"Spectral Mixture Gaussian Processes are Probabilistic Filter Banks":

Unifying Probabilistic Models for Time-Frequency Analysis

William J. Wilkinson[1] **Michael Riis Andersen**[2] **Joshua D. Reiss**[1] **Dan Stowell**[1] **Arno Solin**[2]

w.j.wilkinson@qmul.ac.uk

[1]Centre for Digital Music, Queen Mary University of London

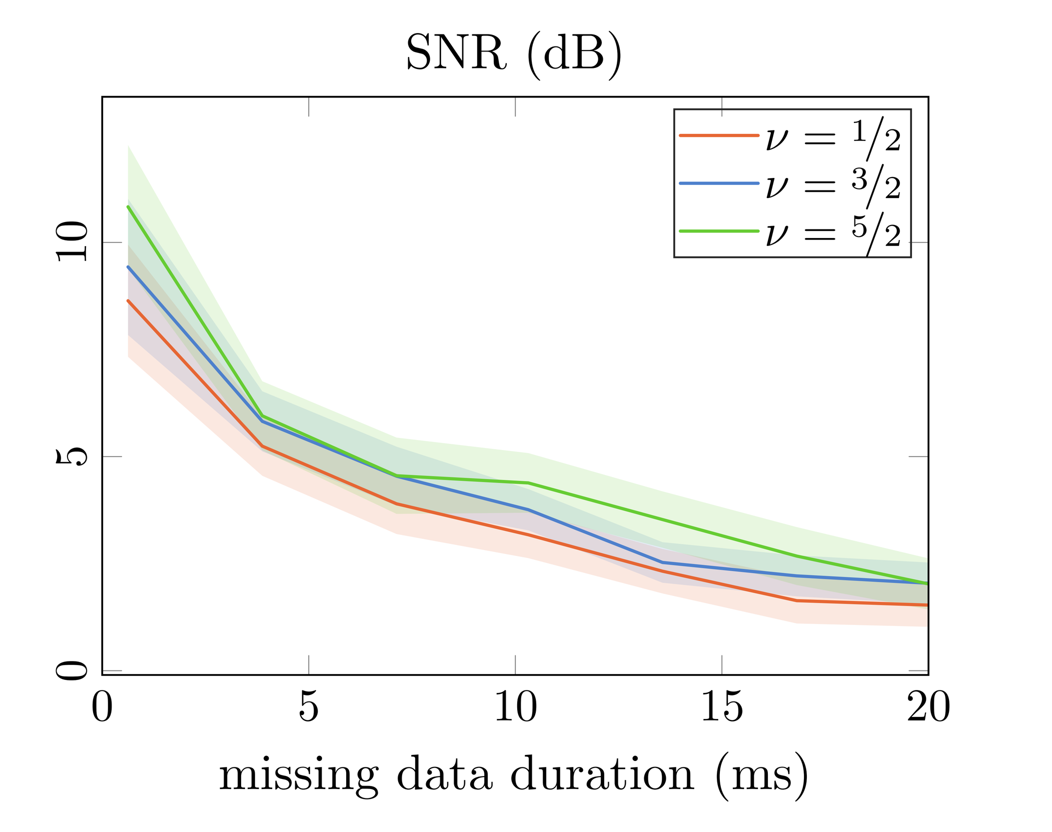
[2]Department of Computer Science, Aalto University

Overview

- **Gaussian processes** (GPs) are a probabilistic machine learning approach that allow us to learn distributions over *functions*. Useful tools for regression, interpolation, extrapolation and noise removal.
- A **spectral mixture GP** (1) models the covariance as a sum of quasi-periodic components [1].
- The **probabilistic phase vocoder** (PPV) (2) is a signal processing method that allows us to fit a filter bank to a signal by leveraging uncertainty [2].
- GPs have equivalent representations as **stochastic differential equations** (SDEs) [3].
- By formulating its SDE representation, we show that the **Matérn spectral mixture GP (1) is exactly equivalent to the PPV (2)**.
- We leverage the best of both worlds for inference on audio signals: fast frequency-domain optimisation, interpretability, easy to switch out kernels, guaranteed stationarity.

Results

- This generative model can handle missing data synthesis, denoising, source separation.
- Swapping the kernel for a higher-order Matérn-ν allows instantaneous frequency to be correlated through time & improves missing data synthesis:



Spectral Mixture GPs

$$f(t) \sim \mathrm{GP}\left(0, \sum_{d=1}^{D} C^{(d)}_{\text{q-per}}(t,t')\right) \qquad (1)$$
$$y_k = f(t_k) + \sigma_{y_k}\varepsilon_k$$

a GP model whose kernel is a sum of quasi-periodic covariance functions:

$$C^{(d)}_{\text{q-per}}(t,t') = \sigma_d^2 \cos(\omega_d(t-t'))\exp(|t-t'|/\ell_d)$$

ℓ_d = lengthscale,
ω_d = frequency,
σ_d^2 = variance.

- ✘ Inference is slow for long time series
- Frequency domain parameter learning possible
- ✓ All model assumptions encoded in the kernel
- ✓ Changing the model is easy

=

Probabilistic Phase Vocoder

$$x_{d,k} = \psi_d e^{i\omega_d} x_{d,k-1} + \rho_d \varepsilon_{d,k}$$
$$y_k = \sum_{d=1}^{D} \mathrm{Re}[x_{d,k}] + \sigma_{y_k}\varepsilon_k \qquad (2)$$

a complex first-order autoregressive process. $x_{d,k}$ is a complex phasor.

ψ_d = process variance,
ω_d = frequency,
ρ_d = noise variance.

- ✓ Fast inference via Kalman filtering
- ✓ Fast frequency domain parameter learning
- Interpreting the model is challenging
- ✘ Changing the model is hard

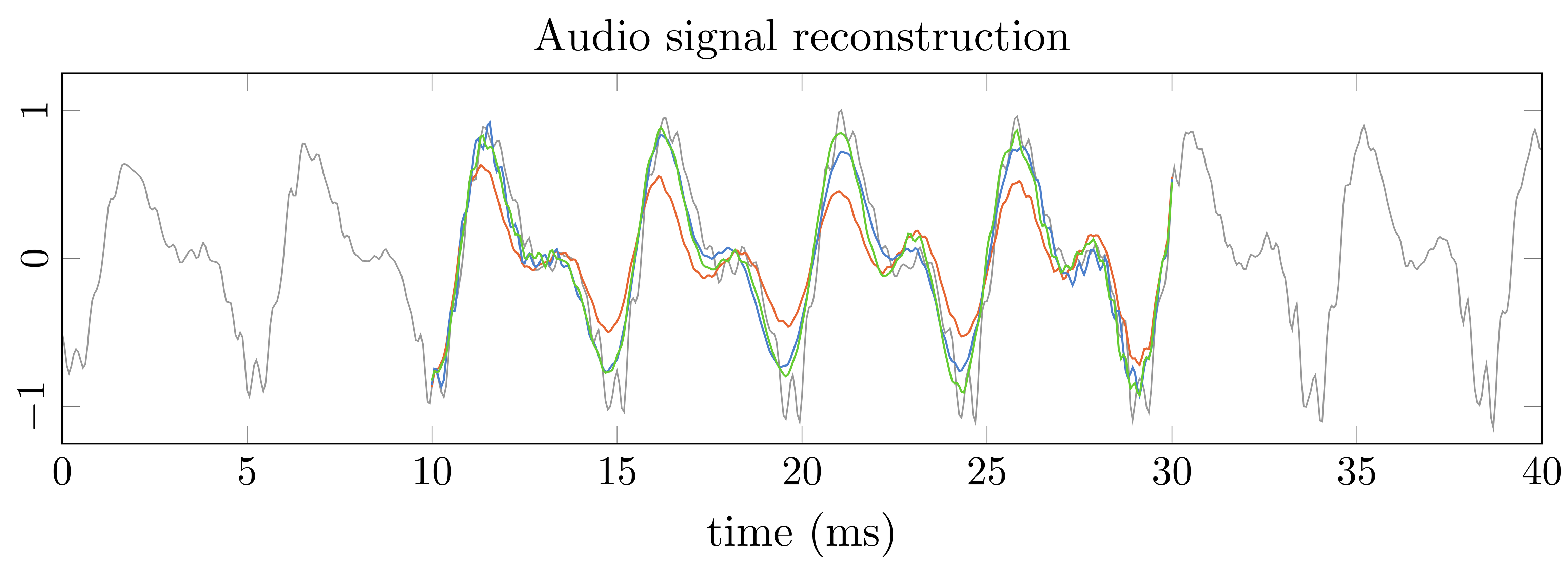


Fig. 1: Missing data synthesis example with the probabilistic filter bank using 3 different Matérn-ν kernels: the Matérn-1/2 (PPV), the Matérn-3/2 and the Matérn-5/2.

Code and resources available:
https://github.com/wil-j-wil/unifying-prob-time-freq

References

[1] A. Wilson and R. Adams (2013). Gaussian process kernels for pattern discovery and extrapolation. *Proceedings of ICML*.
[2] R. E. Turner and M. Sahani (2014). Time-frequency analysis as probabilistic inference. *IEEE trans. on Signal Processing*.
[3] A. Solin (2016). Stochastic differential equation methods for spatio-temporal Gaussian process regression. *Doctoral dissertation*.