Accelerated MCMC for OCO-2’s CO$_2$ Retrieval

Otto Lamminpää, Jenny Brynjarsdottir, Jonathan Hobbs, Amy Braverman, Marko Laine, Hannakaisa Lindqvist, Johanna Tamminen

15th IWGGMS, Sapporo, Japan, June 3rd 2019
Finnish Centre of Excellence in Inverse Modelling and Imaging 2018 – 2025

A research project to combine important application areas and world leading mathematics in inverse problems and uncertainty quantification.

FMI’s interest: remote sensing, climate and weather, carbon cycle, air quality, ozone
Outline

- OCO-2 surrogate model
- Inverse Problem
- MCMC & Dimension Reduction
- Results
Orbiting Carbon Observatory 2

Carbon dioxide monitoring mission by NASA Jet Propulsion Laboratory, launched on and operational since July 2014.
Problem Setup

Surrogate model state vector [Hobbs et al. 2017]

\[ x_1 - x_{20} \quad - \quad \text{CO}_2 \text{ profile} \]
\[ x_{21} \quad - \quad \text{Surface pressure} \]
\[ x_{22} - x_{27} \quad - \quad \text{Surface albedo} \]
\[ x_{28} - x_{39} \quad - \quad \text{Aerosols} \]
Problem Setup

Simulated state vectors [Brynjarsdottir et al 2018] drawn from a normal distribution with empirical mean and covariance given by:


- Surface pressure and aerosols: Modern Era Retrospective Analysis for Research and Applications Aerosol Reanalysis (MERRAero).

- Surface albedo: Moderate Resolution Imaging Spectrometer (MODIS) albedo product
MCMC test case: tmp001
Posterior Correlation Matrix
Problem Setup

Inverse Problem:

\[ y = F(x) + \varepsilon, \quad x_{pr} \sim N(x_0, \Gamma_{pr}), \varepsilon \sim N(0, \Gamma_{\varepsilon}) \]

Bayesian Solution: Posterior distribution

\[ \pi(x | y) \propto \pi_{\varepsilon}(y | x) \pi_{pr}(x) \]

Optimal estimation: \( N(x_{MAP}, \hat{S}) \)

\[ x_{MAP} = \arg\min_x -2 \ln (\pi(x | y)) \]
\[ \hat{S} = (J(x_{MAP})^T \Gamma_{\varepsilon}^{-1} J(x_{MAP}) + \Gamma_{pr}^{-1})^{-1} \]
Problem Setup

Markov Chain Monte Carlo (MCMC):
Propose \( x_t \sim N(x_{t-1}, C) \), accept/reject with
\[
\text{min} \left( \frac{\pi(x_t | y)}{\pi(x_{t-1} | y)}, 1 \right)
\]

Adaptive MCMC:
\[
C_t = \begin{cases} 
C_0, & t < t_0 \\
\mathcal{d}_{\text{cov}} [X_0, \ldots, X_{t-1}] + s_{d\varepsilon} I, & t \geq t_0
\end{cases}
\]

Dimension reduction:
\[
x = P_r x_r + P_\perp x_\perp, \quad \pi_r(x_r | y) \propto \pi_\varepsilon(y | x_r) \pi_{pr}(x_r)
\]
XCO$_2$ from MCMC
Optimal Estimation w/ different first guess

\[ \gamma = 10, \quad \text{Step tol. } = 40 \quad \gamma = 30, \quad \text{Step tol. } = 10^{-4} \]

FINNISH METEOROLOGICAL INSTITUTE
Posterior histograms for Aerosols

- SO Coefficient 1
- DU Coefficient 1
- Ic Coefficient 1
- Wa Coefficient 1
- SO Coefficient 2
- DU Coefficient 2
- Ic Coefficient 2
- Wa Coefficient 2
- SO Coefficient 3
- DU Coefficient 3
- Ic Coefficient 3
- Wa Coefficient 3
2D Posterior plots: aerosols
2D Posterior plots: XCO$_2$ vs. Aerosols
Posterior XCO$_2$ with looser prior

- True XCO$_2$: 393.2693, MCMC mean XCO$_2$: 393.2977
- True XCO$_2$: 394.0594, MCMC mean XCO$_2$: 393.5289
- True XCO$_2$: 393.0985, MCMC mean XCO$_2$: 393.2427
Posterior Aerosols with looser prior
Ongoing / future efforts

- Surrogate vs Full Physics
- Identification of problematic geolocations / aerosol types
- In-depth look at operational prior covariance matrix
References


Thank you!