A Bayesian Generative Model With Gaussian Process Priors For Thermomechanical Analysis Of Micro-Resonators

M. F. Vording@[a], P. O. Okeyo[c], J. J. R. Guillamón[a], P. E. Larsen[b], M. N. Schmidt[a], T. S. Alstrøm[a]

[a] DTU Compute & [b] DTU Health Tech, Technical University of Denmark [c] Department of Pharmacy, University of Copenhagen @ maxvo@dtu.dk

Summary

Motivation

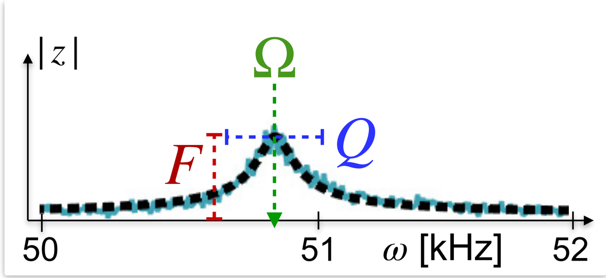
- **Thermal analysis** of resonating micro-electromechanical systems (MEMS)
- **Characterise properties** of drugs & materials in early-phase development

Problem

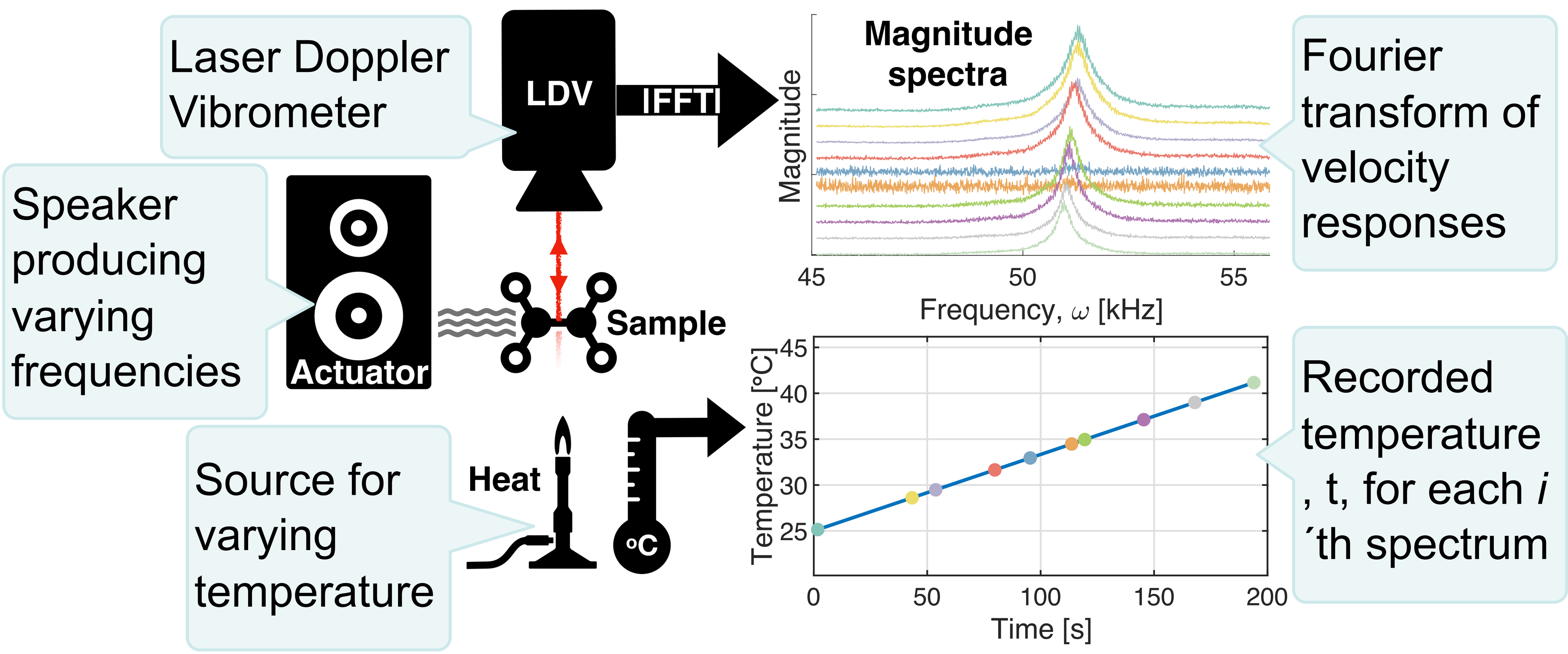
- **Unknown global relation**: resonance frequency vs. temperature over time
- **Low signal-to-noise-ratio** for specific areas of resonator = excluded data
- **Filter + Annotate:** tedious work and only fit high SNR spectra one-by-one

Solution

- **Unsupervised tracking** of peak height (F), position (Ω) and shape (Q)
- **Bayesian generative model** with
- **Warped Gaussian process priors** to
- **Regularise towards non-negative & smooth**
- **Impute peaks** missing or disappearing in noise



Schematic setup for obtaining data



Methods

Governing Function

Magnitude solution for driven-damped vibration on a linear resonator [12]

$$f(\omega_j,\theta_i)=\frac{F_i\Omega_i^2}{\sqrt{Q_i^2(\Omega_i^2-\omega_j^2)^2+\Omega_i^2\omega_j^2}}.$$

$$\theta_i=\{F_i,\Omega_i,Q_i\},$$

Generative Model

Gaussian noise model over j´th frequency and i´th spectrum

$$z_{i,j}=f(\omega_j,\theta_i)+\epsilon_{i,j}.\qquad \epsilon_{i,j}\sim\mathcal{N}(0,\sigma_\epsilon)$$

Normal Likelihood

Joint by factorising over all $\{i, j\}$

$$p(\boldsymbol{Z}|\boldsymbol{\Theta})=\prod_{i=1}^{N}\prod_{j=1}^{M}\mathcal{N}(z_{i,j};f(\omega_j,\theta_i),\sigma_\epsilon^2),$$

$$\boldsymbol{\Theta}=\{\theta_i\}_{i=1}^N=\{\boldsymbol{H}_k\}_{k=1}^{L=3}=\{\boldsymbol{F},\boldsymbol{\Omega},\boldsymbol{Q}\}$$

Maximize Posterior

Parameter estimates for all i
Solved in log and GP prior domain

$$\hat{\boldsymbol{\Theta}}=\underset{\boldsymbol{\Theta}}{\operatorname{argmax}}\,p(\boldsymbol{Z}|\boldsymbol{\Theta})p(\boldsymbol{\Theta}).\qquad \hat{\boldsymbol{\Theta}}=g(h(\hat{\boldsymbol{\eta}}))$$

$$\hat{\boldsymbol{\eta}}=\underset{\boldsymbol{\eta}}{\operatorname{argmax}}\left(\log p(\boldsymbol{Z}|g(h(\boldsymbol{\eta})))+\sum_{k=1}^{L}\boldsymbol{\eta}_k^\top\boldsymbol{\eta}_k\right)$$

Change of variables

To correlate and truncate parameters in non-negative domain [9, 10, 13]

$$\boldsymbol{\Theta}=g(h(\boldsymbol{\eta}))=\left\{\ell_k^{-1}(\boldsymbol{h}_k(\boldsymbol{\eta}_k))\right\}_{k=1}^{L}$$

$$\boldsymbol{h}_k\sim\mathcal{GP}(\boldsymbol{0},\boldsymbol{\Sigma}_k)\qquad \boldsymbol{\Sigma_k}=\boldsymbol{C}_k\boldsymbol{C}_k^\top,\qquad \boldsymbol{h}_k(\boldsymbol{\eta}_k)=\boldsymbol{C}_k\boldsymbol{\eta}_k,$$

$$\boldsymbol{\eta}_k\sim\mathcal{N}(0,\boldsymbol{I}),\qquad \boldsymbol{\eta}=\{\boldsymbol{\eta}_k\}_{k=1}^{L},\qquad \boldsymbol{\eta}_k\in\mathbb{R}^N$$

Inverse Link Function

Map η_k in CDF of marginal GP → then through inverse CDF of the truncated normal to non-negative domain in range [a, b]. →

$$\ell_k^{-1}(\boldsymbol{h}_k)=P_{H_k}^{-1}(P_{h_k}(\boldsymbol{h}_k)),$$

$$P_h(h_i)=\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{h_i}{\sqrt{2}\sigma_i}\right)\right),$$

$$P_H^{-1}(\xi)=\mu+\sigma\sqrt{2}\operatorname{erf}^{-1}\left(\xi\operatorname{erf}(\beta)+(1-\xi)\operatorname{erf}(\alpha)\right),$$

$$\beta=\frac{b-\mu}{\sigma\sqrt{2}}\ \text{and}\ \alpha=\frac{a-\mu}{\sigma\sqrt{2}}.$$

Covariance functions

Smoothing over differences $\Delta_i=(i-i')$
length scales λ over covariates $\phi=\{i,t,x,y\}$

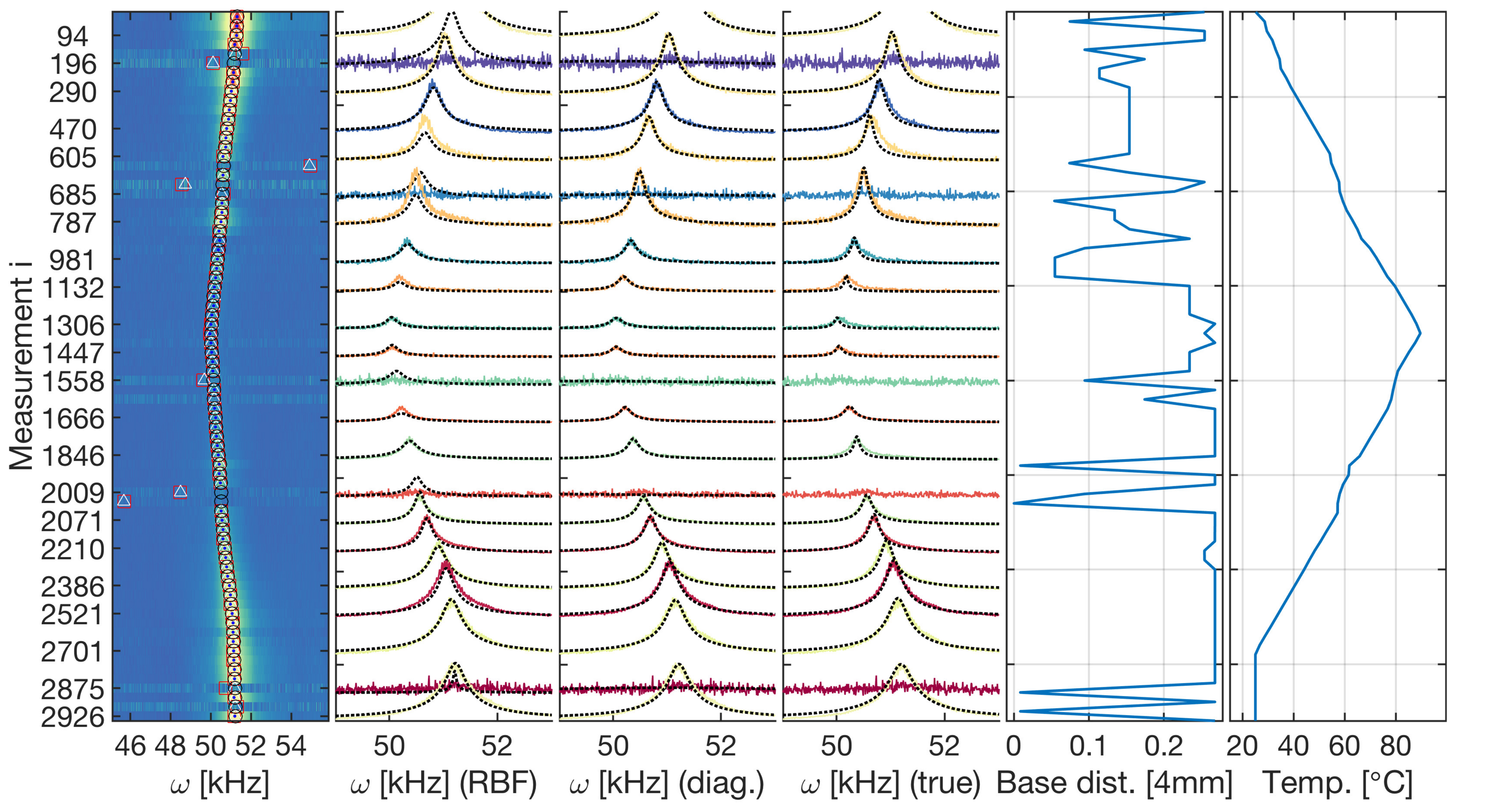
$$c(\phi,\phi')=\exp\left(-\frac{\Delta_i^2}{2\lambda_i^2}-\frac{\Delta_t^2}{2\lambda_t^2}-\frac{\Delta_x^2}{2\lambda_x^2}-\frac{\Delta_y^2}{2\lambda_y^2}\right)\quad \text{RBF}$$

$$c(\phi,\phi')=\delta(\Delta_i)\quad \text{Diagonal}$$

Results

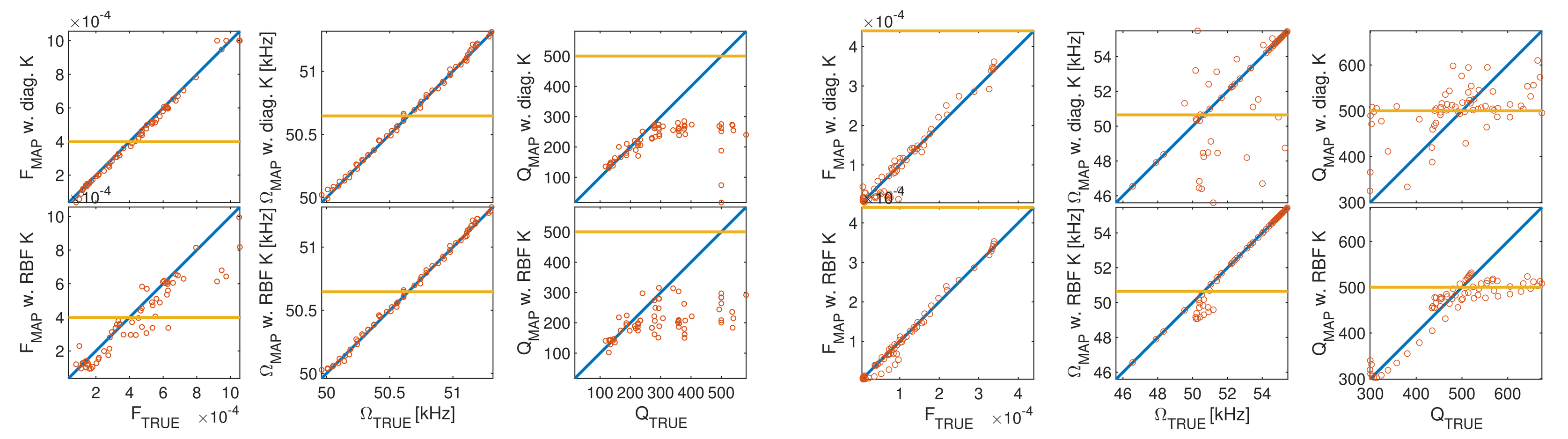
Smooth tracking of peaks in real and synthetic spectra

- **Col. 1:** Initial points [red square], diag. [white triangle], RBF [black circle]
- **Col. 2-4:** Spectra with estimated peaks in dashed black lines
- **RBF:** Regularised to coherent and imputed peaks even in low SNR noise
- **Diagonal:** Closer fit than expert annotations, but miss out low SNR peaks.



Comparison of models: Reproducing ground truth?

- **Diagonal kernel** [upper]: Closer to **expert annotations** on resonator data
- **RBF kernel** [lower]: Closer to **synthetic parameters** and **mode of prior**



(a.) Resonator data (b.) Synthetic data

	GPP w. RBF Σ		GPP w. diag. Σ		GPP w. RBF Σ		GPP w. diag. Σ	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
$\hat{F}$ $[10^{-5}]$	9.701	6.080	**2.339**	**1.736**	1.002(331)	0.673(228)	1.123(394)	0.704(240)
$\hat{\Omega}$ $[10^{2}]$	0.213	0.154	**0.209**	**0.146**	**9.743(3.596)**	**5.085(2.076)**	23.54(7.88)	15.95(6.39)
$\hat{Q}$ $[10^{2}]$	1.297	0.847	**1.281**	**0.722**	2.066(483)	1.451(395)	**1.011(210)**	**0.686(136)**

Conclusion

We find that using warped GP priors can

- **Regularise** parameter estimates to be **smooth** and **non-negative**
- **Impute peaks** with low SNR allowing for higher inclusion rate
- **Lead to underestimation** of peak shape **Q** and height **F**
- Potentially support **fast and efficient characterisation** of drugs

Future work on

- **Modelling shift in noise** σ_ϵ over frequencies and time with GP prior
- **Correcting underestimation of Q** by fitting phase shift slope
- **Metropolis-Hastings with Gibbs sampling** for uncertainty estimates
- **Active learning** to query expert for new annotations in training loop

Bibliography

[1] R. Hilfiker et al., "Polymorphism in the pharmaceutical industry". *Wiley Online Library*, 2006.

[2] J. Scheirs, "Compositional and failure analysis of polymers: a practical approach". *John Wiley & Sons*, 2000.

[3] S. Bose et al., "Micromechanical fast quasi-static detection of and relaxations with nanograms of polymer," *Journal of Polymer Science*, 2015.

[4] S. A. M. Bukhari et al., "Thermomechanical analysis of picograms of polymers using a suspended microchannel cantilever," *RSC Adv.*, 2017.

[5] M. Karl et al., "Ultrasensitive microstring resonators for solid state thermomechanical analysis of small and large molecules," Journal of the American Chemical Society, 2018.

[6] K. Ekinci and M. Roukes, "Nanoelectromechanical systems," *Review of scientific instruments*, 2005.

[7] P. O. Okeyo et al., "Imaging of dehydration in particulate matter using raman line-focus microscopy," *Scientific reports*, 2019.

[8] C. K. Williams and C. E. Rasmussen, "Gaussian processes for machine learning". *MIT Press Cambridge*, 2006.

[9] M. N. Schmidt and H. Laurberg, "Non-negative matrix factorization with gaussian process priors," *Computational Intelligence and Neuroscience*, 2008.

[10] M. N. Schmidt, "Function factorization using warped gaussian processes,", *ICML*, 2009.

[11] M. Zieba et al., "Gaussian process regression for automated signal tracking in step-wise perturbed nuclear magnetic resonance spectra," *Applied Soft Computing*.

[12] S. Schmid, L. G. Villanueva, and M. L. Roukes, "Fundamentals of nano-mechanical resonators", *Springer*, 2016.

[13] J. Burkardt, "The truncated normal distribution," *Department of Scientific Computing Website, Florida State University*, 2014.

DTU Compute
Department of Applied Mathematics and Computer Science

Danmarks Grundforskningsfond
Danish National Research Foundation

VILLUM FONDEN

DEPARTMENT OF PHARMACY AT UNIVERSITY OF COPENHAGEN