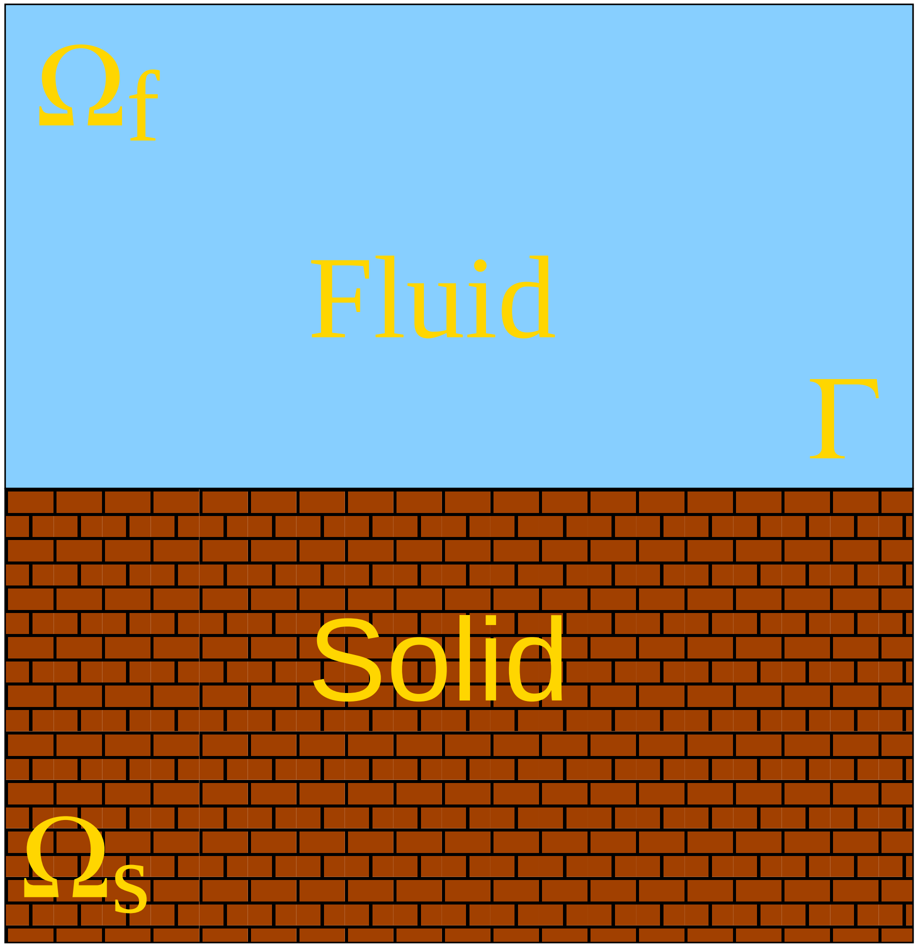
Estimation of the mechanical properties of a solid elastic medium in contact with a fluid medium

Édouard Canot Jocelyne Erhel Nabil Nassif
Samih Zein
SAGE team INRIA/IRISA – (Rennes, France)

Plan

- Forward and inverse problems
- Bayesian inference
- Choice of the estimators
- The MCMC method
- The SPSA method
- Sensitivity analysis

The forward problem



The forward problem

The continuous formulation:

$$(1) \quad \begin{cases} \dfrac{\partial p}{\partial t} + c_f^2 \rho_f \ div\ v_f = 0 \ (\Omega_f) & (1.1) \\ \rho_f \dfrac{\partial v_f}{\partial t} + \nabla p = 0 \ (\Omega_f) & (1.2) \\ A \dfrac{\partial \sigma}{\partial t} - \epsilon(v_s) = 0 \ (\Omega_s) & (1.3) \\ \rho_s \dfrac{\partial v_s}{\partial t} - \ div\ \sigma = 0 \ (\Omega_s) & (1.4) \\ v_s.n = v_f.n \ (\Gamma) & (1.5) \\ \sigma.n = -p.n \ (\Gamma) & (1.6) \end{cases}$$

The forward problem

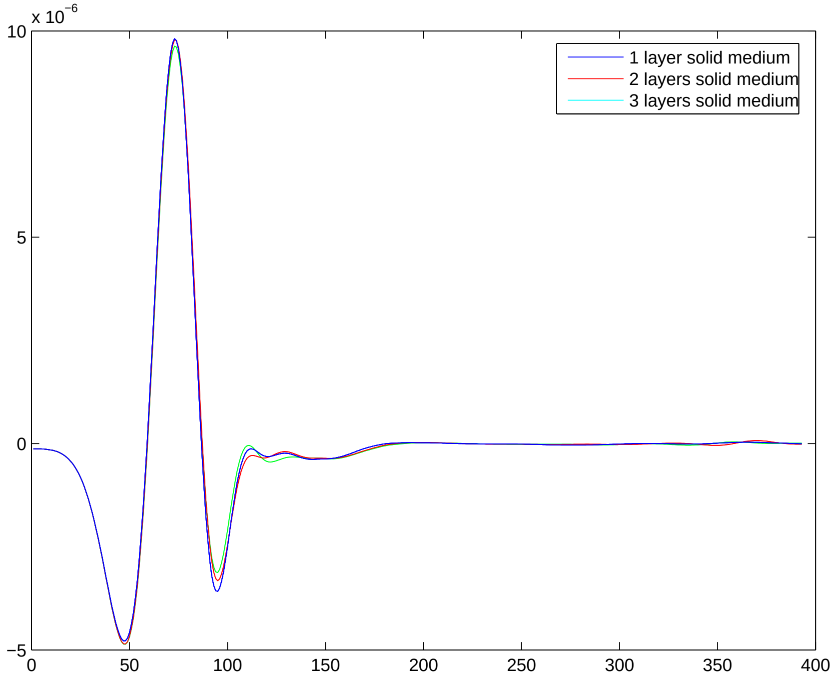
The discrete formulation:

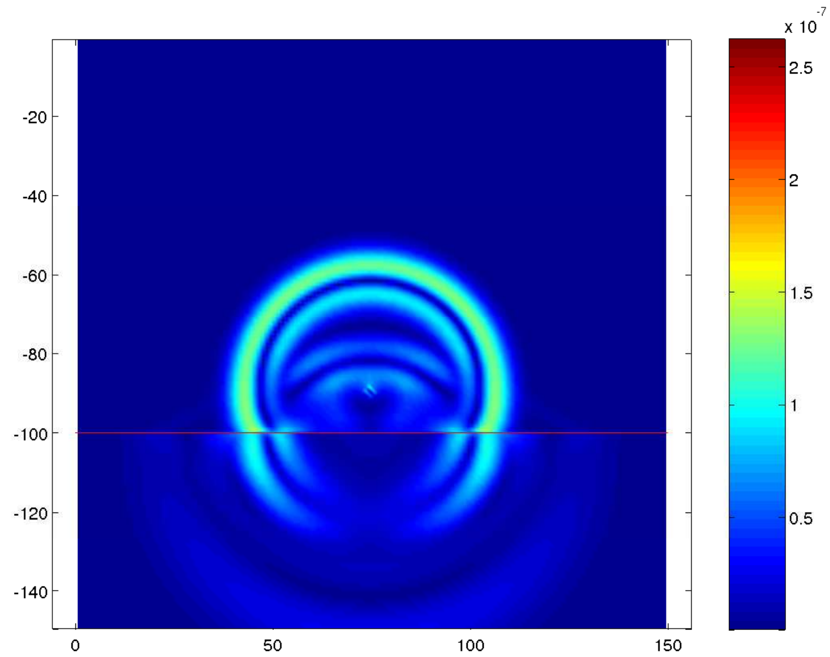
$$
\begin{cases}
M_p \dfrac{P^{n+\frac{1}{2}} - P^{n-\frac{1}{2}}}{\Delta t} + D_f V_f^n = 0 & (2.1)\\
M_f \dfrac{V_f^{n+1} - V_f^n}{\Delta t} - D_f^t P^{n+\frac{1}{2}} + B \dfrac{\Sigma^{n+1} + \Sigma^n}{2} = 0 & (2.2)\\
M_\Sigma \dfrac{\Sigma^{n+1} - \Sigma^n}{\Delta t} + D_s^t V_s^{n+\frac{1}{2}} - B^t \dfrac{V_f^{n+1} + V_f^n}{2} = 0 & (2.3)\\
M_s \dfrac{V_s^{n+\frac{1}{2}} - V_s^{n-\frac{1}{2}}}{\Delta t} + D_s \Sigma^n = 0 & (2.4)
\end{cases}
$$

(J. Diaz and P. Joly 2005)

The forward problem

Simulations results:





The inverse problem

- θ: the coefficients to recover (λ, μ, ρ)

- Pressure measures

$$y = u + \epsilon$$

where $u := \{u_{ij} = u(x_i, t_j)\}$,$\epsilon := \{\epsilon_{ij}\} \sim N(0, s^2)$.

Estimators of θ

1. the expectation with respect to the posterior probability:

$$E(\theta|y) = \int \theta \; p(\theta|y) \; d\theta$$

2. the maximum a posteriori:

$$\theta^* = \text{arg} \max_{\theta \in D_\theta} \; p(\theta|y),$$

The Bayesian model and the inverse problem

From Bayes's formula:

$$p(\theta|y) = \frac{1}{p(y)} p(y|\theta) p(\theta)$$

We consider:

$$p(\theta) \propto \begin{cases} 1 & \text{si } \forall\, i\,, \theta_i \in [\theta_{min}, \theta_{max}] \\ 0 & \text{elsewhere} \end{cases}$$

$$p(y|\theta) \propto exp\Big(-\frac{1}{2} \sum_i \Big(\frac{y_i - u(x_i, T, \theta)}{s} \Big)^2 \Big)$$

$$p(\theta|y) \propto \begin{cases} exp\Big(-\frac{1}{2} \sum_i \Big(\frac{y_i - u(x_i, T, \theta)}{s} \Big)^2 \Big) & \text{si } \forall\, i\,, \theta_i \in [\theta_{min}, \theta_{max}] \\ 0 & \text{elsewhere} \end{cases}$$

Markov Chain Monte Carlo

Estimation of θ:

$$E[\theta|y] = \int \theta p(\theta|y) d\theta$$

This integral is approached by:

$$E[\theta|y] \approx \frac{1}{n} \sum_{k=1}^{n} \theta_k$$

with $\theta \sim p(\theta|y)$ the limiting distribution of a Markov chain (Harold Niederreiter, SIAM, 1992).

The Metropolis-Hasting algorithm

The accelerated version of M-H algorithm with $p^*(.|y)$ a linear interpolation of $p(.|y)$:

1- At θ_k generate a proposal C from $q(\cdot|\theta_k)$.

2- With probability $\alpha_{pred}(C, \theta_k) = min\Big\{\frac{p^*(C|y)}{p^*(\theta_k|y)}, 1\Big\}$ promote C to be a candidate to the standard M-H algorithm. Otherwise, pose $\theta_{k+1} = \theta_k$.

3- With probability $\alpha(C, \theta) = min\Big\{\frac{p(C|y)}{p(\theta_k|y)}, 1\Big\}$ accept $\theta_{k+1} = C$; Otherwise reject C, $\theta_{k+1} = \theta_k$.

(J. Andre's Christen and C. Fox, 2005.)

The variance of the estimator given by MCMC:

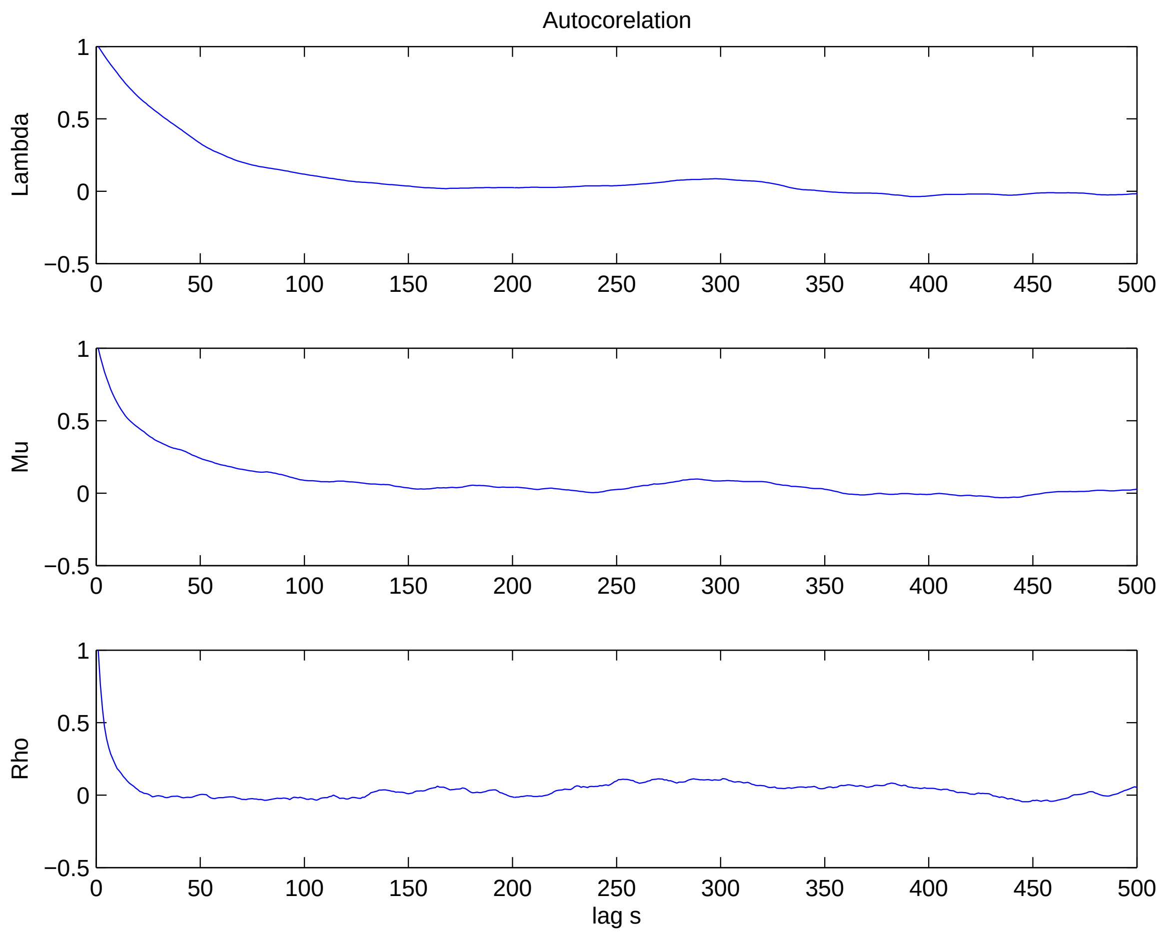
$$var(\overline{\theta}_{MC}) = \tau \frac{var(\theta)}{n}$$

where τ is the integrated autocovariance time (IACT).

$$\overline{\tau} = 1 + 2 \sum_{s=1}^{M} \rho(s).$$

(S. Meyer, N. Christensen and G. Nicholls, 2001)

Markov Chain Monte Carlo



Markov Chain Monte Carlo

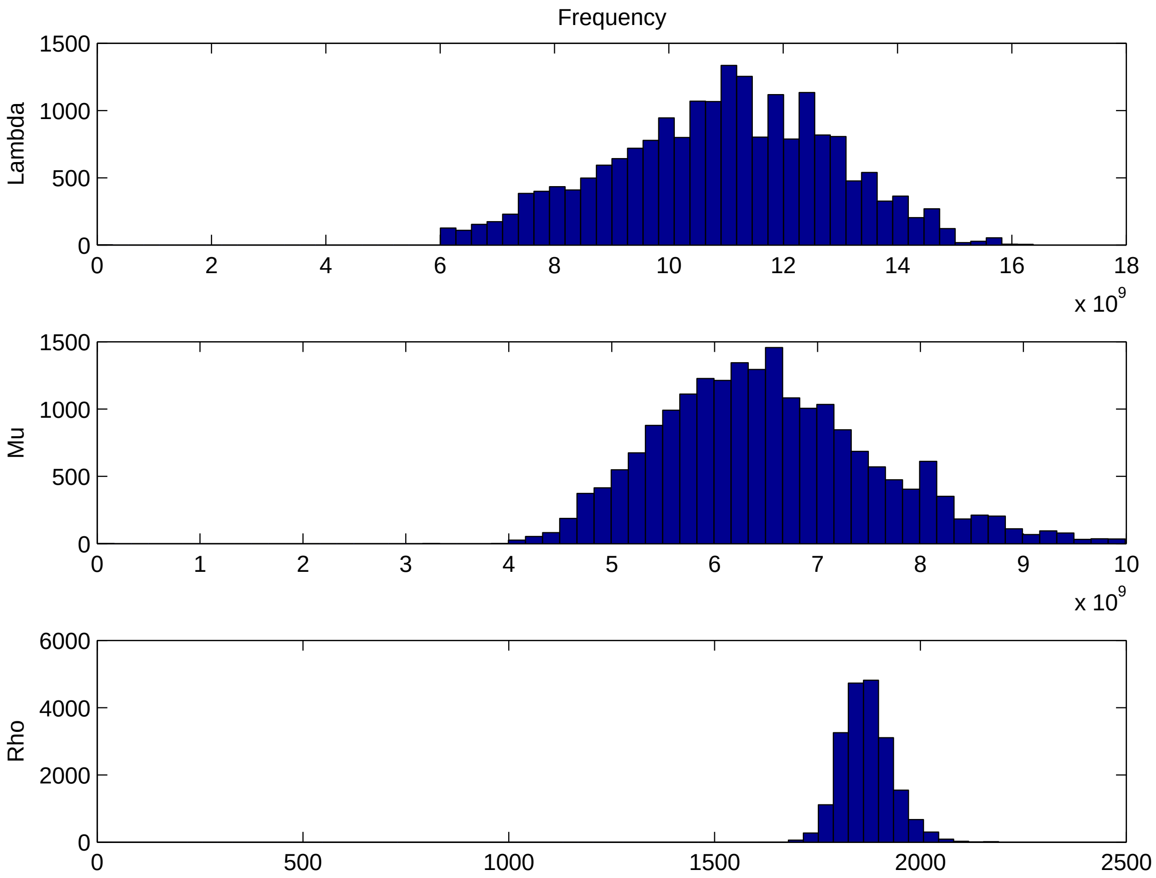
- 19000 samples of the Markov chain with the accelerated version of M-H algorithm (6000 simulations and noise $< 6\%$):

θ	**Exact Value** (SI)	**Conf. Interval**	**% of error**
λ	11.5×10^9	$10.9\times10^9\pm2.6\%$	5.2%
μ	6×10^9	$6.5\times10^9\pm2\%$	8%
ρ	1850	$1867\pm0.15\%$	0.9%

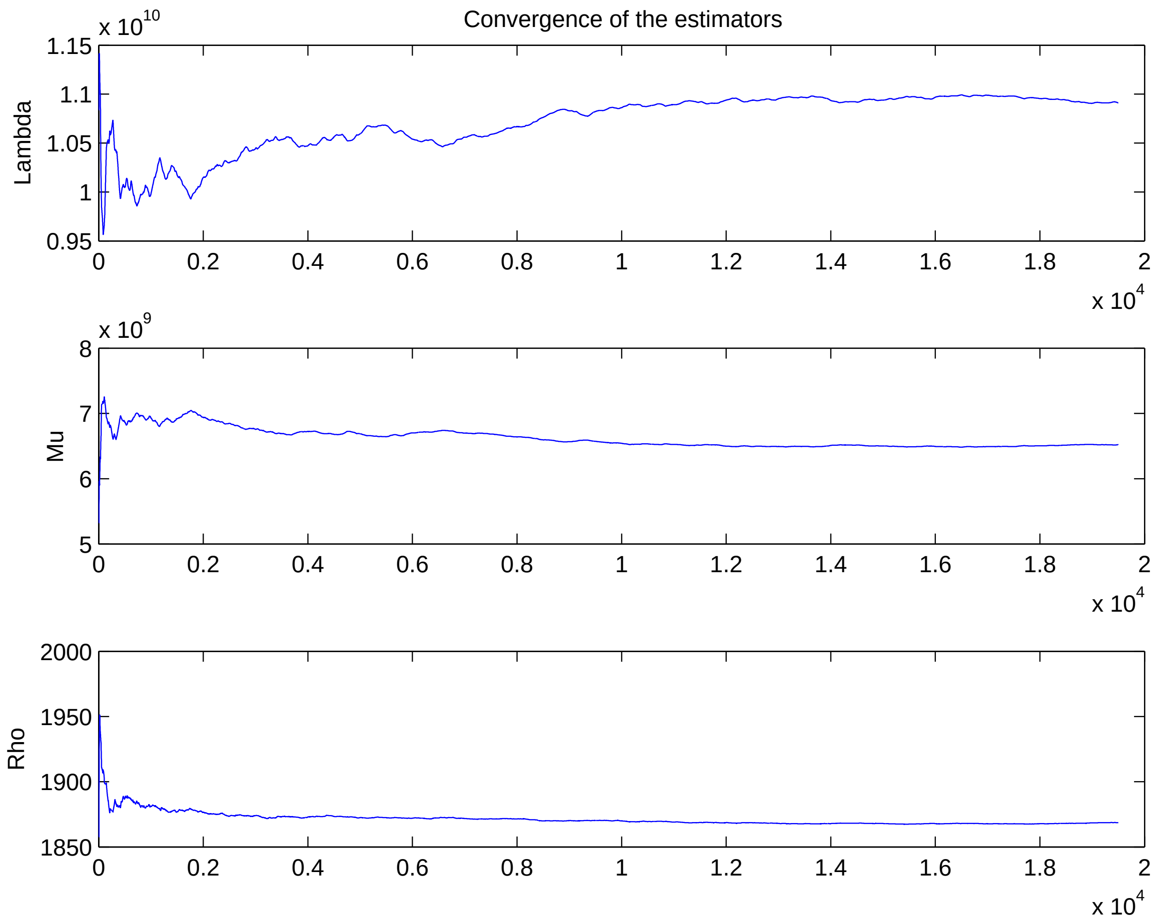
- 19000 samples of the Markov chain with the standard M-H algoritm and different starting points:

θ	$\theta_0 = \theta_{min}$	$\theta_0 = \theta_{max}$
λ	$11.1 \times 10^9 \pm 2.8\%$	$10.8 \times 10^9 \pm 2.7\%$
μ	$5.82 \times 10^9 \pm 2.5\%$	$6.14 \times 10^9 \pm 2.6\%$
ρ	$1827 \pm 0.21\%$	$1911 \pm 0.4\%$

Markov Chain Monte Carlo



Markov Chain Monte Carlo



(C. Robert 1996)

Simultaneous Perturbation Stochastic Approximation

Optimize the log of the posterior probability:

$$L(\theta) = -\log p(\theta|y)$$

In our case, it is a least squares problem

The form of the algoritm is as follows:

$$\widehat{\theta}_{k+1} = \widehat{\theta}_k - a_k \widehat{g}_k(\widehat{\theta}_k)$$

with

$$\widehat{g}_k(\widehat{\theta}_k) = \frac{L(\widehat{\theta}_k + c_k \Delta_k) - L(\widehat{\theta}_k - c_k \Delta_k)}{2c_k} [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \ldots, \Delta_{kp}^{-1}]^T$$

(J.C. Spall, 2003).

Simultaneous Perturbation Stochastic Approximation

Asymptotic normality:

$$k^{\beta/2}(\widehat{\theta}_k - \theta^*) \xrightarrow{\text{dist}} N(0, \Sigma)$$

(J.C. Spall, 2003).

It is possible to have a confidence interval for θ^* by applying the Monte Carlo method:

$$\frac{1}{N}\sum_{k=1}^{N} k^{\beta/2}(\widehat{\theta}_k - \theta^*) = 0$$

$$\Longrightarrow \theta^* = \frac{\sum_{k=1}^{N} k^{\beta/2}(\widehat{\theta}_k)}{\sum_{k=1}^{N} k^{\beta/2}}$$

$$\Sigma = \frac{1}{N}\sum_{k=1}^{N} k^{\beta}(\widehat{\theta}_k - \theta^*)(\widehat{\theta}_k - \theta^*)^T$$

Simultaneous Perturbation Stochastic Approximation

θ	**Exact Values** (SI)	**Confidence Intervals**	**Errors**
λ	11.5×10^9	$11.83\times10^9\pm1.4\%$	2.8%
μ	6×10^9	$5.75\times10^9\pm1.6\%$	4.1%
ρ	1850	$1856\pm0.12\%$	0.03%

Injected noise $< 1\%$ and 600 simulations

θ	**Exact Values** (SI)	**Confidence Intervals**	**Errors**
λ	11.5×10^9	$12.2\times10^9\pm6.12\%$	6.6%
μ	6×10^9	$5.4\times10^9\pm7.2\%$	9.5%
ρ	1850	$1868\pm0.83\%$	1%

Injected noise $< 6\%$ and 700 simulations

Simultaneous Perturbation Stochastic Approximation

θ	**Exact Values** (SI)	**Estimated Values**	**Errors**
$\lambda 1$	11.5×10^9	$11.43\ \times 10^9$	0.6%
$\mu 1$	6×10^9	$6.01\ \times 10^9$	0.1%
$\rho 1$	1700	1702	0.1%
$\lambda 2$	9×10^9	8.93×10^9	0.7%
$\mu 2$	7×10^9	7.02×10^9	0.2%
$\rho 2$	2000	2003	0.1%
$\lambda 3$	11.5×10^9	11.4×10^9	2.1%
$\mu 3$	6×10^9	6.02×10^9	1.8%
$\rho 3$	2400	2405	0.1%

Injected noise $< 1\%$ and 6000 simulations

Sensitivity analysis

$$\epsilon \approx F'(\theta_0)\delta\theta$$

Consider the singular value decomposition (SVD) of the Jacobian matrix ($F'(\theta_0)$).

$$F'(\theta_0) = USV^T$$

One easily verifies that:

$$\delta\theta_k = \frac{\epsilon_k}{s_k}, \ \forall k = 1, ..., p.$$

Sensitivity analysis

A first order analysis yields:

$$0 = F(0) = F(\theta_0) - F'(\theta_0)\theta_0$$

Thus

$$\|u_0\| = \|F(\theta_0)\| \simeq \|F'\theta_0\| < s_1\|\theta_0\|$$
$$\forall k, \frac{|\delta\theta_k^*|}{\|\theta_0\|} \leq \frac{s_1}{s^k}\frac{|\epsilon_k|}{\|u_0\|} \leq \sigma_u$$

Hence, if the accuracies on θ and u, respectively σ_θ and σ_u, verify the inequality:

$$\frac{s_k}{s_1} \geq \frac{\sigma_u}{\sigma_\theta}\ \forall k,$$

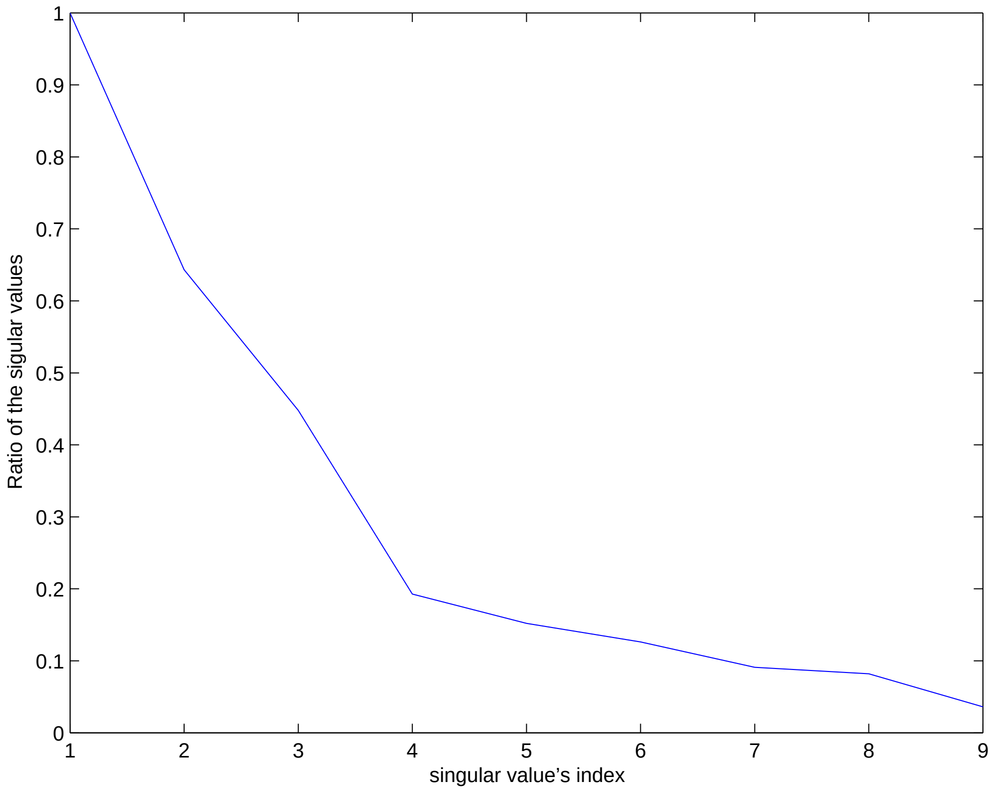
Sensitivity analysis

then:

$$\frac{|\delta\theta_k^*|}{\|\theta_0\|} \leq \sigma_\theta \ \forall k \text{ whenever } \frac{|\epsilon_k|}{\|u_0\|} \leq \sigma_u.$$

(P. Al Khoury, G. Chavant, F. Clément and P. Hervé, 2002).

Sensitivity analysis



Conclusion

- MCMC is more accurate than SPSA but it is harder to implement and to have correct results.
- SPSA is much less expensive in computations then MCMC .
- In the case of multiple layers, SPSA is a more appropriate method.