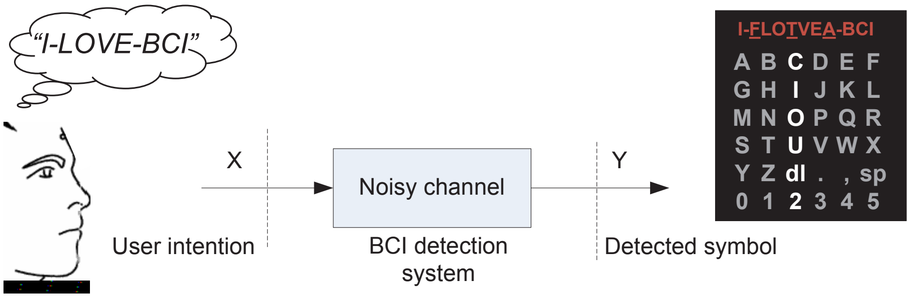

Figure 3.3: BCI system modeled as a communication channel. Simulation of an online experiment with 3 errors in 10 characters (70% accuracy).

made by the BCI system is given by

$$I(X;Y) = H(X) - H(X|Y) \tag{3.1}$$

where $H(X)$ is the source entropy and $H(X|Y)$ is the information lost in the noisy channel, i.e., it represents the classification error rate of the BCI system. Assuming a BCI with N_s possible choices (number of symbols) which are equiprobable, and an online classification accuracy of P_{ac}, then $I(X;Y) \equiv B$, measured in bits/symbol, is given by (see Appendix A.1 for a derivation of this formula)

$$B = \log_2(N_s) + P_{ac}\log_2(P_{ac}) + (1-P_{ac})\log_2\frac{(1-P_{ac})}{(N_s-1)}. \tag{3.2}$$

Taking the rate of possible selections per minute r_s (symbols per min, SPM) then the ITR is expressed by

$$ITR = r_s B. \tag{3.3}$$

The r_s rate is obtained from

$$r_s = \frac{60}{N_{rep} \times (N_{ev} \times SOA) + ITI}. \tag{3.4}$$

This metric is currently widely used by the BCI community as a benchmark metric for performance comparison, particularly in P300-based BCIs. The ITR reflects simultaneously the accuracy, the number of symbols per minute and the amount of encoded information. However, the use of this metric for the assessment of a BCI should always be accompanied with the online accuracy. Metric (3.3) can be fallacious because low levels of accuracy may provide reasonable bit rates while at the same time be unacceptable for an effective communication [Sellers 2006a]. For example, in Fig. 3.3, a simulation of an

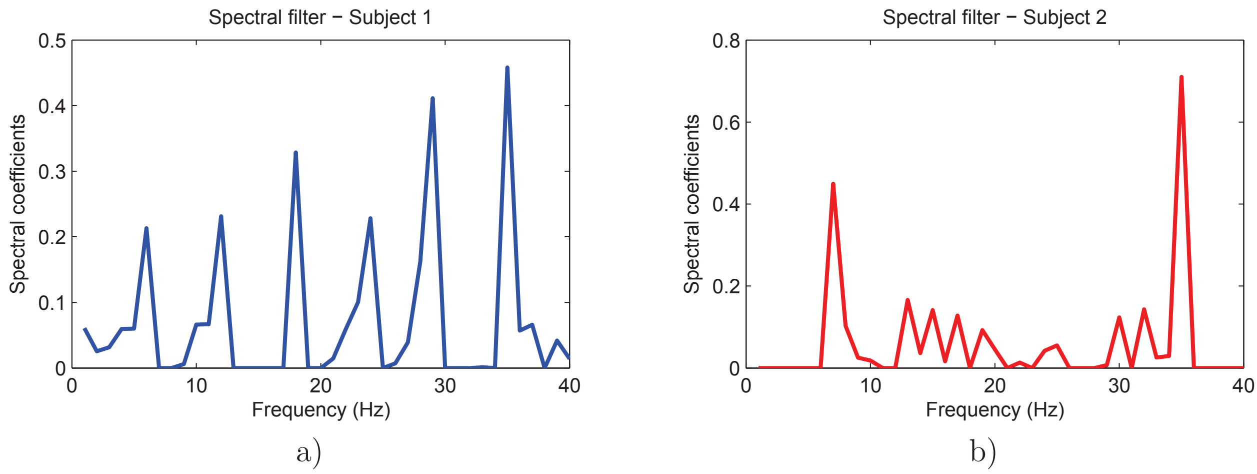

Figure 8.3: Optimal spectral filters $\tilde{\mathbf{h}}$ obtained for subject 1 (a) and subject 2 (b) learned by ASSFCB.

coefficients (spectral peaks) seem related to SSVEP evoked by stimuli frequencies. These frequencies overlap rhythms in the α band, which makes difficult to infer whether spectral coefficients are related to attention or relaxation states. Curiously, a high discriminative frequency around 35 Hz is detected by ASSFCB in both subjects. This deserves a future neurophysiological analysis. To assess the performance of ASSFCB across iterations, we computed the classification error rate for each iteration as shown in Fig. 8.4. Note that the first iteration coincides with the original spatial filter. The subsequent iterations show a significative improvement of the classification, showing the effectiveness of ASSFCB. This improvement is reached just after two or three iterations, and subsequent iterations may reduce the performance. Despite this significative improvement, the classification results were not sufficiently robust for online application.

8.2.5 Posterior probabilities - threshold adjustment

This approach takes the values returned from the posterior probabilities $P(C_i|\mathbf{y})$ computed from (5.30) using the Bayes theorem. Let us consider the three classes C_i with $i \in \{T, NT, NC\}$, then we have $P(C_T|\mathbf{y})$, $P(C_{NT}|\mathbf{y})$ and $P(C_{NC}|\mathbf{y})$ respectively for target, non-target and non-control. These probabilities were computed applying the qNB classifier to spatial FCB projections. The log-transformation was applied to (5.29), obtaining

$$P(C_i|\mathbf{y}) = \log(\pi_i) - \sum_j^{N_f} \log(\sigma_i(j)) - \frac{1}{2}\sum_j^{N_f}\left((\mathbf{y}(j) - \mu_i(j))\sigma_i^{-2}(j)(\mathbf{y}(j) - \mu_i(j))\right). \quad (8.1)$$

Appendix A

Bit rate metrics

A.1 Derivation of the channel capacity of a BCI

The channel capacity is obtained from the mutual information of the channel [Shannon 1948]

$$I(X;Y) = H(X) - H(X|Y) \tag{A.1}$$

where $H(X)$ is the source entropy and $H(X|Y)$ is the information lost in the noisy channel. Assuming N_s possible equiprobable choices (probability $P(x_i) = 1/N_s$) of the BCI paradigm, and the conditional probability $P(x_i|y_i)$ given by the BCI accuracy, p_a, then $I(X;Y)$ is given by

$$\begin{aligned} I(X;Y) &= \sum_{x} P(x_i)\log\frac{1}{P(x_i)} - \sum_{x,y} P(x_i y_j)\log\frac{1}{P(x_i|y_j)} \\ &= \sum_{x} P(x_i)\log\frac{1}{P(x_i)} - \sum_{x,y} P(x_i|y_j)P(y_j)\log\frac{1}{P(x_i|y_j)} \\ &= \sum_{x} P(x_i)\log\frac{1}{P(x_i)} - [P(x_1|y_1)P(y_1)\log\frac{1}{P(x_1|y_1)} + \cdots + \\ &\qquad + \cdots + P(x_1|y_{N_s})P(y_{N_s})\log\frac{1}{P(x_1|y_{N_s})}]N_s \\ &= \log(N_s) - [p_a\log\frac{1}{p_a} + (1-p_a)\log\frac{1}{\frac{1-p_a}{N_s-1}}] \\ &= \log(N_s) + p_a\log(p_a) + (1-p_a)\log\frac{1-p_a}{N_s-1}. \end{aligned} \tag{A.2}$$

The channel capacity $I(X;Y) \equiv B$, represents the amount of information transferred per symbol, measured in bits/symbol.