Speedy Gibbs Sampling for $A^T CA$ Equilibrium Systems

M. Neumayer and C. Fox

Institute of Electrical Measurement and Measurement Signal Processing,
Inffeldgasse 23/II, A-8010 Graz, Graz University of Technology

8.1.2013

Outline

- Short introduction to inverse Problems
- Introduction to $\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ systems
- Computational issues
- Fast conditional sampling
- Example on Electrical Capacitance Tomography (ECT)

Inverse Problems

Determine $\boldsymbol{x} \in \mathbb{R}^N$ from

$$\tilde{\boldsymbol{d}} = P(\boldsymbol{x}) + \boldsymbol{v}, \tilde{\boldsymbol{d}}, \boldsymbol{v} \in \mathbb{R}^M \tag{1}$$

Forward map (model) $F : \boldsymbol{x} \mapsto \boldsymbol{y}$.

$$\boldsymbol{y} = F(\boldsymbol{x}) \tag{2}$$

$$F(\boldsymbol{x}) = P(\boldsymbol{x}) \quad \forall \boldsymbol{x}, \tag{3}$$

Where does the name come from:

$$\boldsymbol{x} = F^{-1}(\tilde{\boldsymbol{d}}), \tag{4}$$

Does not work as inverse problems are ill posed.

- Regularization
- Prior knowledge

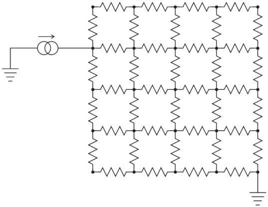
$$\rightarrow \mathscr{F}^{-1}(\tilde{\boldsymbol{d}})$$

$\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ what are they? Why are they important?

- What are they: Systems where the underlying stiffness matrix $\boldsymbol{K}$ can be expressed by the form $\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$. $\boldsymbol{C}$ is diagonal and contains the parameter (material values, etc.). $\boldsymbol{A}$ is a mapping.
  - The stiffness matrix depends linear to $\boldsymbol{C}$
  - The system itself is nonlinear.
- Why are they so important?: The appear in many inverse problems with an underlying electrical, mechanical, heat transfer, hydraulic, etc. problem (Gilbert Strang).
- More general: everything with an underlying resistor network structure $\rightarrow$ so even in finite element discretiations of pde's.

Resistor Network

$$\boldsymbol{K}_{(l,m)} = \begin{cases} -\frac{1}{R_{(l,m)}} & l \neq m \\ \sum_{k=1}^{N} \frac{1}{R_{(l,k)}} & l = m \end{cases} \tag{5}$$



$\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ decomposition:

$$C_{(l,m)} = \begin{cases} \frac{1}{R_{(l,m)}} & (l,m) \in \mathscr{D} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$A_{(l,m)} = \begin{cases} +1 & \text{'+' end of resisor } l \text{ is connected to node } m \\ -1 & \text{'−' end of resisor } l \text{ is connected to node } m \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

How to obtain a full rank system?

- Reduction of the reference node (nodal analysis)

$$\boldsymbol{Y}\boldsymbol{u} = \boldsymbol{i} \tag{8}$$

  Most common method in circuit analysis. Easy to implement - just skip the reference column from $\boldsymbol{A}$.
- Penalty method: Add some large number to the reference node. Simple, but not exact, Problem with the condition number.
- Variational formulation

$$\begin{aligned} \boldsymbol{u} &= \arg\min \boldsymbol{u}^T\boldsymbol{K}\boldsymbol{u} - \boldsymbol{i}^T\boldsymbol{u} && (9) \\ & \text{s.t.} \boldsymbol{u}^T\chi = 0 && (10) \end{aligned}$$

  $\rightarrow$ Constraint optimization problem
- Non-standard penalty method*

*We have not found a better name yet

Non-standard penalty method*

- Approach

$$\boldsymbol{u} = \arg\min \boldsymbol{u}^T \boldsymbol{K} \boldsymbol{u} - \boldsymbol{i}^T \boldsymbol{u} \quad (11)$$

$$\text{s.t.} \boldsymbol{u}^T c = 0 \quad (12)$$

with $\boldsymbol{i}^T \boldsymbol{c} \neq 0$

- Unconstraint optimization problem

$$\boldsymbol{u} = \arg\min(\boldsymbol{u}^T \boldsymbol{K} \boldsymbol{u} - \boldsymbol{i}^T \boldsymbol{u}) + \lambda \boldsymbol{u}^T \boldsymbol{c}^T \boldsymbol{c} \boldsymbol{u} \quad (13)$$

- Exact method for any value of λ (conditioning), the equation system preserves all nice properties (symmetry, poss. def.).
- $\boldsymbol{i}$ has to contain the input currents and the output current at the reference node!

*We have not found a better name yet

Solution with Green's Functions

- For all further derivations we assume that $\boldsymbol{K} = \boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ has full rank.
- Our problem

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{b} \tag{14}$$

  is often a self adjoint ($\boldsymbol{K}$ is symmetric) problem
- Solve for Green's functions $\boldsymbol{g}$
- Solve

$$\boldsymbol{K}\boldsymbol{g}_i = \boldsymbol{e}_i, \tag{15}$$

  where $\boldsymbol{e}_i$ is the i^{th} unit vector
- Solution

$$u_i = \boldsymbol{g}^T\boldsymbol{b} \tag{16}$$

Jacobian operation for $\boldsymbol{K}\boldsymbol{u}=\boldsymbol{b}$

Lets start with $J:\boldsymbol{C}\mapsto\boldsymbol{u}$

$$(\boldsymbol{K}+d\boldsymbol{K})(\boldsymbol{u}+d\boldsymbol{u})=\boldsymbol{b}, \tag{17}$$

can be rearranged to

$$\boldsymbol{K}d\boldsymbol{u}=-d\boldsymbol{K}(\boldsymbol{u}+d\boldsymbol{u}). \tag{18}$$

$$\frac{d\boldsymbol{u}}{dc_j}=-\boldsymbol{K}^{-1}\frac{d\boldsymbol{K}}{dc_j}\boldsymbol{u}. \tag{19}$$

Chain rule

$$d\boldsymbol{u}=\boldsymbol{J}d\boldsymbol{c} \quad = \quad -\sum_j \boldsymbol{K}^{-1}\frac{d\boldsymbol{K}}{dc_j}\boldsymbol{u}dc_j=-\boldsymbol{K}^{-1}\left[\sum_j\frac{d\hat{\boldsymbol{K}}}{du_j}du_j\right]\boldsymbol{u} \tag{20}$$

$$= \quad -\boldsymbol{K}^{-1}\boldsymbol{K}_{d\boldsymbol{C}}\boldsymbol{u}. \tag{21}$$

Jacobian operation 2

The last equation still requires $\hat{\boldsymbol{K}}^{-1} \longrightarrow$ replace by $\boldsymbol{G}$

$$d\boldsymbol{u} = -\boldsymbol{G}^T \boldsymbol{K}_{d\boldsymbol{C}} \boldsymbol{G}. \quad (22)$$

Similar $J^T : \boldsymbol{r} \mapsto \boldsymbol{C}$

$$\boldsymbol{J}^T \boldsymbol{r} = -\text{diag}\left(\left(\boldsymbol{G}^T \boldsymbol{A}^T\right)^T \left(\boldsymbol{R}\boldsymbol{G}^T \boldsymbol{A}\right)\right), \quad (23)$$

with $\boldsymbol{R} = \text{diag}(\boldsymbol{r})$
We have

- a **linear approximation** and can do
- **gradient based optimization**

for free by maintaining Green's functions.

Exact low rank updates for $\boldsymbol{K}\boldsymbol{u} = \boldsymbol{b}$

How does $\boldsymbol{u}$ change when a low number of elements 0f $\boldsymbol{C}$ are updated?
Purpose: line search, conditional sampling
Woodbury Formula:

$$\left(\boldsymbol{K}+\boldsymbol{U}\boldsymbol{W}^T\right)^{-1} = \boldsymbol{K}^{-1} - \underbrace{\boldsymbol{K}^{-1}\boldsymbol{U}\left(\boldsymbol{I}+\boldsymbol{W}^T\boldsymbol{K}^{-1}\boldsymbol{U}\right)^{-1}\boldsymbol{W}^T\boldsymbol{K}^{-1}}_{\text{update term}}, \qquad (24)$$

How can we use this:

$$\boldsymbol{K}_{\text{new}} = \boldsymbol{K}_{\text{old}} + \gamma\boldsymbol{A}^T\boldsymbol{C}_{\Delta}\boldsymbol{A}^T = \hat{\boldsymbol{K}}_{\text{old}} + \gamma\boldsymbol{U}\boldsymbol{W}^T \qquad (25)$$

With $\boldsymbol{G}$ to replace $\boldsymbol{K}_{\text{old}}^{-1}$

$$\Delta\boldsymbol{u} = -\gamma\boldsymbol{G}^T\boldsymbol{U}\left(\boldsymbol{I}+\gamma\boldsymbol{W}_{:,C}^T\boldsymbol{G}_{C,C}\boldsymbol{U}_{C,:}\right)^{-1}\boldsymbol{W}^T\boldsymbol{G}, \qquad (26)$$

Drawback: $\boldsymbol{G}_{C,C}$ (Green's functions for the effected nodes C) is needed

Woodbury for $\boldsymbol{A}^T\boldsymbol{S}\boldsymbol{A}$ systems 1

Can we do better for our system structure?
Assume

$$\boldsymbol{K}_{\mathrm{old}} = \boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}^T, \tag{27}$$

with $\boldsymbol{S}$ being a $m \times m$ diagonal matrix and $\boldsymbol{A}$ being a $m \times n$ matrix.

$$\boldsymbol{K}_{\mathrm{new}} = \boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A} + \boldsymbol{A}^T\boldsymbol{C}_{\Delta}\boldsymbol{A}, \tag{28}$$

With the decomposition: $\boldsymbol{U} = \boldsymbol{A}^T$ and $\boldsymbol{W}^T = \boldsymbol{C}\Delta\boldsymbol{A}$, the Woodbury is given as

$$\boldsymbol{K}_{\mathrm{new}}^{-1} = \boldsymbol{K}_{\mathrm{old}}^{-1} - \boldsymbol{K}_{\mathrm{old}}^{-1}\boldsymbol{A}^T\left(\boldsymbol{I} + \boldsymbol{C}_{\Delta}\boldsymbol{A}\boldsymbol{K}_{\mathrm{old}}^{-1}\boldsymbol{A}^T\right)^{-1}\boldsymbol{C}\Delta\boldsymbol{A}\boldsymbol{K}_{\mathrm{old}}^{-1}, \tag{29}$$

Is there a way to replace $\boldsymbol{K}_{\mathrm{old}}^{-1}$?

Woodbury for $\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ systems 2

Assume

$$\boldsymbol{K}_{\text{old}}^{-1} = \boldsymbol{J}_1^{-1}\boldsymbol{C}^{-1}\boldsymbol{J}_2^{-T}, \quad (30)$$

Thus

$$\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}\boldsymbol{J}_1^{-1}\boldsymbol{C}^{-1}\boldsymbol{J}_2^{-T} = \boldsymbol{I}. \quad (31)$$

- $\boldsymbol{C}\boldsymbol{C}^{-1} = \boldsymbol{I}$ ... $\checkmark$
- $\boldsymbol{A}^T\boldsymbol{J}_2^{-T} = \boldsymbol{I}$ ... $\boldsymbol{J}_2^{-T}$? Pseudo inverse

$$\boldsymbol{J}_2^{-1} = \underbrace{\left(\boldsymbol{A}\boldsymbol{A}^T\right)}_{\text{Laplacian}}{}^{-1}\boldsymbol{A} \quad (32)$$

- $\boldsymbol{A}\boldsymbol{J}_1^{-1} = \boldsymbol{I}$ ... not possible

Woodbury for $\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ systems 3

$\boldsymbol{J}_1^{-1}$ is a problem. However, lets insert it into the inverse of the Woodbury matrix

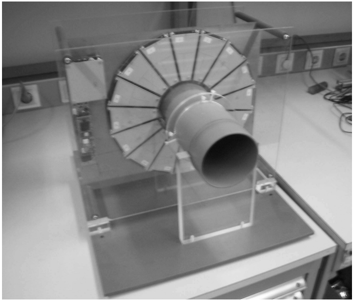
$$\left(\boldsymbol{I}+\boldsymbol{C}_\Delta\boldsymbol{A}\boldsymbol{K}_{\mathrm{old}}^{-1}\boldsymbol{A}^T\right)^{-1} = \left(\boldsymbol{I}+\boldsymbol{C}_\Delta\boldsymbol{A}\boldsymbol{J}_1^{-1}\boldsymbol{C}^{-1}\boldsymbol{J}_2^{-T}\boldsymbol{A}^T\right)^{-1}, \quad (33)$$

$$= \left(\boldsymbol{I}+\boldsymbol{C}_\Delta\boldsymbol{C}^{-1}\boldsymbol{J}_2^{-T}\boldsymbol{A}^T\right)^{-1}, \quad (34)$$

- Instead of $\boldsymbol{G}_{C,C}$ now only the diagonal matrix $\boldsymbol{S}$ has to be inverted. $\boldsymbol{J}_2^{-T}$ can be computed in advance.
- Again, this form can be reduced and Green's functions can be used on the left and right hand side!

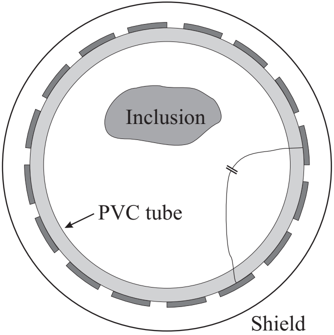
Electrical Capacitance Tomography

- Inverse problem using capacitance measurements.
- Ill-posed $\rightarrow$ regularization or prior knowledge required.
- Suitable for process tomography due to good contrast of the permittivity (i.e. oil/water).
- Noninvasive.
- Cheap instrumentation but low spatial resolution due to soft field.
- Non ionizing $\rightarrow$ safe.



Neumayer et. al.: Current Reconstruction Algorithms in Electrical Capacitance Tomography, Springer 2010

Electrical Capacitance tomography



- Ω: whole domain
- Ω_{ROI}: interior of the pipe.
- $\partial\Omega_{ROI}$: interior of the pipe.

PDE:

$$\nabla \cdot (\varepsilon_0 \varepsilon_r \nabla V) = 0, \qquad (35)$$

BC:

$$V_{\partial\Omega} = 0, \qquad (36)$$
$$V_{\Gamma_i} = V_{\Gamma_i}, \qquad (37)$$
$$V_{\Gamma_j} = 0 \quad \forall j \neq i, \qquad (38)$$

Gauss law:

$$C_{i,j} = -\frac{1}{V_{\Gamma_i}} \oint_{\Gamma_j} \varepsilon_0 \varepsilon_r \nabla V \cdot \vec{\mathbf{n}} d\Gamma. \qquad (39)$$

Gives a $N_{elec} \times N_{elec}$ matrix.

Standard computations: $F : \varepsilon \mapsto \boldsymbol{C}$

Finite element system:

$$\boldsymbol{K} = \sum_{i=1}^{N_e} \varepsilon_i \boldsymbol{K}_{e,i}, \tag{40}$$

With BC:

$$\hat{\boldsymbol{K}} \boldsymbol{v} = \boldsymbol{r}. \tag{41}$$

Charge method:

$$Q_{elec} = \sum_{n_{elec}} (\boldsymbol{K}\boldsymbol{v})_{n_{elec}}, \tag{42}$$

Derivatives:

$$dC_{i,j} = \gamma_j^T \left[\left[\frac{\partial \boldsymbol{r}}{\partial \varepsilon_k} \right] - \left[\frac{\partial \hat{\boldsymbol{K}}}{\partial \varepsilon_k} \right] \boldsymbol{v}_i \right] d\varepsilon_k, \tag{43}$$

Fast material update

Eigenvector decomposition:

$$\boldsymbol{K}_{\mathrm{e}} = \boldsymbol{V}_{\mathrm{e}}\boldsymbol{D}_{\mathrm{e}}\boldsymbol{V}_{\mathrm{e}}^{-1}, \tag{44}$$

If $\boldsymbol{K}_{\mathrm{e}} = \boldsymbol{K}_{\mathrm{e}}^{T}$:

$$\boldsymbol{K}_{\mathrm{e}} = \boldsymbol{A}_{\mathrm{e}}^{T}\boldsymbol{S}_{\mathrm{e}}\boldsymbol{A}_{\mathrm{e}}, \tag{45}$$

For FE system

$$\boldsymbol{K} = \boldsymbol{K}_{ini} + \sum_{l=1}^{p} \boldsymbol{A}_{l}^{T}\mathscr{E}\boldsymbol{A}_{l}, \tag{46}$$

with $\boldsymbol{K}_{ini}$ such that $\boldsymbol{K}$ has full rank (includes the BC).

Charge map: $Q_c : V_{\partial\Omega_{ROI}} \mapsto \Delta \boldsymbol{Q}$

How does the potential on the tube influences the charge on the electrodes? Solve

$$\nabla \cdot (\varepsilon_0 \varepsilon_r \nabla V) = 0 \tag{47}$$

in $\Omega \setminus \Omega_{ROI}$ with BC:

$$V_{\partial\Omega} = 0, \tag{48}$$

$$V_{\partial\Omega_{ROI}} = \delta(x - x_i) \qquad \forall x_i \in \partial\Omega_{ROI} \tag{49}$$

$$V_{\Gamma_j} = 0 \qquad \forall j, \tag{50}$$

Discrete version $\boldsymbol{Q}_c : \boldsymbol{v}_{\partial\Omega_{ROI}} \mapsto \Delta \boldsymbol{Q}$

$$\boldsymbol{Q} = \boldsymbol{Q}_a + \boldsymbol{Q}_c \boldsymbol{V}_{\partial\Omega_{ROI}}, \tag{51}$$

$\boldsymbol{V}_{\partial\Omega_{ROI}}$ contains the nodal potentials on $\partial\Omega_{ROI}$

Solution with Green's functions

$$\hat{\boldsymbol{K}}\boldsymbol{G} = \boldsymbol{E}_{\partial\Omega_{ROI}}, \tag{52}$$

$$\boldsymbol{V}_{\partial\Omega_{ROI}} = \boldsymbol{G}^T\boldsymbol{R}, \tag{53}$$

$\boldsymbol{R}$ contains the right hand side terms for each electrode.

$$\boldsymbol{Q} = \boldsymbol{Q}_a + \boldsymbol{Q}_c \underbrace{\boldsymbol{G}^T\boldsymbol{R}}_{\boldsymbol{V}_{\partial\Omega_{ROI}}}, \tag{54}$$

... is bad, because $\boldsymbol{E}_{\partial\Omega_{ROI}}$ in equation (52) means the solution for all nodes on $\partial\Omega_{ROI}$.
Trick:

$$\hat{\boldsymbol{K}}\boldsymbol{G}\boldsymbol{Q}_c^T = \boldsymbol{E}_k\boldsymbol{Q}_c^T, \tag{55}$$

$$\hat{\boldsymbol{K}}\boldsymbol{G}_Q = \boldsymbol{E}_k\boldsymbol{Q}_c^T = \boldsymbol{R}_Q, \tag{56}$$

Hence:

$$\boldsymbol{Q} = \boldsymbol{Q}_a + \boldsymbol{G}_Q^T\boldsymbol{R}. \tag{57}$$

Jacobian operations for ECT

- Linearization: $J : \varepsilon \mapsto \boldsymbol{Q}$

$$d\boldsymbol{Q} = -\boldsymbol{G}_Q^T \hat{\boldsymbol{K}}_{d\varepsilon} \boldsymbol{G}_Q \quad (58)$$

$$d\boldsymbol{Q} = -\boldsymbol{G}_Q^T \left[\sum_l^p \boldsymbol{A}_l^T d\mathscr{E} \boldsymbol{A}_l \right] \boldsymbol{G}_Q . \quad (59)$$

- Gradient: $J^T : \boldsymbol{Q} \mapsto \varepsilon$

$$\boldsymbol{J}^T \boldsymbol{q} = -\mathrm{diag}\left(\left(\sum_{l=1}^p \left(\boldsymbol{G}_Q^T \boldsymbol{A}_l^T \right)^T \left(\boldsymbol{Q} \boldsymbol{G}_Q^T \boldsymbol{A}_l \right) \right) \right), \quad (60)$$

Low Rank updates for ECT

- With computation of the additional Green's Funcions

$$\Delta \boldsymbol{Q} = -\gamma \boldsymbol{G}_Q^T \boldsymbol{L} \left(\boldsymbol{I} + \gamma \boldsymbol{U}_{:,C} \boldsymbol{G}_{C,C} \boldsymbol{L}_{C,:} \right)^{-1} \boldsymbol{U} \boldsymbol{G}_Q, \tag{61}$$

- $\boldsymbol{A}^T \boldsymbol{C} \boldsymbol{A}$ form

$$\Delta \boldsymbol{Q} = -\boldsymbol{G}_{Q,C_2,:}^T \boldsymbol{A}_{1,:,C}^T \left(\boldsymbol{I} + \boldsymbol{S}_C \boldsymbol{S}_{0,C}^{-1} \boldsymbol{W}_{3,:,C}^T \boldsymbol{W}_{1,:,C} \right)^{-1} \boldsymbol{S} \boldsymbol{W}_{1,:,C}^T \boldsymbol{G}_{Q,C_2,:}, \tag{62}$$

A speed comparison

Operation/Method	t_{mesh1}	t_{mesh2}
	ms	ms
Forward problem standard	96	640
Forward problem new	4.8	26
Standard material update	35.5	230
Fast material update	0.039	2
Matrix inversion	3.3	19
Jacobian by AVM	360	$>$ 2000
Jacobian op. $J : \varepsilon \mapsto \boldsymbol{Q}$	0.48	3.9
Jacobian transp. op. $J^T : \boldsymbol{Q} \mapsto \varepsilon$	3.3	15.5
Exact low rank update (1 elem.)	3.5	16.3
$\boldsymbol{WSW}^T$ Woodbury (1 elem.)	0.66	4.2
Exact low rank update (20 elem.)	11.6	55.4
$\boldsymbol{WSW}^T$ Woodbury (20 elem.)	2.7	14.1
Exact update 1 elem $\times$ 20	5.6	18
$\boldsymbol{WSW}^T$ Woodbury 1 elem $\times$ 20	1.1	3
Domain d. by Schur c.	23.8	66
Schur c. with Cho.	2.9	8
Woodb. for Schur c. with Chol.	0.895	17.1

Fast Deterministic Method

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \left\|F(\boldsymbol{x}) - \tilde{\boldsymbol{d}}\right\|_2^2 + \alpha \boldsymbol{x}^T \boldsymbol{L}^T \boldsymbol{L} \boldsymbol{x}, \tag{63}$$

Gauss Newton method:

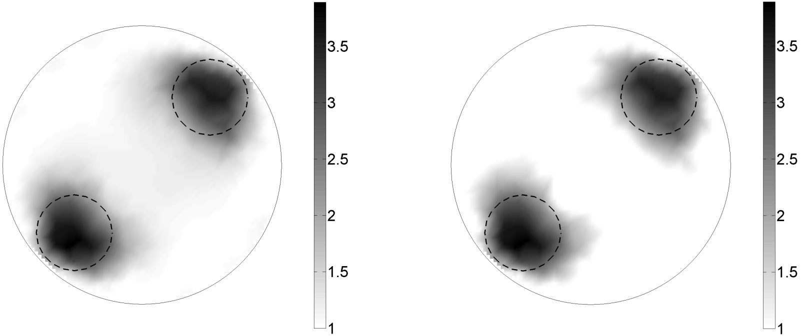
$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - s\left(\mathbf{J}^T\mathbf{J} + \alpha \boldsymbol{L}^T \boldsymbol{L}\right)^{-1}\left(\boldsymbol{J}^T \boldsymbol{r}_k + \boldsymbol{L}^T \boldsymbol{L} \boldsymbol{x}_k\right), \tag{64}$$

BFGS scheme:

1. Evaluate the Newton direction $\boldsymbol{p}_k = -\boldsymbol{H}_k^{-1}\boldsymbol{g}_k$
2. Find s to set $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + s\boldsymbol{p}_k$ and set $\boldsymbol{s}_k = s\boldsymbol{p}_k$
3. Compute $\boldsymbol{y}_k = \boldsymbol{g}(\boldsymbol{x}_{k+1}) - \boldsymbol{g}(\boldsymbol{x}_k)$.
4. Evaluate $\boldsymbol{H}_{k+1}^{-1} = \boldsymbol{H}_k^{-1} + \frac{\boldsymbol{s}_k^T \boldsymbol{y}_k + \boldsymbol{y}_k^T \boldsymbol{H}_k^{-1} \boldsymbol{y}_k}{(\boldsymbol{s}_k^T \boldsymbol{y}_k)^2} \boldsymbol{s}_k \boldsymbol{s}_k^T - \frac{\boldsymbol{H}_k^{-1} \boldsymbol{y}_k \boldsymbol{s}_k^T + \boldsymbol{s}_k \boldsymbol{y}_k^T \boldsymbol{H}_k^{-1}}{(\boldsymbol{s}_k^T \boldsymbol{y}_k)^2}$.

Bayesian Inversion

Deterministic results:



About 50 iterations per second are possible on a standard PC.

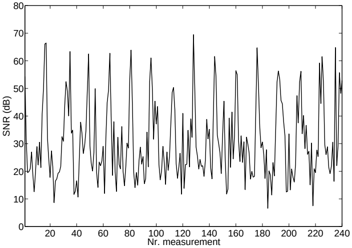
Bayesian Inversion

- Posterior to sample from

$$\pi(\boldsymbol{x}|\tilde{\boldsymbol{d}}) \propto \exp\left(-\frac{1}{2}\boldsymbol{e}^T\boldsymbol{\Sigma}_{\boldsymbol{v}}^{-1}\boldsymbol{e}\right)\exp\left(-\frac{1}{2}\alpha\boldsymbol{x}^T\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{x}\right)I(\boldsymbol{x}). \quad (65)$$

- $\boldsymbol{L}$: smoothing prior

Signal disturbed by additive Gaussian noise.



How to implement a Gibbs sampler

- General Steps

  1. Draw a conditional sample.
  2. If the new sample is accepted, compute the new forward map to obtain the new Green's functions $\boldsymbol{G}_Q$ for the cheap evaluation of the next conditional sample.

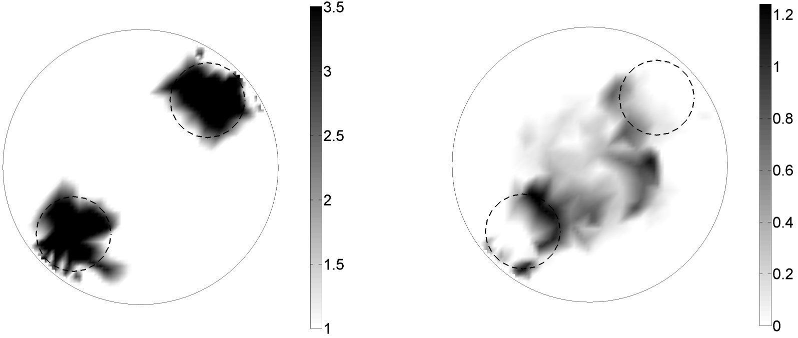
  Versions:

- Bimodal (two valued) material distributions
- General material values

Bimodal Material Distributions

- Computation steps:
  1. Flip one element of $\boldsymbol{x}$ to generate the proposal $\boldsymbol{x}'$.
  2. Compute the likelihood ratio $\alpha = \min\left[1, \frac{\pi(\boldsymbol{x}'|\tilde{\boldsymbol{d}})}{\pi(\boldsymbol{x}|\tilde{\boldsymbol{d}})}\right]$.
  3. Accept $\boldsymbol{x}'$ with probability α.

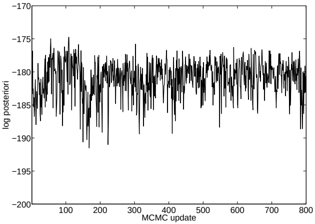
Bimodal Material Distributions

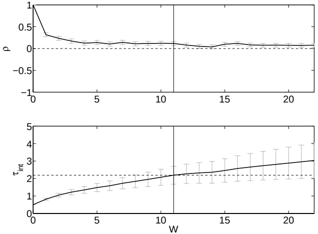


- Mean estimate and standard deviation
- About 3 frames per second are possible using the Jacobian operations

Bimodal Material Distributions

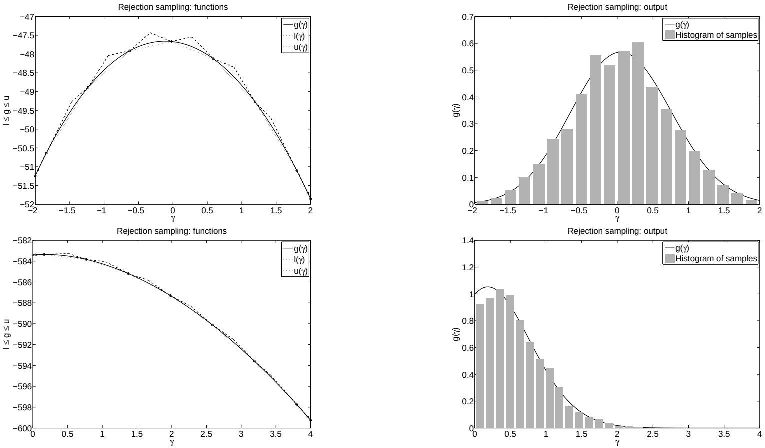
Behavior of the chain



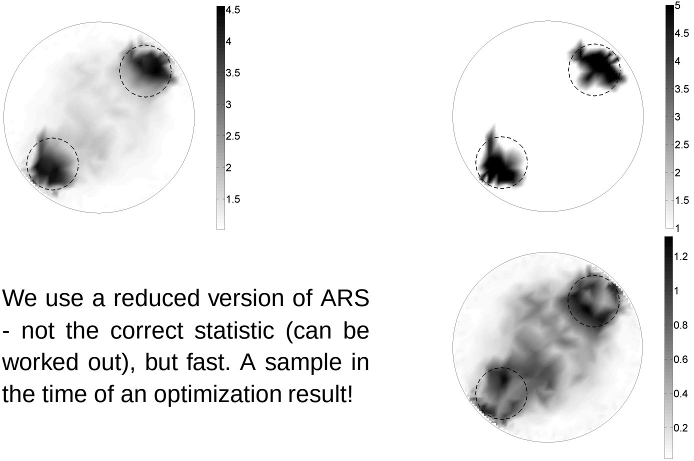


General Gibbs Sampler

Conditional sampling by ARS



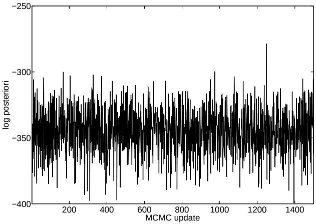
Gibbs Sampler

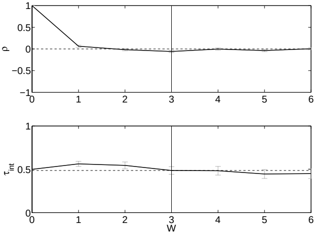


We use a reduced version of ARS - not the correct statistic (can be worked out), but fast. A sample in the time of an optimization result!

Gibbs Sampler

Behavior of the chain





- And this is just the most trivial version of a Gibbs sampler!

Summary

- Inverse problems
- General introduction to $\boldsymbol{A}^T\boldsymbol{C}\boldsymbol{A}$ systems
- Computational aspects
- Fast low rank updates
- Example on ECT

$\rightarrow$ with some tricks we can generate samples in the time of an optimization run.

Neumayer and Fox, JUQ 2013