



HAL
open science

Jitter-Adaptive Dictionary Learning - Application to Multi-Trial Neuroelectric Signals

Sebastian Hitziger, Maureen Clerc, Alexandre Gramfort, Sandrine Sallet, Christian G. B  nar, Th  odore Papadopoulos

► **To cite this version:**

Sebastian Hitziger, Maureen Clerc, Alexandre Gramfort, Sandrine Sallet, Christian G. B  nar, et al.. Jitter-Adaptive Dictionary Learning - Application to Multi-Trial Neuroelectric Signals. International Conference on Learning Representations 2013, May 2013, Phoenix, Arizona, United States. 2013. hal-01094663

HAL Id: hal-01094663

<https://inria.hal.science/hal-01094663v1>

Submitted on 5 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin  e au d  p  t et    la diffusion de documents scientifiques de niveau recherche, publi  s ou non,   manant des   tablissements d'enseignement et de recherche fran  ais ou   trangers, des laboratoires publics ou priv  s.

Jitter-Adaptive Dictionary Learning - Application to Multi-Trial Neuroelectric Signals

Sebastian Hitziger¹, Maureen Clerc¹, Alexandre Gramfort², Sandrine Saitlet³, Christian Bénar³, Théodore Papadopoulo¹

¹ Project-Team Athena, INRIA Sophia Antipolis, France

² Institut Mines-Telecom, Telecom ParisTech CNRS LTCI, France

³ Institut de Neurosciences des Systèmes, UMR 1106 INSERM, Aix-Marseille Université, Faculté de Médecine La Timone, Marseille, France

sebastian.hitziger@inria.fr

https://team.inria.fr/athena/



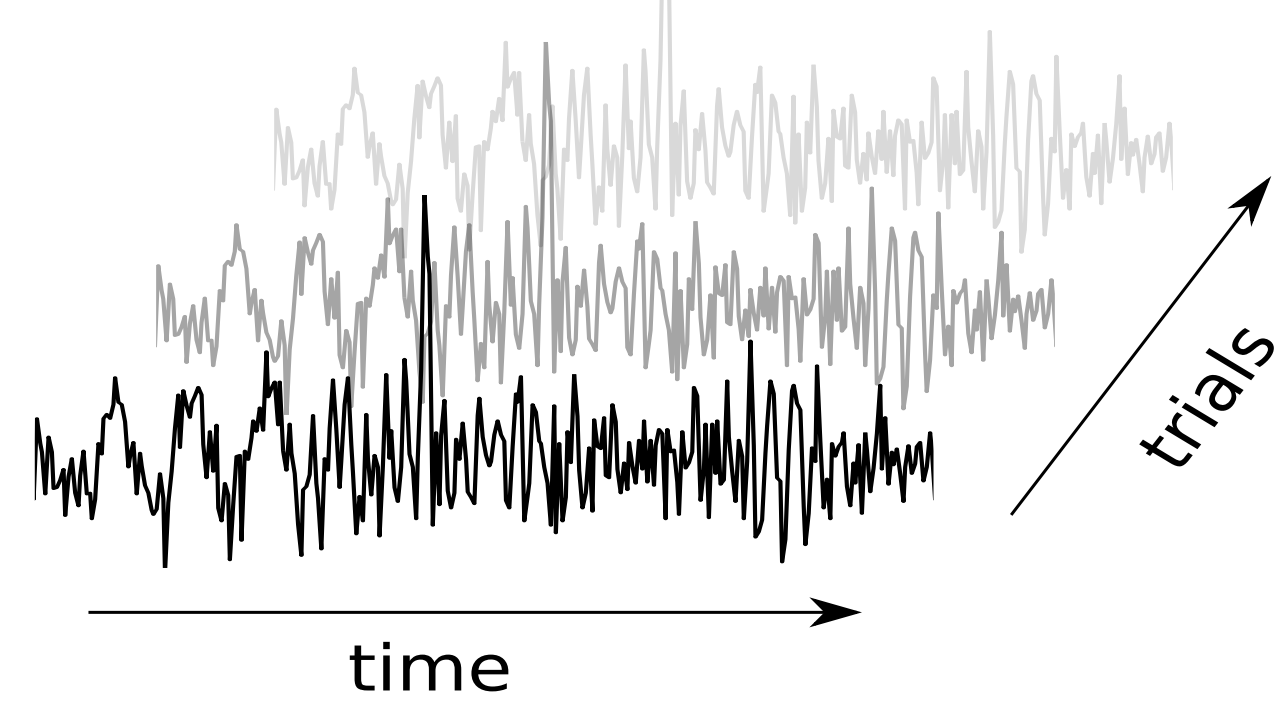
The simultaneous analysis of multiple recordings of neuronal electromagnetic activity is an important task requiring sophisticated and extremely noise robust techniques. A general goal is to find a representation of the similarities (e.g. repeating

waveforms) as well as the differences (e.g. varying shape, latency, phase, or amplitude of waveforms) across the signals. Here, we present an extension to dictionary learning that explicitly accounts for small variations in latency and phase of atoms.

Multi-trial analysis in neuroscience

Trials: recordings of neuronal electromagnetic activity under similar conditions.

Goal: detect similar waveforms and describe how they change across trials.

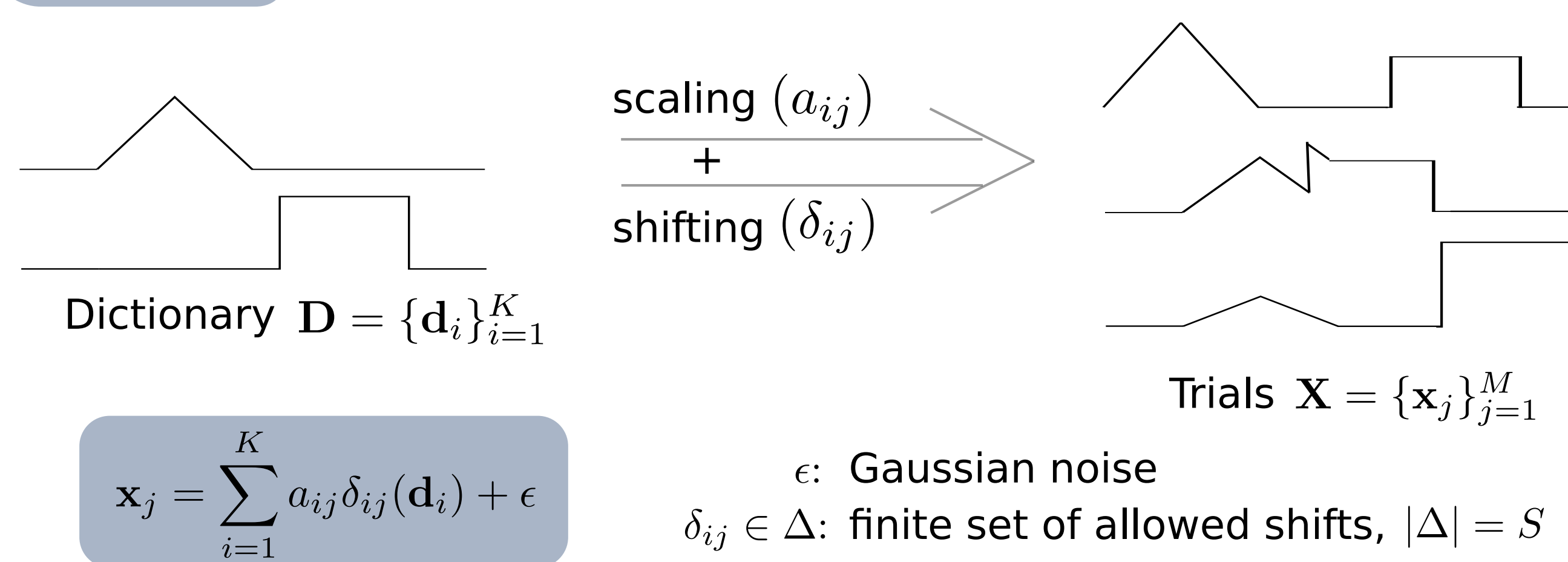


Existing approaches

- Averaging
 - Loses the information present in individual trials.
- Matrix factorization (PCA, ICA, dictionary learning [1])
 - Linear approach, does not account for temporal shifts [2].
- Variants of matching pursuit [3]
 - Do not learn waveforms but require predefined dictionary.

Model

Assumption: trials can be generated from a dictionary.



Jitter-adaptive dictionary learning (JADL)

L1-regularized optimization

$$(1) \quad \{d_i, a_{ij}, \delta_{ij}\} \leftarrow \min \sum_{j=1}^M \left(\frac{1}{2} \left\| x_j - \sum_{i=1}^K a_{ij} \delta_{ij}(d_i) \right\|_2^2 + \lambda \|a_{ij}\|_1 \right) \quad \text{s.t.} \quad \|d_i\|_2 = 1, \quad \delta_{ij} \in \Delta$$

Solve using alternating minimization

Algorithm 1

Require: trials \mathbf{X} , shifts Δ , $K \in \mathbb{N}$, $\lambda \in \mathbb{R}$.

Initialize $\mathbf{D} = \{d_i\}_{i=1}^K$ (e.g. white noise)

repeat

Sparse coding:

fix \mathbf{D} , solve (1) for $\{a_{ij}, \delta_{ij}\}$

Dictionary update:

fix $\{a_{ij}, \delta_{ij}\}$, solve (1) for $\{d_i\}$

until convergence

Dictionary update

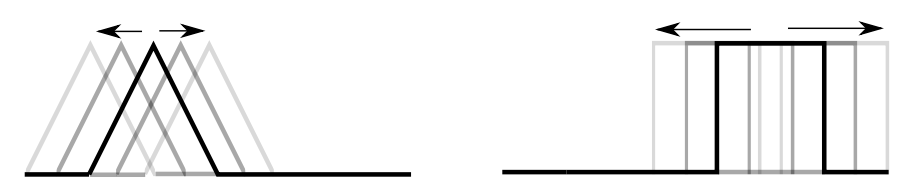
- Atoms can be updated iteratively using block coordinate descent

$$\tilde{d}_k = \sum_{j=1}^M a_{kj} \delta_{kj}^{-1} \left(x_j - \sum_{i \neq k} a_{ij} \delta_{ij}(d_i) \right)$$

where the inverse shift δ_{kj}^{-1} functions as a realignment operator. This is followed by normalization.

Sparse coding: update $\{a_{ij}, \delta_{ij}\}$

- Idea: as $S = |\Delta|$ is finite, we can first apply all possible shifts to \mathbf{D} , yielding the "unrolled" dictionary \mathbf{D}^S .



- Sparse coding can now be performed over \mathbf{D}^S , the non-zero coefficients show which shifts are used.

- A uniqueness constraint on the coefficients ensures, that at most one shifted version of each atom is used; the LARS algorithm [4] can be modified to guarantee this constraint.

Relations to other DL approaches

Dictionary learning (DL)

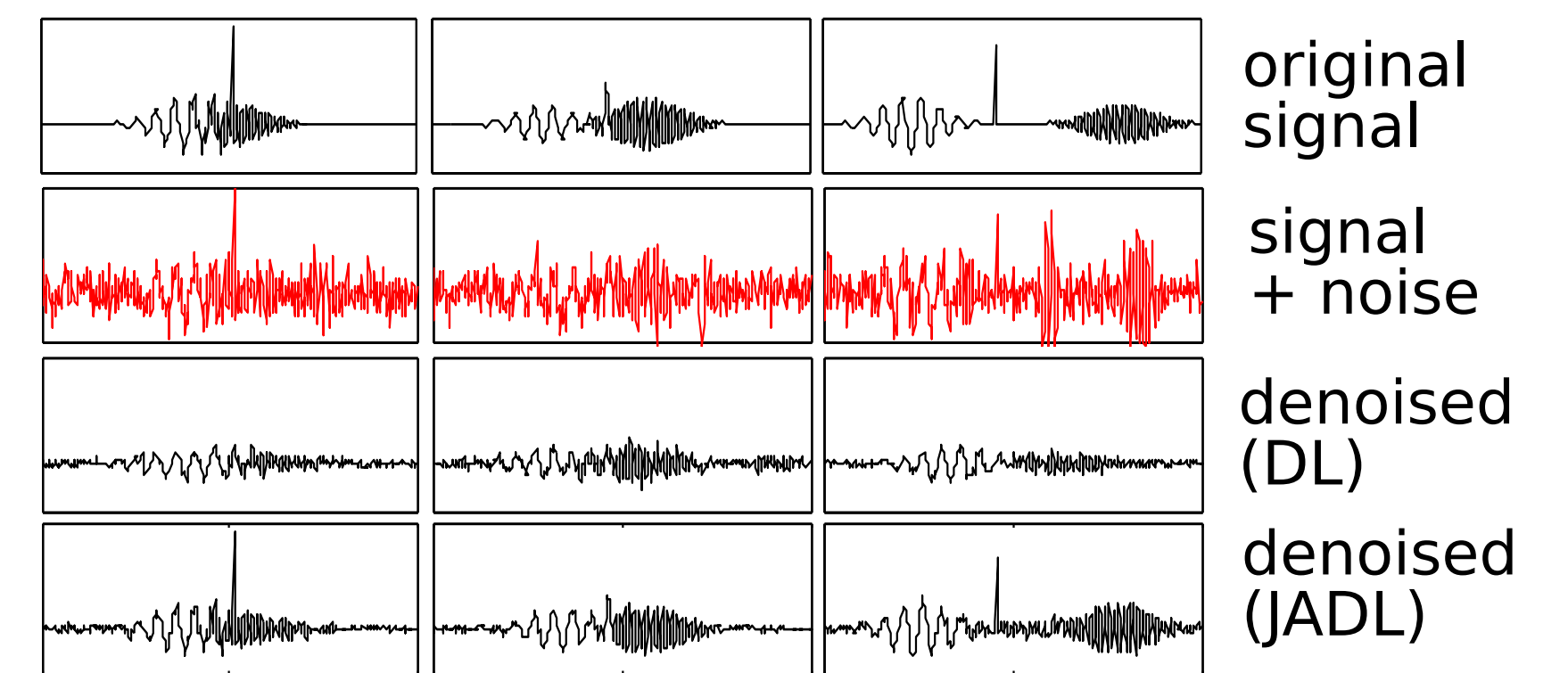
- Minimization (1) becomes DL for $\Delta = \{\mathbf{I}\}$.
- Structure of algorithms of JADL and DL are similar.

Shift-invariant sparse coding (SISC) [5]

- Atoms can shift arbitrarily.
- Atoms typically shorter than signal.
- No uniqueness constraint as in JADL: multiple shifts per atom allowed in each signal.

Synthetic data

200 signals were generated from a synthetic dictionary of $K = 3$ atoms. Random events and Gaussian noise were added. Dictionaries with different numbers of atoms (see table below) were learned for JADL and DL on the noisy signals (200 iterations each). Denoising was then performed by sparse coding over the learned dictionaries. The plots show the denoised signals for best performing K .



Error for different numbers K of reconstructed atoms*

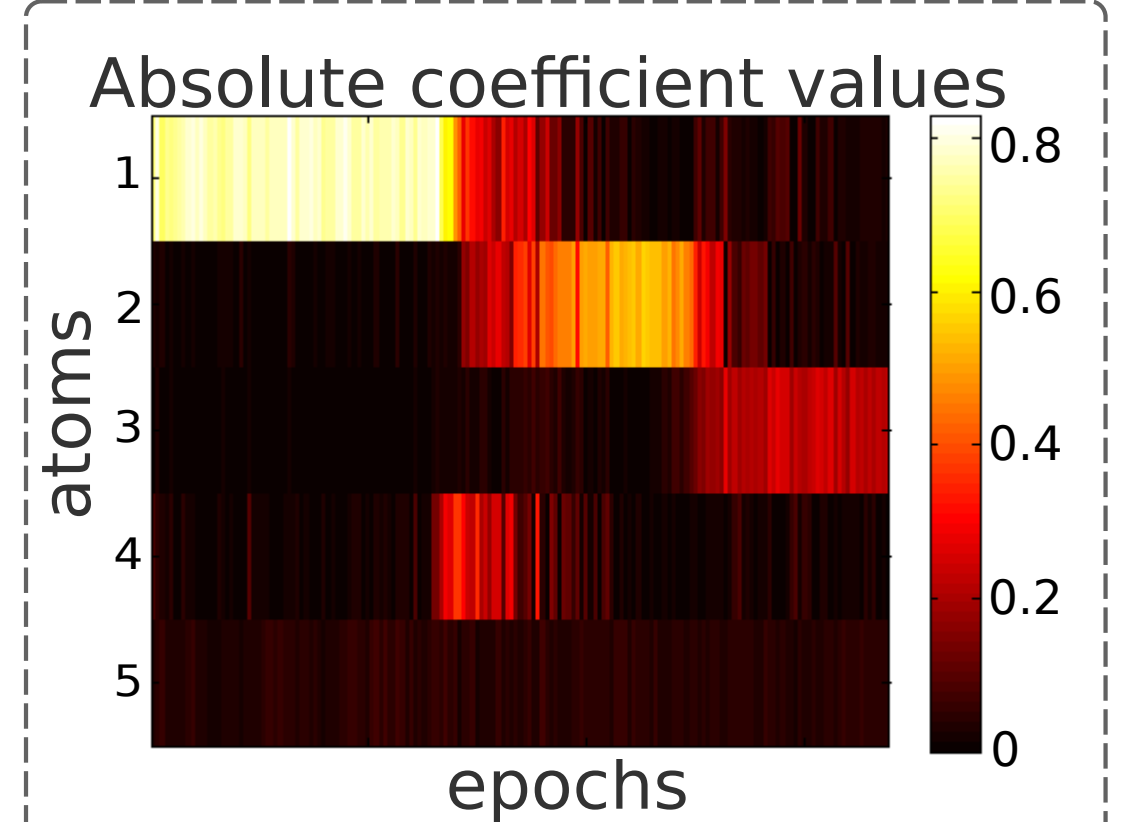
	K=1	K=2	K=3	K=4	K=5	K=6	K=8	K=10	K=12
DL	0.869	0.747	0.635	0.535	0.515	0.498	0.487	0.505	0.539
JADL	0.505	0.283	0.214	0.230	0.277	0.284	0.317	0.325	0.330

*For each K , we selected the parameter λ that gave the smallest error. This was $\lambda = 0.1$ for DL and $K = 8$ and $\lambda = 0.1$ for JADL and $K = 3$.

Real data

In an animal model of epilepsy, local field potentials were recorded during one hour with an intra-cranial electrode in a Wistar-Kyoto rat. Bicuculline (a blocker of inhibition) was injected in the cortex to elicit epileptic-like discharges. 169 of these spikes were then selected visually and segmented into epochs of 10 seconds.

Learned coefficients and shifts provide insight into data

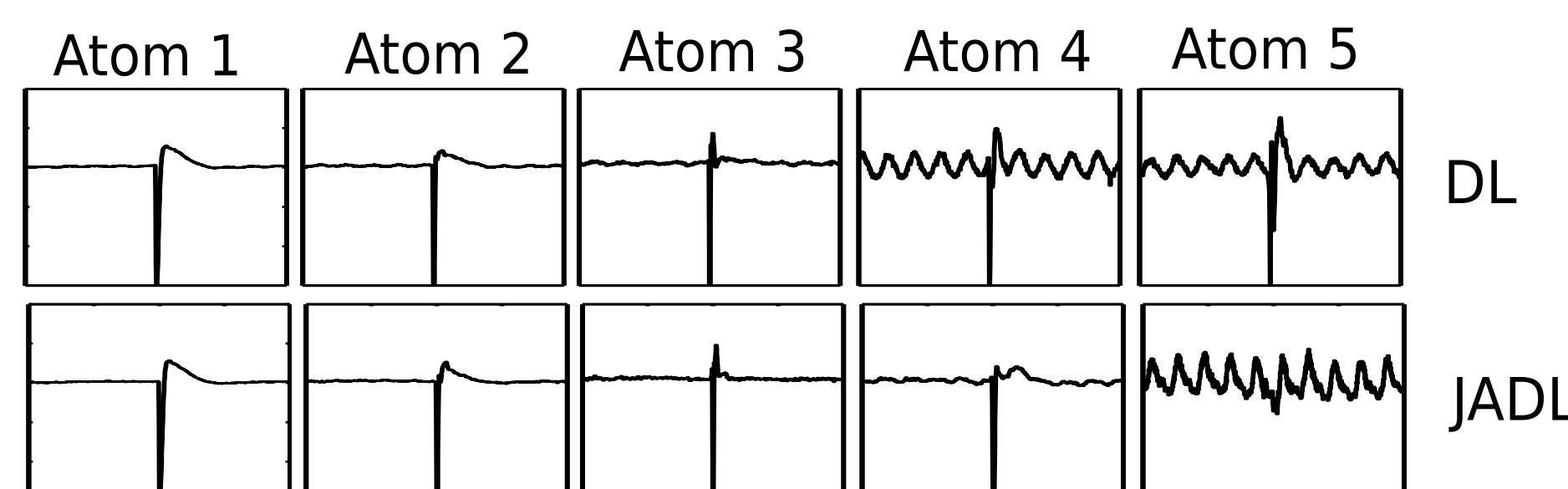


- Epochs can be divided into segments with one dominating spike each.

Epoch 41 Epoch 111 Epoch 161 Average

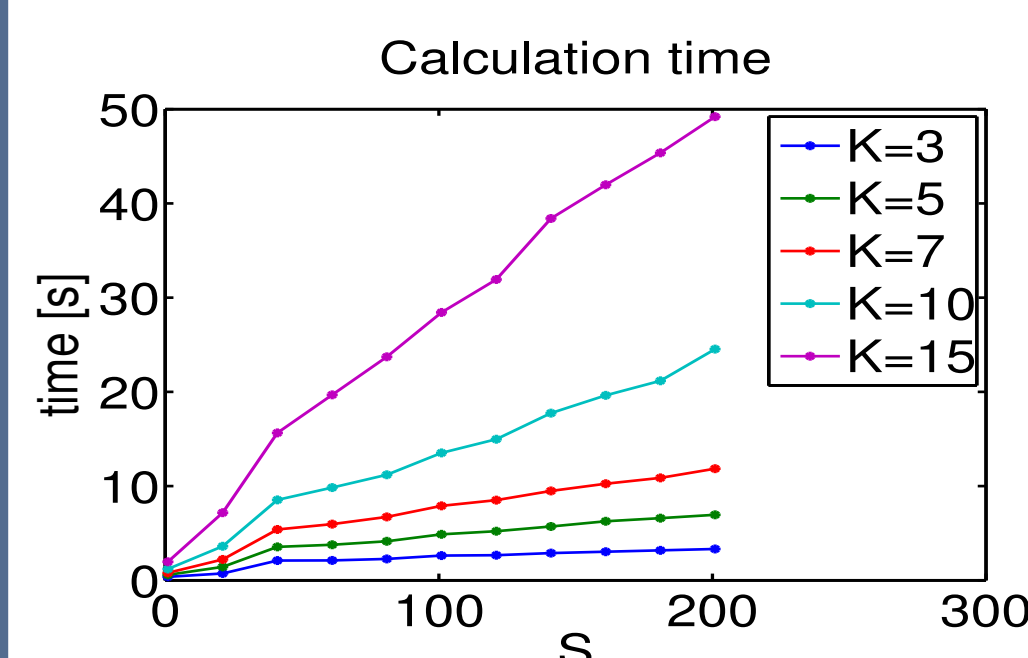
Learned dictionaries

On the 169 epochs, dictionaries were learned using DL and JADL, each algorithm performing 200 iterations.

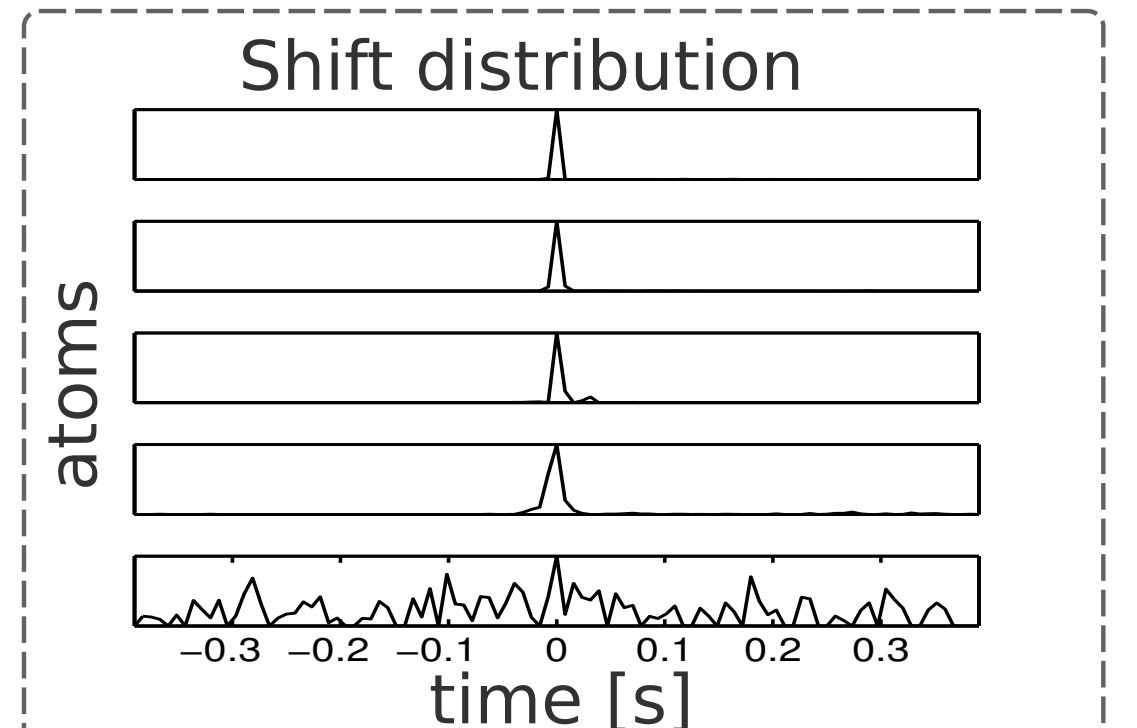


- Only JADL clearly separates spikes from oscillations.
- DL smooths the periodic wave, as its atoms do not adapt to phase shifts.

Performance



While both S (the number of allowed shifts) and K influence the size of the "unrolled" dictionary, the increase in computation time due to S is only linear. This comes from the uniqueness constraint which bounds the complexity as well as the use of fft-based convolution for large S .



- The spikes (first four atoms) are well aligned across trials.
- The periodic atom (five) occurs with uniformly distributed phase.

Conclusion

We presented a new method (JADL) which is an extension to dictionary learning and designed to analyze multi-trial neuroelectric datasets. We evaluated JADL on synthetic and real data and showed its superiority to common dictionary learning. In particular, JADL showed the following qualities:

- Ability to learn main waveforms in trials and to separate them.
- Learned shifts and coefficients give insight into the changes of waveforms (phase, latency, amplitude).
- Computational efficiency, even for high shift-tolerance.
- Robustness and denoising qualities.

References :
 [1] Mairal et al. Online Learning for Matrix Factorization and Sparse Coding. *JMLR*, 11(Aug.), 19-60 (2010)
 [2] Hanselmayr et al. Alpha phase reset contributes to the generation of ERP's. *Cerebral Cortex*, 17, 1-8 (2007)
 [3] Benar et al. Consensus matching pursuit for multi-trial EEG signals. *J. Neurosci. Meth.*, 180(1), 161-170 (2009)
 [4] Efron et al. Least angle regression. *The Annals of Statistics*, 32(2), 407-499 (2004)
 [5] Blumensath et al. Sparse and shift-invariant representations of music. *IEEE Trans Audio Speech Lang Processing*, 14(1), 50-57 (2006)