Quantitative Analysis of Nonlinear Multifidelity Optimization for Inverse Electrophysiology

Fatemeh Chegini, Alena Kopaničáková, Martin Weiser, Rolf Krause

Abstract The electric conductivity of cardiac tissue determines excitation propagation and is important for quantifying ischemia and scar tissue and for building personalized models. Estimating conductivity distributions from endocardial mapping data is a challenging inverse problem due to the computational complexity of the monodomain equation, which describes the cardiac excitation.
For computing a maximum posterior estimate, we investigate different optimization approaches based on adjoint gradient computation: steepest descent, limited memory BFGS, and recursive multilevel trust region methods using mesh hierarchies or heterogeneous model hierarchies. We compare overall performance, asymptotic convergence rate, and pre-asymptotic progress on selected examples in order to assess the benefit of our multifidelity acceleration.

1 Introduction

Reliable cardiac excitation predictions depend not only on accurate geometric and physiological models, usually formulated as PDEs, and our ability to solve those faithfully, but also on the model's correct parameterization. One critical parameter is

Fatemeh Chegini
Zuse Institute Berlin, Germany, and Center for Computational Medicine in Cardiology, Università della Svizzera italiana, Switzerland, e-mail: chegini@zib.de

Alena Kopaničáková
Euler Institute, Università della Svizzera italiana, Switzerland, e-mail: alena.kopanicakova@usi.ch

Martin Weiser
Zuse Institute Berlin, Germany, e-mail: weiser@zib.de

Rolf Krause
Center for Computational Medicine in Cardiology, Euler Institute, Università della Svizzera italiana, Switzerland, e-mail: rolf.krause@usi.ch

the tissue conductivity. Its correct identification from measurement data can provide valuable information about the location and size of scars, which would be beneficial for diagnosis and treatment of several heart diseases [13].

One approach to parameter identification in electrocardiography is minimizing the mismatch between simulated and measured voltages on the heart's inner surface [28]. This inverse problem can be formulated as a PDE constrained optimization problem and has, e.g., been addressed by using BFGS for a reduced problem formulation [32]. A related general framework for multilevel parameter optimization can be found in, e.g., [23].

Solving this optimization problem is, however, a major computational challenge, since the forward models describing the electrical excitation of the heart exhibit very different temporal and spatial scales and therefore require the use of fine meshes and short time steps. Together with a considerable number of optimization iterations, the resulting computational complexity is a major hurdle for widespread practical application. Consequently, several attempts have been made to reduce the computational effort, including model reduction by proper orthogonal decomposition and empirical interpolation [33], Gaussian process surrogate models [11], and topological derivative formulations [2].

A nonlinear multilevel approach based on heterogeneous model hierarchies has recently been proposed by the authors [6]. In the present study, we analyze the performance benefits and relative merits of different hierarchies quantitatively, and obtain insights concerning the behaviour of nonlinear multilevel approaches applied to the inverse problem at hand.

The remainder of the paper is organized as follows. Mathematical models for cardiac electrophysiology are briefly recalled in Section 2, while Section 3 formalizes the inverse problem under consideration. In Section 4, the recursive multilevel trust-region (RMTR) method and the model hierarchies are described. Section 5 contains the numerical results for single-level trust-region, RMTR with multigrid, and RMTR with heterogeneous model hierarchies, using limited memory BFGS.

2 Electrophysiological models

Excitation of cardiac tissue occupying the domain $\Omega \subset \mathbb{R}^d$ in terms of the transmembrane voltage v between intracellular and extracellular domain is usually described by the bidomain model or its monodomain and eikonal simplifications [9]. For simplicity, we will consider only monodomain and eikonal models here.

The monodomain system consists of a nonlinear parabolic reaction-diffusion equation for the transmembrane voltage $v : \Omega \to \mathbb{R}$ and a system of ordinary differential equations (ODEs) describing the dynamics of the ion channels, which regulate the transmembrane current, in terms of gating variables $w : \Omega \to \mathbb{R}$:

$$\begin{aligned} \operatorname{div}(\sigma \nabla v) &= \chi(C_m \dot{v} + I_{\text{ion}}(v, w)) && \text{in } \Omega \times [0, T] \\ \dot{w} &= f(v, w) && \text{in } \Omega \times [0, T] \\ \boldsymbol{n}^T \sigma \nabla v &= 0 && \text{on } \partial\Omega \times [0, T] \\ v|_{t=0} &= v_0 && \text{in } \Omega \\ w|_{t=0} &= w_0 && \text{in } \Omega. \end{aligned} \tag{1}$$

Here, $\boldsymbol{n}$ is the unit outer normal vector to Ω, σ a symmetric positive definite conductivity tensor, χ the membrane surface area per unit volume, and C_m the membrane capacity per unit area. I_{ion} denotes the transmembrane current density and f the gating dynamics, both defined by an electrophysiological membrane model (2).

Many different membrane models have been developed [19, 17]. Here, we use the modified Fitzhugh-Nagumo (FHN) model by [31],

$$\begin{aligned} I_{\text{ion}}(v, w) &= \eta_0 v \left(1 - \frac{v}{v_{\text{th}}}\right)\left(1 - \frac{v}{v_{\text{pk}}}\right) + \eta_1 v w \\ f(v, w) &= \eta_2 \left(\frac{v}{v_{\text{pk}}} - \eta_3 w\right), \end{aligned} \tag{2}$$

with positive coefficients $\eta_0, \eta_1, \eta_2, \eta_3, v_{th}, v_{pk}$. In particular, peak and threshold potential are given by $v_{\text{pk}} > v_{\text{th}}$, respectively.

Eikonal models derived from bidomain or monodomain models [7, 29, 8] consider only the activation time $u(x)$ of the tissue at a particular spatial position x, and recover the transmembrane voltage by the travelling wave ansatz

$$v(x, t) = v_m(t - u(x)), \tag{3}$$

which depends on some fixed activating front shape v_m of usually hyperbolic tangent or sigmoid structure. This ansatz results in a nonlinear elliptic equation for the activation time,

$$c_0 \sqrt{\nabla u \cdot \sigma \nabla u} - \nabla \cdot (\sigma \nabla u) = \tau_m \qquad \text{on } \Omega, \tag{4}$$

where c_0 and τ_m are parameters used for fitting the eikonal model to mono- or bidomain models.

As the eikonal equation is stationary and activation times are significantly smoother than the transmembrane voltage, eikonal solutions can be obtained much faster and on coarser grids than monodomain solutions. Nevertheless, they are a rather good approximation of the more involved models in many cases.

3 Inverse problem of Conductivity Identification

Here we turn to the prototypical inverse problem of estimating a scalar conductivity $\sigma \in H^1(\Omega)$ from N_σ voltages $\hat{v}_i$ given at disjoint open surface patches $\Gamma_i \subset \partial\Omega$

by minimizing the mismatch between simulated voltages $v|_{\Gamma_i}$ and measurements. Writing $\Gamma = \bigcup_i \Gamma_i$, the resulting optimization problem reads

$$\begin{aligned} \min_{v,\sigma} \hat{J}(v,\sigma) &= \frac{1}{2}\|v - \hat{v}\|^2_{L^2(\Gamma\times[0,T])} + R(\sigma,\beta) \\ \text{subject to} \quad & C(v,\sigma) = \mathbf{0} \\ & \sigma \in \mathcal{F} = \{s \in L^2(\Omega) \mid \sigma_{\min} \le s \le \sigma_{\max}\}. \end{aligned} \tag{5}$$

$C(v,\sigma)$ is the monodomain model (1), and $\sigma_{\min}$ and $\sigma_{\max}$ are the lower and upper bounds of the conductivity. As the problem is ill-posed, a regularization term R is added [30] in order to reduce high-frequent solution components amplified by measurement noise. Here, we choose $R(\sigma,\beta) = \frac{1}{2}\|\beta_1(\sigma - \overline{\sigma})\|^2_{L^2(\Omega)} + \frac{1}{2}\|\beta_2\nabla\sigma\|^2_{L^2(\Omega)}$, where $\overline{\sigma}$ is an a priori reference conductivity. The regularization parameters β_i can be determined, e.g., by the L-curve method [5] or Morozov's discrepancy principle. For a more detailed discussion of modeling aspects we refer to [6].

Reduced problem

In order to avoid a large 4D discretization of the space-time problem resulting from the first order necessary optimality conditions, we resort to the reduced problem by eliminating the transmembrane voltage v explicitly as $v(\sigma)$ satisfying $C(v(\sigma),\sigma) = 0$, and obtain

$$\begin{aligned} \min_{\sigma\in H^1(\Omega)} \quad & J(\sigma) = \hat{J}(v(\sigma),s) \\ \text{subject to} \quad & \sigma \in \mathcal{F}. \end{aligned} \tag{6}$$

This bound-constrained problem can then be solved by gradient type algorithms such as steepest descent or quasi-Newton methods. The gradient of the reduced objective J with respect to σ can be obtained efficiently by solving the adjoint equation

$$\begin{aligned} -\chi C_m\dot{\lambda} &= \operatorname{div}(\sigma\nabla\lambda) - \chi I_{\text{ion},v}(v,w)\lambda - f_v(v,w)\eta \\ -\dot{\eta} &= \chi I_{\text{ion},w}(v,w)\lambda + f_w(v,w)\eta \end{aligned} \tag{7}$$

with terminal and boundary conditions

$$\begin{aligned} \lambda(T) = 0, \quad \eta(T) &= 0 \\ \boldsymbol{n}^T\sigma\nabla\lambda &= 0 && \text{on } (\partial\Omega\backslash\Gamma)\times[0,T] \\ \boldsymbol{n}^T\sigma\nabla\lambda &= \hat{v} - v && \text{on } \Gamma\times[0,T] \end{aligned}$$

backwards in time and then computing

$$\nabla J = \int_0^T \nabla\lambda^T\nabla v\, dt + \nabla R. \tag{8}$$

Since the state v, η enters as data into the adjoint equation, the whole 4D trajectory still needs to be stored. This can be done efficiently by error-controlled lossy data compression [14]. When using the eikonal equation for describing cardiac excitation, the reduced gradient ∇J can be computed analogously. Conveniently, the adjoint equation is then again a single and much simpler stationary equation.

Discretization

For the spatial discretization of the conductivity σ, the transmembrane voltage v, the gating variables w, and the adjoint states λ and η, and the activation time u we employ standard linear finite elements on a simplicial grid covering the domain Ω. The time integration is done by a common equidistant implicit-explicit Euler scheme with operator splitting for both the monodomain problem (1) and adjoint equation (7).

Denoting by $\boldsymbol{x} \in \mathbb{R}^N$ the coefficient vector of the conductivity σ, we obtain thus, with a slight abuse of notation, the discretized version of (6) as

$$\begin{aligned} &\min_{\boldsymbol{x}\in\mathbb{R}^N} \; J(\boldsymbol{x}) \\ &\text{subject to} \quad \boldsymbol{x} \in \mathcal{F}. \end{aligned} \tag{9}$$

Due to the use of Lagrangian finite elements, the continuous feasible set $\mathcal{F}$ for σ translates into component-wise bounds on $\boldsymbol{x}$, such that (9) is again a bound-constrained problem.

4 Multilevel Quasi-Newton Trust-region Method

In this section, we discuss how to minimize (9) using a multilevel solution strategy, namely the recursive multilevel trust-region (RMTR) method [15]. The RMTR method combines the global convergence properties of the trust-region method with the efficiency of multilevel methods. In this work, we consider three different approaches for obtaining the multilevel hierarchy: i) multi-resolution, ii) multi-model, and iii) combined (multi-resolution and multi-model) approach.

Quasi-Newton trust-region method

A trust-region method (TR) is an iterative method, which generates a sequence $\{\boldsymbol{x}_i\}$ of iterates converging to a first-order critical point [10]. At each iteration i, the TR method approximates the objective function J by a quadratic model

$$m_i(\boldsymbol{x}_i + \mathbf{p}) = J(\boldsymbol{x}_i) + J'(\boldsymbol{x}_i)\mathbf{p} + \frac{1}{2}\mathbf{p}^T\mathbf{H}_i\mathbf{p}$$

around the current iterate $\boldsymbol{x}_i$. For the Hessian approximation $\mathbf{H}_i$ we employ a memory-efficient quasi-Newton approach known as the L-BFGS(m) [4, 3], where only the m most recent gradients are taken into account in order to update the Hessian $\mathbf{H}_i \approx J(\boldsymbol{x}_i)''$ recursively using a rank-two update formula [24]. For $m \ll n$, significantly less storage is needed compared to dense Hessian approximations used in [32].

Being based on a Taylor-like approximation, the model m_i is considered to be an adequate representation of the objective J only in a certain region, called the trust-region. The trust-region $\mathcal{B}_i := \{\boldsymbol{x}_i + \mathbf{p} \in \mathbb{R}^n \mid \|\mathbf{p}\| \leq \Delta_i\}$ is defined around the current iterate, with a size prescribed by the trust-region radius $\Delta_i > 0$ and a shape defined by the choice of norm. Here, we employ the maximum norm $\|\cdot\|_\infty$, which simplifies the step computation in bound-constrained problems compared to the Euclidean norm. The trial step $\mathbf{p}_i$ is determined by solving the constrained minimization problem

$$\begin{aligned} \min_{\mathbf{p}_i \in \mathbb{R}^n} \; m_i(\boldsymbol{x}_i + \mathbf{p}_i) \quad \text{subject to} \quad & \boldsymbol{x}_i + \mathbf{p}_i \in \mathcal{F}, \\ & \|\mathbf{p}_i\|_\infty \leq \Delta_i. \end{aligned} \tag{10}$$

The first constraint in (10) ensures the feasibility of the iterates throughout the solution process, while the second constraint restricts the size of the trial step $\mathbf{p}_i$. Both constraints are defined component-wise, such that (10) is a bound-constrained problem with easily computable bounds.

To ensure global convergence, it is sufficient to solve the trust-region subproblems (10) approximately, such that an approximate solution $\mathbf{p}_i$ of (10) satisfies the so called sufficient decrease condition (SDC), see [10]. An obtained step $\mathbf{p}_i$ is accepted, if the actual decrease in the objective, $J(\boldsymbol{x}_i) - J(\boldsymbol{x}_i + \mathbf{p}_i)$, agrees sufficiently well with the predicted decrease $m_i(\boldsymbol{x}_i) - m_i(\boldsymbol{x}_i + \mathbf{p}_i)$. This is quantified in terms of the trust-region ratio

$$\rho_i = \frac{J(\boldsymbol{x}_i) - J(\boldsymbol{x}_i + \mathbf{p}_i)}{m_i(\boldsymbol{x}_i) - m_i(\boldsymbol{x}_i + \mathbf{p}_i)}. \tag{11}$$

If ρ_i is close to unity, there is a good agreement between the objective J and the model m_i and it is therefore safe to accept the step $\mathbf{p}_i$. More precisely, the step $\mathbf{p}_i$ is accepted, only if $\rho_i > \eta_1$, where $0 < \eta_1 < 1$. In addition, the trust-region radius has to be adjusted accordingly.

Remark 1 It is important to update the approximation $\mathbf{H}_i$ even if the trial step $\mathbf{p}_i$ is rejected, since the rejection might indicate that the current $\mathbf{H}_i$ is not an adequate approximation of the true Hessian $J''(\boldsymbol{x}_i)$.

Remark 2 Using the L-BFGS method, the implementation of the trust-region algorithm can be realized in a matrix-free way. The operations involving $\mathbf{H}_i$, or its inverse $(\mathbf{H}_i)^{-1}$, can be implemented using the approach proposed in [25] and the two-loop recursion algorithm developed in [27], respectively.

Recursive multilevel trust-region method

The computational cost of the trust-region method is dominated by evaluating the objective in (11) and the reduced gradient $J'(\boldsymbol{x}_i)$ via (8), which incurs the solution of at least two parabolic equations per accepted trial step. Reducing the computational effort requires a decrease of the number of steps, which in turn is only possible if the quadratic models m_i of the objective are replaced or complemented by models J_i^- that approximate J on larger trust-regions, but are nevertheless significantly cheaper to minimize than the original objective. Approximate models that we consider here are (i) the monodomain equation on coarser grids and (ii) an eikonal model.

These models can be defined on the same discretization, i.e. approximation space, for σ of size n, or on a coarser one of size $n^- < n$, in which case a transfer between the original problem and the model is necessary. This affects the transfer of the current iterate $\boldsymbol{x}_i$ (projection) and the gradient ∇J (restriction) to the model J_i^-, and the transfer of the model's minimizer back to the original problem (prolongation). Note that, if the models J_i^- are formulated on the same discretization, all these transfers are trivial. Otherwise, we define both the prolongation $\mathbf{I} : \mathbb{R}^{n^-} \to \mathbb{R}^n$ and the projection $\mathbf{P} : \mathbb{R}^n \to \mathbb{R}^{n^-}$ as pseudo-L^2-projection, as proposed in [18] and successfully applied in [21]. We assemble these transfer operators using the library MOONoLith [22]. As usual, the restriction is defined as the adjoint of the prolongation, i.e. $\mathbf{R} = \mathbf{I}^T$.

Naturally, we intend J_i^- to approximate J well. Therefore, we enforce first-order consistency between both models i.e., the gradients of both models shall coincide locally as far as possible. As common for nonlinear multilevel schemes, the model functions J_i^- can be defined in terms of some computationally cheaper/coarse approximation j^- of the objective J by means of the additive approach [26] as

$$J_i^-(\boldsymbol{x}^-) = j^-(\boldsymbol{x}^-) + (\boldsymbol{x}^- - \mathbf{P}\boldsymbol{x}_i)^T \underline{(\mathbf{R}\nabla J(\boldsymbol{x}_i) - \nabla j^-(\mathbf{P}\boldsymbol{x}_i))}. \tag{12}$$

Alternatively, we can exploit a multiplicative approach [1, 20] and define the models

$$J_i^-(\boldsymbol{x}^-) = \beta(\boldsymbol{x}_i, \boldsymbol{x}^-) j^-(\boldsymbol{x}^-) \tag{13}$$

with

$$\beta(\boldsymbol{x}, \boldsymbol{x}^-) = \frac{J(\boldsymbol{x})}{j^-(\mathbf{P}\boldsymbol{x})} + (\boldsymbol{x}^- - \mathbf{P}\boldsymbol{x})^T \underline{\left(\frac{1}{j^-(\mathbf{P}\boldsymbol{x})} \mathbf{R}\nabla J(\boldsymbol{x}) - \frac{J(\boldsymbol{x})}{(j^-(\mathbf{P}\boldsymbol{x}))^2} \nabla j^-(\mathbf{P}\boldsymbol{x}) \right)}.$$

Both approaches employ a so called coupling term (underlined), which takes into account the difference between restricted original gradient $\mathbf{R}\nabla J(\boldsymbol{x}_i)$ and initial coarse gradient $\nabla j^-(\mathbf{P}\boldsymbol{x}_i)$. The use of this coupling term guarantees that the first-order behavior of J and J^- is locally coherent in the neighborhood of $\boldsymbol{x}_i$ and $\mathbf{P}\boldsymbol{x}_i$, respectively [26].

At each iteration i, the trial step $\mathbf{p}_i \in \mathbb{R}^{n^-}$ is obtained either by approximately solving the quadratic trust-region subproblem (10) or the coarse subproblem

$$\min_{\mathbf{p}\in\mathbb{R}^{n^-}} \; J_i^-(\mathbf{P}\boldsymbol{x}_i + \mathbf{p}), \quad \text{subject to} \quad \boldsymbol{x}_i + \mathbf{I}\mathbf{p} \in \mathcal{F}, \quad \|\mathbf{I}\mathbf{p}_i\|_\infty \leq \Delta_i. \tag{14}$$

As common for trust-region methods, it is not necessary to solve the problem (14) exactly. Indeed, it is sufficient that an approximate minimizer $\mathbf{p}$ of (14) satisfies the SDC condition. Here, we solve the nonlinear problem (14) iteratively by employing few steps of the trust-region method. This gives rise to a recursive multilevel trust-region (RMTR) scheme [16]. A line search-based alternative would be the multilevel model correction (MMC) method [23].

Potentially, we can utilize a hierarchy of multiple coarse models $\{j^l\}_{l=1}^L$, where $L > 1$, which gives rise to a truly multilevel method. In this work, we obtain models $\{j^l\}_{l=1}^L$ by exploring the following alternatives:

1. **Multi-resolution:** We uniformly coarsen finite element grids (by factor of 2) in order to discretize the monodomain equation entering the reduced objective J. Consequently, the coarse-level models are computationally cheaper to optimize. *Note that a certain mesh resolution is required to reasonably resolve the monodomain model, such that mesh coarsening is limited.*
2. **Multi-model:** The eikonal model is used instead of monodomain on coarser grids. This model is significantly cheaper and a better global approximation model for the monodomain model compared to the standard quadratic model.
3. **Combined:** Combinations of multi-resolution and multi-model variants are also possible. For instance, one can obtain a hierarchy of models $\{j^l\}_{l=1}^{l=L}$ by first coarsening the spatial-resolution and then changing the model complexity.

At the end, we highlight the fact that the overall efficiency of the multilevel algorithm is determined by how many times the respective coarse and fine level models are minimized. For instance, in the multi-resolution approach, it is crucial to alternate between both models, such that the components of the error associated with a given level are effectively eliminated.

5 Numerical results

Here, we focus on numerical results for different algorithmic configurations on a simple 2D geometry. For the numerical tests, different synthetic transmembrane voltage data $\hat{v}$ have been created by simulations on a finer mesh. For illustration, we also present some reconstruction results for scar tissue on a 3D ventricular geometry. For a more detailed discussion of reconstruction quality we refer to [6].

5.1 Patient-specific geometry

As an example from clinical practice, we use the ventricular geometry of a patient with a nontransmural scar located on the left endocardium. In Fig. 1, scar tissue is shown in blue, sparse endocardial measurement locations by yellow spheres, and the reconstructed conductivity color-coded on the right. The reconstruction quality depends on the quantity and location of available cardiac mapping data. Due to the stability of excitation propagation, reliable results can in general only be expected in the vicinity of measurement locations. In this case, the small number of measured data on the left endocardium is not enough to reconstruct the scar shape accurately.

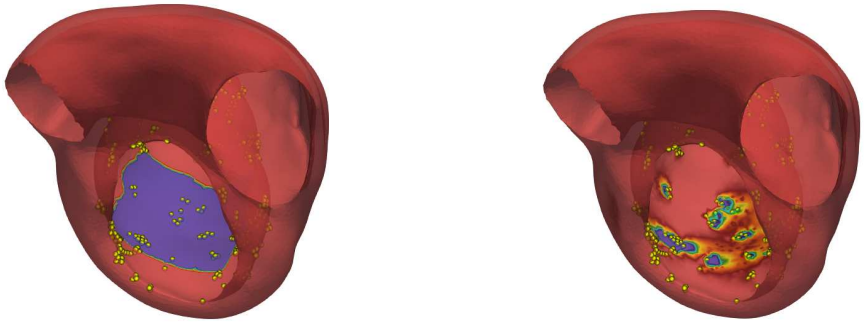

Fig. 1: *Left:* Target conductivity with marked measured data. *Right:* Solution with marked measured data.

5.2 Convergence study

In this section, we compare the convergence behavior of single-level trust-region methods with several RMTR variants on a simpler idealized 2D cross-section of a left ventricle. We use L-BFGS(m) with $m = 1$ or $m = 8$ secant pairs, and a termination criterion $\|\mathcal{P}(\boldsymbol{x} - \nabla J(\boldsymbol{x})) - \boldsymbol{x}\| < 10^{-4}$ based on the projected gradient expressed in terms of an orthogonal projection $\mathcal{P}$ onto the feasible set $\mathcal{F}$. The arising quadratic trust-region subproblems (10) are solved using the MPRGP method [12]. The RMTR method is configured with additive coarse level models (12) for multi-resolution variants, while multiplicative coarse level models (13) are employed for multi-model variants. Solution strategies are implemented as part of the open-source library UTOPIA [34], while the implementation of inverse problems, including monodomain and eikonal models, is part of our framework HEART. All simulations have been run using 10 nodes (XC50, 12 cores) of the Piz Daint supercomputer (CSCS, Switzerland).

To provide a more robust insight, four different sets of simulated measurement data $\hat{v}$ have been used: generated with monodomain on a finer mesh, with additional Gaussian noise, with slightly changed membrane area χ per volume, and generated with eikonal on a finer grid. We provide averaged iteration counts and run times in Tab. 1 for single level trust-region with different L-BFGS memory size, and for RMTR with two or three levels of monodomain on coarser grids or eikonal models.

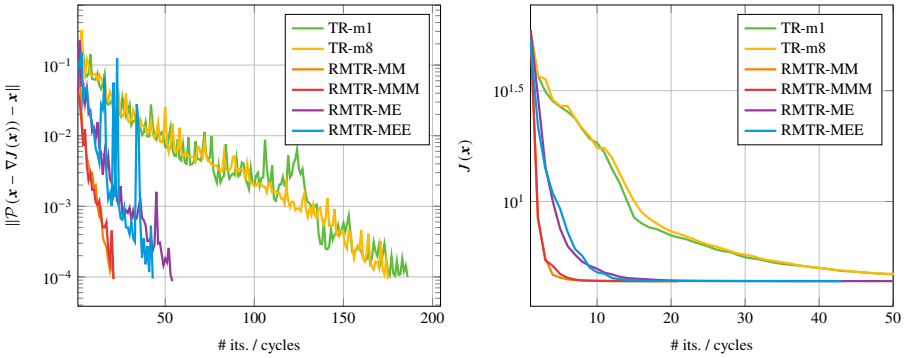

Fig. 2: The convergence history in terms of the projected gradient (left) and the objective function J. The measurement data were generated using the monodomain model on a finer mesh.

The convergence results suggest that both monodomain multigrid and heterogeneous monodomain-eikonal multilevel methods lead to a significant reduction of iteration count by a factor between 3 and 6. For the used grid resolution, monodomain multigrid is more effective by a factor 1.5 to 2 in reducing iteration counts. Since the coarse level subproblems are more expensive to solve, the heterogeneous multilevel approach is almost as efficient. We can also observe a slight convergence rate deterioration of the heterogeneous approach in the asymptotic phase, probably due to a less accurate Hessian approximation of the eikonal model. The three-level multigrid approach appears to be less effective than the two-level method, probably because the monodomain model deteriorates quickly for coarser grids.

6 Conclusion

Identifying tissue conductivities using monodomain models from surface measurements is computationally expensive and calls for acceleration. Multilevel methods can be effective in two ways: First, classical multigrid based on a Galerkin projection of the Hessian improves the convergence rate of steepest descent or similar smoothers, which suffer from ill-conditioning. Second, nonlinear multilevel methods aim at improving the objective reduction also in the pre-asymptotic phase, where the progress of first or second order methods is limited due to high nonlinearity.

	models	meshes	m	# its/cycles	time (minutes)
TR	mono (TR-m1)	$\mathcal{T}^3$	1	179 ± 29	149 ± 35
	mono (TR-m8)	$\mathcal{T}^3$	8	148 ± 19	127 ± 44
RMTR	mono-mono (RMTR-MM)	$\mathcal{T}^3, \mathcal{T}^2$	8	26 ± 5*	80 ± 49
	mono-mono-mono (RMTR-MMM)	$\mathcal{T}^3, \mathcal{T}^2, \mathcal{T}^1$	8	31 ± 11*	88 ± 56
	mono-eiko (RMTR-ME)	$\mathcal{T}^3, \mathcal{T}^3$	8	51 ± 6	97 ± 35
	mono-eiko-eiko (RMTR-MEE)	$\mathcal{T}^3, \mathcal{T}^3, \mathcal{T}^2$	8	47 ± 8	85 ± 32

Table 1: The average computational cost required by trust-region and RMTR method. The results are obtained by averaging over four datasets. The symbol * indicates that for one dataset the termination criterion was not satisfied within 500 cycles.

The numerical results suggest that the RMTR method used here is effective in both regimes and leads to a clear reduction of iterations. Due to the overhead of the subproblems, the reduction of run time is not as large, but still significant.

Acknowledgements Funding by EU and BMBF via the JU Euro-HPC project MICROCARD (grant agreement No 955495), Swiss National Science Foundation project ML^2 (197041), and CCMC (Fidinam, Horten), is gratefully acknowledged.

References

1. Alexandrov, N.M., Lewis, R.M.: An overview of first-order model management for engineering optimization. Optimization and Engineering **2**(4), 413–430 (2001)
2. Beretta, E., Cavaterra, C., Ratti, L.: On the determination of ischemic regions in the monodomain model of cardiac electrophysiology from boundary measurements. Nonlinearity **33**(11), 5659–5685 (2020)
3. Byrd, R., Lu, P., Nocedal, J., Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing **16**(5), 1190–1208 (1995)
4. Byrd, R., Nocedal, J., Schnabel, R.: Representations of quasi-Newton matrices and their use in limited memory methods. Mathematical Programming **63**(1-3), 129–156 (1994)
5. Calvetti, D., Lewis, B., Reichel, L.: GMRES, L-curves, and discrete ill-posed problems. BIT Numerical Mathematics **42**(1), 44–65 (2002)
6. Chegini, F., Kopaničáková, A., Krause, R., Weiser, M.: Efficient identification of scars using heterogeneous model hierarchies. EP Europace **23**, i113–i122 (2021)
7. Colli Franzone, P., Guerri, L., Rovida, S.: Wavefront propagation in an activation model of the anisotropic cardiac tissue: asymptotic analysis and numerical simulations. Journal of mathematical biology **28**(2), 121–176 (1990)
8. Colli Franzone, P., Guerri, L., Taccardi, B.: Modeling ventricular excitation: axial and orthotropic anisotropy effects on wavefronts and potentials. Mathematical biosciences **188**(1-2), 191–205 (2004)
9. Colli Franzone, P., Pavarino, L., Scacchi, S.: Mathematical cardiac electrophysiology. Springer (2014)
10. Conn, A., Gould, N., Toint, P.: Trust region methods. SIAM (2000)
11. Dhamala, J., Arevalo, H.J., Sapp, J., Horacek, M., Wu, K.C., Trayanova, N.A., Wang, L.: Spatially adaptive multi-scale optimization for local parameter estimation in cardiac electrophysiology. IEEE Transactions on Medical Imaging **36**(9), 1966–1978 (2017)

12. Dostál, Z.: Optimal quadratic programming algorithms: with applications to variational inequalities, vol. 23. Springer Science & Business Media (2009)
13. Fernández-Armenta, J., Berruezo, A., Mont, L., Sitges, M., Andreu, D., Silva, E., Ortiz-Pérez, J., Tolosana, J., de Caralt, T., Perea, R., Calvo, N., Trucco, E., Borrás, R., Matas, M., Brugada, J.: Use of myocardial scar characterization to predict ventricular arrhythmia in cardiac resynchronization therapy. EP Europace **14**(11), 1578–1586 (2012)
14. Götschel, S., Chamakuri, N., Kunisch, K., Weiser, M.: Lossy compression in optimal control of cardiac defibrillation. J. Sci. Comp. **60**(1), 35–59 (2014)
15. Gratton, S., Mouffe, M., Toint, P., Weber-Mendonca, M.: A recursive trust-region method for bound-constrained nonlinear optimization. IMA Journal of Numerical Analysis **28**(4), 827–861 (2008)
16. Gratton, S., Sartenaer, A., Toint, P.: Recursive trust-region methods for multiscale nonlinear optimization. SIAM Journal on Optimization **19**(1), 414–444 (2008)
17. Greenstein, J.L., Winslow, R.L.: An integrative model of the cardiac ventricular myocyte incorporating local control of Ca2+ release. Biophysical journal **83**(6), 2918–2945 (2002)
18. Groß, C., Krause, R.: A recursive trust-region method for non-convex constrained minimization. In: Domain Decomposition Methods in Science and Engineering XVIII, pp. 137–144. Springer (2009)
19. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its appl. to conduction and excitation in nerve. The Journal of physiology **117**(4), 500–544 (1952)
20. Kopaničáková, A.: Multilevel minimization in trust-region framework: algorithmic and software developments. Ph.D. thesis, Università della Svizzera italiana (2020)
21. Kopaničáková, A., Krause, R.: A recursive multilevel trust region method with application to fully monolithic phase-field models of brittle fracture. Computer Methods in Applied Mechanics and Engineering **360**, 112720 (2020)
22. Krause, R., Zulian, P.: A parallel approach to the variational transfer of discrete fields between arbitrarily distributed unstructured finite element meshes. SIAM Journal on Scientific Computing **38**(3), C307–C333 (2016)
23. Li, J., Zou, J.: A multilevel model correction method for parameter identification Inverse Problems **23**(5), 1759–1786, 2007
24. Liu, D., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Mathematical programming **45**(1-3), 503–528 (1989)
25. Mahidhara, D., Lasdon, L.: An SQP algorithm for large sparse nonlinear programs. Austin, MSIS Department–School of Business Administration, University of Texas (1991)
26. Nash, S.: A multigrid approach to discretized optimization problems. Optimization Methods and Software **14**(1-2), 99–116 (2000). DOI 10.1080/10556780008805795
27. Nocedal, J.: Updating quasi-Newton matrices with limited storage. Mathematics of computation **35**(151), 773–782 (1980)
28. Pullan, A., Cheng, L., Nash, M., Ghodrati, A., MacLeod, R., Brooks, D.: The inverse problem of electrocardiography. In: Comprehensive Electrocardiology, pp. 299–344. Springer (2010)
29. Pullan, A., Tomlinson, K., Hunter, P.: A finite element method for an eikonal equation model of myocardial excitation wavefront propagation. SIAM J. Appl. Math. **63**(1), 324–350 (2002)
30. Tikhonov, A.: On the solution of ill-posed problems and the method of regularization. In: Doklady Akademii Nauk, vol. 151, pp. 501–504. Russian Academy of Sciences (1963)
31. Xu, A., Guevara, M.: Two forms of spiral-wave reentry in an ionic model of ischemic ventricular myocardium. Chaos: An Interdisciplinary Journal of Nonlinear Science **8**(1), 157–174 (1998)
32. Yang, H., Veneziani, A.: Estimation of cardiac conductivities in ventricular tissue by a variational approach. Inverse Problems **31**(11), (2015)
33. Yang, H., Veneziani, A.: Efficient estimation of cardiac conductivities via POD-DEIM model order reduction. Applied Numerical Mathematics **115**, 180–199 (2016)
34. Zulian, P., Kopaničáková, A., Nestola, M.C.G., Fink, A., Fadel, N., Rigazzi, A., Magri, V., Schneider, T., Botter, E., Mankau, J., Krause, R.: Utopia: A C++ embedded domain specific language for scientific computing. Git repository. https://bitbucket.org/zulianp/utopia (2016)