Modelling Non-Smooth Signals with Complex Spectral Structure

Wessel P. Bruinsma[12], Martin Tegnér*[34], and Richard E. Turner[1]
[1]University of Cambridge, [2]Invenia Labs, [3]University of Oxford, [4]Oxford-Man Institute

GP Convolutional Model (GPCM) [1]

The GPCM as a **linear system**:

$$x \sim \mathcal{GP}(0, \delta(t-t')), \quad h \sim \mathcal{GP}(0, k_h),$$

$$f(t)\,|\,h, x = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)\,\mathrm{d}\tau$$

Equivalently, this is a GP with **random kernel**:

$$h \sim \mathcal{GP}(0, k_h),$$

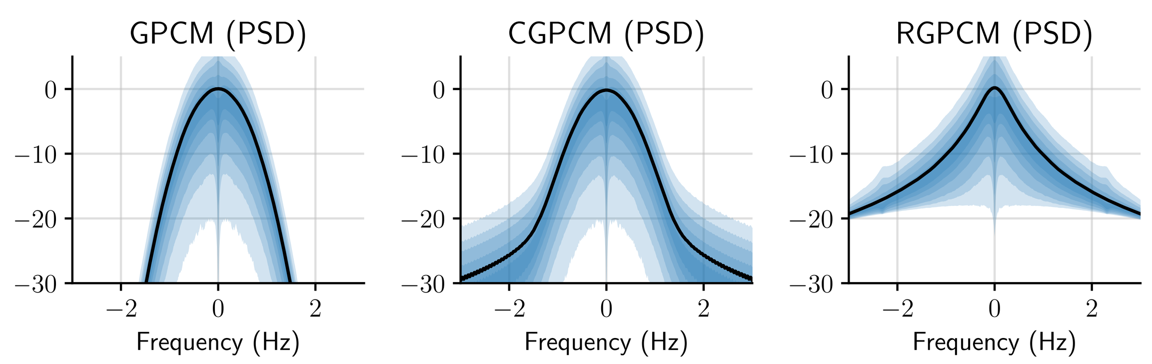
$$f\,|\,h \sim \mathcal{GP}\left(0, \int_{-\infty}^{\infty} h((t-t')+\tau)h(\tau)\,\mathrm{d}\tau\right)$$

GPCM: a flexible time-series model class based on a GP with **nonparametric kernel** learned from data by **probabilistic inference**.

Our contribution

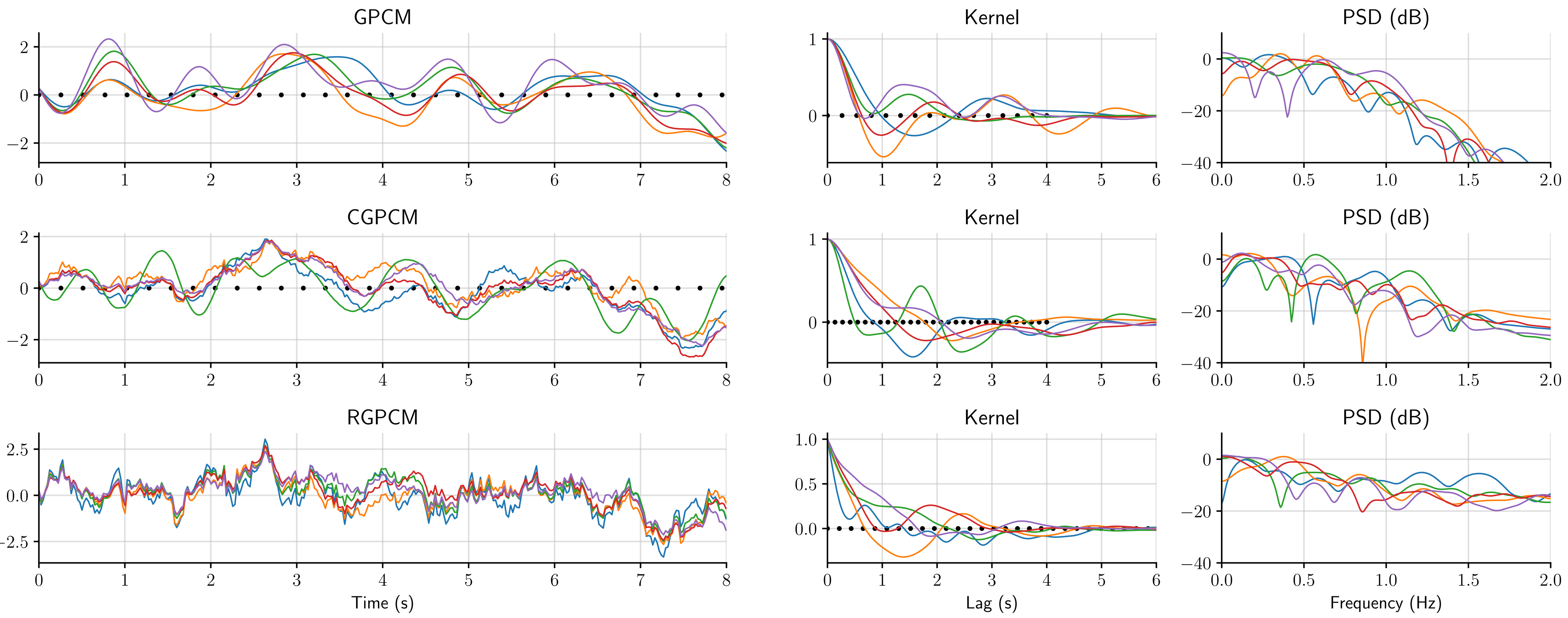
GPCM models **smooth** signals with **rapidly decaying spectrum**. Inference is **mean-field** with poor uncertainty estimates and tedious optimisation of large covariance matrices.

We propose **Causal** and **Rough GPCM**: relaxed smoothness assumptions for a **richer spectrum**. We also relax the mean-field assum. and circumvent variational optimisation.



The Causal and Rough GPCM

- **GPCM [1]:**
  $$f(t)\,|\,h, x = \int_{-\infty}^{\infty} e^{-\alpha(t-\tau)^2} h(t-\tau)x(\tau)\,\mathrm{d}\tau, \quad h \sim \mathcal{GP}(0, e^{-\gamma(t-t')^2}), \quad x \sim \mathcal{GP}(0, \delta(t-t'))$$
- **Causal GPCM (CGPCM):**
  $$f(t)\,|\,h, x = \int_{-\infty}^{t} e^{-\alpha(t-\tau)^2} h(t-\tau)x(\tau)\,\mathrm{d}\tau, \quad h \sim \mathcal{GP}(0, e^{-\gamma(t-t')^2}), \quad x \sim \mathcal{GP}(0, \delta(t-t'))$$
- **Rough GPCM (RGPCM):**
  $$f(t)\,|\,h, x = \int_{-\infty}^{t} e^{-\alpha|t-\tau|} h(t-\tau)x(\tau)\,\mathrm{d}\tau, \quad h \sim \mathcal{GP}(0, \delta(t-t')), \quad x \sim \mathcal{GP}(0, e^{-\lambda|t-t'|})$$



Left: Prior samples of f to demonstrate smoothness (GPCM, top) and non-differentiability (C/RGPCM, middle and bottom). **Right:** Corresponding samples from kernel and spectrum.

Variational inference beyond mean-field

Structured approximate posterior:

$$p_\theta(h, x, \mathbf{u}, \mathbf{z}|\mathbf{y}) \approx p_\theta(h\,|\,\mathbf{u})p_\theta(x\,|\,\mathbf{z})q(\mathbf{u}, \mathbf{z})$$

with $\mathbf{u}$ and $\mathbf{z}$ interdomain inducing points; VFFs [2] for $\mathbf{z}$. **Optimal** q^* from evidence lower bound:

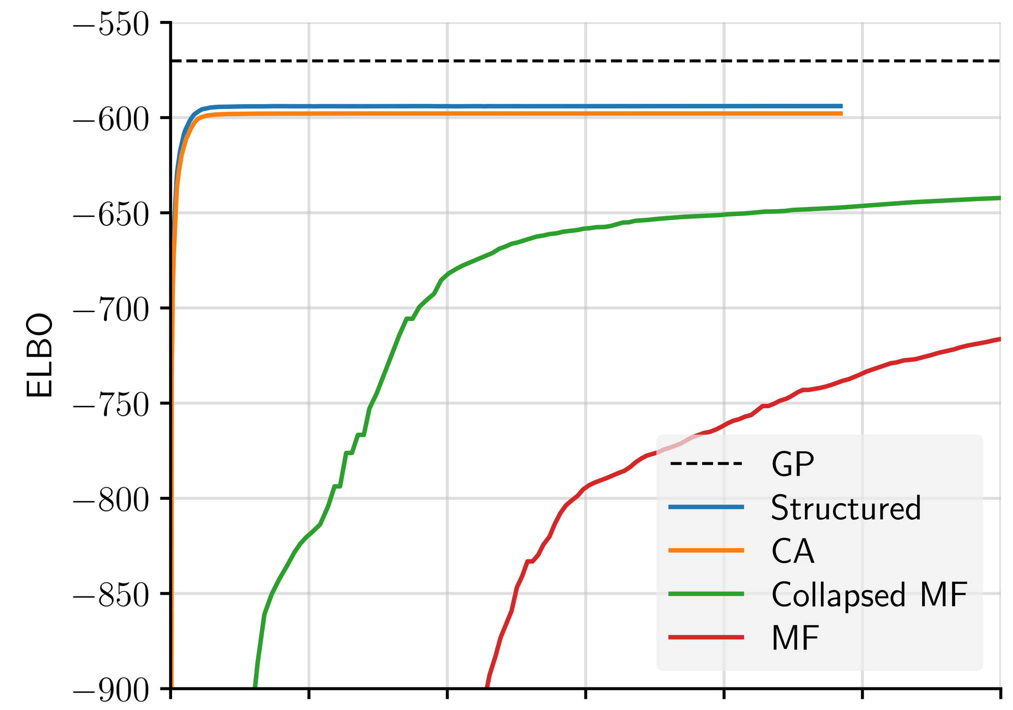
$$q^*(\mathbf{u}, \mathbf{z}) = \operatorname{argmax}_q \mathcal{F}_\theta(q) \le \log p_\theta(\mathbf{y}).$$

Insight: $q^*(\mathbf{z}\,|\,\mathbf{u})$ and $q^*(\mathbf{u}\,|\,\mathbf{z})$ are Gaussian!

Gibbs sampler: initial $\mathbf{u}^{(0)} \sim p(\mathbf{u})$ and iterate

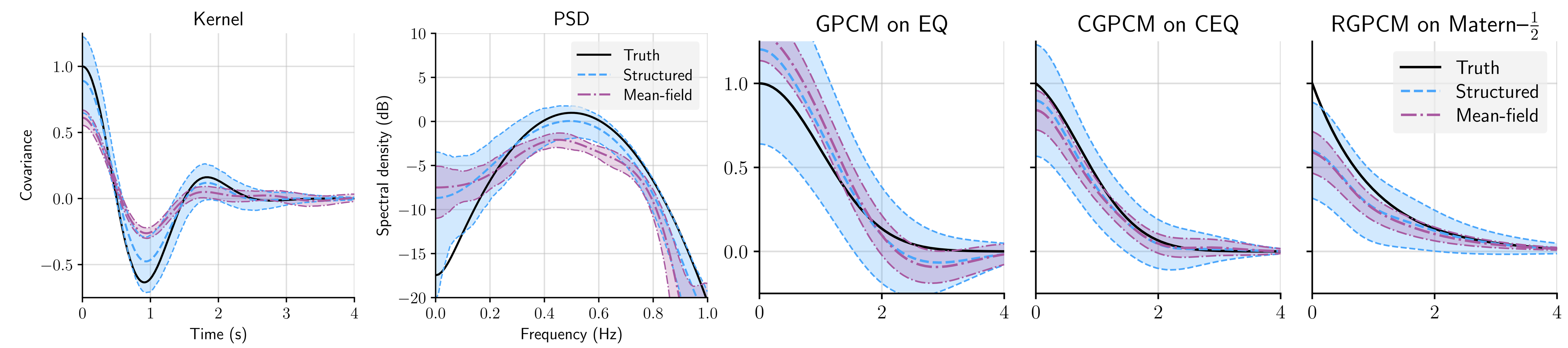
$$\mathbf{z}^{(i)} \sim q^*(\mathbf{z}\,|\,\mathbf{u}^{(i-1)}),$$
$$\mathbf{u}^{(i)} \sim q^*(\mathbf{u}\,|\,\mathbf{z}^{(i)}).$$

Optimise θ with SGD on $\mathcal{F}_\theta(q^*(\mathbf{u}, \mathbf{z}))$.



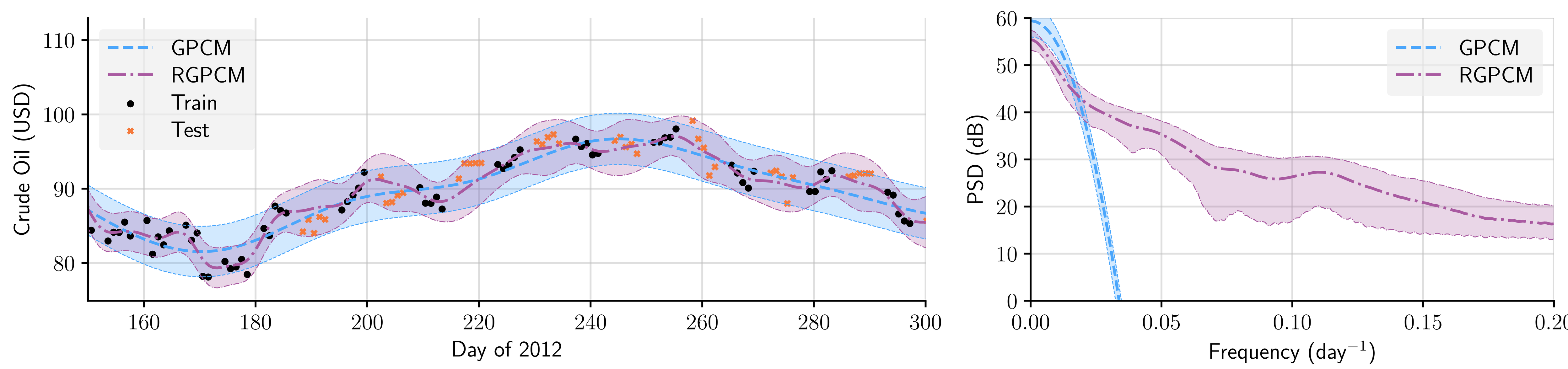
Reconstruction of known model: $\mathcal{F}$ during optimisation of GPCMs. Mean-field (MF), collapsed MF, coordinate-ascent (CA) and our Gibbs sampler (structured); (GP) is truth.

Approximation quality of inference schemes



⟹ Structured approximation produces significantly improved uncertainty estimates.

Prediction of crude oil prices



⟹ GPCM oversmooths, whereas RGPCM captures signal and predicts expressive spectral content.

Links and references

Code:
`github.com/wesselb/gpcm`

[1] Felipe Tobar, Thang D. Bui, and Richard E. Turner. Learning stationary time series using Gaussian processes with nonparametric kernels. *Advances in Neural Information Processing Systems*, 29:3501–3509, 2015.

[2] James Hensman, Nicolas Durrande, and Arno Solin. Variational fourier features for Gaussian processes. *Journal of Machine Learning Research*, 18(151):1–52, 2018.

*Currently with the University of Copenhagen and Oxford-Man Institute, University of Oxford.

25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022