Suitably impressive thesis title

Lennart Golks

Department of Physics
University of Otago

A thesis submitted for the degree of
Doctor of Philosophy

November 2025

Abstract

Your abstract text goes here. Check your departmental regulations, but generally this should be less than 300 words. See the beginning of Chapter **??** for more.

Suitably impressive thesis title

Lennart Golks

Department of Physics

University of Otago

A thesis submitted for the degree of

Doctor of Philosophy

November 2025

Acknowledgements

Personal

I would like to thank Alex Elliott for his wonderful help and support. None of this would be possible otherwise.

Institutional

If you want to separate out your thanks for funding and institutional support, I don't think there's any rule against it. Of course, you could also just remove the subsections and do one big traditional acknowledgement section.

Abstract

Your abstract text goes here. Check your departmental regulations, but generally this should be less than 300 words. See the beginning of Chapter **??** for more.

Contents

List of Figures

List of Abbreviations

i.i.d. . . . . . . independent and identically distributed

MRF . . . . . . Markov Random Field

GMRF . . . . Gaussian Markov Random Field

MTC . . . . . Marginal Then Conditional sampler

GOMOS . . . Global Ozone Monitoring by Occultation of Stars

MCMC . . . . Markov Chain Monte-Carlo

MH . . . . . . Metropolis-Hastings

columnwidth 421.10046pt

1
Introduction

1.1 What is going on?, 3 facts, What is new in this thesis?

- hierachical Bayesian model, sampling to TT approx
- RTE as an example
- nonLinear to Linear Affine funciton (affine RTO)

1.2 What has been published?

2
Building a physics based hierarchical Linear model

- two hyperparameters, marginal and then conditonal, MTC, use as a building block
- sampling, Gibbs-MH, t-walk
- increase hyperparameters, Temperature and pressure, tt-approx

2.1 Linear model

- define model
- explain parameters
- how pressure to height and temperature

2.2 two hyperameters, fast sampling paper

- similar to MTC paper, ref
- can compare to regularization and can integrate easily with less solves to regularization or RTO, ref

- define priors
- how many steps, integrated autocorrealtion time , ref
- relative error
- how I found max curvature, ref
- how many taylor series

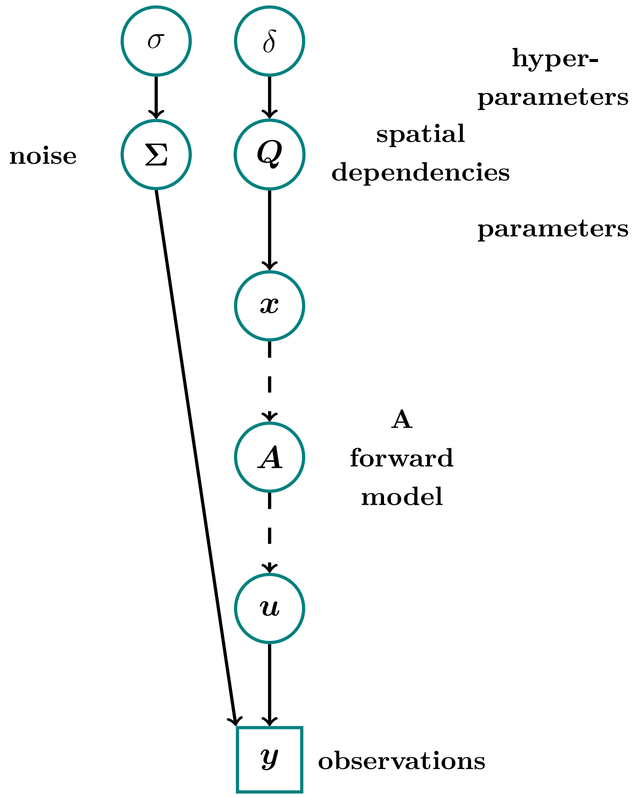

Figure 2.1

2.3 four hyperameters, t-walk, TT approx, and RTO

- t-walk ref
- motivation why more hyper-parameters, explain parabula
- how sample from them, compared to TT approx

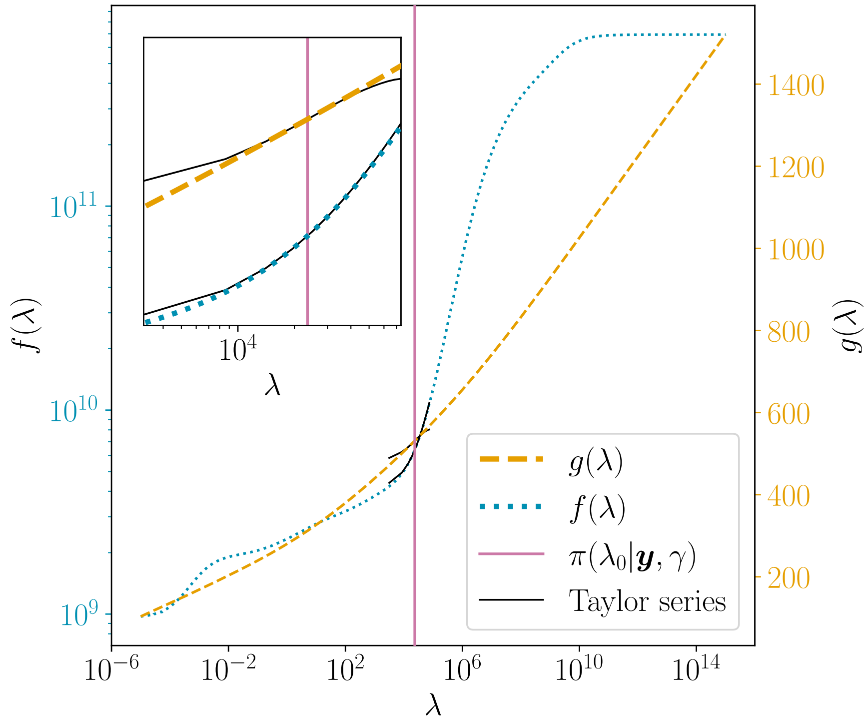

Figure 2.2: Functions $f(\lambda)$, dotted, and $g(\lambda)$, dashed, of the marginal posterior distribution for the specific forward model used in this study. Both functions are well-behaved over a large range of λ. In the support region of the MWG the pink square refers to the mode of the marginal posterior. Additionally, we plot the Taylor series of fourth order for $f(\lambda)$ and $g(\lambda)$ around the mode, see black line.

- explain Rosenblatt and SIRT trasnport

- RTO

- define priors

- how many steps, integrated autocorrealtion time

- relative error

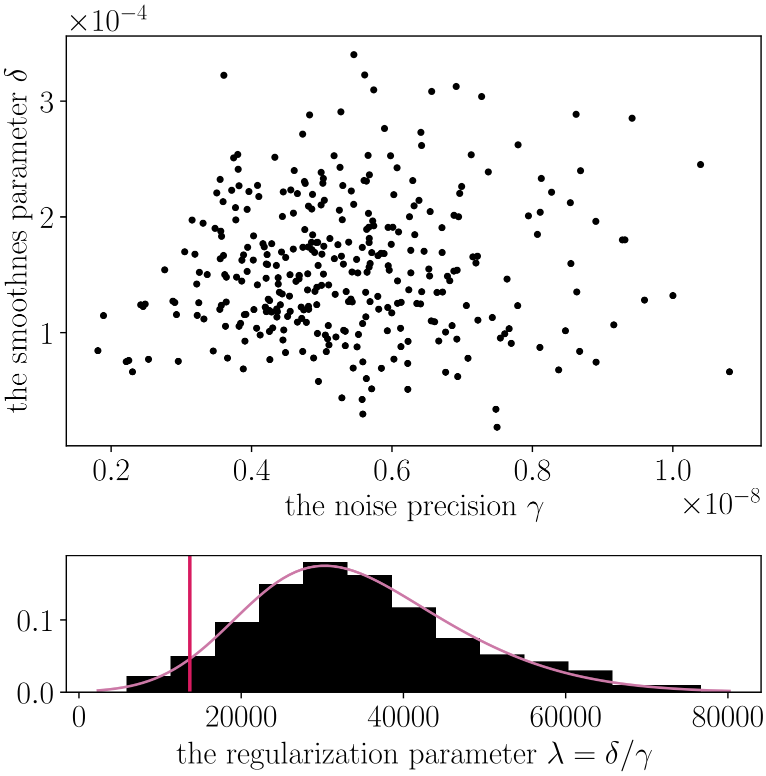

Figure 2.3: The scatter plot shows independent samples of δ and γ as the result of the MWG algorithm. The histogram displays independent samples of $\lambda \sim \pi(\lambda|\boldsymbol{y}, \gamma)$. The vertical line corresponds to the optimal regularization parameter.

2.3.1 Sampling

2.3.2 Rosenblatt Transform

2.3.3 (Squared) Inverse Rosenblatt Transform

2.3.4 Tensor-Train Approximation

2.4 Temperature and pressure hyperameters, tt-approx

- updating scheme

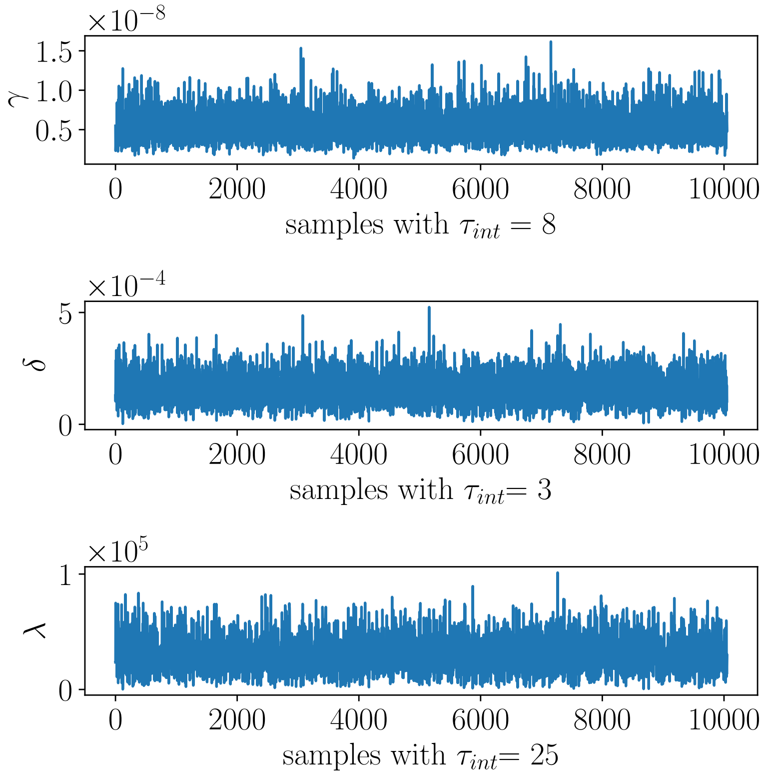

Figure 2.4: text

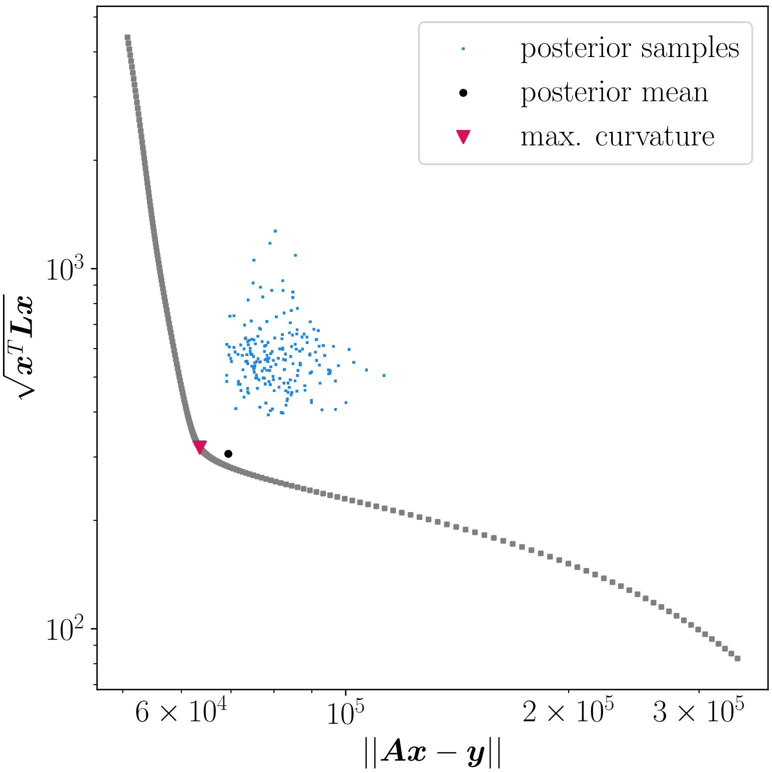

Figure 2.5: For varying λ we plot the seminorm $\sqrt{\boldsymbol{x}_\lambda^T \boldsymbol{L} \boldsymbol{x}_\lambda}$ against data misfit $\|\boldsymbol{A}\boldsymbol{x}_\lambda - \boldsymbol{y}\|$ of the regularised profiles. The triangle marks the point of maximum curvature closest to the origin of the L-curve. We plot the seminorm and the data misfit of the conditional posterior samples as well as of the posterior mean.

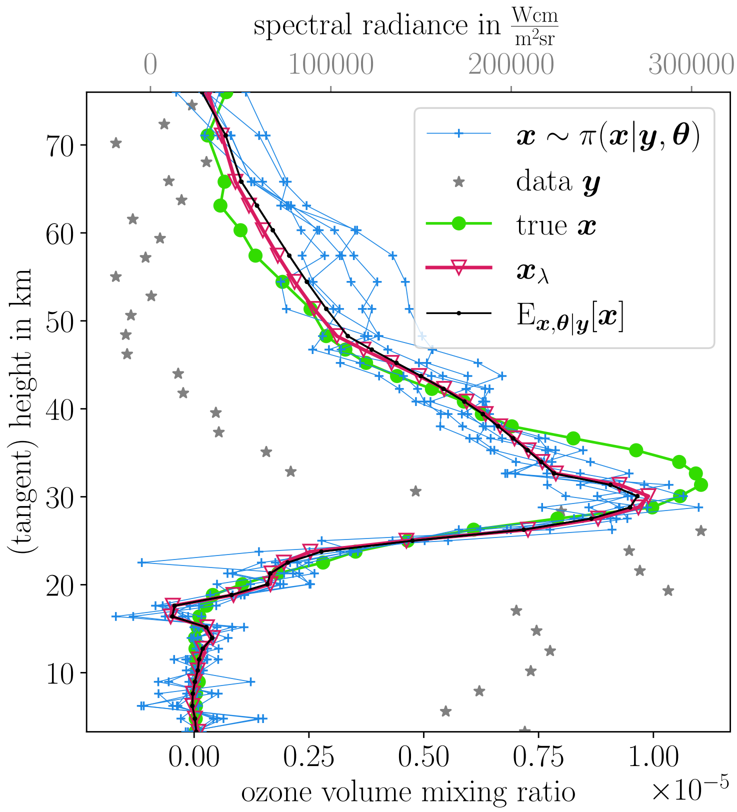

Figure 2.6: Plot of the true ozone profile (●), posterior samples (+), and posterior mean (●). We display the optimal regularised solution (▽) and the simulated data (∗) in spectral radiance.

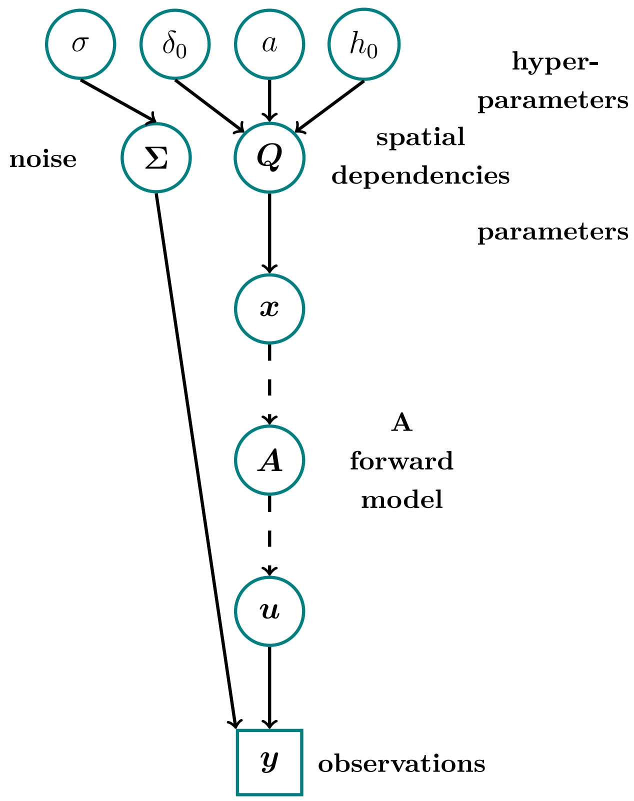

Figure 2.7: text

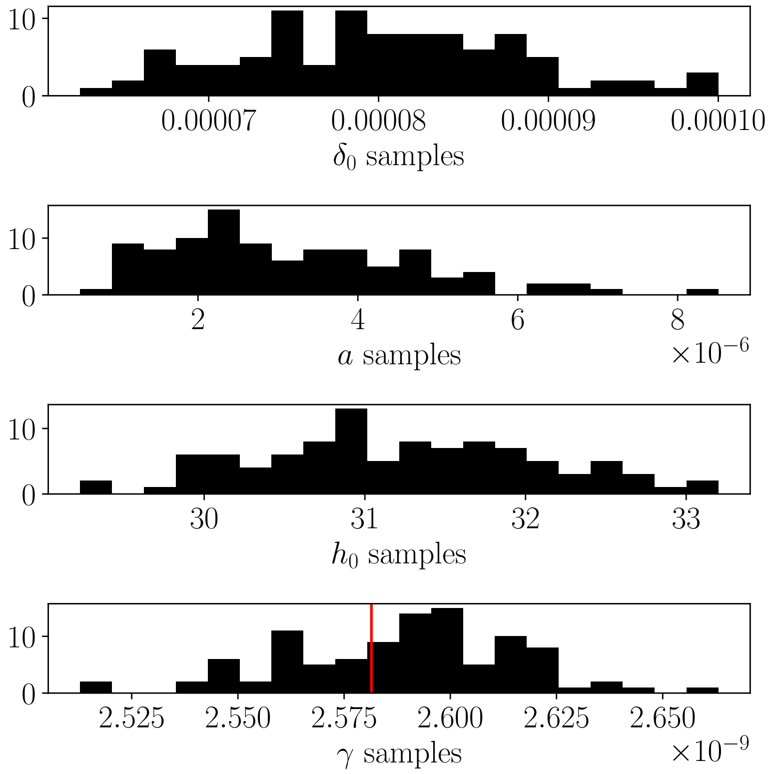

Figure 2.8: text

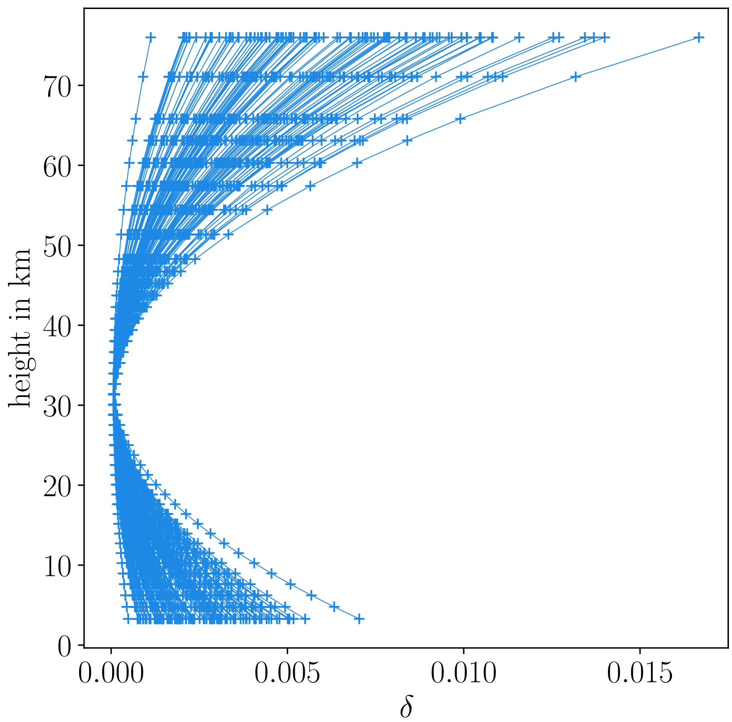

Figure 2.9: text

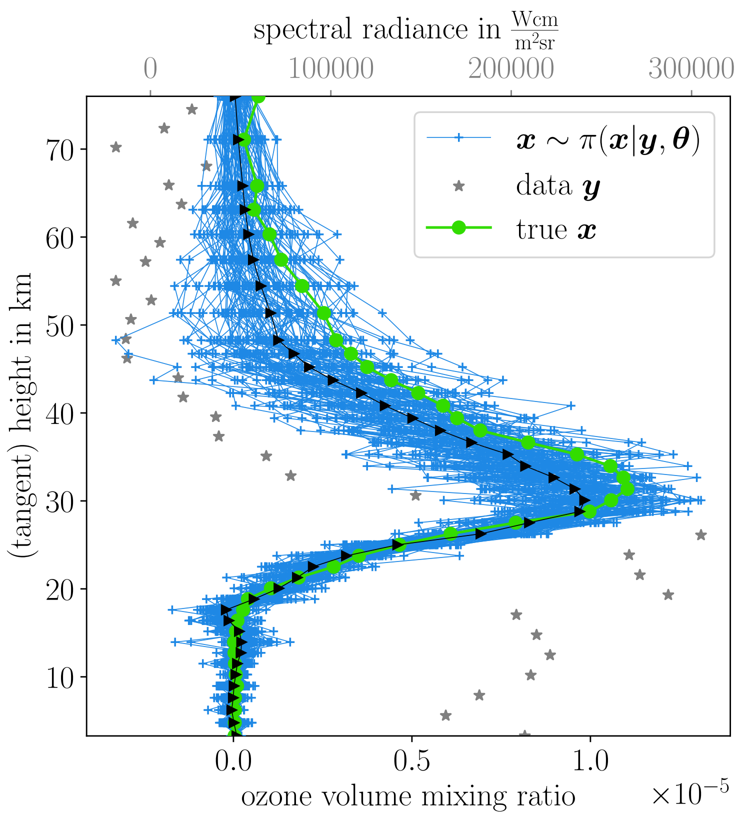

Figure 2.10: text

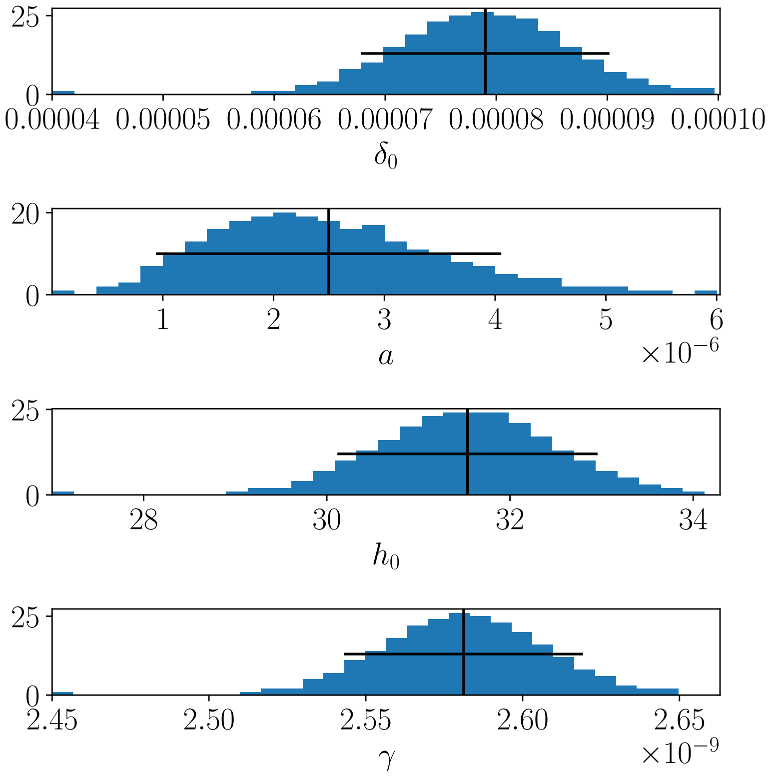

Figure 2.11: text

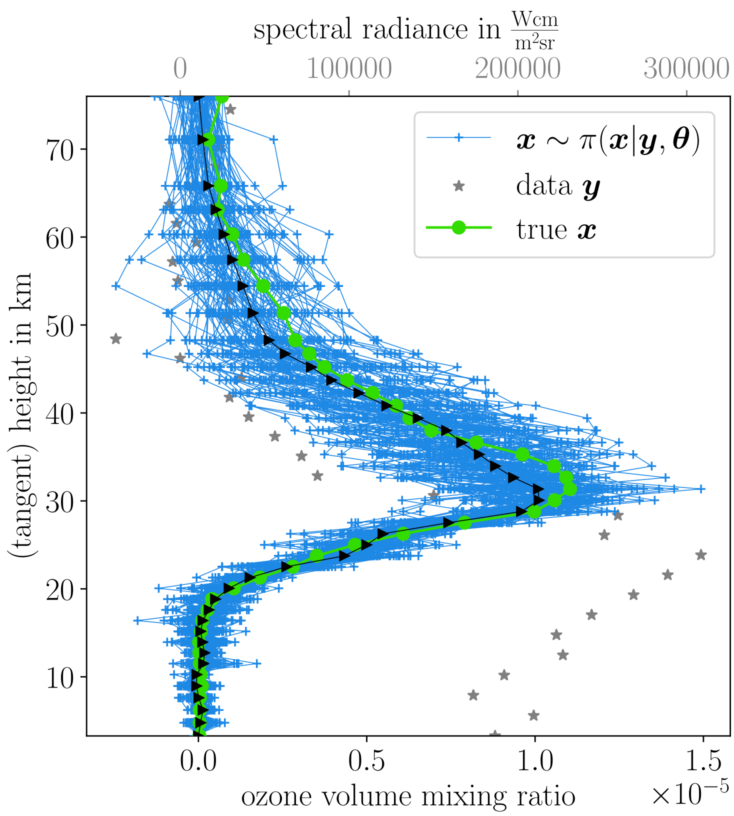

Figure 2.12: text

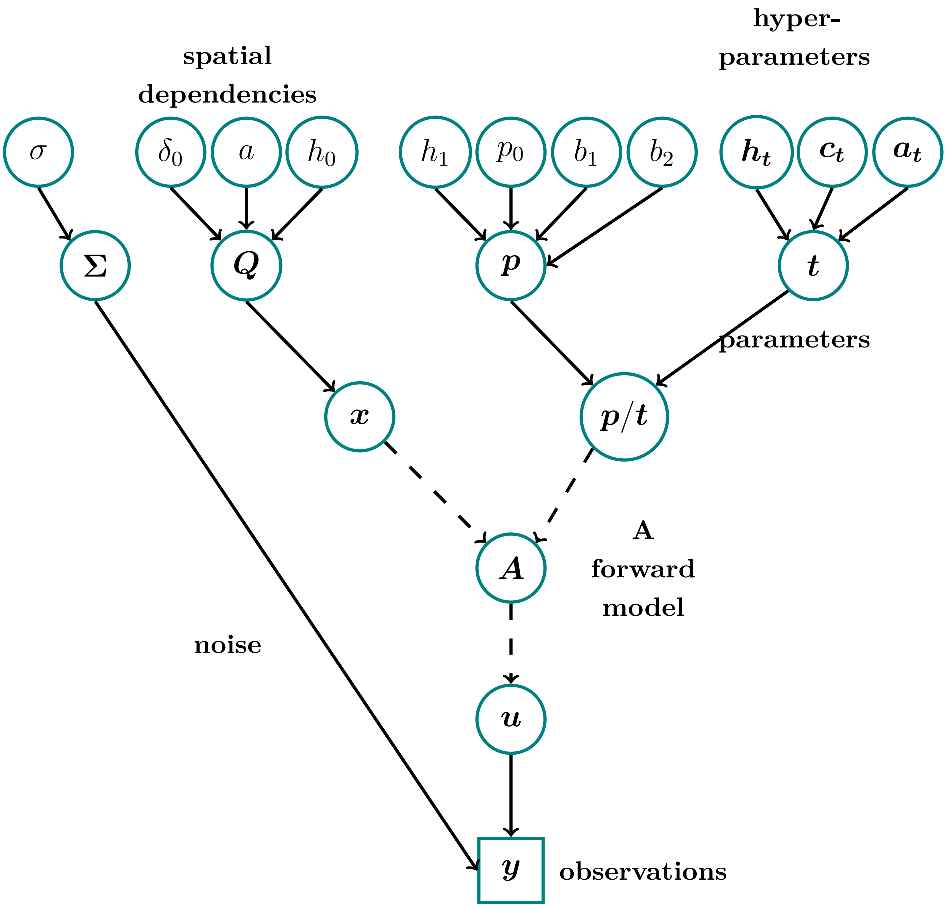

Figure 2.13: text

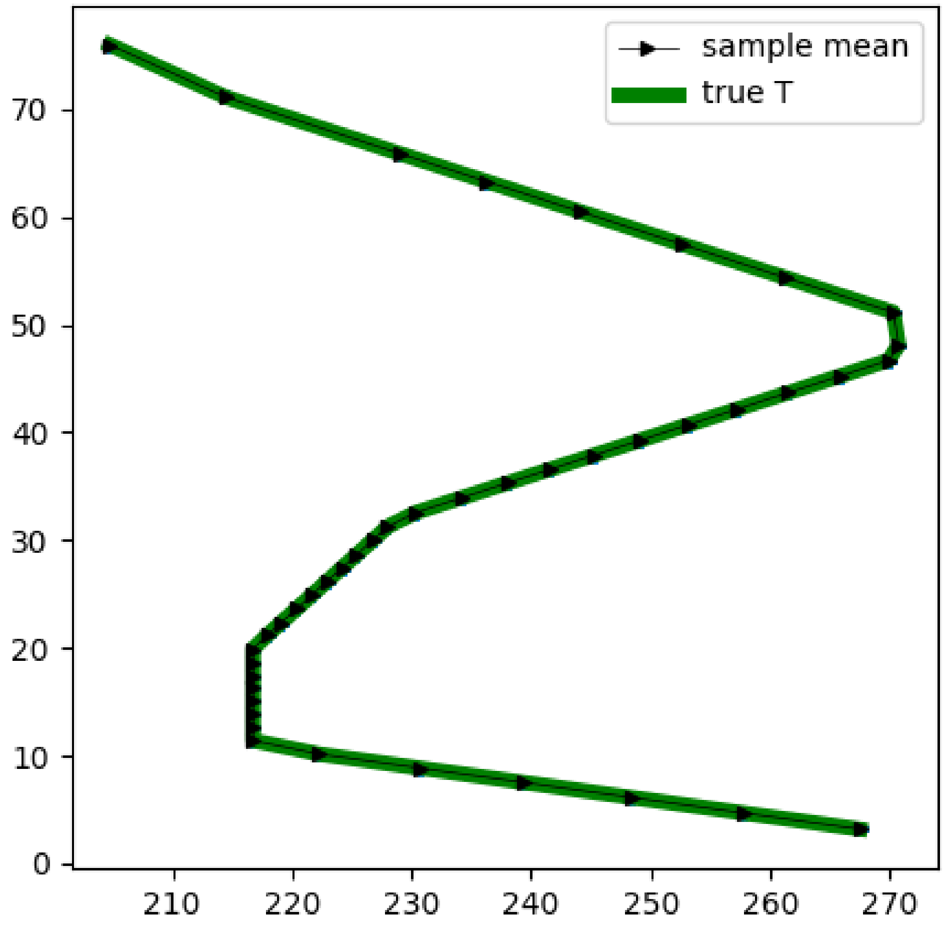

Figure 2.14: text

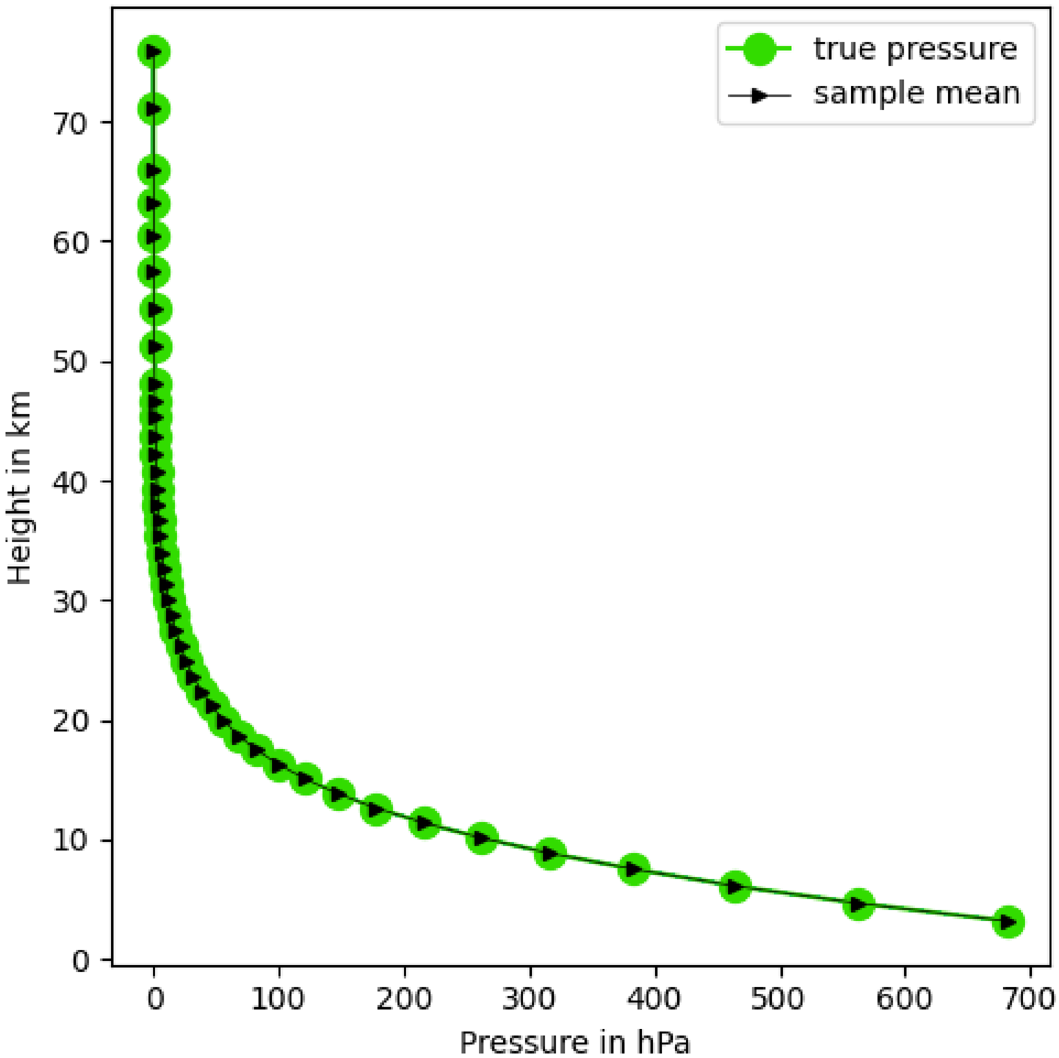

Figure 2.15: text

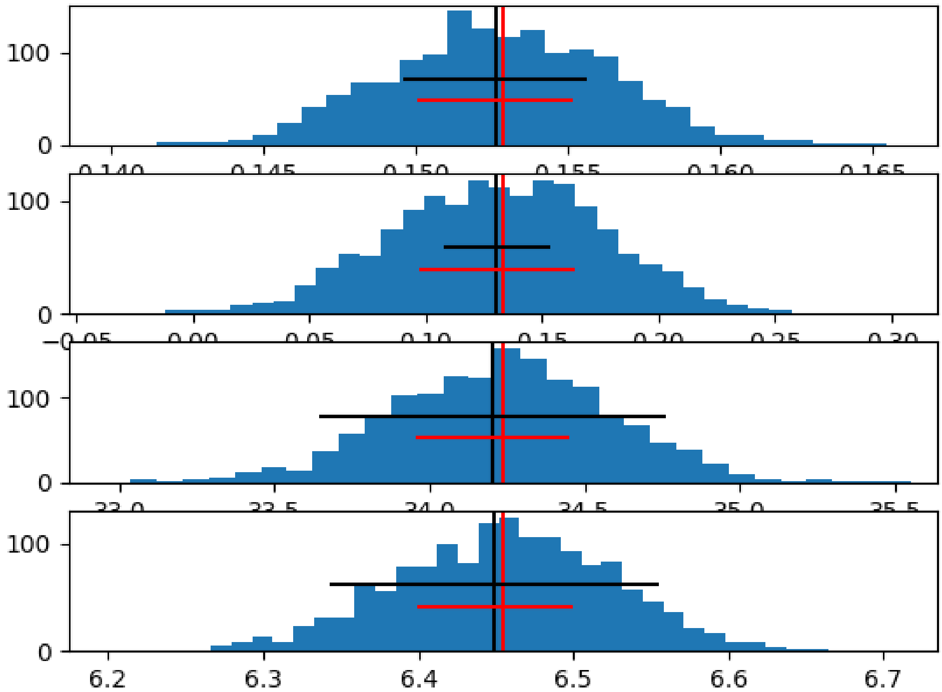

Figure 2.16: text

3

Nonlinear Forward model

- updating scheme, slow
- local linear map, strategy, schematic
- affine function, RTO

3.1 Sampling

3.2 local linear Map and strategy

- one data vector
- strategy to find convergence and local linear map

3.2.1 Machine learning vs Gaussian elimination

- linear solve
- machine learning class optimizer which package

3.3 affine RTO

- does it sample from the correct distribution

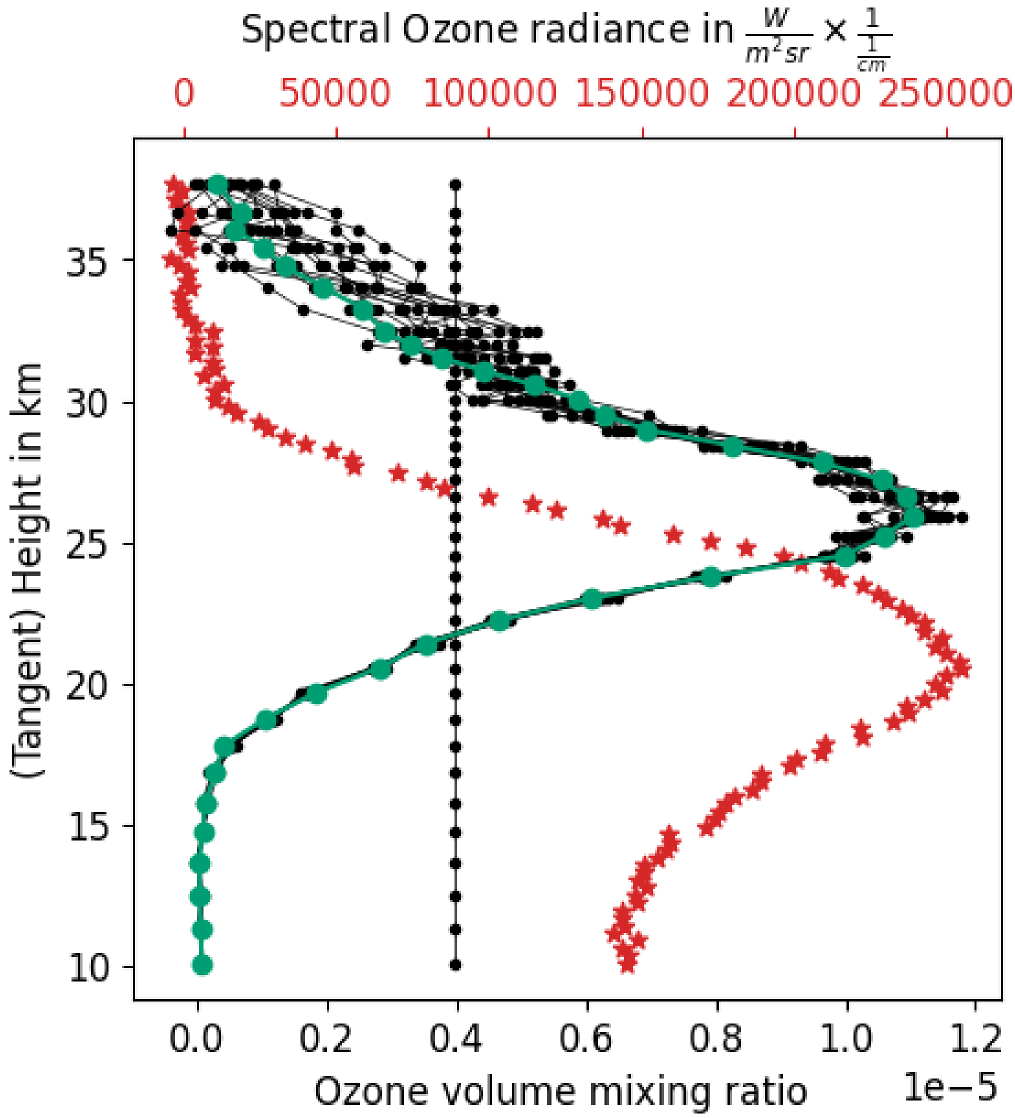

Figure 3.1: text

4
Introduction

4.1 What is going on?, 3 facts, What is new in this thesis?

- hierachical Bayesian model, sampling to TT approx
- RTE as an example
- nonLinear to Linear Affine funciton (affine RTO)

4.2 What has been published?

Appendices

A
Posterior of Bayesian Hierachical model

Here we show how to obtain the posterior covariance and mean of our hierarchical Bayesian model in **??** - **??**. We do not consider the hyper-parameters and start with the joint probaility distribution of $(\boldsymbol{x}^T, \boldsymbol{y}^T)^T$, where $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{y} \in \mathcal{Y}$ do not intersect. For more details we refer to Chapter 2 in [21] and to the book of Rue and Held [1].

The exponent of the normal Gaussian can be rewritten into:

$$-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{Q}(\boldsymbol{x}-\boldsymbol{\mu}) = -\frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{x}^T \boldsymbol{Q}\boldsymbol{\mu} + \text{const.} \tag{A.1}$$

We like to bring the joint distribution into a similar form so that we can compare the linear and second order terms and find the precision matrix and mean of the joint distribution.

In general the joint distribution to find the experssino for the postiror dostrbution

We can express this posterior through the likelihood and prior probability by Bayesian theorem, with a constant and positive normalization constant:

$$\pi(\boldsymbol{x}|\boldsymbol{y}) \propto \pi(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x}) \tag{A.2}$$

Taking the logarithmic function of this formulation we can find an expression for the the posterior covariance, with the $\mathrm{Var}(\boldsymbol{x}) = \boldsymbol{Q}_{\boldsymbol{x}}^{-1}$ and $\mathrm{Var}(\boldsymbol{y}) = \boldsymbol{Q}_{\boldsymbol{y}}^{-1}$.

$$\begin{aligned}
\ln \pi(\boldsymbol{x}|\boldsymbol{y}) \propto & \ln \pi(\boldsymbol{y}|\boldsymbol{x}) + \ln \pi(\boldsymbol{x}) && \text{(A.3)}\\
= & -\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{Q}_{\boldsymbol{x}}(\boldsymbol{x}-\boldsymbol{\mu}) - \frac{1}{2}(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x})^T \boldsymbol{Q}_{\boldsymbol{y}}(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}) && \text{(A.4)}\\
= & -\frac{1}{2}\Big[\boldsymbol{x}^T\Big[\boldsymbol{Q}_{\boldsymbol{x}} + \boldsymbol{A}^T\boldsymbol{Q}_{\boldsymbol{y}}\boldsymbol{A}\Big]\boldsymbol{x} + \boldsymbol{x}^T\Big[-\boldsymbol{A}^T\boldsymbol{Q}_{\boldsymbol{y}}\Big]\boldsymbol{y} && \text{(A.5)}\\
& + \boldsymbol{y}^T\Big[-\boldsymbol{Q}_{\boldsymbol{y}}\boldsymbol{A}\Big]\boldsymbol{x} + \boldsymbol{y}^T\Big[\boldsymbol{Q}_{\boldsymbol{y}}\Big]\boldsymbol{y} - 2\boldsymbol{x}^T\boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{\mu}\Big] + \text{const.} && \text{(A.6)}
\end{aligned}$$

Hence we deal with a Gaussian distribution, we consider second order terms only and rearrange to the precision matrix.

$$-\frac{1}{2}\Big[\boldsymbol{x}^T\Big[\boldsymbol{Q}_{\boldsymbol{x}} + \boldsymbol{F}^T\boldsymbol{Q}_{\boldsymbol{y}}\boldsymbol{F}\Big] + \boldsymbol{y}^T\Big[-\boldsymbol{Q}_{\boldsymbol{y}}\boldsymbol{F}\Big] \quad \boldsymbol{y}^T\Big[\boldsymbol{Q}_{\boldsymbol{y}}\Big] + \boldsymbol{x}^T\Big[-\boldsymbol{F}^T\boldsymbol{Q}_{\boldsymbol{y}}\Big]\Big]\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \tag{A.7}$$

$$= \begin{bmatrix}\boldsymbol{x}^T & \boldsymbol{y}^T\end{bmatrix}\underbrace{\begin{bmatrix}\boldsymbol{Q}_x + \boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{F} & -\boldsymbol{F}^T\boldsymbol{Q}_y\\ -\boldsymbol{Q}_y\boldsymbol{F} & \boldsymbol{Q}_y\end{bmatrix}}_{\text{precision matrix}}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{bmatrix} \tag{A.8}$$

We denote the precision matrix of the joint field as:

$$\boldsymbol{Q}_{xy} = \begin{bmatrix}\boldsymbol{Q}_{aa} & \boldsymbol{Q}_{ab}\\ \boldsymbol{Q}_{ba} & \boldsymbol{Q}_{bb}\end{bmatrix} = \begin{bmatrix}\boldsymbol{Q}_x + \boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{F} & -\boldsymbol{F}^T\boldsymbol{Q}_y\\ -\boldsymbol{Q}_y\boldsymbol{F} & \boldsymbol{Q}_y\end{bmatrix} \tag{A.9}$$

The mean is defined through the linear term.

$$\frac{-2\boldsymbol{x}^T\boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{\mu}}{-2} = \begin{bmatrix}\boldsymbol{x}^T & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{\mu}\\ 0\end{bmatrix} \tag{A.10}$$

Comparing to the linear term of Equation A.1 we can formulate an expression for the joint mean:

$$\Rightarrow \boldsymbol{\mu}_{\boldsymbol{xy}} = \boldsymbol{Q}_{\boldsymbol{xy}}^{-1}\begin{bmatrix}\boldsymbol{Q}_{\boldsymbol{x}}\boldsymbol{\mu}\\ 0\end{bmatrix} \tag{A.11}$$

The mean of the conditional distribution $\boldsymbol{x}|\boldsymbol{y}$ is given by:

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu}_{\boldsymbol{x}} + \boldsymbol{Q}_{ba}^{-1}\boldsymbol{Q}_{ab}(\boldsymbol{x} - \boldsymbol{\mu}_{\boldsymbol{y}}) \tag{A.12}$$

$$\boldsymbol{\mu}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{\mu} + (\boldsymbol{Q}_x + \boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{F})^{-1}\boldsymbol{F}^T\boldsymbol{Q}_y(\boldsymbol{x} - \boldsymbol{F}\boldsymbol{\mu})\,, \tag{A.13}$$

and the covariance of $\boldsymbol{x}|\boldsymbol{y}$ is given by:

$$\boldsymbol{Q}_{\boldsymbol{x}|\boldsymbol{y}} = \boldsymbol{Q}_{aa} = \boldsymbol{Q}_x + \boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{F}\,, \tag{A.14}$$

as illustrated through Theorem 2.5 in [1].

B
Convergence of the Metropolis-Hastings

If we show that the detailed balance condition holds and that the state space is irreducible and aperiodic under the transition matrix $\boldsymbol{P}$, we generate a Markov chain with a unique stationary distribution proportional to $\pi(\boldsymbol{x}, \boldsymbol{\theta}|\boldsymbol{y})$. Since the posterior is strictly positive $\pi(\boldsymbol{x}, \boldsymbol{\theta}|\boldsymbol{y}) \geq 0$ on the finite state space $\Omega(\mathcal{X}, \theta)$ the generated chain is irreducable. Further, it is possible to reject any proposed state and stay in the current state, which leads to aperiodicity. The detailed balance holds for the case that $\boldsymbol{j} = \boldsymbol{i}$, but if $\boldsymbol{j} \neq \boldsymbol{i}$ it is not trivial. In case we accept $\{\boldsymbol{x}, \boldsymbol{\theta}\}^{(n+1)} = \boldsymbol{j}$ as the new state we have $\pi(\boldsymbol{j}|\boldsymbol{y})g(\boldsymbol{i}|\boldsymbol{j}) > \pi(\boldsymbol{i}|\boldsymbol{y})g(\boldsymbol{j}|\boldsymbol{i})$. This gives us $\alpha(\boldsymbol{j}|\boldsymbol{i}) = 1$ and $\alpha(\boldsymbol{i}|\boldsymbol{j}) = \frac{\pi_{\boldsymbol{i}} g(\boldsymbol{j}|\boldsymbol{i})}{\pi_{\boldsymbol{j}} g(\boldsymbol{i}|\boldsymbol{j})}$ and satisfies the detailed balance:

$$\cancel{\pi_{\boldsymbol{j}}} \frac{\pi_{\boldsymbol{i}}}{\cancel{\pi_{\boldsymbol{j}}}} g(\boldsymbol{j}|\boldsymbol{i}) = \pi_{\boldsymbol{i}} g(\boldsymbol{j}|\boldsymbol{i}) \quad .$$

If $\pi(\boldsymbol{j}|\boldsymbol{y})g(\boldsymbol{i}|\boldsymbol{j}) < \pi(\boldsymbol{i}|\boldsymbol{y})g(\boldsymbol{j}|\boldsymbol{i})$ then $\alpha(\boldsymbol{i}|\boldsymbol{j}) = 1$ and $\alpha(\boldsymbol{j}|\boldsymbol{i}) = \frac{\pi_{\boldsymbol{j}} g(\boldsymbol{i}|\boldsymbol{j})}{\pi_{\boldsymbol{i}} g(\boldsymbol{j}|\boldsymbol{i})}$, this satisfies the detailed balance as well.

In conclusion the Metropolis-Hastings algorithm samples from a unique distribution proportional to the posterior distribution.

C
Randomize then Optimize - RTO

$$\pi(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\theta}) \propto \pi(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\theta})\pi(\boldsymbol{x}|\boldsymbol{\theta}) \tag{C.1}$$

$$\propto \exp\Big[(\boldsymbol{F}\boldsymbol{x}-\boldsymbol{y})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{F}\boldsymbol{x}-\boldsymbol{y}) + (\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{Q}(\boldsymbol{x}-\boldsymbol{\mu})\Big] \tag{C.2}$$

$$= \exp\|\hat{\boldsymbol{F}}\boldsymbol{x}-\hat{\boldsymbol{y}}\|^2 \tag{C.3}$$

where

$$\hat{\boldsymbol{F}} = \begin{bmatrix} \boldsymbol{\Sigma}^{-1/2}\boldsymbol{F} \\ \boldsymbol{Q}^{1/2} \end{bmatrix}, \quad \hat{\boldsymbol{y}} = \begin{bmatrix} \boldsymbol{\Sigma}^{-1/2}\boldsymbol{y} \\ \boldsymbol{Q}^{1/2}\boldsymbol{\mu} \end{bmatrix} \tag{C.4}$$

One sample from the posterior can be computed by minimizing the following with respect to $\boldsymbol{x}$

$$\boldsymbol{x} = \arg\min_{\hat{\boldsymbol{x}}} \|\hat{\boldsymbol{F}}\hat{\boldsymbol{x}} - (\hat{\boldsymbol{y}}+\boldsymbol{\eta})\|^2, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0},\mathbf{I}) \tag{C.5}$$

We can solve this and rewrite to

$$\frac{\partial}{\partial \boldsymbol{x}}\Big[(\hat{\boldsymbol{F}}\boldsymbol{x} - (\hat{\boldsymbol{y}}+\boldsymbol{\eta}))^T(\hat{\boldsymbol{F}}\boldsymbol{x} - (\hat{\boldsymbol{y}}+\boldsymbol{\eta})\Big] = 0 \tag{C.6}$$

$$\Leftrightarrow \boldsymbol{x}^T\hat{\boldsymbol{F}}^T\hat{\boldsymbol{F}} + \hat{\boldsymbol{F}}^T\hat{\boldsymbol{F}}\boldsymbol{x} - \hat{\boldsymbol{F}}^T(\hat{\boldsymbol{y}}+\boldsymbol{\eta}) - (\hat{\boldsymbol{y}}+\boldsymbol{\eta})^T\hat{\boldsymbol{F}}\boldsymbol{x} = 0 \tag{C.7}$$

We can argue through the symmetry of the inner product that and the symmetry of the precision matrix

$$\hat{\boldsymbol{F}}^T\hat{\boldsymbol{F}}\boldsymbol{x} = \hat{\boldsymbol{F}}^T(\hat{\boldsymbol{y}}-\boldsymbol{\eta}) \tag{C.8}$$

$$\Leftrightarrow (\boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{F}+\boldsymbol{Q})\boldsymbol{x} = \boldsymbol{F}^T\boldsymbol{Q}_y\boldsymbol{y} + \boldsymbol{Q}\boldsymbol{\mu} - \hat{\boldsymbol{F}}^T\boldsymbol{\eta} \tag{C.9}$$

If we substitute $-\hat{\boldsymbol{F}}^T\boldsymbol{\eta} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ we end up with

$$(\boldsymbol{F}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{F} + \boldsymbol{Q})\boldsymbol{x} = \boldsymbol{F}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y} + \boldsymbol{Q}\boldsymbol{\mu} + \boldsymbol{v}_1 + \boldsymbol{v}_2 \tag{C.10}$$

where $\boldsymbol{v}_1 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{F}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{F})$ and $\boldsymbol{v}_2 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ are independent random variables. mayeb introduce... x^2 time nomral variubale

D

Inverting Matrices - QR factorization

E
Taylor expansion of $g(\lambda)$

We Taylor expand the function $g(\lambda)$ around $\lambda = \lambda' - \Delta\lambda$

$$g(\lambda) = \ln\det \underbrace{(\boldsymbol{F}^T \boldsymbol{F} + \lambda \boldsymbol{L})}_{\boldsymbol{B}} \tag{E.1}$$

$$\begin{aligned} g(\lambda') - g(\lambda) &= \ln\det(\boldsymbol{F}^T \boldsymbol{F} + \lambda' \boldsymbol{L}) - \ln\det(\boldsymbol{F}^T \boldsymbol{F} + \lambda \boldsymbol{L}) && \text{(E.2)} \\ &= \ln\det\left[\frac{(\boldsymbol{F}^T \boldsymbol{F} + (\lambda + \Delta\lambda)\boldsymbol{L})}{(\boldsymbol{F}^T \boldsymbol{F} + \lambda \boldsymbol{L})}\right] && \text{(E.3)} \\ &= \ln\det\left[1 + \frac{\Delta\lambda \boldsymbol{L}}{\boldsymbol{B}}\right] && \text{(E.4)} \\ &= \sum_{r=1}^{\infty} \frac{(-1)^{r+1}}{r!} \mathrm{tr}((\boldsymbol{B}^{-1}\boldsymbol{L})^r)(\Delta\lambda)^r && \text{(E.5)} \end{aligned}$$

, where we use the identity from [22] at page 29. So the derivatives of $g(\lambda)$ are:

$$\begin{aligned} g^{(r)}(\lambda) =& (-1)^{r+1}\,\mathrm{tr}\Big((\boldsymbol{B}^{-1}\boldsymbol{L})^r\Big) && \text{(E.6)} \\ \approx& (-1)^{r+1} \sum_{k=1}^{p} \boldsymbol{z}_{\boldsymbol{k}}^T (\boldsymbol{B}^{-1}\boldsymbol{L})^r \boldsymbol{z}_{\boldsymbol{k}} && \text{(E.7)} \end{aligned}$$

Here we use a Monte Carlo estimate and draw p vectors $\boldsymbol{z}_{\boldsymbol{k}} \in \mathbb{R}^n$, where each vector element $z_i \overset{\text{i.i.d.}}{\sim} \mathcal{U}(\{-1, 1\})$ and $i = 1, \ldots, n$.

F
Radiation transfer and absorption line shape

G
whispering gallery resonator

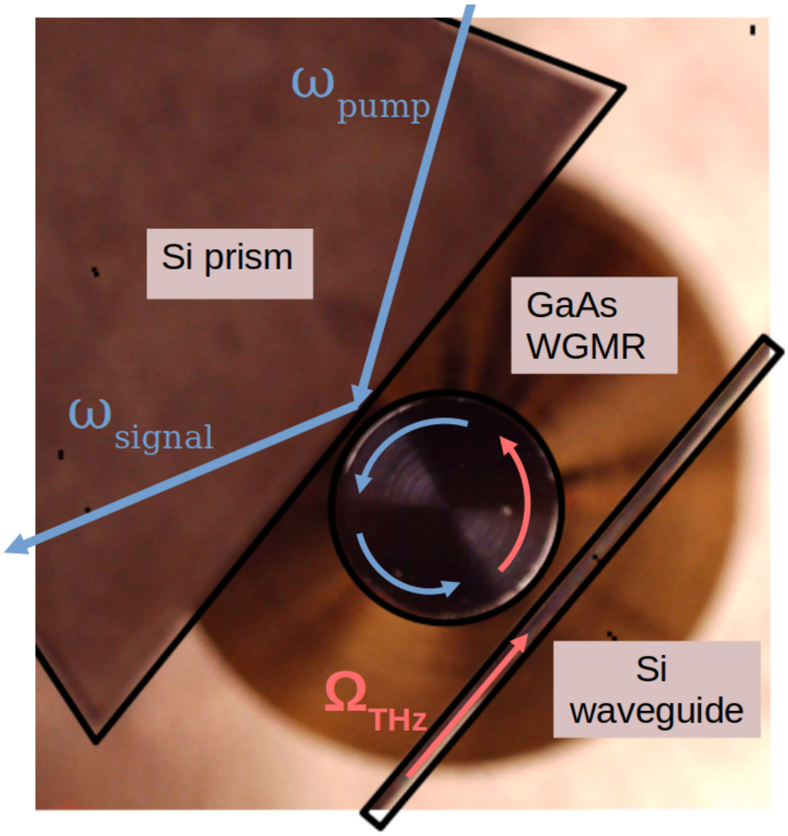

Figure G.1: whispering gallery resonator

References

[1] Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications*. Chapman and Hall/CRC, 2005.

[2] Dave Higdon. “A primer on space-time modeling from a Bayesian perspective”. In: *Monographs on Statistics and Applied Probability* 107 (2006), p. 217.

[3] Pierre Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Vol. 31. Springer Science & Business Media, 2013.

[4] Julian Besag. “Spatial interaction and the statistical analysis of lattice systems”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 36.2 (1974), pp. 192–225.

[5] Colin Fox and Richard A Norton. “Fast sampling in a linear-Gaussian inverse problem”. In: *SIAM/ASA Journal on Uncertainty Quantification* 4.1 (2016), pp. 1191–1218.

[6] Daniel Simpson, Finn Lindgren, and Håvard Rue. “Think continuous: Markovian Gaussian models in spatial statistics”. In: *Spatial Statistics* 1 (2012), pp. 16–29.

[7] Nicholas Metropolis and Stanislaw Ulam. “The monte carlo method”. In: *Journal of the American statistical association* 44.247 (1949), pp. 335–341.

[8] John Michael Hammersley and David Christopher Handscomb. “General principles of the Monte Carlo method”. In: *Monte Carlo Methods*. Springer, 1964, pp. 50–75.

[9] Paula A Whitlock and MH Kalos. *Monte Carlo Methods*. Wiley, 1986.

[10] AA Markov. “Extension of the law of large numbers to quantities, depending on each other (1906). Reprint.” In: *Journal Électronique d’Histoire des Probabilités et de la Statistique [electronic only]* 2.1b (2006), Article–10.

[11] Colin Fox, Geoff K Nicholls, and Sze M Tan. “An Introduction to Inverse Problems”. In: *Course notes for ELEC* 404 (2010).

[12] W Keith Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: (1970).

[13] Nicholas Metropolis et al. “Equation of state calculations by fast computing machines”. In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.

[14] Johnathan M Bardsley. “MCMC-based image reconstruction with uncertainty quantification”. In: *SIAM Journal on Scientific Computing* 34.3 (2012), A1316–A1332.

[15] Johnathan M Bardsley et al. “Randomize-then-optimize for sampling and uncertainty quantification in electrical impedance tomography”. In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 1136–1158.

[16] D S. Oliver, Nanqun He, and A C. Reynolds. *Conditioning permeability fields to pressure data*. 1996, cp–101.

[17] François Orieux, Olivier Féron, and J-F Giovannelli. “Sampling high-dimensional Gaussian distributions for general linear inverse problems”. In: *IEEE Signal Processing Letters* 19.5 (2012), pp. 251–254.

[18] H Fischer et al. “Envisat-Mipas, the Michelson Interferometer for Passive Atmospheric Sounding; An instrument for atmospheric chemistry and climate research”. In: *ESA SP* 1229 (2000).

[19] J Andrés Christen and Colin Fox. “A general purpose sampling algorithm for continuous distributions (the t-walk)”. In: (2010).

[20] Ulli Wolff. *Matlab function UWerr.m Version6 described in the paper ‘Monte Carlo errors with less errors’*. `https://www.physik.hu-berlin.de/de/com/ALPHAsoft`. [Online; accessed 22-August-2023]. 2003.

[21] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*. Vol. 4. 4. Springer, 2006.

[22] Israel Gohberg, Seymour Goldberg, and Nahum Krupnik. *Traces and determinants of linear operators*. Vol. 116. Birkhäuser, 2012.