

Segmentation de signaux physiologiques par optimisation globale

Charles TRUONG^{1,2}, Laurent OUDRE^{3,2}, Nicolas VAYATIS^{1,2}

¹CMLA - ENS Cachan, CNRS - 61 Avenue du Président Wilson, 94230 Cachan

²COGNAC G - Université Paris Descartes, CNRS - 45, rue des Saints-Pères, 75006 Paris

³L2TI - Université Paris 13 - 99 Avenue Jean Baptiste Clément, 93430 Villetaneuse

charles.truong@cmla.ens-cachan.fr, laurent.oudre@univ-paris13.fr, nicolas.vayatis@cmla.ens-cachan.fr

Résumé – Cet article s’intéresse au problème concret de la segmentation de signaux physiologiques grâce à des méthodes d’optimisation globale. Plusieurs fonctions de coût, parfois inédites dans ce contexte, sont proposées pour détecter des changements dans la moyenne ou la variance, mais aussi dans la périodicité éventuelle des signaux. Ces fonctions seront évaluées et comparées dans un cadre unifié sur des signaux synthétiques et réels.

Abstract – In this article, we address the problem of segmentation of physiological signals through global optimization. Several cost functions are tested (some of them new in this context) in order to detect changes not only in the mean or variance but also in the periodicity of the signals. Results are provided in an unified framework on real and artificial signals.

1 Introduction

Cet article s’intéresse à la segmentation de séries temporelles multivariées en plusieurs phases homogènes. C’est une étape nécessaire, par exemple lorsque l’on cherche à extraire automatiquement de ces signaux des descripteurs pour réaliser des tâches de classification. Dans le meilleur des cas, on peut supposer que le signal est une suite de phases plus ou moins homogènes, qui correspondent à des phénomènes successifs différents. Une étape préalable à tout traitement de ces données consiste à isoler automatiquement ces phases (en estimer le début et la fin). Ce n’est qu’une fois cette segmentation réalisée que les descripteurs extraits sur chaque section peuvent s’avérer pertinents.

Dans cet article, nous nous intéressons à des signaux d’accélérométrie enregistrés dans le bas du dos grâce à des accéléromètres/gyromètres et échantillonnés à 100 Hz. Ces signaux ont été acquis durant des exercices de marche réalisés par des sujets sains et des patients souffrant de diverses pathologies orthopédiques. Ils sont multivariés et hétérogènes car issus de capteurs différents. Un exemple est présenté sur la figure 1. Les données collectées correspondent à l’exercice suivant : phase de repos, marche aller sur une certaine distance à la vitesse de confort, demi-tour, puis marche retour sur la même distance. L’objectif est ici d’isoler automatiquement les phases de repos, de marche et de demi-tour à partir des signaux, de la façon la plus robuste possible. Les ruptures entre phases sont parfois visibles à l’œil sur une ou plusieurs dimensions du signal. Néanmoins, la relative liberté des sujets lors de l’exercice introduit une grande variabilité : deux patients ont des accélérations d’ordres de grandeur très différents, les phases ont des durées très disparates, les vitesses de marche varient beaucoup entre les sujets, etc. Par ailleurs, les capteurs sont sujets à des chocs

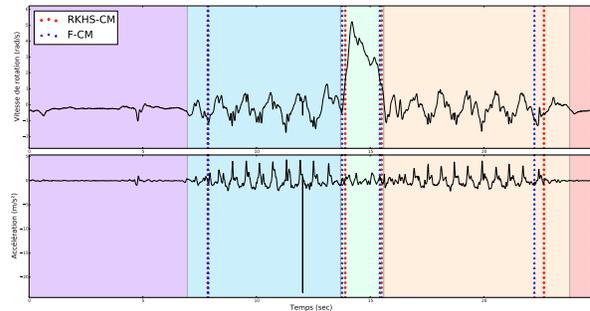


FIGURE 1 – Vitesse de rotation (haut) et accélération (bas). La véritable segmentation est marquée par les traits pleins et les changements de couleur. On distingue cinq phases successives : arrêt, marche, demi-tour, marche et arrêt. Les ruptures calculées avec F-CM et RKHS-CM (voir section 2.2) sont marquées par les traits pointillés.

qui créent du bruit et des points aberrants. Les contraintes issues de ce contexte clinique créent des signaux ayant une forme complexe. Pour répondre à ce problème, la méthode de segmentation par optimisation globale est la plus précise. On lui oppose souvent les méthodes à fenêtres glissantes (une rupture est placée quand un changement est détecté dans la fenêtre), les méthodes de partition *bottom-up* (le signal est segmenté en de nombreux segments, qui sont agrégés jusqu’à ce qu’un certain critère soit atteint) et *top-down* (le signal est segmenté en deux segments, puis l’opération est répétée sur les deux segments qui en résultent). Ces méthodes permettent d’étendre des algorithmes de détection de rupture unique. Elles rajoutent malheureusement des paramètres (taille de fenêtre, critère d’arrêt, d’élagage, ...) dont la calibration nécessite l’avis d’experts ou une optimisation complexe en grande dimension. De plus, elles ne permettent pas de facilement spécifier un *prior* pourtant naturel : le nombre de ruptures. C’est pourquoi certaines techniques ne sont pas adaptées à notre problème, comme la détection par maximum de vraisemblance d’une rupture dans

les coefficients ARMA [1], pourtant performante dans d'autres contextes. Bien que coûteuse en temps de calcul, l'optimisation globale permet d'obtenir la meilleure segmentation possible tandis que les autres techniques de la littérature n'offrent bien souvent que des approximations de la segmentation optimale [2]. D'autre part, l'utilisation du signal global permet de s'affranchir de paramètres additionnels, comme une taille de fenêtre ou des seuils empiriques [3]. En pratique, le problème est résolu dans un temps raisonnable car les signaux 2D considérés ne dépassent pas, en général, une minute (soit 6000 échantillons).

L'objet de l'article est l'étude de l'influence des fonctions de coût sur la segmentation par optimisation globale. Dans cette optique, plusieurs fonctions de coût, parfois jamais utilisées dans ce contexte, seront présentées, ainsi qu'un cadre permettant leur évaluation et leur comparaison sur des signaux synthétiques et réels. Ces fonctions permettent de prendre en compte des phénomènes complexes, comme des phases presque périodiques, qui existent dans les signaux multivariés réels. Elles étendent ainsi le cadre de la littérature où sont détectés des changements dans la moyenne et la variance de signaux univariés.

2 Méthode

2.1 Principe de la segmentation par optimisation globale

Le but est de segmenter un signal discret en portions homogènes, en utilisant tout le signal disponible. Il est nécessaire pour cela de définir une fonction caractéristique de l'uniformité ou de l'homogénéité d'une portion, communément appelée fonction de coût ou de contraste. Etant donné un intervalle entier $\llbracket s, e \rrbracket$, et une fonction de coût $\Delta(\cdot)$, on note $\Delta(s, e)$ la valeur de cette fonction de coût sur la portion de signal compris entre s et e . Le problème est ici résolu de façon globale : la meilleure segmentation possible est définie comme celle minimisant la somme des coûts sur les différents segments. Il faut pour cela explorer l'espace de toutes les segmentations possibles du signal et trouver celle qui minimise notre critère. Cette optimisation exhaustive peut s'exprimer facilement sous la forme d'un problème de programmation dynamique.

Notons $\mathcal{M}_K(s, e)$ l'ensemble des segmentations à K segments (donc $K - 1$ ruptures) de l'intervalle $\llbracket s, e \rrbracket$. Une segmentation $m \in \mathcal{M}_K(s, e)$ peut s'écrire $m = \{\llbracket s_1, e_1 \rrbracket, \dots, \llbracket s_K, e_K \rrbracket\}$, où, par convention, $s_1 = s$, $e_K = e$ et $s_{i+1} = e_i + 1$. Les intervalles $\llbracket s_i, e_i \rrbracket$ forment une partition disjointe de $\llbracket s, e \rrbracket$.

La segmentation revient alors à résoudre le problème de minimisation suivant :

$$\min_{m \in \mathcal{M}_K(s, e)} \sum_{i=1}^K \Delta(s_i, e_i) \quad (1)$$

Notons $I_K(s, e)$ la valeur minimum et $m_K^*(s, e)$ l'élément de $\mathcal{M}_K(s, e)$ pour lequel ce minimum est atteint. Il suffit de remarquer que :

$$I_K(s, e) = \min_{n \in \llbracket s, e-K+1 \rrbracket} (\Delta(s, n) + I_{K-1}(n+1, e)) \quad (2)$$

pour retrouver la formulation *top-down* d'un problème de programmation dynamique, qui a depuis été appliquée plusieurs fois à la détection de ruptures [2]. On pose par convention que $I_1(s, e) = \Delta(s, e)$ et $m_1^*(s, e) = \{\llbracket s, e \rrbracket\}$.

2.2 Choix de la fonction de coût

Soit $\{x_n\}_{n=1}^N$ un signal multivarié à valeurs dans \mathbb{R}^p ($x_n = (x_n^1, \dots, x_n^p)^T$) dont les dimensions ont été mises à l'échelle. Considérons une transformation Φ (potentiellement non linéaire) de \mathbb{R}^p dans un espace Ω . On définit alors :

- $y_n = \Phi(x_n)$ la transformation de x_n par Φ
- $\psi_s^e(y_n)$ une approximation du signal y_n sur l'intervalle $\llbracket s, e \rrbracket$

Les fonctions de coût présentées dans la suite de cet article seront toutes de la forme

$$\Delta(s, e) = \sum_{i=s}^e \|y_i - \psi_s^e(y_i)\|_{\Omega}^2. \quad (3)$$

$\Delta(s, e)$ correspond donc à l'erreur quadratique de modélisation de la transformation Φ de x_n . Résoudre (2) avec Δ comme en (3) revient à faire une régression des moindres carrés sur chacun des K segments des $\binom{N}{K}$ segmentations possibles (où $\binom{b}{a} = \frac{b!}{a!(b-a)!}$)

Identité - Constante par morceaux (I-CM). La première fonction de coût présentée utilise une transformation Φ égale à l'identité ($y_n = x_n$) et une approximation par une fonction constante par morceaux. On a donc :

$$\forall n \in \llbracket s, e \rrbracket, \psi_s^e(y_n) = \frac{1}{e-s+1} \sum_{i=s}^e y_i \quad (4)$$

Cette fonction de coût est donc sensible aux changements dans la moyenne du signal. Elle a déjà été utilisée dans la littérature [3]. Elle est néanmoins aveugle à des changements complexes, comme des ruptures dans les moments d'ordre supérieur à deux ou la périodicité. La méthode I-CM est équivalente à l'algorithme *k-means* en forçant les *clusters* à être composés de point contigus.

RKHS - Constante par morceaux (RKHS-CM). Une extension de l'algorithme *k-means* est le *kernel k-means*. Il peut séparer des *clusters* aux formes complexes (pas seulement des ensembles concentrés autour d'un centroid) [2]. De la même façon, la méthode I-CM est étendue grâce à une transformation Φ de *mapping*, classiquement employée dans les problèmes d'apprentissage statistique et basée sur l'utilisation de noyaux reproduisants. Considérant un noyau positif semi-défini $\mathcal{K} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, il définit un espace \mathcal{H} qu'on appelle l'espace de Hilbert associé au noyau \mathcal{K} (RKHS) et une fonction $\Phi : \mathbb{R}^d \rightarrow \mathcal{H}$ telle que $\langle \Phi(x) | \Phi(y) \rangle_{\mathcal{H}} = \mathcal{K}(x, y)$. L'idée est donc ici d'utiliser cette fonction Φ comme transformation ($y_n = \Phi(x_n)$), suivie d'une modélisation constante par morceaux (4). Un choix classique pour \mathcal{K} est un noyau gaussien dont la largeur est adaptée pour chaque signal selon l'heuristique de la médiane : $\mathcal{K}(x, y) = \exp(-\frac{\|x-y\|^2}{\sigma^2})$ et $\sigma = \text{médiane}_{i,j}(\|x_i - x_j\|)$. La transformée du signal n'est pas à calculer explicitement grâce

au *kernel trick*, car le coût sur un segment ne dépend que des produits scalaires $\langle \Phi(x_i) | \Phi(x_j) \rangle_{\mathcal{H}} = \mathcal{K}(x_i, x_j)$.

Différences finies - Constant (DF-C). Pour un signal quasi périodique de période T , les quantités $\|x_n - x_{n+T}\|$ sont presque nulles. On considère donc une transformation $y_n = x_n - x_{n+T}$ et on cherche à minimiser, sur un ensemble de périodes T prédéterminé, l'erreur quadratique en approchant ces quantités par zéro ($\psi_s^e(y_n) = 0$) :

$$\Delta(s, e) = \min_T \sum_{n=s}^{e-T} \|x_n - x_{n+T}\|^2 \quad (5)$$

A notre connaissance, les deux fonctions de coût suivantes n'ont jamais été utilisées dans le cadre de la détection de ruptures.

Fourier - Constante par morceaux (F-CM). Etant donné la présence de phases pseudo-périodiques, il est pertinent de s'intéresser au contenu fréquentiel du signal x_n . Pour cela, la transformation utilisée sera une transformée de Fourier à temps discret estimée sur une fenêtre de taille N_w et sur un ensemble de q fréquences f_1, \dots, f_q . Le vecteur y_n est alors défini comme la concaténation des profils spectraux (valeur absolue) obtenus pour chacune des p composantes du signal. Comme précédemment, le signal est approché par une fonction constante par morceaux (4) : l'idée est donc de distinguer des changements dans la périodicité du signal, c'est-à-dire des changements de localisation de l'énergie spectrale.

Identité - Cosinus déphasé (I-CD). L'approche fréquentielle peut aussi être illustrée dans le domaine temporel. La transformation Φ est prise égale à l'identité ($y_n = x_n$). En revanche, chaque composante de la série temporelle est approchée par une sinusoïde [4] : $\forall k \in \llbracket 1, p \rrbracket, \forall n \in \llbracket s, e \rrbracket$

$$\psi_s^e(y_n^k) = A \cos\left(2\pi n \frac{f}{F_s} + \theta\right) + B \quad (6)$$

où A, B, f, θ minimisent l'erreur quadratique de modélisation sur l'intervalle $\llbracket s, e \rrbracket$ pour chaque composante p .

3 Evaluation

Nous allons maintenant illustrer les performances des fonctions de coût décrites sur des données synthétiques, puis des signaux réels.

3.1 Données

Données synthétiques. En tout, mille signaux à valeurs dans \mathbb{R}^2 sont simulés. Ils durent chacun vingt secondes, échantillonnés à $F_s = 100$ Hz et ils contiennent un nombre aléatoire de régimes compris entre trois et sept. Sur chaque phase homogène, chacune des deux dimensions est modélisée soit par une constante tirée dans $[-1; 1]$, soit par une fonction sinusoïdale d'amplitude tirée dans $[0; 1]$ et de fréquence tirée dans $[0 \text{ Hz}; 5 \text{ Hz}]$. On y ajoute un bruit blanc gaussien de variance aléatoire, constante sur chaque phase et comprise dans $[0; 0, 1]$. Ces signaux de synthèse peuvent être vus comme une extension du modèle couramment testé (changement dans la moyenne et

la variance), en introduisant en plus des phases sinusoïdales. Il est à noter que l'on retrouve cette alternance de phases périodiques bruitées et de phases constantes bruitées dans des séries temporelles réelles, comme celle présentée sur la figure 1. En effet, la phase marchée se caractérise par une quasi périodicité (la répétition d'un motif de pas) tandis que le demi-tour et le repos sont plus proches d'une constante bruitée.

Données réelles. Comme décrit dans la section 1, les données sont des signaux d'accélérométrie/gyrométrie enregistrés durant des périodes de marche. Plus précisément, l'accélération verticale et la vitesse de rotation autour de l'axe vertical d'un capteur placé sur le bassin d'un sujet seront utilisées. En tout, 311 enregistrements du même exercice ont été faits. Les débuts et fin de phases ont été annotés manuellement par des médecins pour pouvoir évaluer les performances de notre algorithme.

Sur la figure 1 est représenté un exemple de segmentation de nos signaux. Au premier abord, les deux dimensions du signal sont hétérogènes entre elles. Par conséquent, les ruptures s'y manifestent différemment (une plus grande vitesse de rotation mais une accélération plus faible). De plus, elles sont non stationnaires et difficiles à modéliser : même la relative périodicité durant la marche évolue au cours du temps. En outre, les capteurs introduisent du bruit et des points aberrants (voir l'accélération sur la figure 1) dus à des petits chocs. Les exercices étant relativement libres, les phases ont des durées très disparates et sont parfois polluées par des mouvements des patients (comme des petits pas pendant le repos).

3.2 Métriques d'évaluation

Pour comparer la véritable segmentation et la segmentation calculée, trois métriques seront utilisées. Considérons donc deux partitions m et m' appartenant à $\mathcal{M}_K(1, N)$, définissant chacune K segments, $m = \{\llbracket s_1, e_1 \rrbracket, \dots, \llbracket s_K, e_K \rrbracket\}$ et $m' = \{\llbracket s'_1, e'_1 \rrbracket, \dots, \llbracket s'_K, e'_K \rrbracket\}$.

Distance de Hamming [5]. La distance est égale à la proportion de points à supprimer pour que les partitions m et m' soient égales :

$$H(m, m') = 1 - \frac{1}{N} \sum_{k=1}^K \text{card}\{\llbracket s_k, e_k \rrbracket \cap \llbracket s'_k, e'_k \rrbracket\} \quad (7)$$

Notons que $H(m, m') \in [0; 1]$ et vaut zéro ssi $m = m'$.

Métrique spectrale [6]. Notons A_m la matrice $N \times N$ de similarité de m , définie comme suit :

$$A_m(i, j) = \begin{cases} \frac{1}{e_k - s_k + 1} & \text{s'il existe } \llbracket s_k, e_k \rrbracket \in m \text{ tel} \\ & \text{que } i, j \in \llbracket s_k, e_k \rrbracket \\ 0 & \text{sinon} \end{cases} \quad (8)$$

La métrique spectrale entre m et m' vaut alors :

$$J(m, m') = \|A_m - A_{m'}\|_F^2 / (2K - 2) \quad (9)$$

où $\|\cdot\|_F$ désigne la norme de Frobenius. Notons que $J(m, m') \in [0; 1]$ et vaut zéro ssi $m = m'$. Contrairement à (7), J pénalise plus une erreur faite sur un segment de petite taille.

Indice de Rand ajusté [6]. C'est une mesure populaire pour

Synths	I-CM	RKHS-CM	F-CM	I-CD	DF-C
Hamm	0,25 (0,27)	0,14 (0,23)	0,17 (0,23)	0,09 (0,18)	0,37 (0,23)
Met Spec	0,15 (0,09)	0,10 (0,08)	0,15 (0,06)	0,09 (0,07)	0,23 (0,07)
Rand	0,83 (0,20)	0,91 (0,14)	0,91 (0,08)	0,94 (0,13)	0,60 (0,19)

Réels	I-CM	RKHS-CM	F-CM	I-CD	DF-C
Hamm	0,66 (0,06)	0,15 (0,16)	0,19 (0,17)	0,32 (0,22)	0,41 (0,16)
Met Spec	0,25 (0,02)	0,17 (0,04)	0,18 (0,03)	0,21 (0,03)	0,26 (0,02)
Rand	0,59 (0,06)	0,84 (0,09)	0,84 (0,06)	0,73 (0,09)	0,47 (0,11)

TABLE 1 – Evaluation sur les signaux synthétiques (haut) et sur les signaux réels (bas) sous la forme moyenne (écart-type). La distance de Hamming (Hamm) et la métrique spectrale (Met Spec) doivent être les plus proches de 0 possible et l’indice de Rand ajusté (Rand) doit être le plus proche possible de 1.

valider des algorithmes de *clustering*. L’indice de Rand simple vaut :

$$RI(m, m') = \sum_{k,l} \binom{\delta_{m,m'}(k, l)}{2} / \binom{N}{2} \quad (10)$$

où $\delta_{m,m'}(k, l) = \text{card}\{\llbracket s_k, e_k \rrbracket \cap \llbracket s'_l, e'_l \rrbracket\}$. L’indice de Rand ajusté (noté $ARI(m, m')$) est une version normalisée, telle que $ARI(m, m') = 1$ ssi $m = m'$ et $ARI(m, m') \approx 0$ en moyenne si m' est tirée aléatoirement dans l’ensemble des partitions de $\llbracket 1, N \rrbracket$. Contrairement aux métriques précédentes, on cherche donc une valeur ARI la plus proche possible de 1.

4 Résultats

Les moyennes et écart-types des métriques obtenues sur les deux bases avec les différentes méthodes sont présentées sur la table 1.

Données synthétiques. La méthode I-CD donne les meilleurs résultats avec les trois métriques d’évaluation, ce qui est consistant avec le modèle choisi de génération des signaux. Le test de Friedman (qui permet de comparer des résultats expérimentaux) indique que, du point de vue de la distance de Hamming et de la métrique spectrale, elle donne des résultats statistiquement différents des autres méthodes. La technique RKHS-CM donne également des résultats très proches de la véritable segmentation, au point que le test de Friedman indique que du point de vue de l’indice de Rand ajusté, elle n’est pas statistiquement différente de I-CD sur ces signaux. I-CM place correctement les ruptures lorsqu’il y a une rupture dans la moyenne mais, comme attendu, n’est pas sensible à des changements de périodicité ou d’amplitude. Ainsi, selon les signaux considérés, ses performances varient beaucoup : ceci explique pourquoi I-CM a l’écart type le plus grand de toutes les méthodes, pour toutes les métriques. Concernant les phases sinusoïdales, il arrive qu’une sinusoïde à basse fréquence soit interprétée par RKHS-CM et I-CM comme une suite de constantes. Ainsi, ces méthodes détectent parfois un segment dans la partie haute de la sinusoïde et un autre dans la partie basse, et donc une rupture entre les deux. A l’opposé, I-CD est par construction optimale sur les phases sinusoïdales, mais l’on observe qu’elle détecte aussi très bien les phases constantes, car un changement de moyenne rajoute une composante continue qui fait grandement augmenter le coût du segment. Le haut niveau de bruit des signaux synthétiques rajoute dans le spectre de Fourier de l’énergie uniformément, ce qui nuit à la méthode F-CM. Quant à la méthode DF-C, la différenciation discrète est très instable

et tend même à augmenter le niveau de bruit. Cela explique pourquoi elle est moins bonne que toutes les autres techniques.

Données réelles. La méthode RKHS-CM est ici plus performante que toutes les autres pour les trois métriques d’évaluation. Il est intéressant de noter que ses résultats baissent assez peu comparés à ceux obtenus sur les signaux synthétiques. C’est aussi le cas pour F-CM qui reste assez constante. Le test de Friedman n’indique d’ailleurs pas de différence statistique entre les résultats des deux méthodes, pour toutes les métriques. En revanche, I-CD a des résultats nettement moins bons que sur les signaux simulés, ce qui tend à montrer que le modèle est ici trop contraignant pour s’adapter à des signaux réels. I-CD souffre en réalité du fait que pendant la marche, les motifs qui se répètent ne sont pas parfaitement sinusoïdaux ni périodiques. F-CM est moins contraignant sur ces deux aspects : il suffit qu’il y ait une pseudo-périodicité pour qu’on ait un pic d’énergie dans le domaine de Fourier. La méthode DF-C donne des résultats très moyens (elle ne surpasse I-CM que du point de vue de la distance de Hamming). La raison principale est la périodicité qui n’est pas parfaite même dans les phases de marche. Les différences finies ne suppriment pas une saisonnalité qui évolue, même lentement.

La figure 1 illustre bien les avantages de F-CM. Elle est robuste aux points aberrants et les mouvements parasites, en particulier les sursauts d’accélération pendant le repos. Les ruptures les moins bien placées sont très souvent celles entre une période de repos et de marche : le premier ou dernier pas a une amplitude faible et est confondu avec le bruit de la phase de repos. Les ruptures marquées, comme celle du demi-tour, sont détectées avec précision, malgré la brièveté de la phase (une seconde en moyenne). Les phases de marche isolées par F-CM sont extrêmement homogènes du point de vue fréquentiel. En effet, la marche est constituée d’une période transitoire (quelques pas d’appoint) puis une phase de croisière plus régulière. F-CM est particulièrement efficace pour détecter cette dernière. Contrairement à RKHS-CM, qui place une rupture dès qu’un mouvement est décelé, F-CM néglige automatiquement la phase transitoire.

En conclusion, nous avons présenté plusieurs fonctions de coût qui permettent de segmenter un signal physiologique multivarié. Deux méthodes, RKHS-CM et F-CM, se détachent. Les partitions associées à chacune d’entre elles ont un sens clinique légèrement différent mais également intéressant.

Références

- [1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes : Theory and Application*. Prentice Hall Englewood Cliffs, 1993, vol. 104.
- [2] Z. Harchaoui and O. Cappe, “Retrospective Multiple Change-Point Estimation with Kernels,” in *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Workshop on Statistical Signal Processing*, Madison, Wisconsin, USA, 2007, pp. 768–772.
- [3] F. Desobry, M. Davy, and C. Doncarli, “An online kernel change detection algorithm,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2961–2974, 2005.
- [4] D. Percival and a. T. Walden, *Spectral Analysis for Physical Applications*. 1st ed. Cambridge University Press, 1993.
- [5] R. W. Hamming, “Error detecting and error correcting codes,” *Bell System technical journal*, vol. 29, no. 2, pp. 147–160, 1950.
- [6] L. Hubert and P. Arabie, “Comparing partitions,” *Journal of classification*, vol. 2, no. 1, pp. 193–218, 1985.