

Linear Periodic Discriminant Analysis of Multidimensional Signals

Dounia Mulders[1,2(✉)], Cyril de Bodt[1], Nicolas Lejeune[2], André Mouraux[2], and Michel Verleysen[1]

[1] ICTEAM Institute, Université catholique de Louvain, Place du Levant 3, 1348 Louvain-la-Neuve, Belgium
dounia.mulders@uclouvain.be
[2] IONS Institute, Université catholique de Louvain, Avenue Mounier 53, 1200 Woluwe-Saint-Lambert, Belgium

Abstract. Extracting relevant information from noisy multidimensional signals has tremendous impacts in numerous applications, ranging from audio separation to electrophysiological recording analysis. Linear filters are often considered to reconstruct and interpret the latent sources generating the data. Known properties of the sources can be used to guide their separation. In neuroscience, the cortical processes underlying perception in different modalities (visual, auditory, ...) is often studied using electroencephalography (EEG) during periodic stimulation, eliciting periodic activity in neural sources, some of which being specific to the considered modality. Whereas current approaches extract sources either periodic or discriminative, none of them accounts for both aspects at once. This paper proposes several methods extracting periodic sources specific between two classes, hence termed as Linear Periodic Discriminant Analysis methods. They are validated on synthetic data and EEG recordings of subjects to whom periodic stimulation from two modalities is applied. The methods highlight modality-specific periodic responses.

Keywords: Linear filtering · Specific periodic components
Generalized Rayleigh Quotient · Source separation · Steady-states
EEG

1 Introduction

Multidimensional signals are encountered in a broad range of situations including music editing, speech recognition or physiological monitoring. In many cases, such observed signals result from the noisy mixing of several components or sources of interest [6]. So-called source separation algorithms are therefore considered to recover the latent sources. The mixing is often linearly approximated, allowing an easy interpretation of the sources in the original sensor space. Independent Component Analysis (ICA) methods are probably the most well-known approaches to this kind of problems, only relying on the minimal assumption of

© Springer Nature Switzerland AG 2018
L. Cheng et al. (Eds.): ICONIP 2018, LNCS 11306, pp. 476–487, 2018.
https://doi.org/10.1007/978-3-030-04224-0_41

independence between the source signals [8]. ICA has proven to be very useful in many contexts, including the extraction of single speech signal from mixed recordings [7] or artifacts removal from electrophysiological signals [12]. In several applications however, specific structures of the latent sources are expected, which can be used to guide the definition of source separation algorithms. In audio signal processing for instance, diverse forms of guidance have already been considered, using either priors on the mixing processes or the temporal dynamics of the sources [20]. Such temporal structures are also exploited in the study of biological signals, for instance to enhance periodic components of a given multidimensional signal [14]. Other works have extended this kind of methods to extract quasi-periodic signals such as the electrocardiogram [17].

In this line, structured, and in particular periodic, signals also often arise in neuroscience. For instance, periodic stimulation is increasingly proposed to probe sensory perception in humans [3]. This kind of stimulus elicit a periodic activity at the stimulation frequency in some neural populations, known as a steady-state response (SSR), which can be partly recorded with electroencephalography (EEG). Each source contributing to the SSR associated to stimuli from one modality (e.g. visual, auditory, tactile,...) can be common to several studied modalities, or specific to a particular one. As an example, if cool and warm periodic stimuli are applied to a subject, part of the SSR can be expected to be common to the two classes, warm and cool, since both stimuli share some characteristics (they both activate the spinothalamic system). To study the specific cortical processing of each class of stimulus however, one is more interested in finding the most discriminant components of the SSR. Although linear filtering methods aiming to extract either periodic or discriminant components exist and are well-documented in the literature, the optimization of filters defining periodic components which are specific to one class of signals among two has surprisingly not yet been studied. Such a method would better capture the mechanisms governing specific sensory perception in humans, thereby helping to understand and subsequently diagnose sensory disorders. In this context, this paper presents methods to extract such sources. The approach is inspired, on the one hand, by the source separation methods aiming to maximize periodicity of the extracted components [14] and, on the other hand, by the Common Spatial Pattern (CSP) algorithm largely studied and adapted in Brain-Computer Interfacing (BCI) researches [2,11]. The specific periodic components are obtained by optimizing interpretable linear filters, hence the name of the presented approach: Linear Periodic Discriminant Analysis (LPDA). Several objective functions are proposed, compared and assessed on synthetic and real data. In particular, results on both simulated and real data are convincing, the presented methods enabling the discovery of modality-specific periodic structures. Our methodology, combining periodicity and discrimination maximization, outperforms state-of-the-art approaches accounting for one of both aspects at once.

The paper is organized as follows. Section 2 introduces and motivates the definition of several LPDA objective functions. Experiments are conducted in Sect. 3, both on synthetic and EEG data. Section 4 finally concludes the work.

2 Linear Periodic Discriminant Analysis Methods

This section defines and justifies the proposed approaches to extract periodic components specific to one among two multidimensional signals, belonging to two different classes. The zero-mean (over time) signals from the first and second classes, respectively denoted by $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t) \in \mathbb{R}^C$, are assumed to result from the linear mixing of latent periodic sources. For instance, they can be C-channel EEG signals recorded on one subject in two different conditions, e.g. exposed to periodic warm vs. cool stimulation. The sources, of period $T_\star$ and fundamental frequency $f_\star = 1/T_\star$ for both classes 1 and 2, can either be common to both signals or specific to $\mathbf{x}_1$ or $\mathbf{x}_2$. This leads to the noiseless latent model

$$\mathbf{x}_i(t) = A_u \mathbf{z}_u(t) + A_{s,i} \mathbf{z}_{s,i}(t), \tag{1}$$

where $i \in \{1, 2\}$ and s (resp. u) indexes the *specific* (resp. *unspecific*) mixing matrices $A_{s,i} \in \mathbb{R}^{C \times N_{s,i}}$ (resp. $A_u \in \mathbb{R}^{C \times N_u}$) and periodic sources $\mathbf{z}_{s,i}(t) \in \mathbb{R}^{N_{s,i}}$ (resp. $\mathbf{z}_u(t) \in \mathbb{R}^{N_u}$), $N_{s,i}$ (resp. N_u) being the number of specific (resp. unspecific) sources. In practice, different sources of noise will also affect the observed signals. Also, each specific source has a distinct level of specificity expressed by its variance. For convenience, the sources in a vector are ordered in decreasing order of specificity; for instance, $\mathbf{z}^1_{s,1}(t)$ is the most specific source for class 1, where the exponent indicates the component of the vector $\mathbf{z}_{s,1}(t)$. Each column of a mixing matrix A is a spatial pattern of the associated source [16]. Under this framework, both the specific spatial patterns $A_{s,i}$ and the specific source time courses $\mathbf{z}_{s,i}(t)$ should be recovered using $\mathbf{x}_1$ and $\mathbf{x}_2$. To this end, spatial filters $\mathbf{w} \in \mathbb{R}^C$ are optimized to obtain filtered signals $s_i(t) := \mathbf{w}^T \mathbf{x}_i(t)$ such that $s_1(t) \approx \mathbf{z}^1_{s,1}(t)$. Similarly, the filters can be optimized for class 2 to obtain $s_2(t) \approx \mathbf{z}^1_{s,2}(t)$ instead. Once a first optimal filter $\mathbf{w}_1$ for class i is found by optimizing a cost function F, a second optimal filter can be obtained, leading to a component in the orthogonal subspace of the first one. A matrix $W \in \mathbb{R}^{C \times C}$ is hence recursively built, whose columns $\mathbf{w}_k$ are the filters ranked according to F. The first $N_{s,i}$ columns of the pseudo-inverse of W^T then estimate the spatial patterns of the extracted components for class i: $\left[(W^T)^{-1}\right](:, 1 : N_{s,i}) = \widehat{A_{s,i}}$.

In order to obtain optimized filters, one must first define relevant cost functions characterizing the class-specific periodicity of the filtered signals. In the following, we propose several such objective functions, based on generalized Rayleigh quotients capturing periodicity differences, to obtain the filters for the specific activity of class 1, without loss of generality; both classes can be swapped to obtain the specific activity of class 2.

• **Method 1 (M1).** Periodic Component Analysis [14,19] aims to extract periodic sources from a single data set based on the periodicity measure $(\sum_t (s(t + T_\star) - s(t))^2 / (\sum_t (s(t))^2)$ of the signal $s(t)$. In the studied two-classes setting, a ratio of periodicity inspired by this measure can be optimized to find components $s_i(t) = \mathbf{w}^T \mathbf{x}_i(t)$, being periodic in class 1 and not necessarily in

class 2:

$$\mathbf{w}_1 = \arg\min_{\mathbf{w}} \left\{ F_1(\mathbf{w}) = \frac{\sum_t (s_1(t+T_\star) - s_1(t))^2}{\sum_t (s_2(t+T_\star) - s_2(t))^2} = \frac{\mathbf{w}^T R_1(T_\star)\mathbf{w}}{\mathbf{w}^T R_2(T_\star)\mathbf{w}} \right\}, \quad (2)$$

where $R_i(T_\star) = \mathbb{E}_t\{(\mathbf{x}_i(t+T_\star) - \mathbf{x}_i(t))\,(\mathbf{x}_i(t+T_\star) - \mathbf{x}_i(t))^T\}$. It can be noted that $F_1(\mathbf{w}) \geq 0$ and that $F_1(\mathbf{w}) = 0$ for a perfectly periodic signal $s_1(t)$. However, $F_1(\mathbf{w}) = 0$ as well when $s_1(t) = 0$ even though a flat signal may probably not be related to the sources of interest. This objective should hence be used with caution when some channels are correlated, in which case $s_1(t)$ could be 0.

• **M2.** Alternatively, the periodicity of the filtered signals can be emphasized by a simple averaging over the fundamental periods: $s_{\mu i}(t) := \frac{1}{K} \cdot \sum_{k=0}^{K-1} s_i(t+kT_\star) := \mathbf{w}^T \mathbf{x}_{\mu i}(t)$, with K the number of periods in the available signals. As in the CSP algorithm [2], the variance ratio of these averaged signals can be optimized as

$$\mathbf{w}_1 = \arg\max_{\mathbf{w}} \left\{ F_2(\mathbf{w}) = \frac{\sum_t (s_{\mu 1}(t))^2}{\sum_t (s_{\mu 2}(t))^2} = \frac{\mathbf{w}^T C_{\mu 1}\mathbf{w}}{\mathbf{w}^T C_{\mu 2}\mathbf{w}} \right\}, \quad (3)$$

where $C_{\mu i} = \mathbb{E}_t\{\mathbf{x}_{\mu i}(t)\mathbf{x}_{\mu i}(t)^T\}$.

• **M3.** Another approach is based on Canonical Correlation Analysis (CCA), which was already proposed to extract periodic components [15]. In the present work however, the method maximizes the correlation between the differences of the filtered signals in both classes and an arbitrary periodic signal $\mathbf{w}_y^T \mathbf{y}(t)$ of the same length as s_i. This periodic signal is defined from the Fourier series of a periodic signal of fundamental frequency $f_\star$: $\mathbf{y}(t) = (\sin(2\pi f_\star t)\ \cos(2\pi f_\star t)\ \sin(2\pi 2 f_\star t)\ \ldots \sin(2\pi N_h f_\star t)\ \cos(2\pi N_h f_\star t))^T$, where N_h indicates the number of accounted harmonics. Since the correlation $\mathbb{E}_t\{\mathbf{w}^T(\mathbf{x}_1(t) - \mathbf{x}_2(t))\mathbf{y}(t)^T\mathbf{w}_y\}$ must be maximized with scale invariance, the optimization problem is formulated as

$$(\mathbf{w}_1, \mathbf{w}_{1,y}) = \arg\max_{\mathbf{w},\mathbf{w}_y} \left\{ F_3(\mathbf{w}, \mathbf{w}_y) = \frac{\mathbf{w}^T(C_{1;\mathbf{y}} - C_{2;\mathbf{y}})\mathbf{w}_y}{\sqrt{\mathbf{w}^T C_{12}\mathbf{w} \cdot \mathbf{w}_y^T C_{\mathbf{y}}\mathbf{w}_y}} \right\}, \quad (4)$$

where $C_{i;\mathbf{y}} = \mathbb{E}_t\{\mathbf{x}_i(t)\mathbf{y}(t)^T\}$, $C_{12} = \mathbb{E}_t\{(\mathbf{x}_1(t) - \mathbf{x}_2(t))(\mathbf{x}_1(t) - \mathbf{x}_2(t))^T\}$ and $C_{\mathbf{y}} = \mathbb{E}_t\{\mathbf{y}(t)\mathbf{y}(t)^T\}$. This method is only well suited when the common activity is assumed to be highly similar (in terms of amplitude and phase) in both classes, leading to its cancellation through the point-by-point difference $\mathbf{x}_1(t) - \mathbf{x}_2(t)$.

• **M4.** Inspired by the Spectral Contrast Maximization (SCM) approach [18], the ratio of spectral content at the fundamental frequency $f_\star$ and its harmonics in both classes can be optimized. With $S_i(f) := \mathcal{F}_f\{s_i(t)\} = \mathbf{w}^T\mathcal{F}_f\{\mathbf{x}_i(t)\} = \mathbf{w}^T X_i(f)$ the Fourier transform of the filtered signals at frequency f, a fourth objective can be formulated in the following way:

$$\mathbf{w}_1 = \arg\max_{\mathbf{w}} \left\{ F_4(\mathbf{w}) = \frac{\mathbb{E}_{f\in\nu}\{|S_1(f)|^2\}}{\mathbb{E}_{f\in\nu}\{|S_2(f)|^2\}} = \frac{\mathbf{w}^T S_{\mathbf{x}_1}\mathbf{w}}{\mathbf{w}^T S_{\mathbf{x}_2}\mathbf{w}} \right\}, \quad (5)$$

with $\nu := \{\pm f_\star, \pm 2f_\star, \ldots, \pm N_h f_\star\}$ the set of frequencies of interest and $S_{\mathbf{x}_i} := \mathbb{E}_{f\in\nu}\{X_i(f)X_i(f)^*\}$. This objective is expected to be maximized by concentrating the frequency content of $s_1(t)$ to these frequencies. It should however be stressed that it can also be maximized if its denominator tends toward 0 while keeping the numerator bounded.

• **M5.** The quantity $\sum_t (s_i(t+T_\star) - s_i(t))^2$ is used in (2) to describe the periodicity of the signal $s_i(t)$. An alternative periodicity measure which can be considered is $\sum_t s_i(t+T_\star)\cdot s_i(t)$ [17], leading to the optimization problem

$$\mathbf{w}_1 = \arg\max_{\mathbf{w}} F_5(\mathbf{w}), \text{ with} \tag{6}$$

$$F_5(\mathbf{w}) = \frac{\sum_t s_1(t+T_\star)\cdot s_1(t) - s_2(t+T_\star)\cdot s_2(t)}{\sum_t (s_1(t))^2 + (s_2(t))^2} = \frac{\mathbf{w}^T(C_1(T_\star) - C_2(T_\star))\mathbf{w}}{\mathbf{w}^T(C_1(0) + C_2(0))\mathbf{w}},$$

where $C_i(\tau) = \mathbb{E}_t\{\mathbf{x}_i(t+\tau)\mathbf{x}_i(t)^T\}$. Note that a difference of periodicity between the two classes is considered instead of a ratio similar to (2), because the matrices $C_i(T_\star)$ are not necessarily positive definite contrarily to $R_i(T_\star)$.

• **M6.** Another variant of (2) can be defined using a difference instead of a ratio of periodicity. A distinct normalization ensures the scale-invariance of the filters:

$$\mathbf{w}_1 = \arg\min_{\mathbf{w}} \left\{ F_6(\mathbf{w}) = \frac{\mathbf{w}^T(R_1(T_\star) - R_2(T_\star))\mathbf{w}}{\mathbf{w}^T(C_1(0) + C_2(0))\mathbf{w}} \right\}. \tag{7}$$

• **M7.** A similar change can be applied to (5), leading to a last formulation:

$$\mathbf{w}_1 = \arg\max_{\mathbf{w}} \left\{ F_7(\mathbf{w}) = \frac{\mathbf{w}^T(S_{\mathbf{x}_1} - S_{\mathbf{x}_2})\mathbf{w}}{\mathbf{w}^T(C_1(0) + C_2(0))\mathbf{w}} \right\}. \tag{8}$$

All the objective functions above, except the third one, are expressed as generalized Rayleigh quotients of the form $(\mathbf{w}^T A\mathbf{w})/(\mathbf{w}^T B\mathbf{w})$ and can be solved by the generalized eigenvalue decomposition (GEVD) of the matrix pair (A, B), yielding a set of filters $\{\mathbf{w}_k\}_{k=1}^C$ corresponding to the generalized eigenvectors ranked in decreasing or increasing order of cost function values, depending on whether the objective function is maximized or minimized. When A is symmetric and B positive-definite, the decomposition is real and hence interpretable [5]. The formulation (4) is also solved analytically, as for classical CCA [9].

It is here assumed that the filters are ranked in decreasing order of specificity for class 1, hence with $\mathbf{w}_1$ being the most specific filter for class 1. The successive extracted sources are denoted by $s_i^k(t) := \mathbf{w}_k^T\mathbf{x}_i(t)$.

3 Experimental Results

This section assesses the methods introduced in Sect. 2. Their behavior and performances are first studied and compared in Sect. 3.1 on synthetic data, for which the latent sources are known. Then, preliminary results on a real EEG data set are presented in Sect. 3.2, for which the specific patterns are obtained. As to the quality assessment, different criteria are proposed in Sect. 3.1.

3.1 Synthetic Data

As stated in Sect. 1, the studied methods aim to extract class-specific periodic components from two multidimensional signals. Therefore, synthetic data are generated by linearly mixing specific and common components for the two classes, with additional noise. Denoting by $X_i(t) \in \mathbb{R}^C$ the simulated data from class i at time sample t, $Z_{u,i}(t) \in \mathbb{R}^{N_u}$ and $Z_{s,i}(t) \in \mathbb{R}^{N_s}$ the common (i.e. unspecific to the class) and specific periodic factors respectively, for $i \in \{1, 2\}$, data are simulated according to the model

$$X_i(t) = \underbrace{A_u \cdot Z_{u,i}(t)}_{\text{periodic interference}} + \underbrace{A_s \cdot Z_{s,i}(t)}_{\text{signal}} + \underbrace{N_i(t)}_{\text{noise}}, \tag{9}$$

with the common and specific mixing matrices $A_u \in \mathbb{R}^{C \times N_u}$ and $A_s \in \mathbb{R}^{C \times N_s}$, $N_i(t) \in \mathbb{R}^C$ being additive random noise. The common activity can be seen as interferences in this problem. Although the mixing matrices are common to both classes, the factor (or source) time courses are not. In particular:

• the specific factors $Z_{s,1}$ and $Z_{s,2}$ are mutually independent, with distinct specificity levels expressed by their variances which are proportional to $(N_s - 1,\ N_s - 2,\ \ldots,\ 1,\ 0)$ and $(0,\ 1,\ \ldots,\ N_s - 2,\ N_s - 1)$ [21]. These sources are ordered in decreasing order of specificity for class 1, the first (resp. last) $\lfloor \frac{N_s}{2} \rfloor$ being specific to class 1 (resp. 2). If $N_s = 1$, there is no specific factor for class 2.

• The common factors $Z_{u,i}$ are equal in both classes, up to an arbitrary time lag, as the common activity can have distinct latencies, e.g. due to differences in fiber conduction velocities in the case of EEG signals.

For each simulation, the entries of the mixing matrices are uniformly distributed in $[0, 1]$ and the factors are random periodic signals of period $T_\star$. However, since the performances of the optimized filters highly depend on the relative position in $\mathbb{R}^C$ of the columns of A_u with respect to the ones of A_s, the columns of A_s can be altered to reach some pre-defined angles with respect to the ones of A_u.

Regarding the assessment, different types of metrics can be considered.

• The reconstruction quality of the mixing matrix is evaluated using the canonical angle (CA) between the subspace spanned by the columns of A_s and its estimate $\widehat{A_s}$ computed from the optimized filters. The CA is the maximum angle found between orthogonal bases of these subspaces [5], and is therefore invariant to permutation and scaling of the columns of each matrix. Focusing on the most discriminant periodic source, the angle between the first true and estimated spatial patterns ($A_s^1 := A_s(:, 1)$ and $\widehat{A_s^1} := \widehat{A_s}(:, 1)$) is also evaluated.

• The time-correlation $\text{corr}(s_1^1(t), \mathbf{z}_{s,1}^1(t))$ between the first true and estimated sources, $\mathbf{z}_{s,1}^1(t)$ being the first component of $\mathbf{z}_{s,1}(t)$, is insensitive to the pseudo-inversion required to obtain the spatial patterns from the optimized filters.

• The periodicity difference between $s_1^1(t)$ and $s_2^1(t)$ assesses the relevance of the first filter $\mathbf{w}_1$ in terms of periodicity and discrimination. The following measure

is employed to quantify the periodicity of a signal $y(t) \in \mathbb{R}$ at a frequency $f_\star$:

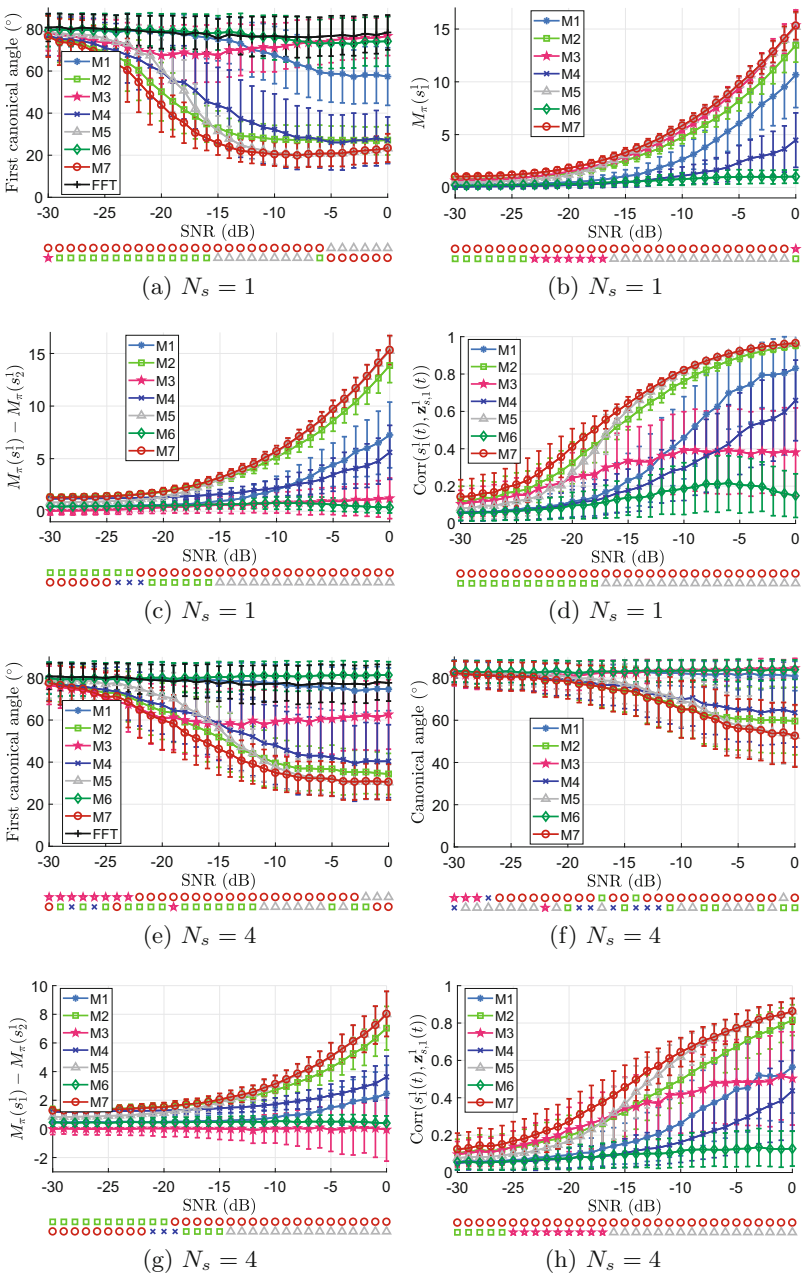
$$M_\pi(y) = 100 \cdot \frac{\sum_{k=1}^{\lfloor f_s/(2\cdot f_\star)\rfloor} Y_{NS}(k \cdot f_\star)}{\sum_f |Y(f)|}, \tag{10}$$

where f_s is the sampling frequency, $Y(f) = \mathcal{F}_f\{y(t)\}$ and Y_{NS} is the *noise-subtracted spectrum*. At frequency f, $Y_{NS}(f) \in \mathbb{R}$ is obtained by subtracting the average amplitude at 10 neighboring frequencies (5 higher and 5 lower) from the frequency amplitude $|Y(f)|$. Additional details are provided in [13,14]. The components extracted using the methods from Sect. 2 should maximize $M_\pi(s_1^1) - M_\pi(s_2^1)$, while providing topographical patterns of the extracted components which can be interpreted thanks to the linear filtering. Only this criterion can be used on real data as it does not rely on the 'true' sources.

In the following, each signal is generated with $C = 20$ channels, a sampling rate of 500 Hz, $f_\star = 0.5$ Hz and a duration of 20 seconds. The signal variances are defined over the channels and time samples [18]. Denoting $\sigma_u^2 := \text{var}(A_u \cdot \mathbf{z}_{u,i}(t))$, $\sigma_s^2 := \text{var}(A_s \cdot \mathbf{z}_{s,i}(t))$ and $\sigma_N^2 := \text{var}(\mathbf{n}_i(t))$, the two first quantities control the signal-to-interference ratio (SIR), while the signal-to-noise ratio (in dB) is defined as $\text{SNR} = 10 \cdot \log_{10}(\sigma_s^2/\sigma_N^2)$. In order to simulate realistic EEG data, which are studied in Sect. 3.2, pink additive noise is chosen for $N_i(t)$, i.e. with a frequency power spectrum proportional to $1/f$ [1]. Unless stated otherwise, there are $N_u = 10$ unspecific factors and the SIR is fixed with $\sigma_u = 1$ and $\sigma_s = 1$. The canonical angles between A_u and A_s are $\geq 70°$ and there is a time lag of 0.3 seconds between $Z_{u,1}(t)$ and $Z_{u,2}(t)$. The performances of the methods are commented with either 1 or 4 specific pattern(s) and then with varying unspecific activity according to N_u and σ_u.

• The performances of the methods with $N_s = 1$ specific factor are depicted in Figs. 1a to d. The angular position of the first specific pattern A_s^1 is also compared to the C-length vector of the Fourier frequency coefficients of the mixed signals at $f_\star$ in Fig. 1a (curve labeled 'FFT'). Although this pattern is usually employed in neuroscience studies to describe the periodic cortical activity elicited by periodic stimuli, it can be highly different from the spatial pattern of the source of interest. All the quality measures considered indicate the superiority of M7 to recover the specific pattern, followed by M2 and M5. Comparing Figs. 1b and c indicates that M3 probably amplifies common instead of specific activity, leading to both high $M_\pi(s_1^1)$ and $M_\pi(s_2^1)$. As this method uses a point-by-point difference between the two classes of signals, it is sensitive to a non-zero time lag between the common activity in both data sets.

• When $N_s = 4$, the results in Figs. 1e to h confirm the overall previous conclusions. The canonical angles between all the true and estimated specific patterns shown in Fig. 1f are larger than the angles computed only using the most specific pattern (Fig. 1e), indicating that the methods indeed recover more easily the first specific pattern. All methods are affected by the increased number of specific sources, and this effect is more pronounced for M1.



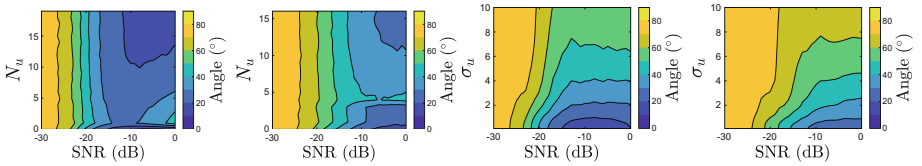
(a) $N_s = 1$ (b) $N_s = 1$

(c) $N_s = 1$ (d) $N_s = 1$

(e) $N_s = 4$ (f) $N_s = 4$

(g) $N_s = 4$ (h) $N_s = 4$

Fig. 1. Methods performances (mean ± std of 500 runs) on synthetic data as a function of the SNR. The two rows of markers below each figure indicate (1) the best method on average for each SNR and (2) the best method which is significantly outperformed by the best one (t-tests, Holm-Bonferroni corrected).



(a) $N_s = 1$, $\sigma_u = 1$ (b) $N_s = 4$, $\sigma_u = 1$ (c) $N_s = 1$, $N_u = 10$ (d) $N_s = 4$, $N_u = 10$

Fig. 2. Angle between the first specific true and estimated patterns, for M7 on synthetic data as a function of the SNR and either N_u (a, b) or σ_u (c, d). Results are averaged over 500 signal generations and $\sigma_s = 1$.

• The effects of σ_u and N_u on the M7 performances (the seemingly most efficient method) are shown in Fig. 2. For a fixed SIR, the number of common factors N_u only affects the performances when the SNR gets closer to 0, leading to improved results as N_u increases. On the other hand, increasing σ_u deteriorates the recovery of the most specific pattern, especially when $N_s = 4$. Let us note that, unlike the angle, the correlation between the first true and estimated sources is almost independent of σ_u (not shown due to a lack of space), suggesting that increasing σ_u does not affect the first optimal filter but rather the following ones, which in turn influences the first pattern after inverting the filter matrix.

The results presented in this section are obtained on data with 20 channels. Similar conclusions can be drawn with different dimensions, although on average all the angles tend to increase with the dimension, as a consequence of the curse of dimensionality [10]. Results are also similar with different canonical angles between A_u and A_s, with overall decreasing performances when they tend to be small as it becomes impossible for a filter to be orthogonal to the interference patterns A_u. M2 is especially affected by the reduction of these angles, while M5 and M7 are more robust (results not shown due to a lack of space).

3.2 Electroencephalogram (EEG) Data

This section shows brief results of the LPDA methods and of state-of-the-art filters on a real EEG data set, with 64 channels and sampled at 1000 Hz, which was recorded on 15 healthy subjects to whom periodic stimulation from four classes was applied: painful warm or non-painful cool, each applied on a variable or fixed skin surface (denoted resp. by w1, w2, c1 and c2). Each subject received each kind of stimulus with a frequency $f_\star = 0.2$ Hz and a duration $T = 75$ s (15 cycles). From a neuroscientific point of view, interest lies in finding the specific components of the condition w1 compared to c1 and w2 compared to c2. It would allow to disentangle the neural processing of warm and cool stimuli, thereby providing cues on the specific brain responses to painful stimuli.

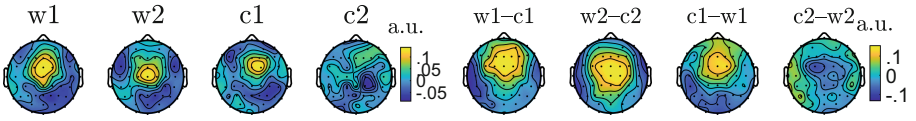
As to the compared methods, to the best of our knowledge, there is no existing approach optimizing both the periodicity and the specificity of an extracted component at once. The results of our methods are hence compared to one method maximizing the periodicity (SCM [14]), another one optimizing the dis-

crimination (CSP [11]) and Principal Component Analysis (PCA), a widely used non-guided filtering method.

LPDA methods are applied on the EEG data of each subject to obtain, for each ordered pair of conditions among w1–c1, w2–c2, c1–w1, c2–w2, the most specific source for the first condition of the pair and its associated spatial pattern and filter. Each optimal filter $\mathbf{w}_1$ can be applied to both signals of the pair to obtain $s_1^1(t)$ and $s_2^1(t)$. Table 1 confirms the superiority of method M7, followed by M5, in terms of difference of periodicity in any two compared classes, as it was observed on synthetic data, with the periodicity measure (10). M2, M3 and M6 results are of the same order of magnitude as the ones of M4 and are not shown due to a lack of space. Except for M7, the standard deviations are roughly of the same order as the mean results, indicating a high variability across subjects. It should be stressed that the performances of all methods vary coherently with respect to the subjects since M7 always statistically significantly outperforms the other schemes. These results are encouraging as they show that periodic components specific to one class can be extracted, M7 overcoming the three baselines. Further validations should be undertaken to disentangle physiologically interpretable patterns from potentially slightly overfitted results.

Table 1. For each pair of conditions (in row), mean(std) for the 15 subjects of the difference of periodicity M_π (10) between the filtered signals in both conditions using the most specific filter for the first condition of the pair. The best performances per row are in bold. Non-italic characters indicate that the corresponding signal is significantly less periodic than the best one of the same row (paired t-tests with Holm-Bonferroni correction). For M4 and M7, $N_h = 10$.

	LPDA methods				Baseline methods		
	M1	M4	M5	M7	SCM	CSP	SVD
w1–c1	0.55(0.42)	0.35(0.43)	0.55(0.94)	**3.25(1.04)**	2.84(1.08)	0.25(0.56)	0.22(0.32)
w2–c2	0.31(0.46)	0.09(0.15)	0.68(1.23)	**3.25(1.05)**	2.98(0.98)	0.15(0.29)	0.17(0.37)
c1–w1	0.10(0.53)	0.30(0.27)	0.31(0.39)	**2.43(0.67)**	1.59(0.86)	−0.12(0.28)	−0.19(0.28)
c2–w2	0.11(0.46)	0.19(0.14)	0.42(0.47)	**2.44(0.55)**	1.78(0.69)	−0.01(0.36)	−0.12(0.43)



(a) FFT topographies (b) Specific spatial patterns (M7)

Fig. 3. Group-level average spatial patterns of (a) the frequency amplitudes at $f_\star = 0.2$ Hz and (b) the most specific components extracted by M7 ($N_h = 10$) for the first class in each pair.

Finally, Fig. 3 shows the spatial pattern of the frequency spectrum at $f_\star$, usually analyzed in neuroscience, and the most specific spatial pattern extracted

for each class of signals. The specific patterns of the two warm conditions are relatively similar but different form those of the cool conditions. The specific warm patterns have a larger central positive cluster than the corresponding FFT topographies, while the specific cool patterns are more diffuse, reflecting a larger inter-subjects variability. Cold receptors being known to habituate rapidly to repeated stimuli [4], this could affect the stability of the SSR.

4 Conclusions and Perspectives

This paper proposes several linear filtering methods for multidimensional signals aiming to extract periodic components which are specific to one class among two. Seven objective functions are defined, combining discrimination and periodicity maximization in simple formulations. The methods are discussed and their results compared on synthetic data sets, allowing to study their behavior and reliability in a controlled context. They are also applied on real EEG recordings to highlight modality-specific spatial patterns of brain activity elicited by periodic cool and warm stimulation. The results are encouraging, as the extracted patterns are consistent with the nature of the considered stimuli and indeed exhibit a markedly increased periodicity for the considered class.

Further investigations will be carried regarding the proper validation of the extracted specific patterns to enable their fair interpretation in the neuroscience context. This validation is however not trivial due to the very limited number of trials available for each subject and condition. While the current results are assessed by quantifying the difference of periodicity between each compared pair of modalities, the discriminative power of the filters could also be analyzed.

Acknowledgments. DM and CdB are Research Fellows of the FNRS. The authors gratefully thank Prof. Christian Jutten for insightful discussions.

References

1. Bedard, C., Kroeger, H., Destexhe, A.: Does the 1/f frequency scaling of brain signals reflect self-organized critical states? Phys. rev. let. **97**(11), 118102 (2006)
2. Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., Muller, K.R.: Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process. Mag. **25**(1), 41–56 (2008)
3. Colon, E., Legrain, V., Mouraux, A.: Steady-state evoked potentials to study the processing of tactile and nociceptive somatosensory input in the human brain. Clin. Neurophysiol. **42**(5), 315–323 (2012)
4. De Keyser, R., van den Broeke, E.N., Courtin, A., Dufour, A., Mouraux, A.: Event-related brain potentials elicited by high-speed cooling of the skin: a robust and non-painful method to assess the spinothalamic system in humans. Clin. Neurophysiol. **129**(5), 1011–1019 (2018)
5. Golub, G.H., Van Loan, C.F.: Matrix Computations, vol. 3. JHU Press, Baltimore (2012)

6. Hyvärinen, A.: Independent component analysis: recent advances. Phil. Trans. R. Soc. A **371**(1984), 20110534 (2013)
7. Hyvärinen, A., Oja, E.: Independent component analysis: algorithms and applications. Neural Netw. **13**(4–5), 411–430 (2000)
8. Jutten, C., Herault, J.: Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture. Signal Process. **24**(1), 1–10 (1991)
9. Krzanowski, W.: Principles of Multivariate Analysis, vol. 23. OUP Oxford, Oxford (2000)
10. Leskovec, J., Rajaraman, A., Ullman, J.D.: Mining of Massive Datasets. Cambridge University Press, New York (2014)
11. Lotte, F., Guan, C.: Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. IEEE TBME **58**(2), 355–362 (2011)
12. Makeig, S., Bell, A.J., Jung, T.P., Sejnowski, T.J.: Independent component analysis of electroencephalographic data. In: Advances in Neural Information Processing Systems, pp. 145–151 (1996)
13. Mouraux, A., Iannetti, G.D., Colon, E., Nozaradan, S., Legrain, V., Plaghki, L.: Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. J. Neurosci. **31**(16), 6079–6087 (2011)
14. Mulders, D., de Bodt, C., Lejeune, N., Mouraux, A., Verleysen, M.: Spatial filtering of EEG signals to identify periodic brain activity patterns. In: Deville, Y., Gannot, S., Mason, R., Plumbley, M.D., Ward, D. (eds.) LVA/ICA 2018. LNCS, vol. 10891, pp. 524–533. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93764-9_48
15. Nakanishi, M., Wang, Y., Wang, Y.T., Mitsukura, Y., Jung, T.P.: A high-speed brain speller using steady-state visual evoked potentials. Int. J. Neural Syst. **24**(06), 1450019 (2014)
16. Samadi, S., Amini, L., Cosandier-Rimélé, D., Soltanian-Zadeh, H., Jutten, C.: Reference-based source separation method for identification of brain regions involved in a reference state from intracerebral EEG. IEEE Trans. Biomed. Eng. **60**(7), 1983–1992 (2013)
17. Sameni, R., Jutten, C., Shamsollahi, M.B.: Multichannel electrocardiogram decomposition using periodic component analysis. IEEE Trans. Biomed. Eng. **55**(8), 1935–1940 (2008)
18. Sameni, R., Jutten, C., Shamsollahi, M.B.: A deflation procedure for subspace decomposition. IEEE Trans. Signal Process. **58**(4), 2363–2374 (2010)
19. Saul, L.K., Allen, J.B.: Periodic component analysis: an eigenvalue method for representing periodic structure in speech. In: Advances in Neural Information Processing Systems, pp. 807–813 (2001)
20. Vincent, E., Bertin, N., Gribonval, R., Bimbot, F.: From blind to guided audio source separation: how models and side information can improve the separation of sound. IEEE Signal Process. Mag. **31**(3), 107–115 (2014)
21. Wu, W., Chen, Z., Gao, S., Brown, E.N.: A probabilistic framework for learning robust common spatial patterns. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009, pp. 4658–4661 (2009)