and apply certain modifications to equation 3.1.1:

$$\left(|1 + \lambda\gamma|\right)^2 < 1^2 \tag{3.1.3a}$$

$$1 + 2\lambda\gamma + \lambda^2\gamma^2 = 1 \tag{3.1.3b}$$

$$\lambda\gamma(2 + \lambda\gamma) < 0 \tag{3.1.3c}$$

Equation 3.1.3c is a condition for adaptive step length selection. Analysis of equation 3.1.3c shows, that in order to fulfill this condition, $\lambda\alpha$ should be in the following interval:

$$\lambda\gamma \in (-2, 0)$$

An algorithm for implementation of dynamic damping based Newton's method is the following:

Algorithm 1: Dynamic damping based Newton's method

input : $[U^{n+1}]^{\nu} \rightarrow$ previous iteration value

output: $[U^{n+1}]^{\nu+1} \rightarrow$ updated value for the iterate $\nu + 1$

while $DT <= T_{final}$ **do**
  NITER = 0;
  while $NITER < MAXITER$ *and* $||R\left(U^{n+1}; U^{n}, \Delta t\right)|| > TOL$ **do**
    calculate γ;
    if $\alpha > 0$ **then**
      $\lambda = 1$;
      else
      $\lambda = -2/\gamma$;
    end
    $\lambda = \min\left(\lambda, 1\right)$;
    solve $\left([U^{n+1}]^{\nu+1}; [U^{n+1}]^{\nu}, DT, \lambda\right)$;
    $NITER\ ++$;
  end
  update DT;
  update $[U^{n+1}]^{\nu}$;
end

As we can see from equation 3.1.3c, calculation of the factor λ is not computationally expensive, but is a very important feature of the method. This approach is hybrid in terms of Newton step length selection, which we hope will give us advantage in terms of number of Newton iterations in comparison to fixed damping based method. In order to clarify it, let us compare these two approaches proposed in the thesis. In order to be consistent, the same cases, with some additional examples, will be tested below.

3.1.2 Results and comparison with nonlinear solvers

First, we start with already tested **Case 1**, where the standard Newton's method does not converge ($S_{in} = 0$, $S_{guess} = 0$, $\Delta t = 1.1$ and the flux function parameter $M = 1$). The objective for the dynamic method is to decrease number of Newton iterations as

much as possible. Figure 3.1 shows, that the number of iterations, needed for obtaining the same solution, decreased four times (from 20 to 5). This is a big gain in the reduction of computational time, and can be even higher for more complicated cases. In comparison to fixed damping, dynamic damping based Newton's method is flexible in Newton step length size selection, as can be seen from Figure 3.2. It damped the Newton step length only at the first iteration down to 0.55, which is almost the same damping factor calculated for the fixed method ($\lambda = 0.56$). This happens because the contraction factor number was higher than unity, which does not guarantee the sequence will convergence. So, dynamic method dampen Newton updates in the regions, only where contraction for the standard method way above unity. On regions where the contraction factor for the standard method is less than one, the contraction mapping of the dynamic method matches to standard. Also, Figure 3.2 shows, that for those regions, the damping factor λ is equal to one. The only chop at the beginning was enough to satisfy the contraction condition and the algorithm converges to the solution.

For validation purposes, it is important to make sure, that the dynamic based method gives the same solution as the standard method, when standard successes, and compare number of iterations. For this purpose, the first case has been modified to **Case 5** by changing time step size up to $\Delta t = 1.8$. Figure 3.3 indicates, that the dynamic method, in some cases (e.g. **Case 5**) when the standard method converges, has even better performance. Note: saturation values below zero and more than one are cut from graph for convenience. The dynamic method applied chops only twice, on first and second iterations, while the fixed damping method applied cuts during all thirty steps (Figure 3.4). This example also confirms that dynamic damping methods applies chopping only for the cases when the contraction factor number is higher than unity; otherwise, it acts as a standard Newton method. This allows one to avoid unnecessary chops in the regions where the standard method is a contraction mapping and allows for potential quadratic convergence.

Case 2 and **Case 3** are sensitivity analysis of viscosity ratio, M, for specific time step sizes. Instead, this section tests big range of different time step sizes and compares

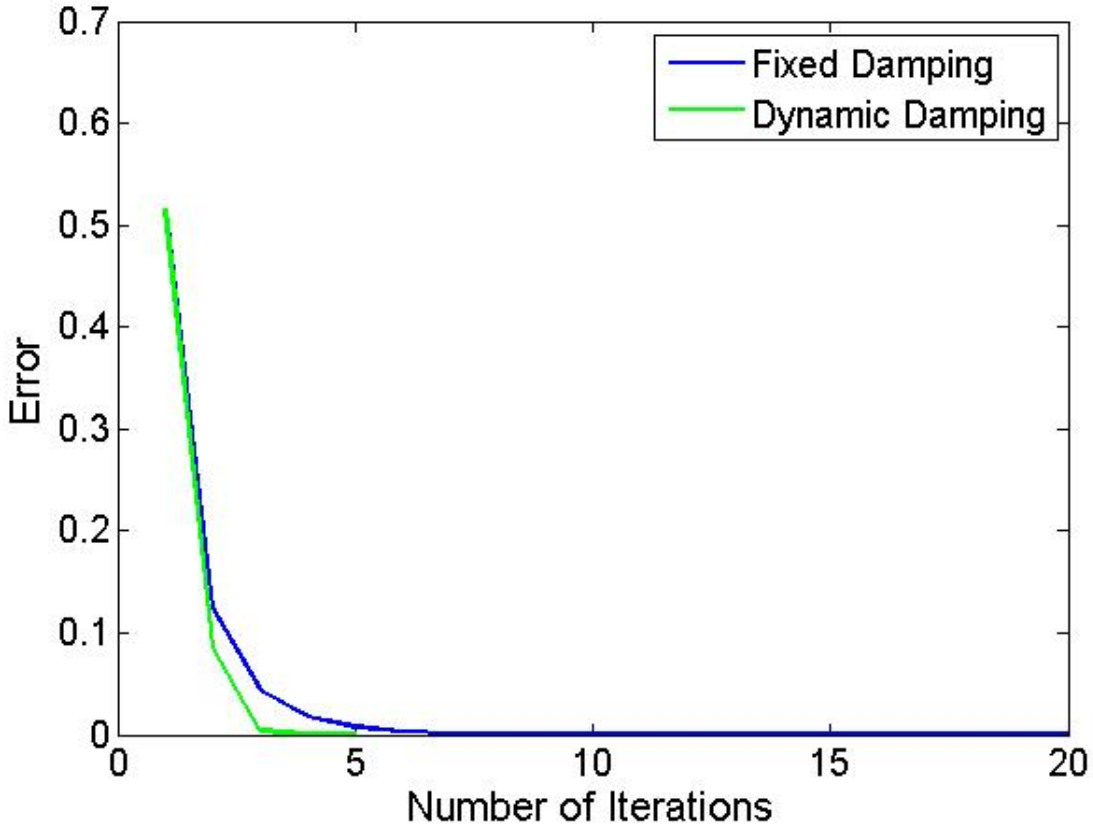

Figure 3.1: Comparison of dynamic and fixedl damping based methods for Case 1.

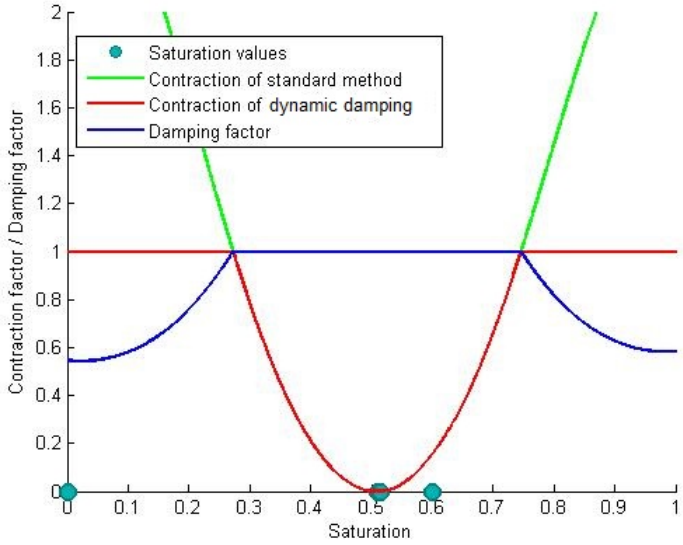

Figure 3.2: Calculation of adaptive Newton update step length selection for Case 1.

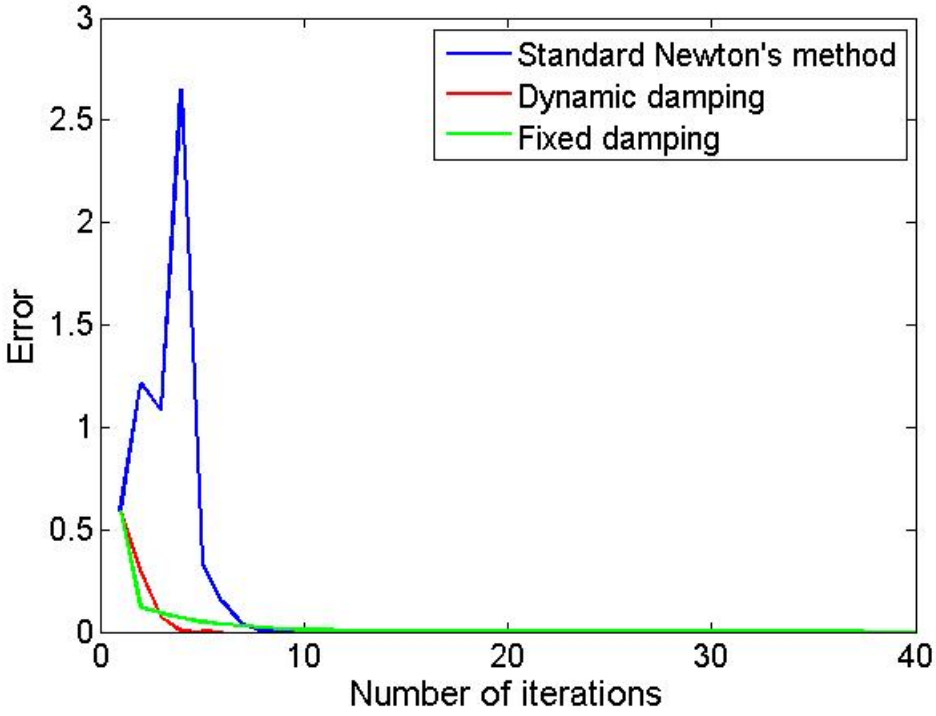

Figure 3.3: Error comparison of dynamic method with standard, along with the fixed method for Case 5.

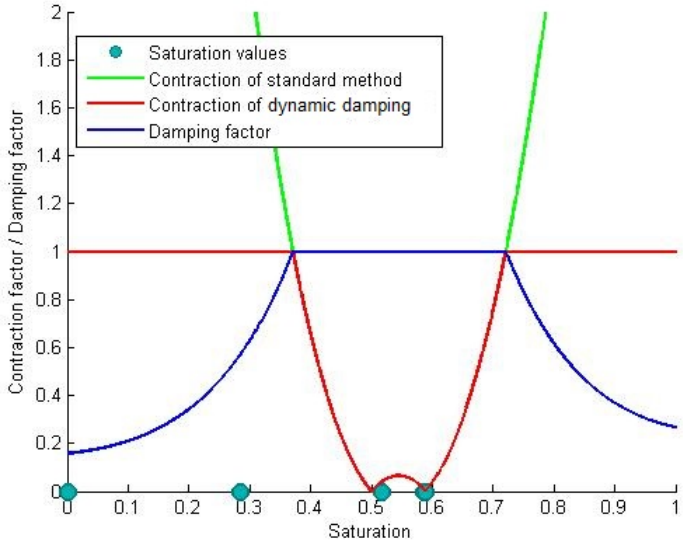

Figure 3.4: Calculation of adaptive Newton update step length selection for Case 5.

dynamic damping based Newton's method with standard and state-of-the-art safeguarding approaches. The main focus of this analysis is the number of nonlinear iterations and solution differences of these methods. Further examples will demonstrate damping behaviors for particular cases.

Case 6: $S_{initial} = S_{guess} = 0$, $\Delta t = [0.0 : 0.1 : 15.0]$ and $M = 1.0$

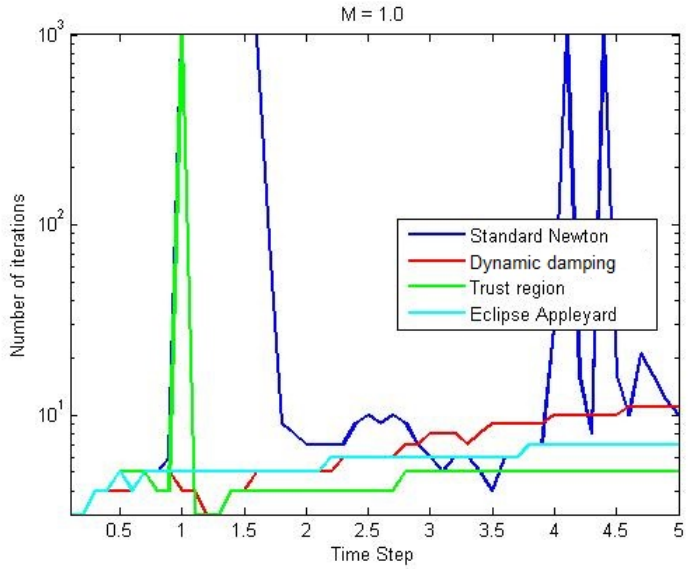

Figure 3.5: Comparison of number of nonlinear iterations for M = 1.0.

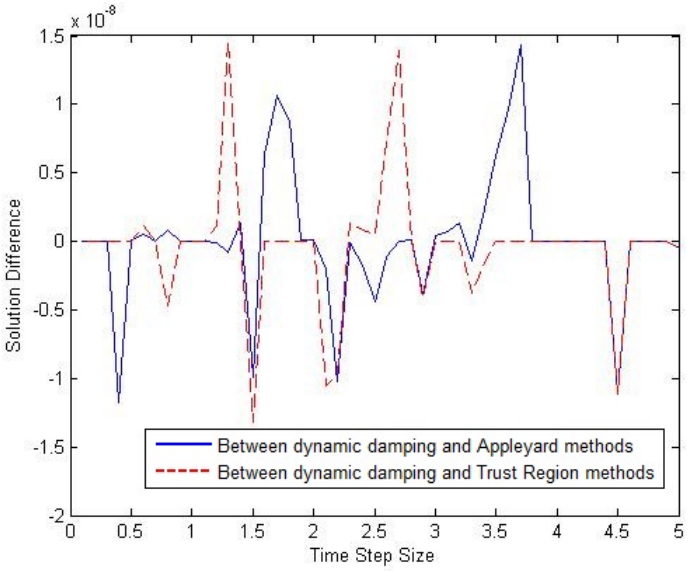

Figure 3.6: Solution differences for all time step sizes for M = 1.0.

As we can see from the Figure 3.5, dynamic damping based Newton's method con-

verges for all time step sizes within a reasonable number of iterations. However, trust region based Newton's method fails to converge within certain time step size range (number of iterations hits maximum value). In the dynamic method, as the time step increases, the number of nonlinear iterations increases monotonically, due to the fact that more and more chops are needed at the beginning iterations in order to guarantee convergence. But the general trend for all dynamic based chopping strategies is that, when the iterative value approaches to the solution within some tolerance, damping values λ becomes equal to one, i.e. behaves as standard Newton's method, which helps to reduce the total number of nonlinear iterations. Figure 3.6 shows that solutions for all time step sizes are very close to each other, within 10^{-8} tolerance. The Eclipse Appleyard method takes a few more iterations than the trust region method, but this is not always true, as will be shown during analysis of next test cases. Note: particular case ($\Delta t = 1.0$), when the trust region based solver diverges, its graph is excluded.

Case 7: $S_{initial} = S_{guess} = 0.5$, $\Delta t = [0.0 : 0.1 : 15.0]$ and $M = 5.0$.

This test case has been implemented in order to visualize convergence performance of all methods for particular viscosity ratio. The value M being equal to 5.0 means that oil's viscosity is five times less than water viscosity and velocity is expected to be higher. Injection saturation S_{inj} has not been changed for any tested cases and is equal to 1.0. It is very crucial to notice, that by changing viscosity ratio, the **sup** of the second order derivative of the flux function increases to some higher point. Hence, fulfillment of some saturation values to contraction condition changes. Figure 3.7 highlights some of the features of each method: standard Newton's method is unstable in the region where time step sizes are between 1 and 3; moreover, it shows poor performance when $\Delta t > 3$, because it converges to nonphysical solutions (where $S > 1$). On the other hand, the dynamic damping based method does not have any convergence failures, and in some particular cases takes even fewer iterations than trust region or EA approaches. Comparison of solutions, as in the previous case has been evaluated which showed the highest difference in answers to be $2{\cdot}10^{-8}$).

Case 8: $S_{initial} = 0.3, S_{guess} = 0.9$, $\Delta t = [0.0 : 0.1 : 15.0]$ and $M = 0.5$

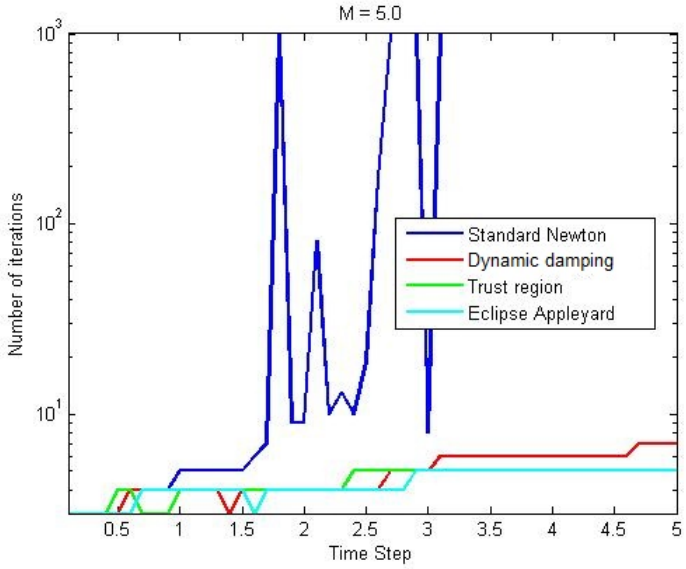

Figure 3.7: Nonlinear solvers' performance for M = 5.0.

As it was mentioned before, usually initial guesses in reservoir simulation problems are chosen to be equal to the value from previous time step. This time we make them different and test for the case, when oil viscosity is twice that of the viscosity of water. From Figure 3.8 we observe that initial guess plays an important role in performance of all methods, especially in the standard Newton's method, where converge are could be obtained for all time step sizes within five iterations. Good initial guesses are reflected in the convergence rate of the dynamic damping based method as well, where the maximum number of iterations is equal to 9. The difference in the number of iterations is the price which we need to pay in order to guarantee convergence. At the same time, at time step size equal to around 2, the number of iterations taken for the dynamic based method is even less than all other methods in our comparison. Comparison of answers with other nonlinear solvers has been demonstrated having the same behavior as in previous examples - maximum difference of is very small and equal to 10^{-8}. Along with it, as it was mentioned before, EA can take fewer iteration to converge in comparison to the trust region method, which makes behavior of other methods also problem dependent.

The adaptive and flexible damping strategy of the dynamic damping based Newton's method, based on analytical calculations, is a crucial feature of the approach to investigate.

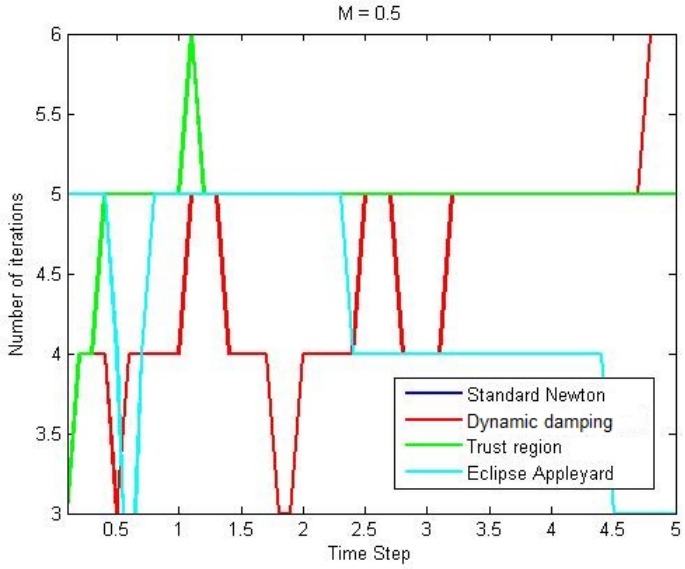

Figure 3.8: Nonlinear solvers' performance for M = 0.5.

As it can be seen from Figure 3.9, damping strategies vary according to input data, such as time step size. Dynamic method can start as standard Newton's method with $\lambda = 1$, apply strict chop equal to 0.81, and back to standard as in the case with $\Delta t = 1$, or remain as standard Newton's method only. Experiments shows, that in most of the cases, it starts with strict chopping, and damping factors value increase monotonically as iteration values gets close enough to the solution. The bigger the time step size, the later the damping factor will be modified to one. All of these strategies converge to the same answer as current safeguarded techniques applied in reservoir simulation.

3.2 Multi cell dynamic damping based Newton's method

3.2.1 General approach

Dynamic damping based approaches for two cell problem has the same strategy as for the single cell, where instead we consider vector spaces and generate numerical scheme again. Norm selection is also remains the same. Adaptive step length selection implemented for both cells separately, by evaluating each row of G', as it is defined for infinity the norm. Two test cases were studied for analysis of dynamic damping based methods in multi cell

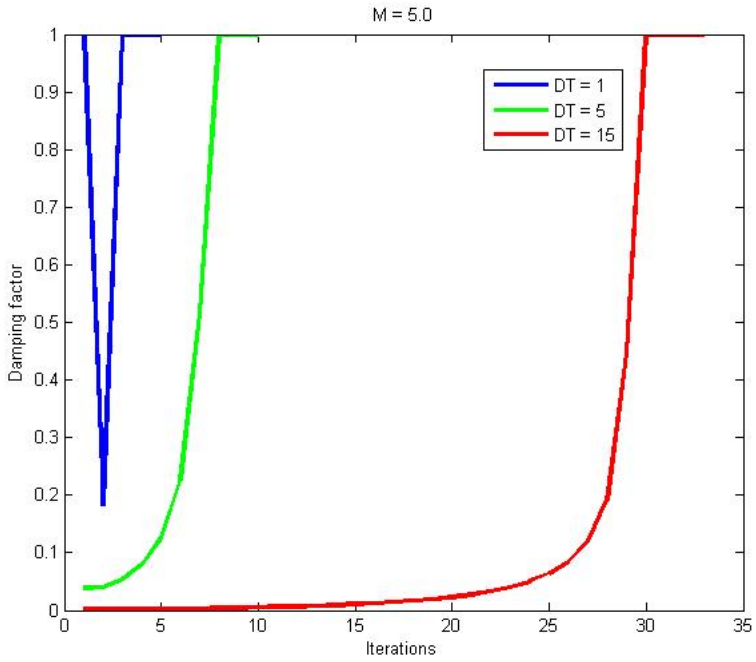

Figure 3.9: Calculation of adaptive damping factors for M = 5.0

problem.

3.2.2 Results and comparison with current safeguarding strategies

Case 9. $S_1^n = S_2^n = 0$ and $S_1^0 = S_2^0 = 0$ with M = 1.0;

This case is chosen for comparison of the number of nonlinear iterations, showing the behavior of damping factors for each grid cell and solution obtained from dynamic damping based method. The number of nonlinear iterations, for both dynamic damping based and Eclipse Appleyard methods, is below ten for all time step sizes, which indicates identical performance of the methods (Figure 3.10). On the other hand, dynamic damping based iteration numbers are slightly lower in comparison with Eclipse Appleyard method, which gives a small advantage to the method. At the same time, trust region based Newton's method fails to converge for times step sizes larger than $\Delta t > 0.3$. Chops to inflection point is happening at each iteration, at least for one of the grid cells. This example demonstrates, that even if solution and iterative values are in the same convex or concave region, as one of the part of S-shaped flux function, it is not guaranteed to converge all the time. If we look

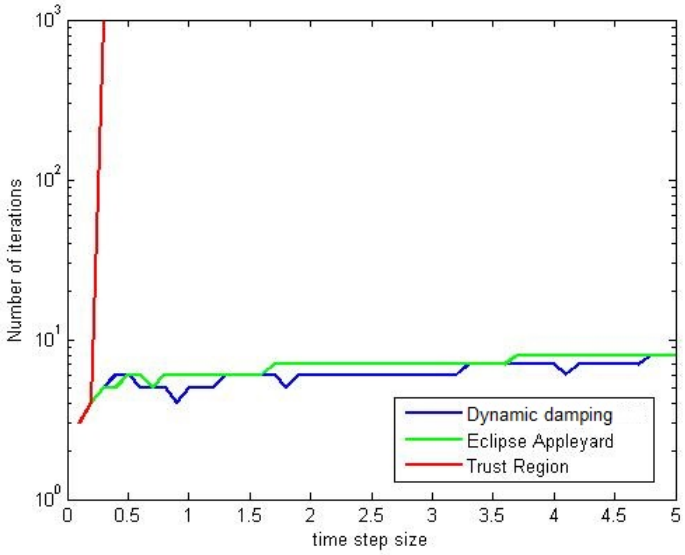

Figure 3.10: Number of nonlinear iterations for two cell problem (Case 9).

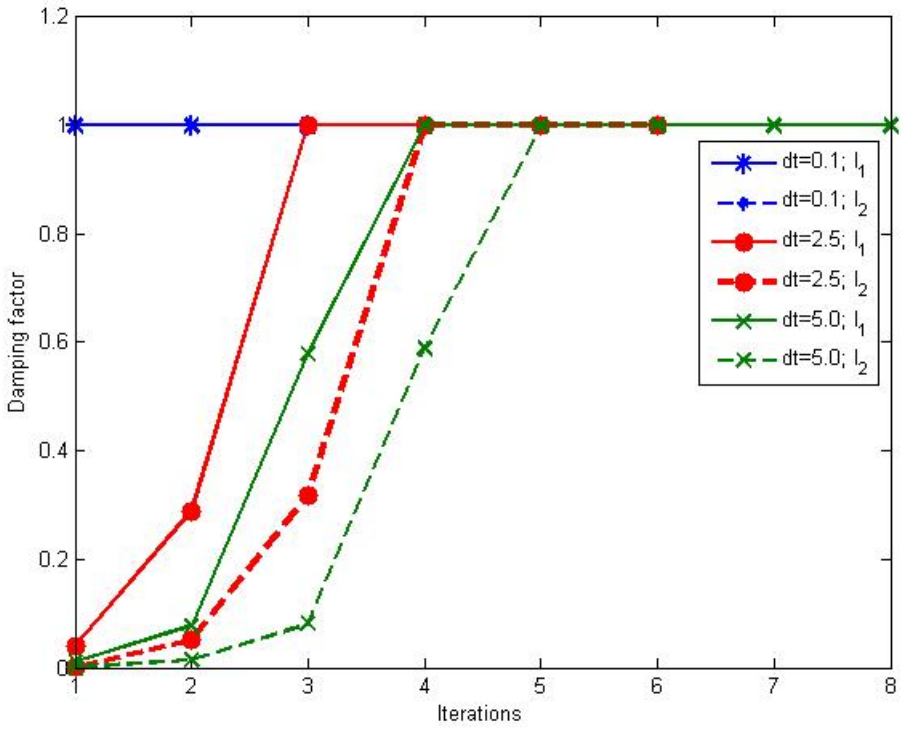

Figure 3.11: Damping factor behavior for both cells with increasing time step size (Case 9).

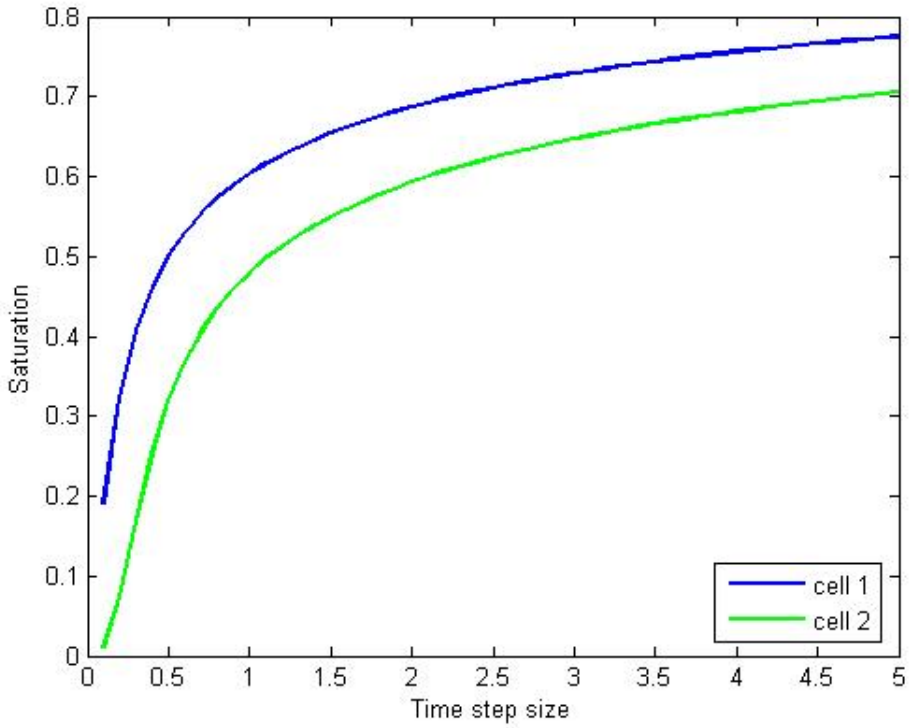

Figure 3.12: Solution of dynamic damping based method for different time step sizes (Case 9).

at this problem from the contraction mapping prospective, divergence can happen due to the fact that contraction condition does not hold at the inflection point with damping factor equal to unity, which may lead to divergence of the sequence.

Figure 3.11 shows the behavior of the damping factor for dynamic method of particular time step sizes: $\Delta t = 0.1$, $\Delta t = 2.5$ and $\Delta t = 5.0$. As it can be noticed from this figure, for small time step size like $\Delta t = 0.1$, the damping factor λ is equal to unity for both grid cells, which means that standard Newton's method can handle this time step size. However, after increasing time step size up to $\Delta t = 2.5$ and $\Delta t = 5.0$, damping factors, for both grid cells, become less with time step size increase.

The solution, obtained by dynamic damping based Newton's method, is shown in Figure 3.12. Comparison with Eclipse Appleyard method demonstrates the same results, while comparison with trust region method can not be estimated, due to the failures of the method.

The next case tests slower physics, when viscosity ratio decreased and solution of all methods will be compared at the same time.

Case 10. $S_1^n = S_2^n = 0.05$ and $S_1^0 = S_2^0 = 0.9$ with M = 0.01;

The transport equation, where water-oil front propagates with lower velocity, due to the high oil viscosity value (e.g. 100 cP), can be interpreted as slow physics. Initial guess is modified along with initial state, because it is not know a priori if a particular value is a sufficiently good guess for Newton's method to converge. Figure 3.13 shows that all methods converge to the solution in four up to nine iterations. Dynamic damping based Newton's method exhibits better performance for time step sizes less than two, and behaves the same as trust region method for the rest of time. This example shows efficiency of the proposed method, which is not slower in terms of the number of nonlinear iterations. Confirming robustness in all previous examples, it always acquires accurate results as well, as can be seen from the Figure 3.14, where dynamic damping based solutions for all time steps compared to Eclipse Appleyard and trust region methods' solutions respectively. The largest difference in solutions of state-of-the-art solvers versus damping based method is $2 \cdot 10^{-8}$. Solutions for first and second grid cells in the following ranges respectively: [0.12,0.36],[0.08,0.29]. Calculation of adaptive damping factor for both grid cells shows identical performance, where $\lambda_1 = \lambda_2 = 1$. Hence, the graph is not shown in this example. Because of identical behavior in the cases with faster physics, where viscosity ratio is higher than one (e.g. M = 2), results are not presented in the section. As in the previous cases, proposed algorithm's efficiency is problem dependent, with the same accuracy and unconditional robustness. Calculated contraction factors for first and second grid cells are depicted in Figures 3.15.

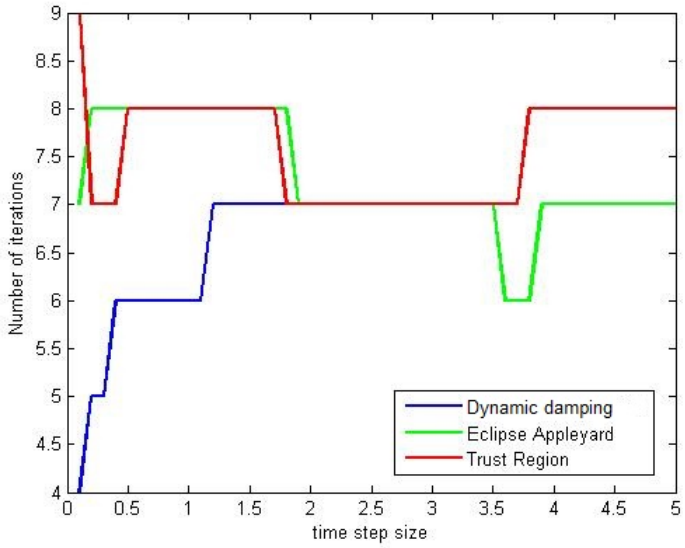

Figure 3.13: Number of nonlinear iterations of slow physics two cell problem (Case 10).

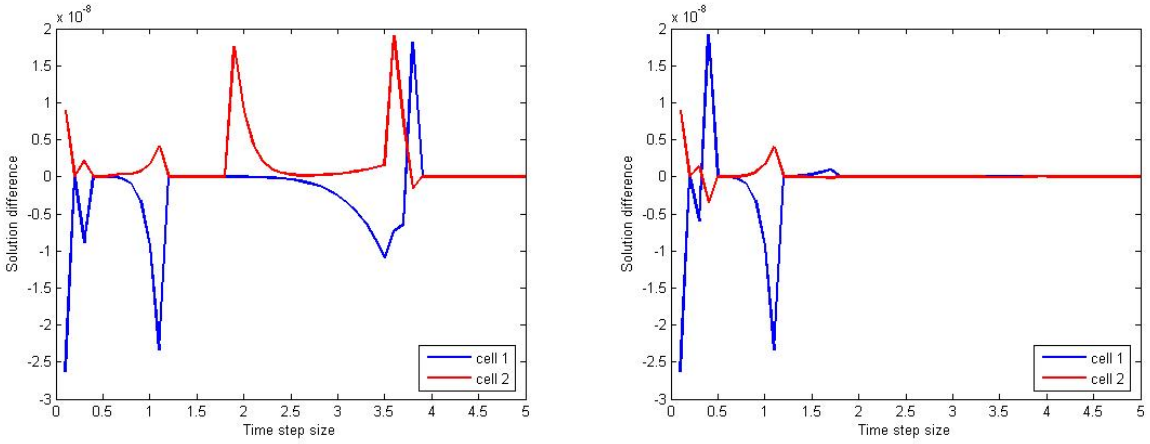

Figure 3.14: Solution differences between: left - dynamic vs. Eclipse Appleyard methods; right - dynamic vs. trust region methods. (Case 10).

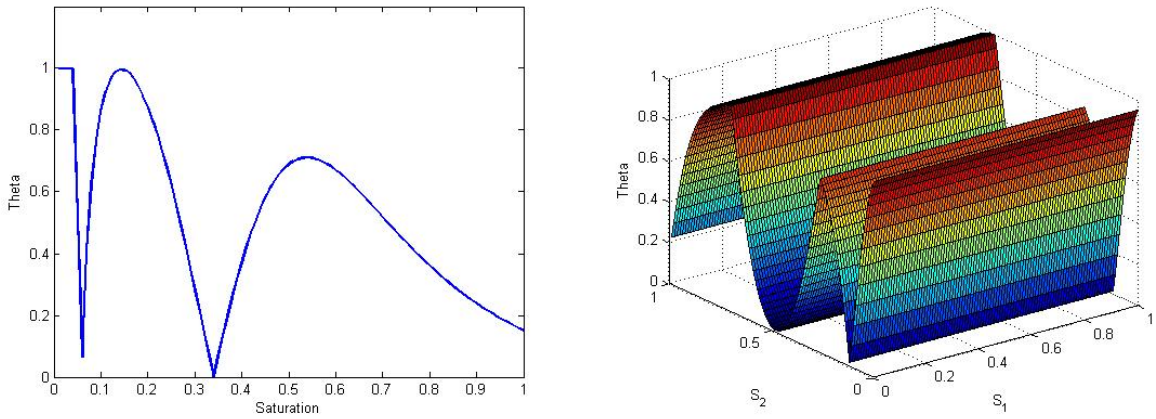

Figure 3.15: Contraction mappings for first and second cells (Case 10).

CHAPTER 4

DISCUSSION

Previous chapters showed analytical safeguarding techniques for Newton's method, which has been developed for single cell and one dimensional problems with two grid cells, where the aim is to avoid convergence failures of transport equation, and hence decrease computational time, by eliminating time step chops throughout the simulation. Analysis shows that complexity is embedded in the formula for the flux function, which leads to convergence difficulties. All the results shown in the thesis have been implemented for S-shaped flux functions with different parameters. Nevertheless, contraction mapping principle is applicable for any type of fractional flow curve, which can be represented even in the form of a table.

First, fixed damping based Newton's method has been derived based on the contraction mapping theorem. This approach can be characterized as following: the damping factor is calculated for particular time step size and initial state, which makes the contraction factor value less than one in order to satisfy the condition that guarantees convergence. Calculated value remain the same for all iterations, until iterative values reaches the solution. As can be seen from Figure 2.7, the fixed damping based method takes many iteration to converge. The reason is unnecessary damping of the Newton update even in the regions, where the contraction condition holds for standard Newton's method. In other words, there is no need to dampen Newton update in already contractive regions, especially when iterative values are "sufficiently" close to the solution. The same can be observed for the two cell problem as well, where a numerical scheme has been developed for implementation of a new method. This inefficiency led to development of another approach, called the dynamic damping based Newton's method.

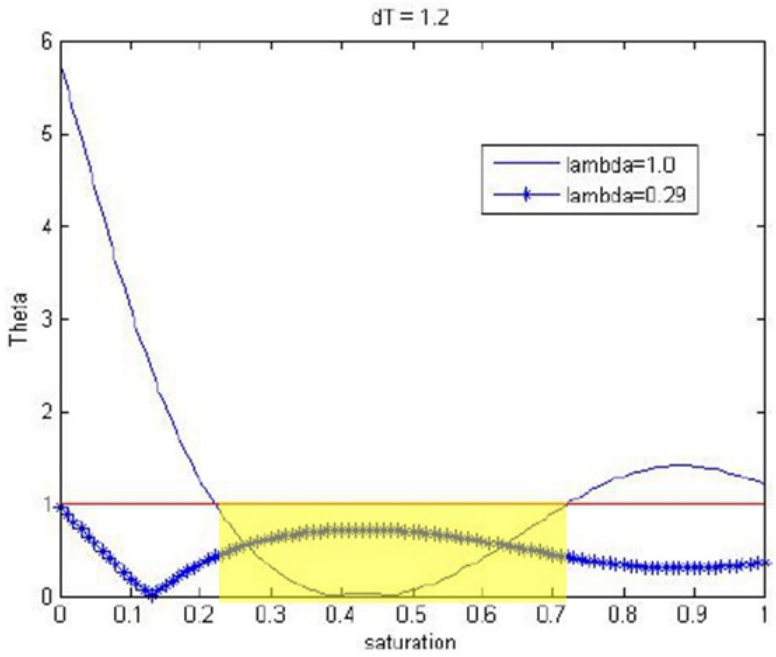

Figure 4.1: Contraction factors for standard and global methods of arbitrary case.

In comparison with the fixed damping method, dynamic damping based Newton's method dampens Newton updates in the regions where the contraction mapping principle does not hold. Otherwise, it behaves as standard Newton's method with damping factor λ equal to unity. For example, shaded region on the Figure 4.1 does not need restriction for updates. Contraction factor values for local method in this region match with values of standard method, as on the Figure 3.2. Also, adaptive damping is not necessarily monotonically increasing towards solution: it can start as standard Newton at first, dampens on the next iteration, and performs back as standard method during the next iterations (Figure 3.9). Robustness and accuracy of this method remains the same as for the fixed method, while efficiency increases dramatically, which gives big advantage in computational time (e.g. Figure 3.3). Another advantage of the method is guaranteed convergence, which is not provided by other safeguarding strategies that we compared. Adaptive damping factor strategy allows the method to perform better or the same as other safeguarding methods. For particular cases it may demonstrate poorer results, which makes efficiency problem dependent.

After obtaining single cell results for dynamic method, contraction mapping theory has been implemented for a one dimensional problem with two grid cells. As for the fixed

approach, one dimensional problem is modified from algebraic equation into vector representation. Definitions of matrix norms have been also provided, which finds direct application for one dimensional problem. Governing transport equation, which is in continuous PDE form, has been spatiotemporally discretized in order to form implicit numerical scheme. Each cell has been treated separately, i.e. adaptive damping factor calculations have been performed to fulfill contraction condition. The transport equation considers viscous forces only, as in the cases with single cell. Robustness, accuracy and efficiency demonstrated relatively same behavior as in single cell problems as well. It's worth mentioning that computation of contraction factor for the second grid cell has influence of the first grid cell's parameters. In other words, every grid cell's contraction condition is dependent on parameters of all previous grid cells. That is one of the possible reasons, why damping strategy for the second cell is more conservative than for previous one, as demonstrated on the Figure 3.11. But the general trend of these values is the same: in a sufficiently close neighborhood to the solution both values adapt and performs as standard method.

As one of the extensions for the work described, generalization of analytical expressions into n cell one dimensional problem is promising, where equations for three, and even for four grid cells problem need to be derived for construction of finite series, which would allow one to create sufficient conditions for each cell automatically.

Discretization in time of the governing PDE and keeping the space in continuous form results in an *infinite dimensional Newton's method*, which becomes first order ODE equation, represented in APPENDIX A. Application of contraction mapping for this approach is very challenging, but would allow one to get higher accuracy for Newton's method. Also, by theory of *asymptotic mesh independence* developed by Allgower et al. [3, 2], it is proved that sufficiently fine discretized Newton's method behaves the same as the infinite dimensional method, implying that convergence difficulties does not arise due to discretization technique.

Classical proof of contraction mapping theorem, which provides guaranteed convergence, highlights that convergence speed of the applied method is at least linear, which means that, there are cases when convergence rate can be even faster, for example with

quadratic speed. L. V. Kantorovich and G. P. Akilov developed sufficient conditions to guarantee quadratic convergence in Banach spaces (see APPENDIX B). The only restriction of this approach is requirement of a priori good initial guess value, which is quite difficult to estimate.

New, theoretically proven and later optimized safeguarding approaches of Newton's method have been developed for the transport equation, for any possible flux function shape. While development of the methods, behavior and complexities in convergence of standard Newton's method has been investigated for various S-shaped flux function parameters. Comparison with state-of-the-art nonlinear solvers confirms that heuristic approaches lack of robustness, leading to restrictions in time step sizes. Contraction mapping theorem can be applied in multiphase, multicomponent and multidimensional sequential simulations with capillary and buoyancy forces consideration.

BIBLIOGRAPHY

[1] Ravi P. Agarwal, Marcia Meehan, and Donal O'Regan. *Fixed Point Theory and Applications*. Cambridge University Press, first edition, 2001.

[2] E. L. Allgower and K. Bohmer. A mesh-independence principle for operator equations and their discretizations. *SIAM Journal on Numerical Analysis*, 24(6):1335–1351, 1987.

[3] E. L. Allgower, K. Bohmer, F. A. Potra, and W. C. Rheinboldt. A mesh-independence principle for operator equations and their discretizations. *SIAM Journal on Numerical Analysis*, 23(1):160–169, 1986.

[4] K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Elsevier Applied Science, 1979.

[5] R.H. Brooks and A.T. Corey. Hydraulic properties of porous media. Technical report, Colorado State University, 1964.

[6] Z. Chen, G. Huan, and Y. Ma. *Computational Methods for Multiphase Flows in Porous Media*. Computational Science and Engineering. Society of Industrial Applied Mathematics, 2006.

[7] A.T. Corey. The interrelation between gas and oil relative permeabilities. *Producers Monthly*, 19(November):38–41, 1954.

[8] P. Deuflhard. *Newton Methods for Nonlinear Problems; Affine Invariance and Adaptive Algorithms*. Springer, 2004.

[9] P. Deuflhard and G. Heindl. Affine invariant convergence theorems for newton's method and extensions to related methods. *SIAM Journal on Numerical Analysis*, 16(1):1–10, 1979.

[10] Peter Deuflhard and Florian A Potra. Asymptotic mesh independence of newton-galerkin methods via a refined mysovskii theorem. *SIAM Journal on Numerical Analysis*, 29(5):1395–1412, 1992.

[11] T. Ertekin, Jamal H. Abou-Kassem, and Gregory R. King. *Basic Applied Reservoir Simulation*. Society of Petroleum Engineers, first edition, 2001.

[12] Qing Fang, Takuya Tsuchiya, and Tetsuro Yamamoto. Finite difference, finite element and finite volume methods applied to two-point boundary value problems. *Journal of Computational and Applied Mathematics*, 139(1):9 – 19, 2002.

[13] Geoquest, Schlumberger, http://www.sis.slb.com/content/software/simulation/eclipse-blackoil.asp. *Eclipse 100 Technical Description 2000A*.

[14] J. R. Giles. *Introduction to the Analysis of Normed Linear Spaces*. Cambridge University Press., 2000.

[15] W.D. Gropp, D.K. Kaushik, D.E. Keyes, and B.F. Smith. High performance parallel implicit CFD. *Parallel Computing*, 27:337–362, 2001.

[16] Patrick Jenny, Hamdi A. Tchelepi, and Seong H. Lee. Unconditionally convergent nonlinear solver for hyperbolic conservation laws with s-shaped flux functions. *Journal of Computational Physics*, 228(20):7497 – 7512, 2009.

[17] L. V. Kantorovich and G. P. Akilov. *Functional analysis in normed spaces*. Pergamon Press, first edition, 1964.

[18] D.E. Keyes. Terascale implicit methods for partial differential equations. *Proceedings of The Barrett Lectures, AMS Contemporary Mathematics, Providence*, 306:29–84, 2002.

[19] F. Kwok and H. Tchelepi. Potential-based reduced newton algorithm for nonlinear multiphase flow in porous media. *Journal of Computational Physics*, 227:706–727, 2007.

[20] Felix Kwok and Hamdi Tchelepi. Convergence of implicit monotone schemes with applications in multiphase flow in porous media. *SIAM Journal of Numerical Analysis*, 46(1):2662 – 2687, 2008.

[21] L.W. Lake. *Enhanced Oil Recovery*. Prentice Hall, first edition, 1996.

[22] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics., 2001.

[23] M. Muskat and M.W. Meres. The flow of heterogeneous fluids through porous media. *Physics*, 7(9):346–363, 1936.

[24] P.F. Naccache. A fully-implicit thermal reservoir simulator. In *Proceedings of The 14th SPE Symposium on Reservoir Simulation, Dallas*, pages 97–103. SPE, June 1997.

[25] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. Academic Press, first edition, 1970.

[26] J.M. Ortega and W.C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*. New York Academic Press, 1970.

[27] Xiaochen Wang and Hamdi A. Tchelepi. Trust-region based solver for nonlinear transport in heterogeneous porous media. *Journal of Computational Physics*, 253(0):114 – 137, 2013.

[28] Martin Weiser, Anton Schiela, and Peter Deuflhard. Asymptotic mesh independence of newton's method revisited. *SIAM journal on numerical analysis*, 42(5):1830–1845, 2005.

[29] R.M. Younis, H.A. Tchelepi, and K. Aziz. Adaptively localized continuation-newton method–nonlinear solvers that converge all the time. *SPE Journal*, 15(2):526 – 544, 2010.

APPENDIX A

INFINITE DIMENSIONAL NEWTON'S METHOD

Infinite dimensional Newton's method would allow us to obtain higher accuracy by eliminating spacial discretization and investigate asymptotic mesh independence principle mentioned before. As for a discrete system, objective is to solve following equation in Banach spaces:

$$R(S(x)) = 0; \tag{A.0.1}$$

Operator for this approach remains in the same form:

$$S^{\nu+1}(x) = G\left(S(x)\right) = S^{\nu}(x) - \lambda \left[R'(S^{\nu}(x))\right]^{-1} R\left(S^{\nu}(x)\right) \tag{A.0.2}$$

with damping factor $\lambda \in (0, 1]$, which should be calculated to make G contractive, i.e.

$$\Theta = \|G'(S(x))\|_{\infty} < 1$$

The residual equation for the infinite dimensional method has following form:

$$R(S(x)) = S(x) - c + \Delta t \frac{d}{dx} f(S(x)); \tag{A.0.3}$$

where $c > 0$ is a constant representing old saturation state, and $f(S(x))$ is positive, differentiable and monotonically increasing function. Fréchet derivative of the residual equation is

$$R'(S(x))\delta(x) = \delta(x) + \Delta t \frac{d}{dx} \left[f'(S(x))\delta(x)\right] \tag{A.0.4}$$

Fréchet operator acting on residual equation $R(S(x))$ is formulated as:

$$[R'(S(x))]^{-1} \cdot R(S(x)) = -\frac{1}{\Delta t f'(S(x)) e^{\frac{1}{\Delta t}\int_0^x \frac{1}{f'(S(u))}du}} \cdot \int_0^x e^{\frac{1}{\Delta t}\int_0^u \frac{1}{f'(S(\mu))}d\mu} R(S(u))du \quad \text{(A.0.5)}$$

After introducing the following notation $\beta(x) = e^{\frac{1}{\Delta t}\int_0^x \frac{1}{f'(S(u))}du}$, we differentiate the operator $G(S(x))$.

$$\begin{aligned} G'(S(x))\delta(x) = \delta(x) + \frac{\lambda}{\Delta t}\Bigg[& \frac{1}{\beta(x)f'(S(x))}\int_0^x \Bigg[\beta(u)R(u)\left(\frac{1}{\Delta t}\int_0^u -\frac{f''(S(\nu))}{f'(S(\nu))^2}\delta(\nu)d\nu\right) \\ & + \beta(u)\delta(u) + \Delta t\beta(u)\frac{d}{du}(f'(S(u))\delta(u))\Bigg]du \\ & - \frac{\beta(x)f''(x)\delta(x) + \beta'(x)\delta(x)f'(x)}{(\beta(x)f'(x))^2}\int_0^x \beta(u)R(u)du\Bigg] \end{aligned}$$

We want to find operator norm (or an upper bound), defined as in [14], for:

$$\Theta = \|G'(S(x))\delta(x)\|_{op} = sup\frac{\|G'(S(x))\delta(x)\|}{\|\delta(x)\|} = \|G'(S(x))\|_\infty < 1$$

In other words, we need to extract all δ values from expression, which can be further canceled by definition of operator norm. However, this operation is not straightforward, due to the signs of fractional flow values, etc.

APPENDIX B

THE MODIFIED NEWTON-KANTOROVICH THEOREM

This appendix presents modified Newton - Kantarovich's theory, developed by Deuflhard - Heindl [9], which gives guaranteed convergence in Banach spaces of Newton-like method for comparable initial guesses.

Newton's method for this theorem is modified to:

$$x_\nu = x_\nu - [A(x_k)]^{-1} R(x_\nu), \;\; \nu = 0, 1, 2..., \tag{B.0.6}$$

where $A(x)$ denotes an invertible, bounded linear operator, which is an approximation to $R'(x)$ at the same time. For example, for single cell problem this operator is equal to:

$$A(x) = \lambda(x) \left[1 + \Delta t f'(x)\right] \tag{B.0.7}$$

Formulation of the theorem states that if following conditions hold, convergence of the sequence is guaranteed:

$$\begin{aligned}
&\left\|[A(x_0)]^{-1}(R'(x) - R'(y))\right\| \leqslant K\|x - y\|, \;\; K > 0 \\
&\left\|[A(x_0)]^{-1}(A(x) - A(x_0))\right\| \leqslant L\|x - x_0\| + l, \;\; L \geqslant 0, l \geqslant 0 \\
&\left\|[A(x_0)]^{-1}(R(x) - A(x))\right\| \leqslant M\|x - x_0\| + m, \;\; M \geqslant 0, m \geqslant 0 \\
&l + m < 1, \;\; \sigma = max\left(1, \frac{L + M}{K}\right) \\
&\left\|[A(x_0)]^{-1} R(x_0)\right\| \leqslant \eta, \;\; h = \frac{\sigma K \eta}{(1 - l - m)^2} \leqslant \frac{1}{2}.
\end{aligned}$$

For cases when standard Newton's method fails, it is possible to find a damping factor λ,

such that conditions of the theorem are satisfied. Also, if that is the case, convergence will be at least quadratic.

Application of this approach is quite restrictive, because of requirement for a sufficiently close initial guess.