Consensus ADMM for Inverse Problems Governed by Multiple PDE Models

Luke Lozenski, Umberto Villa
Washington University in St. Louis
Department of Electrical and Systems Engineering

FEniCS Conference
March 26 2021

This work was partially supported by the National Science Foundation under Grant No ACI-1550593.

Motivation

- PDEs are often dependent on unknown (or difficult to measure) parameters associated with physical systems and can be estimated via an inverse problem
- Inverse problems are often-illposed: there's not enough data to recover the parameter
- Regularization selects one solution among many possible solutions
- Non-smooth regularization reinforces certain "nice" properties in solutions: TV enforces sharp edges
- ADMM provides a natural way of splitting these inverse problems into smaller problems.
  - The subproblems related to the PDEs can be solved efficiently using INCG, which requires a smooth objective term
  - The term related to the regularization can be solved for separately using other proximal methods
- FEniCS is used for efficient discretization of these variational problems

ADMM Description

- Equality between solutions of subproblems is reinforced with a consensus term
- ADMM will only reach moderate accuracy in a few iterations and requires many following iterations for high-precision convergence[1]
- This is sufficient for most large-scale applications including
  - Machine learning
  - Continuum mechanics[2]
  - Imaging[3]

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein,Distributed Optimization andStatistical Learning via the Alternating Direction Method of Multipliers, Foundation andTrends in Machine Learning, Vol. 3, No. 1 (2010).

[2] D. Gabay and B. Mercier,A dual algorithm for the solution of nonlinear variational problems via finite element approximations, Computers and Mathematics with Applications, Vol. 2,No. 1 (1976)

[3] Y. Wang, J. Yang, W. Yin, and Y. Zhang,A New Alternating Minimization Algorithm forTotal Variation Image Reconstruction, SIAM Journal on Imaging Sciences, (2007)

Setting

Consider the minimization problem

$$\min_{m \in \mathcal{M},} \quad \mathcal{L}(m) + \mathcal{R}(m)$$

- $\mathcal{M}$ is a possibly infinite-dimensional Hilbert space.
- $\mathcal{L} : \mathcal{M} \mapsto \mathbb{R}$ is twice differentiable, may be expensive to evaluate
- $\mathcal{R} : \mathcal{M} \mapsto \mathbb{R}$ is assumed convex and non-smooth

Introduce a consensus variable $z \in \mathcal{M}$

$$\begin{aligned} \min_{m,z \in \mathcal{M},} \quad & \mathcal{L}(m) + \mathcal{R}(z), \\ s.t. \quad & m - z = 0 \end{aligned}$$

Consensus ADMM

We introduce the *augmented Lagrangian* for some $\rho > 0$

$$L_\rho(m, z, y) = \mathcal{L}(m) + \mathcal{R}(z) + \langle y, m - z \rangle + \frac{\rho}{2}||m - z||^2$$

Algorithm 1: Consensus ADMM

Begin with starting points (m^0, z^0, y^0)

while *While convergence criterion is not met* **do**

$$m^{k+1} = \operatorname{argmin}_m L_\rho(m, z^k, y^k)$$
$$z^{k+1} = \operatorname{argmin}_z L_\rho(m^{k+1}, z, y^k)$$
$$y^{k+1} = y^k + \rho(m^{k+1} - z^{k+1})$$

end

[4]

[4] L. Lozenski, U. Villa, "Consensus ADMM for Inverse Problems Governed by Multiple PDE Models", in preparation 2021

Inverse problems governed by PDEs

- Goal: Estimate a parameter m given a measurement $\boldsymbol{d} \in \mathcal{D}$ where

$$\boldsymbol{d} = \mathcal{F}(m) + \boldsymbol{e},$$

- $\mathcal{F}$ is the composition of a PDE solver and observation operator $\mathcal{B} : \mathcal{U} \to \mathcal{D}$.
- Introduce the *state variable* $u \in \mathcal{U}$ s.t.

$$\mathcal{F}(m) = B(u(m)), \; r(m, u) = 0$$

$$\min_{m \in \mathcal{M}} \quad \mathcal{J}(m) = \tfrac{1}{2} \|\mathcal{B}(u(m)) - \boldsymbol{d}\|^2 + \mathcal{R}(m)$$

For a Newton type solution method

- Using the Lagrangian formalism, gradient computation requires solving two PDEs: the forward & adjoint problems
- Each Hessian action requires solving two linearized PDEs: the incremental forward & incremental adjoint problems

The proposed consensus ADMM

Algorithm 2: The mean based scaled ADMM for parameter inversion with multiple PDEs

Let q be the number of PDEs
Begin with starting points $(\{m_i^0\}_{i=1}^q, z^0, \{y_i^0\}_{i=1}^q)$
while *While convergence criterion is not met,* $k = 1, \dots$ **do**
  for $i = 1, \dots, q$ **do**
    $m_i^{k+1} = \text{argmin}_{m_i} \frac{1}{2q}||\mathcal{F}_i(m_i) - \boldsymbol{d}_i||^2 + \frac{\rho^k}{2q}||m_i - z^k + y_i^k||^2$
  end
  Set $\bar{m} = \frac{1}{q}\sum_{i=1}^q m_i^{k+1}$ and, $\bar{y} = \frac{1}{q}\sum_{i=1}^q y_i^{k+1}$
  $z^{k+1} = \text{argmin}_z \mathcal{R}(z) + \frac{\rho}{2}||\bar{m} - z + \bar{y}||^2$
  for $i = 1, \dots, q$ **do**
    $y_i^{k+1} = y_i^k + (m_i^{k+1} - z^{k+1})$
  end
  Update ρ^{k+1} adaptively
end

hIPPYlib: Inverse Problem PYthon

- An extensible software framework for PDE-constrained determinsitic and Bayesian inverse problems
- Implements state of the art scalable adjoint based algorithms
- Built on FEniCS for discretization of PDEs and PETSc for scalable and effecient linear algebra
- Employs use of advanced structure-exploiting algorithms and approximations
- Maintains consistency with underlying infinite-dimensional problem
- Facilitates expirementation with different priors, observation operators, noise covariance models, model parameter representations, etc.

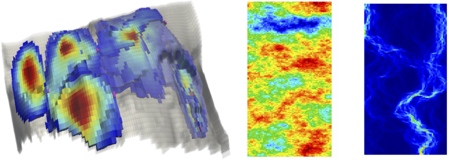
https://hippylib.github.io/[5]

[5] Villa et al., (2018). hIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems. Journal of Open Source Software, 3(30), 940, https://doi.org/10.21105/joss.00940

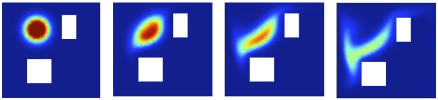
hIPPYlib features

- Friendly, compact, near-mathematical FEniCS notation to express PDE and likelihood in weak form.
- Automatic generation of efficient code.
- Scalable algorithms
  - MAP point computation
  - Low rank representaiton of posterior covariance via randomized algorithms
  - Scalable sampling from prior and posterior
  - Forward/inverse propagation of Uncertainty Quantification

Inference for reservoir modeling



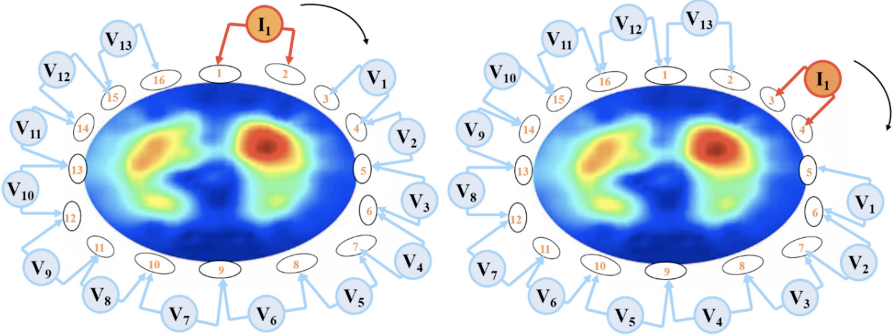
Localization of atmospheric pollution sources



Electrical Impedance Tomography(EIT)

Electrical Impedance Tomography (EIT) is an imaging modality in which

- An electrical current is introduced on the boundary of an object
- The electric potential is measured on the boundary.
- The potential measurements are used to reconstruct for conductivity



time 0 *time 2* [6]

6 https://www.mdpi.com/2077-0383/8/8/1176/htm

Formulating EIT in the continuous setting

Goal to minimize

$$\frac{1}{q}\sum_{i=1}^{q}\mathcal{L}_i(m)+\mathcal{R}(m),\ \ \mathcal{L}_i(m)=\frac{1}{2}\int_{\Gamma_i}(u_i-\boldsymbol{d}_i)^2 ds$$

The regularization used was a combination of TV and L^2
The potential u_i solves the electrostatic Maxwell equation

$$\begin{cases} -\nabla\cdot e^m\nabla u_i = 0 & x\in\Omega \\ \frac{\partial}{\partial\eta}u_i = g_i & x\in\Gamma_N^i \\ u_i = 0 & x\in\Gamma_D^i \end{cases}$$

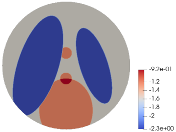
where $\sigma := e^m$ is the conductivity domain and u_i is the electric potential resulting from introducing the current g_i

Discretization

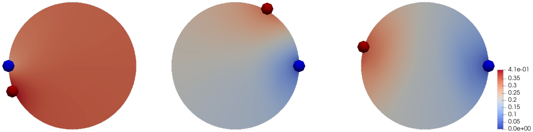
For discretization we applied the finite element method(FEM) used in FEniCS

- $\Omega = D^2$
- Coarsest mesh had 8044 degrees of freedom on $\mathcal{M}$ and $\mathcal{U}$
- Parameter updates were accomplished using the INCG algorithm in hIPPYlib
- Consensus updates were found using the PETScTAOSolver built into Fenics

Ground Truth

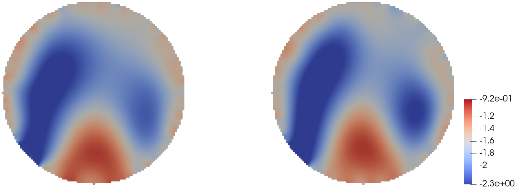


True parameter

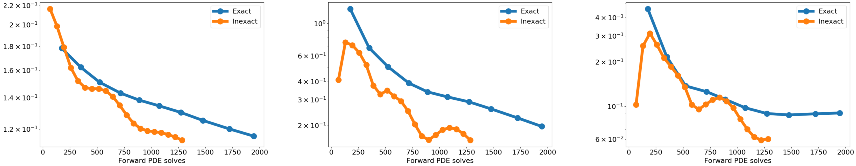


True states 1,11,16 for EIT problem with $q = 16$

H^1 reconstruction with inexact subproblem solutions



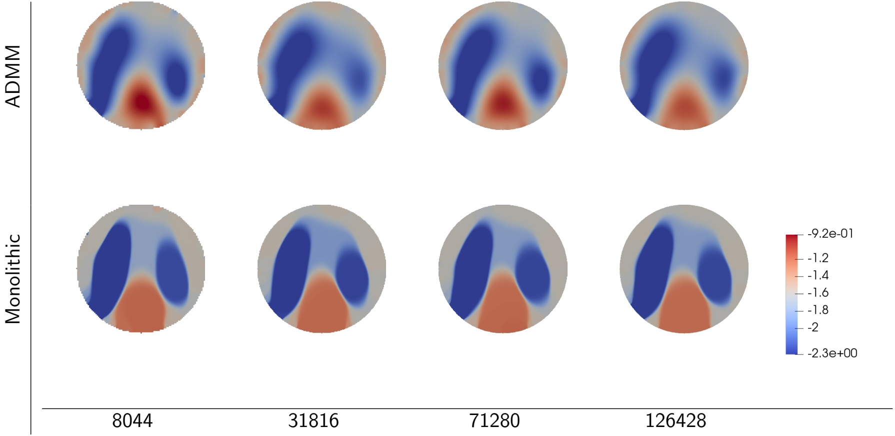
Inverted consensus for EIT problem using exact and inexact *m* solves



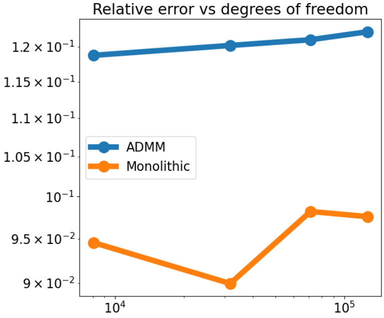
The relative error, primal and dual residuals wrt Foward PDE solves

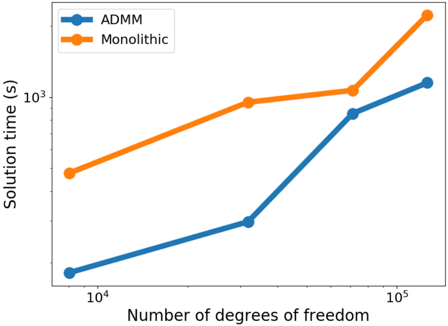
Scalability with respect to problem size

Fix $q = 16$ and sequences of uniformly refined meshes



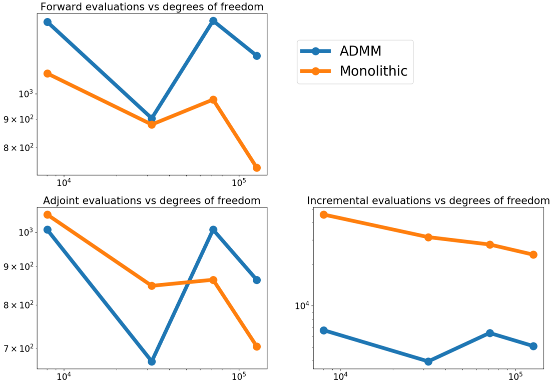
Accuracy with respect to problem size





Relative error and state misfit for ADMM and monolithic approaches vs number of degrees of freedom

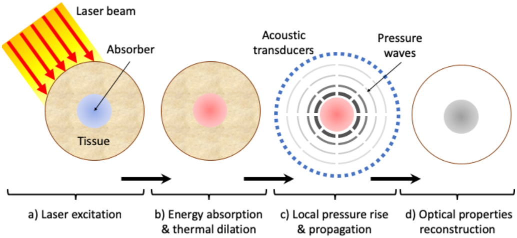
Scalability with respect to problem size (number of PDE solves)



- Similar results hold for scaling by number of forward models

Quantitative photoacoustic tomography(qPACT)

1. A fast laser pulse is sent into an object
2. Underlying material absorbs this energy generating heat and a local increase pressure distribution
3. Pressure distribution transitions into acoustic waves and measured on boundary



Formulation of the qPACT problem

We focused on reconstructing for optical properties given the initial pressure distribution

$$\text{Observation operator } d = \frac{p_0}{\Gamma} = \mu_a \phi + \boldsymbol{e}$$

Diffusion approximation to radiative transport

$$-\nabla \cdot \frac{1}{3(\mu_a + \mu_s')} \nabla \phi + \mu_a \phi = 0 \qquad x \in \Omega$$

with Robin boundary condition

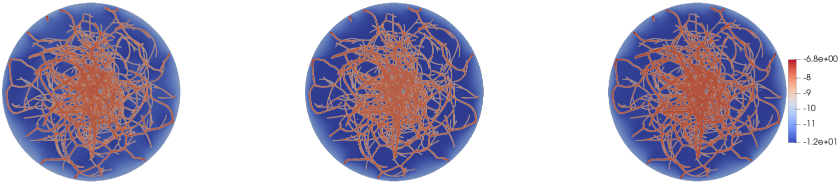
$$\frac{1}{3(\mu_a + \mu_s')} \frac{\partial \phi}{\partial \eta} + \frac{1}{2}\phi = \frac{1}{2}\phi_0 \qquad x \in \partial\Omega$$

Form the data fidelity term

$$\frac{1}{q} \sum_{i=1}^{q} \mathcal{L}_i(s, c_{thb}, \mu_s') = \frac{1}{q} \sum_{i=1}^{q} || \ln(\mu_{a,i} \phi_i) - \ln(d_i) ||^2$$

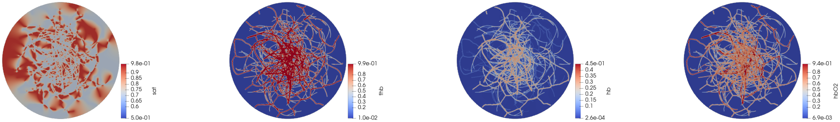
and use regularization with a mixture of Tikhonov, TV, and L1

Forward Results

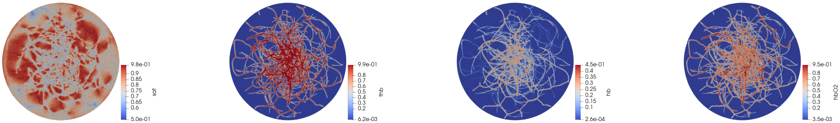


Measurements corresponding to 757, 800, 850 nm

Reconstruction Results



True s, c_{thb}, c_{hb}, and c_{hbO_2}



Reconstructed s, c_{thb}, c_{hb}, and c_{hbO_2}

Conclusions

We presented a framework for solving inverse problems governed by PDE forward models using ADMM

- ADMM is well suited for solving problems involving several large-scale PDE models with nonsmooth regularization
- ADMM solution method significantly reduced computational costs while still achieving satisfactory accuracy

In the future, we plan to improve upon this framework by

- Implementing a primal-dual solver for updating the consensus variable
- Implementing the ADMM process on several processors, with each PDE model being handled by its own set of processors.

The code and EIT example will be included in hIPPYlib

This work was partially supported by the National Science Foundation under Grant No ACI-1550593.