

# Spectral Methods in Gaussian Modelling

## Topic 4: Spectrum Estimation

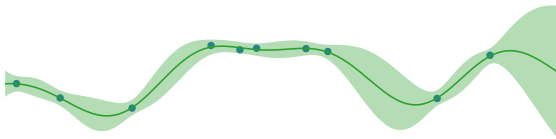
James Requiema and Wessel Bruinsma

University of Cambridge and Invenia Labs

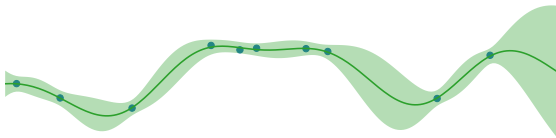
9 June 2019

- PSD is of wide interest.

- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.

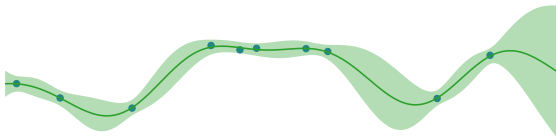


- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



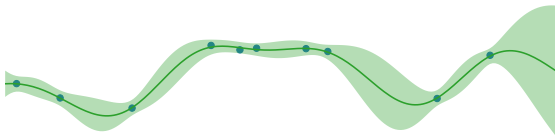
- Estimators:

- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



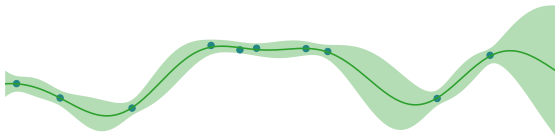
- Estimators:
  - parametric methods

- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



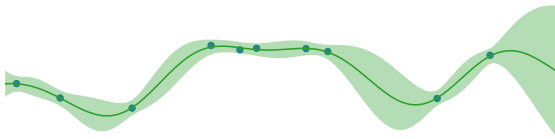
- Estimators:
  - parametric methods (SSA, SMK)

- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



- Estimators:
  - parametric methods (SSA, SMK),
  - nonparametric methods

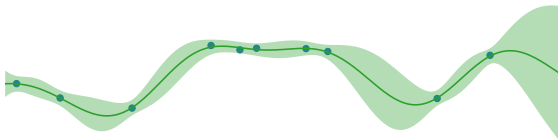
- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



- Estimators:
  - parametric methods (SSA, SMK),
  - nonparametric methods (GPCM).



- PSD is of wide interest.
- Estimation difficult: samples limited, noisy, and nonuniform.



- Estimators:
  - parametric methods (SSA, SMK),
  - nonparametric methods (GPCM).
- Novel model by Tobar (2018).

- Generative model:

$$f \sim \mathcal{GP}(0, k(t - t')), \quad s(\xi) \mid f = |\hat{f}(\xi)|^2.$$

- Generative model:

$$f \sim \mathcal{GP}(0, k(t - t')), \quad s(\xi) \mid f = |\hat{f}(\xi)|^2.$$

- Estimator:  $\hat{s}(\xi) = \mathbb{E}[s(\xi) \mid e]$ .

- Generative model:

$$f \sim \mathcal{GP}(0, k(t - t')), \quad s(\xi) | f = |\hat{f}(\xi)|^2.$$

- Estimator:  $\hat{s}(\xi) = \mathbb{E}[s(\xi) | e]$ .
- Does not work: samples from  $f$  not integrable!

- Generative model:

$$f \sim \mathcal{GP}(0, k(t - t')), \quad s(\xi) | f = |\hat{f}(\xi)|^2.$$

- Estimator:  $\hat{s}(\xi) = \mathbb{E}[s(\xi) | e]$ .
- Does not work: samples from  $f$  not integrable!
- Windowed version of  $f$ :

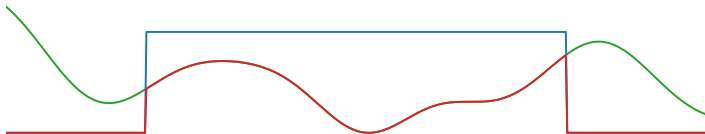
$$f_w(t) | f = f(t)w(t).$$

- Generative model:

$$f \sim \mathcal{GP}(0, k(t - t')), \quad s(\xi) | f = |\hat{f}(\xi)|^2.$$

- Estimator:  $\hat{s}(\xi) = \mathbb{E}[s(\xi) | e]$ .
- Does not work: samples from  $f$  not integrable!
- Windowed version of  $f$ :

$$f_w(t) | f = f(t)w(t).$$

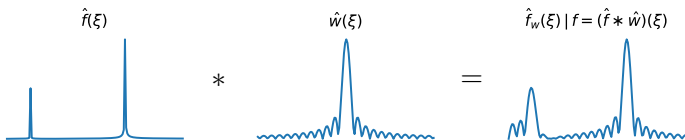


- Effect of window:

$$\hat{f}_w(\xi) | f = (\hat{f} * \hat{w})(\xi).$$

- Effect of window:

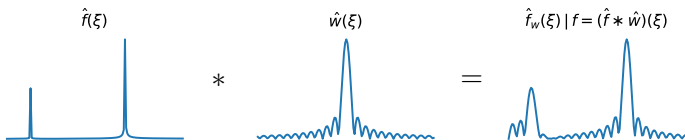
$$\hat{f}_w(\xi) | f = (\hat{f} * \hat{w})(\xi).$$





- Effect of window:

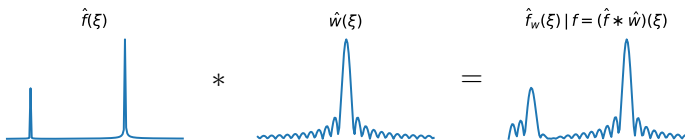
$$\hat{f}_w(\xi) | f = (\hat{f} * \hat{w})(\xi).$$



- Window functions:
  - algorithmic purposes (here: ensure integrability),

- Effect of window:

$$\hat{f}_w(\xi) | f = (\hat{f} * \hat{w})(\xi).$$



- Window functions:
  - algorithmic purposes (here: ensure integrability),
  - acquisition devices (e.g., sampling:  $w = \text{III}$ ).

- Generative model:

$$\begin{aligned}f &\sim \mathcal{GP}(0, k(t - t')), \\f_w(t) \mid f &= f(t)w(t), \\s_w(\xi) \mid f_w &= |\hat{f}_w(\xi)|^2.\end{aligned}$$

- Generative model:

$$\begin{aligned}f &\sim \mathcal{GP}(0, k(t - t')), \\f_w(t) \mid f &= f(t)w(t), \\s_w(\xi) \mid f_w &= |\hat{f}_w(\xi)|^2.\end{aligned}$$

- $\hat{f}_w$  called local spectrum.

- Generative model:

$$\begin{aligned}f &\sim \mathcal{GP}(0, k(t - t')), \\f_w(t) \mid f &= f(t)w(t), \\s_w(\xi) \mid f_w &= |\hat{f}_w(\xi)|^2.\end{aligned}$$

- $\hat{f}_w$  called **local spectrum**.
- Estimator:

$$\hat{s}_w(\xi) = \mathbb{E}[s_w(\xi) \mid e] = \mathbb{E}[(\operatorname{Re} \hat{f}_w(\xi))^2 \mid e] + \mathbb{E}[(\operatorname{Im} \hat{f}_w(\xi))^2 \mid e].$$

- Generative model:

$$\begin{aligned}f &\sim \mathcal{GP}(0, k(t - t')), \\f_w(t) \mid f &= f(t)w(t), \\s_w(\xi) \mid f_w &= |\hat{f}_w(\xi)|^2.\end{aligned}$$

- $\hat{f}_w$  called **local spectrum**.
- Estimator:

$$\hat{s}_w(\xi) = \mathbb{E}[s_w(\xi) \mid e] = \mathbb{E}[(\operatorname{Re} \hat{f}_w(\xi))^2 \mid e] + \mathbb{E}[(\operatorname{Im} \hat{f}_w(\xi))^2 \mid e].$$

- Choices:
  - $w(t) = \exp(-\alpha\pi^2 t^2)$ , (tractability)

- Generative model:

$$\begin{aligned}f &\sim \mathcal{GP}(0, k(t - t')), \\f_w(t) \mid f &= f(t)w(t), \\s_w(\xi) \mid f_w &= |\hat{f}_w(\xi)|^2.\end{aligned}$$

- $\hat{f}_w$  called **local spectrum**.
- Estimator:

$$\hat{s}_w(\xi) = \mathbb{E}[s_w(\xi) \mid e] = \mathbb{E}[(\operatorname{Re} \hat{f}_w(\xi))^2 \mid e] + \mathbb{E}[(\operatorname{Im} \hat{f}_w(\xi))^2 \mid e].$$

- Choices:
  - $w(t) = \exp(-\alpha\pi^2 t^2)$ , (tractability)
  - $k = \text{SMK or EQ in simple cases.}$

- Notation:

- $k_{fg}(t, t') = \mathbb{E}[f^*(t)g(t')]$  with  $k_{ff} = k_f$ ,



- Notation:

- $k_{fg}(t, t') = \mathbb{E}[f^*(t)g(t')] \text{ with } k_{ff} = k_f,$

- $\mathcal{F}_t\{f(t)\}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i \xi t} dt.$

- Notation:

- $k_{fg}(t, t') = \mathbb{E}[f^*(t)g(t')] \text{ with } k_{ff} = k_f,$

- $\mathcal{F}_t\{f(t)\}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i \xi t} dt.$

- We desire

$$k_{\text{Re } \hat{f}_w}(\xi, \xi')$$

$$k_{y(\text{Re } \hat{f}_w)}(t, \xi)$$

$$k_{\text{Im } \hat{f}_w}(\xi, \xi')$$

$$k_{y(\text{Im } \hat{f}_w)}(t, \xi)$$

- Notation:

- $k_{fg}(t, t') = \mathbb{E}[f^*(t)g(t')] \text{ with } k_{ff} = k_f,$

- $\mathcal{F}_t\{f(t)\}(\xi) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i \xi t} dt.$

- We desire

$$k_{\text{Re } \hat{f}_w}(\xi, \xi') = \frac{1}{2}(k_{\hat{f}_w}(\xi, \xi') + k_{\hat{f}_w}(\xi, -\xi')),$$

$$k_{y(\text{Re } \hat{f}_w)}(t, \xi) = \text{Re } k_{y\hat{f}_w}(t, \xi),$$

$$k_{\text{Im } \hat{f}_w}(\xi, \xi') = \frac{1}{2}(k_{\hat{f}_w}(\xi, \xi') - k_{\hat{f}_w}(\xi, -\xi')),$$

$$k_{y(\text{Im } \hat{f}_w)}(t, \xi) = \text{Im } k_{y\hat{f}_w}(t, \xi).$$

## Calculation of Posterior Moments (2)

7/13

$$k_{\hat{f}_w}(\xi, \xi')$$

## Calculation of Posterior Moments (2)

7/13

$$k_{\hat{f}_w}(\xi, \xi') = \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')]$$

## Calculation of Posterior Moments (2)

7/13

$$\begin{aligned}k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi)\hat{f}_w(\xi')] \\ &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')]\end{aligned}$$

$$\begin{aligned}k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi)\hat{f}_w(\xi')] \\&= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\&= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'),\end{aligned}$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')] \\
 &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\
 &= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'), \\
 &= (\hat{k}(u)\delta(u+u') * \hat{r}_w(u, u'))(-\xi, \xi').
 \end{aligned}$$



$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')] \\
 &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\
 &= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'), \\
 &= (\hat{k}(u)\delta(u+u') * \hat{r}_w(u, u'))(-\xi, \xi').
 \end{aligned}$$

$$\hat{r}_w(\xi + u, \xi' + u) = \mathcal{N}(\xi - \xi'; 0, \alpha) \mathcal{N}(u; \frac{1}{2}(\xi + \xi'), \frac{1}{4}\alpha).$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')] \\
 &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\
 &= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'), \\
 &= (\hat{k}(u)\delta(u+u') * \hat{r}_w(u, u'))(-\xi, \xi').
 \end{aligned}$$

$$\hat{r}_w(\xi + u, \xi' + u) = \mathcal{N}(\xi - \xi'; 0, \alpha) \mathcal{N}(u; \frac{1}{2}(\xi + \xi'), \frac{1}{4}\alpha).$$

$$k_{\hat{f}_w}(\xi, \xi') = \int \hat{k}(u)\delta(u+u')\hat{r}_w(-\xi-u, \xi'-u') \, du \, du'$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')] \\
 &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\
 &= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'), \\
 &= (\hat{k}(u)\delta(u+u') * \hat{r}_w(u, u'))(-\xi, \xi').
 \end{aligned}$$

$$\hat{r}_w(\xi + u, \xi' + u) = \mathcal{N}(\xi - \xi'; 0, \alpha) \mathcal{N}(u; \frac{1}{2}(\xi + \xi'), \frac{1}{4}\alpha).$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \int \hat{k}(u)\delta(u+u')\hat{r}_w(-\xi-u, \xi'-u') \, du \, du' \\
 &= \int \hat{k}(u)\hat{r}_w(\xi+u, \xi'+u) \, du
 \end{aligned}$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \mathbb{E}[\hat{f}_w^*(\xi) \hat{f}_w(\xi')] \\
 &= \mathbb{E}[\mathcal{F}_{t,t'}\{f(t)f(t')w(t)w(t')\}(-\xi, \xi')] \\
 &= (\mathcal{F}_{t,t'}\{k(t-t')\}(u, u') * \mathcal{F}_{t,t'}\{w(t)w(t')\}(u, u'))(-\xi, \xi'), \\
 &= (\hat{k}(u)\delta(u+u') * \hat{r}_w(u, u'))(-\xi, \xi').
 \end{aligned}$$

$$\hat{r}_w(\xi + u, \xi' + u) = \mathcal{N}(\xi - \xi'; 0, \alpha) \mathcal{N}(u; \frac{1}{2}(\xi + \xi'), \frac{1}{4}\alpha).$$

$$\begin{aligned}
 k_{\hat{f}_w}(\xi, \xi') &= \int \hat{k}(u)\delta(u+u')\hat{r}_w(-\xi-u, \xi'-u') \, du \, du' \\
 &= \int \hat{k}(u)\hat{r}_w(\xi+u, \xi'+u) \, du \\
 &= \mathcal{N}(\xi - \xi'; 0, \alpha) \left( \hat{k}(u) * \mathcal{N}(u; 0, \frac{1}{4}\alpha) \right) \left( \frac{1}{2}(\xi + \xi') \right).
 \end{aligned}$$

$$k_{y\hat{f}_w}(t, \xi)$$

$$k_{y\hat{f}_w}(t, \xi) = \mathbb{E}[y^*(t)\mathcal{F}_{t'}\{f(t')w(t')\}(\xi)]$$

$$\begin{aligned}k_{y\hat{f}_w}(t, \xi) &= \mathbb{E}[y^*(t)\mathcal{F}_{t'}\{f(t')w(t')\}(\xi)] \\ &= \mathcal{F}_{t'}\{k(t-t')w(t')\}(\xi)\end{aligned}$$

$$\begin{aligned}k_{y\hat{f}_w}(t, \xi) &= \mathbb{E}[y^*(t)\mathcal{F}_{t'}\{f(t')w(t')\}(\xi)] \\&= \mathcal{F}_{t'}\{k(t-t')w(t')\}(\xi) \\&= (\hat{k}(u)e^{-2\pi\iota ut} * \mathcal{F}\{w\}(u))(\xi)\end{aligned}$$



$$\begin{aligned}
 k_{y\hat{f}_w}(t, \xi) &= \mathbb{E}[y^*(t)\mathcal{F}_{t'}\{f(t')w(t')\}(\xi)] \\
 &= \mathcal{F}_{t'}\{k(t-t')w(t')\}(\xi) \\
 &= (\hat{k}(u)e^{-2\pi i u t} * \mathcal{F}\{w\}(u))(\xi) \\
 &= \left(\hat{k}(u)e^{-2\pi i u t} * \mathcal{N}(u; 0, \tfrac{1}{2}\alpha)\right)(\xi).
 \end{aligned}$$

$$k_{\hat{f}_w}(\xi, \xi') = \mathcal{N}(\xi - \xi'; 0, \alpha) \left( \hat{k}(u) * \mathcal{N}(u; 0, \frac{1}{4}\alpha) \right) \left( \frac{1}{2}(\xi + \xi') \right)$$

$$k_{y\hat{f}_w}(t, \xi) = \left( \hat{k}(u) e^{-2\pi i u t} * \mathcal{N}(u; 0, \frac{1}{2}\alpha) \right) (\xi)$$

$$\begin{aligned} k_{\hat{f}_w}(\xi, \xi') &= \mathcal{N}(\xi - \xi'; 0, \alpha) \left( \hat{k}(u) * \mathcal{N}(u; 0, \tfrac{1}{4}\alpha) \right) \left( \tfrac{1}{2}(\xi + \xi') \right) \\ &\approx \mathcal{N}(\xi - \xi'; 0, \alpha) \hat{k}\left(\tfrac{1}{2}(\xi + \xi')\right), \end{aligned}$$

$$k_{y\hat{f}_w}(t, \xi) = \left( \hat{k}(u) e^{-2\pi i u t} * \mathcal{N}(u; 0, \tfrac{1}{2}\alpha) \right) (\xi)$$

$$\begin{aligned} k_{\hat{f}_w}(\xi, \xi') &= \mathcal{N}(\xi - \xi'; 0, \alpha) \left( \hat{k}(u) * \mathcal{N}(u; 0, \tfrac{1}{4}\alpha) \right) \left( \tfrac{1}{2}(\xi + \xi') \right) \\ &\approx \mathcal{N}(\xi - \xi'; 0, \alpha) \hat{k}\left(\tfrac{1}{2}(\xi + \xi')\right), \end{aligned}$$

$$\begin{aligned} k_{y\hat{f}_w}(t, \xi) &= \left( \hat{k}(u) e^{-2\pi i u t} * \mathcal{N}(u; 0, \tfrac{1}{2}\alpha) \right) (\xi) \\ &\approx \hat{k}(\xi) e^{-2\pi i \xi t}. \end{aligned}$$

$$\mathbb{E}[\hat{f}_w(\xi) | e] = \int \hat{k}(u) \left( \sum_{i=1}^N e^{-2\pi i u t_i} (K_e^{-1} e)_i \right) \mathcal{N}(u; \xi, \tfrac{1}{2}\alpha) \, du.$$

$$\mathbb{E}[\hat{f}_w(\xi) | e] = \int \hat{k}(u) \left( \sum_{i=1}^N e^{-2\pi i u t_i} (K_e^{-1} e)_i \right) \mathcal{N}(u; \xi, \frac{1}{2}\alpha) \, du.$$

- Interpretation: **DFT** of **whitened** observations, **weighted** by prior, then **smoothed** due to window.

$$\mathbb{E}[\hat{f}_w(\xi) | e] = \int \hat{k}(u) \left( \sum_{i=1}^N e^{-2\pi i u t_i} (K_e^{-1} e)_i \right) \mathcal{N}(u; \xi, \frac{1}{2}\alpha) \, du.$$

- Interpretation: **DFT** of **whitened** observations, **weighted** by prior, then **smoothed** due to window.
- If prior uninformative,  $K_e \approx I$ , then weighted DFT in the limit:

$$\lim_{\alpha \rightarrow 0} \mathbb{E}[\hat{f}_w(\xi) | e] \approx \hat{k}(\xi) \sum_{i=1}^N e^{-2\pi i \xi t_i} e_i.$$

## Comparison with Lomb–Scargle

11/13

- Commonly used for nonuniform data.



- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Then estimates power at  $\xi$  from  $f(t)$ .

- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Then estimates power at  $\xi$  from  $f(t)$ .

BNSE:

+ Probabilistic

Lomb–Scargle:

— Deterministic

- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Then estimates power at  $\xi$  from  $f(t)$ .

BNSE:

- + Probabilistic
- + Nonparametric

Lomb–Scargle:

- Deterministic
- Parametric

- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Then estimates power at  $\xi$  from  $f(t)$ .

BNSE:

- + Probabilistic
- + Nonparametric
- + Closed-form estimate

Lomb–Scargle:

- Deterministic
- Parametric
- Optimisation per  $\xi$

- Commonly used for nonuniform data.
- For each  $\xi$ , LS fits a sine using least squares:

$$f(t) = A \sin(2\pi\xi t) + B \cos(2\pi\xi t).$$

- Then estimates power at  $\xi$  from  $f(t)$ .

BNSE:

- + Probabilistic
- + Nonparametric
- + Closed-form estimate

Lomb–Scargle:

- Deterministic
- Parametric
- Optimisation per  $\xi$

- Gaussian prior on  $A$  and  $B$ : LS recovers BNSE in the limit.

## Experiment: Two Heart Rate Signals

12/13

- $k = \text{SMK}$ .

## Experiment: Two Heart Rate Signals

12/13

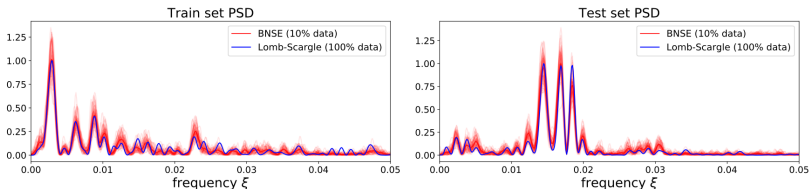
- $k = \text{SMK}$ .
- Trained on first, tested on second.



# Experiment: Two Heart Rate Signals

12/13

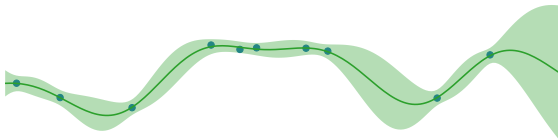
- $k = \text{SMK}$ .
- Trained on first, tested on second.



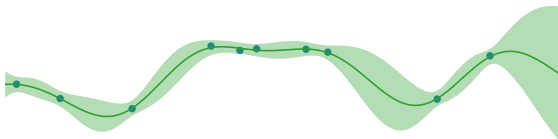
(Figure taken from Tobar (2018).)

- BNSE (Tobar, 2018) novel nonparametric model for SE.

- BNSE (Tobar, 2018) novel nonparametric model for SE.
- Bayesian in nature: handles uncertainty automatically.

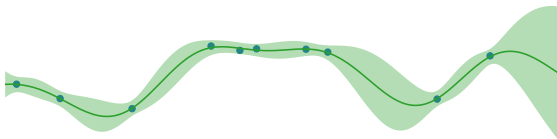


- BNSE (Tobar, 2018) novel nonparametric model for SE.
- Bayesian in nature: handles uncertainty automatically.



- Closed-form estimate of PSD.

- BNSE (Tobar, 2018) novel nonparametric model for SE.
- Bayesian in nature: handles uncertainty automatically.



- Closed-form estimate of PSD.
- ⇒ Can optimise to find periodicities.

# Appendix

## References

Tobar, F. (2018). Bayesian nonparametric spectral estimation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in neural information processing systems 31*, Curran Associates, Inc. eprint: <https://arxiv.org/abs/1809.02196>