DRIVER ESTIMATION IN NON-LINEAR AUTOREGRESSIVE MODELS

Tom Dupré la Tour[1,2], Yves Grenier[1,2], Alexandre Gramfort[1,2,3]

1. Télécom ParisTech, 2. Université Paris-Saclay, 3. Inria Saclay

tom.duprelatour@telecom-paristech.fr

Abstract

In non-linear autoregressive models, the time dependency of coefficients is often driven by a particular time-series which is not given and thus has to be estimated from the data.

To allow model evaluation on a validation set, we describe a parametric approach for such driver estimation. After estimating the driver as a weighted sum of potential drivers, we use it in a non-linear autoregressive model with a polynomial parameterization.

Using gradient descent, we optimize the linear filter extracting the driver, outperforming a typical grid-search on predefined filters.

We apply this method on electrophysiological signals to better describe phase-amplitude couplings.

1. Driven autoregressive (DAR) models

Linear AR model

$$y(t) + \sum_{i=1}^{p} a_i y(t-i) = \varepsilon(t)$$

Driven AR model, with a polynomial parametrization

$$a_i(t) = \sum_{j=0}^{m} a_{ij} x(t)^j, \qquad \log(\sigma(t)) = \sum_{j=0}^{m} b_j x(t)^j$$

Maximum likelihood estimate (MLE) :

- Linear system for the AR coefficients a_{ij}
- Newton-Raphson for the gain coefficients b_j

Model likelihood

$$L = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma(t)^2}} \exp\left(-\frac{\varepsilon(t)^2}{2\sigma(t)^2}\right)$$

2. Driver estimation

We parameterize the driver as a weighted sum:

$$x(t) = \sum_{n=1}^{N} \alpha_n x_n(t)$$

We derive the gradient of the log-likelihood:

$$\frac{\partial \log L}{\partial \alpha_n} = -\sum_{t\in\Theta}\left(\frac{\varepsilon(t)}{\sigma(t)^2}\frac{\partial \varepsilon(t)}{\partial \alpha_n} + \left(1 - \frac{\varepsilon(t)^2}{\sigma(t)^2}\right)\frac{\partial \log \sigma(t)}{\partial \alpha_n}\right)$$

As a special case, we use delayed versions of an exogenous signal containing the driver:

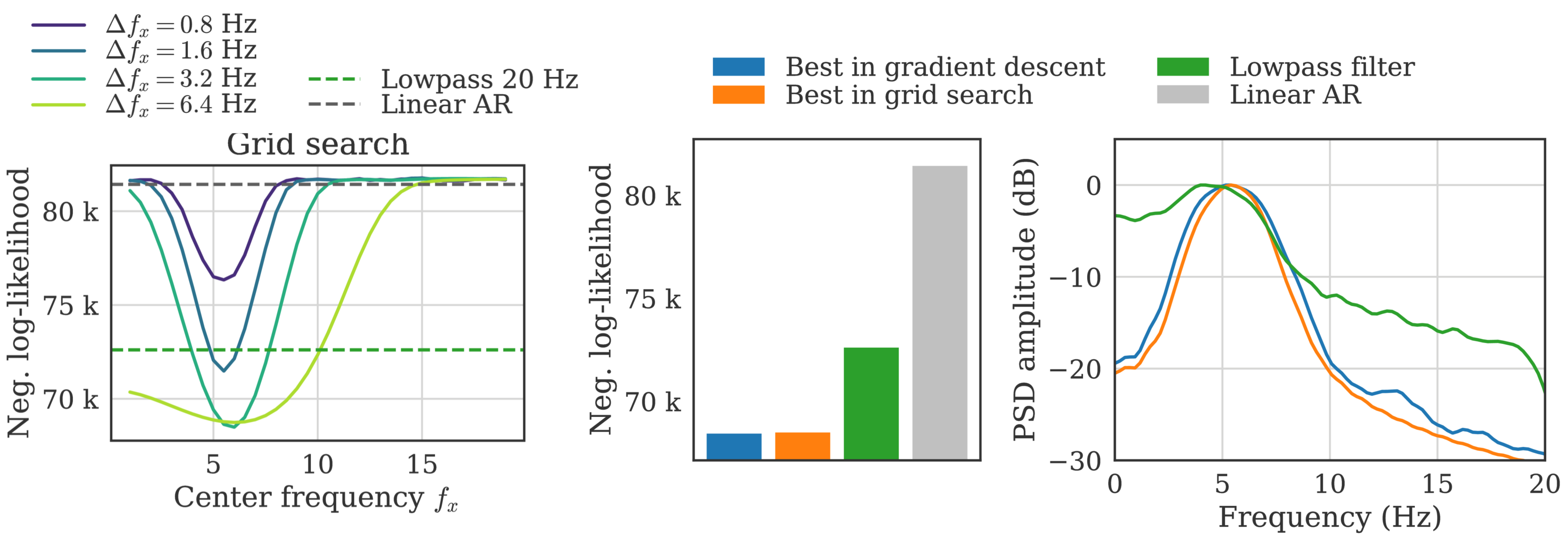
$$x_n(t) = z(t-n) \qquad -M \le n \le M$$

The weighted sum thus corresponds to a linear filter, that we learn from the data with a gradient descent.

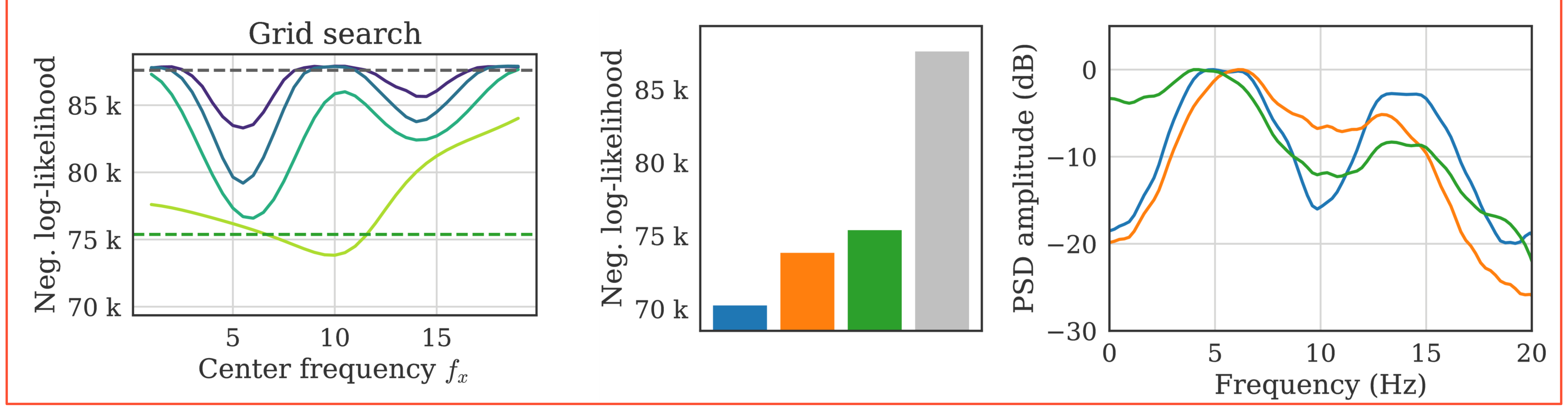
Performances can be evaluated with the model likelihood, through cross-validation.

3. Simulations

Simulations with unimodal power spectral density (PSD) drivers:
The gradient descent strategy reaches the same performances as the grid search.



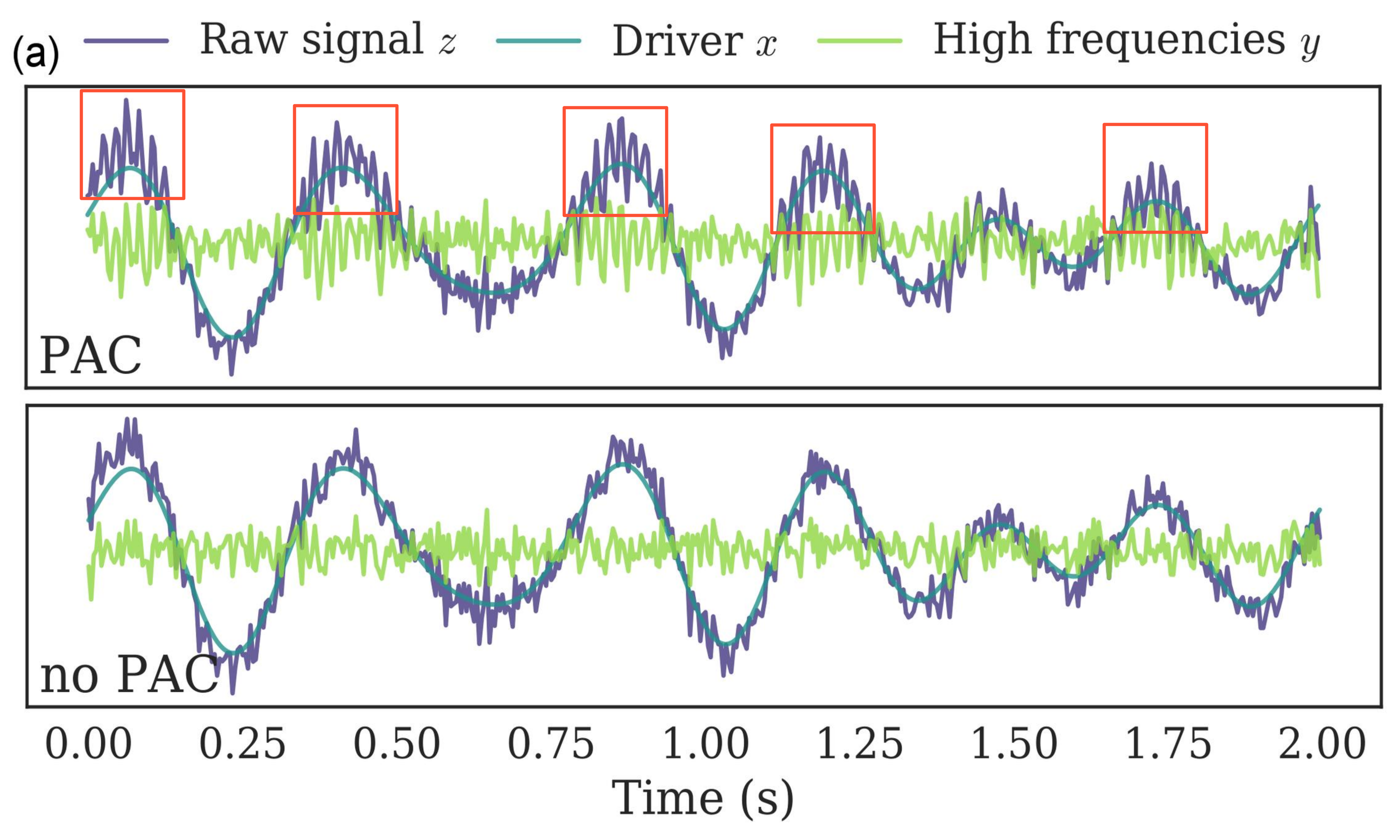
Simulations with bimodal power spectral density (PSD) drivers:
The gradient descent strategy outperforms the grid search strategy.



4. Phase-amplitude coupling

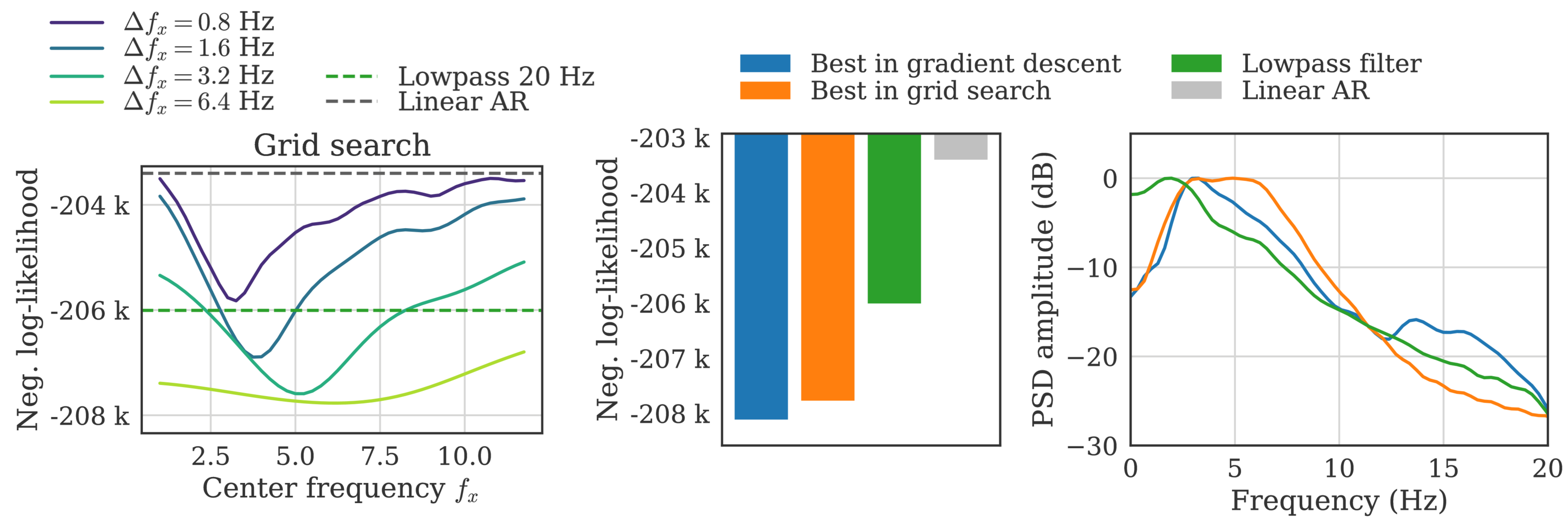
It is a coupling between:

- The phase of a slow oscillation
- The amplitude of high frequencies



5. Human auditory electrocorticogram (ECoG)

The learned filter better extracts an asymmetrical spectral shape of the driver. This asymmetry is also observed in the grid-search, since the minimum shifts to the right as the bandwidth increases.



References

Canolty, et al *"High gamma power is phase-locked to theta oscillations in human neo-cortex"* Science (2006)
Dupré la Tour, et al *"Parametric estimation of spectrum driven by an exogenous signal"* ICASSP (2017)
Dupré la Tour, et al *"Non-linear auto-regressive models for cross-frequency coupling in neural time series"* PLOS Computational Biology (2017)