Sampling Expansions for AFB Signals via Change Point Detection

Michael Baron
Department of Mathematics and Statistics,
American University, Washington, DC 20016 USA
Email: *baron@american.edu*

Stephen D. Casey
Department of Mathematics and Statistics,
American University, Washington, DC 20016 USA
Email: *scasey@american.edu*

***Abstract*—This article outlines sampling expansions designed specifically for adaptive frequency band (AFB) signals. Optimal sampling of AFB requires changing sampling rates. The proposed expansion, the *Projection Method* gives a more computationally efficient way to sample, transmit, and then reconstruct AFB signals. The entire system is developed with flexible parameters, perfect for AFB. To detect changing bandwidth, we use statistical methods of *change-point detection*. We propose a two-component system – using techniques of statistical change point detection as a front end to detect changing signal bandwidth, and then using the flexibility built into the Projection system to adjust to the changing bandwidth.**

I. Introduction

One of the biggest challenges in wireless communication is energy use. With more and more expectations placed upon these systems and the effect of increased capability of microchips, we run into the problem of how to efficiently power these systems. Our "Age of Information" presents us with new challenges, and, in particular, with respect to sampling. We need to develop systems that can communicate information more efficiently in terms of energy use. Digital circuitry has provided dramatically enhanced digital signal processing operation speeds, but there has not been a corresponding dramatic energy capacity increase in batteries to operate these circuits; there is no Moore's Law for batteries. Sampling theory provides insight on how to best process signals for these new systems, and remains as timely as it ever was.

This article outlines sampling expansions designed specifically for adaptive frequency band (AFB) signals. We call this procedure the *Projection Method*. It gives a more computationally efficient way to sample, transmit, and then reconstruct AFB signals, which require rapidly changing sampling rates. The entire system is developed with flexible parameters, perfect for AFB. But when do you change? We use the statistical *Change-Point Detection* methodology to determine when there has been a change in bandwidth. We propose a two-component system, using techniques of statistical change-point detection as a front end to detect changing signal bandwidth, and then using the flexibility already built into the systems to then adjust to the changing bandwidth.

II. Change Point Detection

We first construct a *sequential change-point detection algorithm* that signals an abrupt change in the range of observed frequencies. We focus on such changes that cause an expansion of the original bandwidth. That is, based on sequential, real-time observation of a signal, we are interested in detecting the moments when its frequencies start exceeding the upper end of the original bandwidth, its lower end, or both ends. This represents three types of *change-point detection*, and we show how the proposed detection algorithm can be tuned to each of these scenarios. Furthermore, it is a change-point detection problem with *nuisance parameters* because besides the time of the change, the original pre-change bandwidth, the new post-change bandwidth, and actually, the signal itself, remain unknown to the observer.

First, we notice that expansion of the original range of frequencies inevitably causes occurrence of record frequencies, the highest or the lowest observed to the moment, more often than it would be expected without change-points. Indeed, only unexpectedly many record frequencies signal an expansion of a bandwidth. In other words, if the distribution of very high or very low frequencies appears consistent with the signal that has no change-points, there would be no reason to decide that a change has ever occurred in the signal.

Next, we note that in a sequence of independent and identically distributed random variables (with a continuous and unbounded support), the n-th observation is a high record, the highest value observed so far, with probability $p_n = 1/n$. That is because all permutations of the first n observations are equally likely, and therefore, each observation can appear as record value with the same probability ($p_n = 1/n$ for the detection of a lower-end bandwidth expansion and record low frequencies, and $p_n = 2/n$ for the detection of a bandwidth expansion in any or both directions).

Although we do not claim that frequencies observed in a signal have to be independent and identically distributed, we have no prior information on when the record frequencies may occur, and $p = 1/n$ ($p = 2/n$ for a two-sided bandwidth expansion) provides a benchmark that the observed distribution of record frequencies can be compared with. Essentially, the proposed change-point detection tools score the observed occurrences of recent records against the benchmark and signal a change when records are observed too often, compared with the completely random sequence of frequencies.

Change-point detection has been studied in statistical science literature for at least seven decades. Detailed overviews of the developed methodologies can be found in [1], [4], [8], [18], [21]. In particular, for a non-Bayesian (without a prior distribution) sequential (in real time) change-point detection, the *cumulative-sum*, or *CUSUM* detection algorithms, introduced by [17], proved to be optimal in several ways [13], [15], [19].

Unlike these classical works, our change-point problem is *nonparametric*, because we make no assumptions about the signal and its pre-change and post-change behavior. Thus, we need a nonparametric solution such as the one proposed in [2], which we adapt here to the distribution of record frequencies.

From now on, we focus on one-sided high-frequency detection of bandwidth expansion. Adaptation to the other two change-point detection problems is straightforward, tracing the suitable types of records and using the benchmark probability $1/n$ or $2/n$ for one-tail or two-tail changes.

Suppose that the distribution, the spectral distribution in our case, changes from F to G after ν samplings, so that the frequencies $x_1, \ldots, x_\nu$ come from distribution F, and $x_{\nu+1}, \ldots$ from G. Given a change at ν, the probability $p_{k\nu}$ of the k-th sampling breaking a record is

$$p_{k\nu} = \left\{ \begin{array}{lll} 1/k & \text{for} & k \leq \nu, \\ \int_{-\infty}^{\infty} F^{\nu}(x) G^{k-\nu-1}(x) g(x) d\mu(x) & \text{for} & k > \nu, \end{array} \right. \tag{1}$$

where g is the density of G with respect to a reference measure μ. The integral in (1) represents the probability for the currently observed frequency to exceed all the pre- and post-change frequencies observed earlier. When $F \equiv G$ meaning no change, this integral reduces to $1/k$.

Following the general approach of [2], we use the histogram density estimation to estimate the distributions F and G for any potential change-point k and substitute the resulting estimates into the CUSUM process Λ_n. Then, the algorithm signals a change (of the intended type) when the CUSUM process exceeds a detection threshold. For such density estimation, partition the bandwidth into bins or bandwidths and count N_{km}, the number of observed frequencies in the m-th bandwidth among the first k samplings. Given a change-point at ν, the distribution functions $F(x)$ and $G(x)$ for x in the m-th bandwidth are then estimated from the first k observations as $\widehat{F}(x) = \nu^{-1} \sum_{m' \leq m} N_{\nu m'}$ and $\widehat{G}(x) = (k-\nu)^{-1} \sum_{m' \leq m} (N_{km'} - N_{\nu m'})$. Using these estimates in (1), with μ being the counting measure, yields a nonparametrically estimated probability of a record value x_k,

$$\widehat{p}_{\nu k} = \sum_{m} \left(\frac{\sum_{m' \leq m} N_{\nu m'}}{\nu} \right)^{\nu} \left(\frac{\sum_{m' \leq m} (N_{km'} - N_{\nu m'})}{k - \nu} \right)^{k-\nu-1} \times \left(\frac{N_{km} - N_{\nu m}}{k - \nu} \right). \tag{2}$$

Finally, we apply probabilities (2) to construct the CUSUM process Λ_n, that is based on the occurrence of record frequencies, $y_k = \chi\{k\text{-th value is a record}\}$. Being indicators, the sequence y_k is Bernoulli, with the parameter changing according to probabilities (1) that are estimated by (2). The resulting CUSUM process is

$$\begin{aligned} \Lambda_n &= \max_{k \leq n} \sum_{j=k+1}^{n} \log \frac{\widehat{p}_{kj}^{y_j} (1 - \widehat{p}_{kj})^{1-y_j}}{(1/j)^{y_j} (1 - 1/j)^{1-y_j}} \\ &= \max_{k \leq n} \sum_{j=k+1}^{n} \left\{ y_j \log(j \widehat{p}_{kj}) + (1 - y_j) \log \frac{1 - \widehat{p}_{kj}}{1 - 1/j} \right\}. \end{aligned} \tag{3}$$

We can see in (3) how the CUSUM statistic Λ_n compares the estimated probabilities of records $\widehat{p}_{kj}$ with the corresponding benchmark probabilities $(1/j)$, agreeing with our intuition. It will signal detection of a change-point when Λ_n exceeds a threshold, defining the stopping time

$$\tau = \inf\{n \,:\, \Lambda_n \geq h\}.$$

This threshold h is determined based on the desired condition on the rate of false alarms or the detection speed, similarly to the probabilities of Type I and Type II errors in statistical hypothesis testing. With *known* distributions, and consequently, known probabilities of records, the CUSUM process Λ_n has a negative drift and most values near zero before the change, and it has a positive drift and an increasing trend after the change. The post-change expected value of its increments is the Kullback-Leibler information distance $K_{G,F}$, in our case, between the post- and pre-change probabilities of record values,

$$K_{G,F}(\nu, n) = p_{\nu n} \log(n p_{\nu n}) + (1 - p_{\nu n}) \log\{(1 - p_{\nu n}) \tfrac{n}{n-1}\}.$$

Then, in order to detect a change after an expected number of k post-change samplings and n samplings overall, the threshold h should be set at

$$h = \textstyle\sum_{j=1}^{k} K_{G,F}(n - k, n - k + j). \tag{4}$$

Unknown probabilities p_{0k} used in (4) would be replaced by the probabilities computed given the minimal change in the frequency bandwidth that we are determined to detect.

As proven in [2], the nonparametric CUSUM process estimator in (3) has an asymptotic behavior similar to the process with known pre- and post-change distributions, and thus, threshold (4) is still valid for our proposed procedure. Nevertheless, incorporating any additional information should improve detection sensitivity over the completely nonparametric scheme.

As an example, suppose that the frequencies occurring in the signal are uniformly distributed over a certain bandwidth $[a, b]$, representing a white noise within this range. After a change, the interval expands by a factor of $r > 1$ to $[a, c]$, extending the range of high frequencies and causing more frequent records. The change-point detection problem becomes *parametric*. The post-change probabilities of records will then be equal

$$\begin{aligned} p_{\nu k} &= \int_a^b \left(\tfrac{x-a}{b-a}\right)^{\nu} \left(\tfrac{x-a}{c-a}\right)^{k-\nu-1} \tfrac{dx}{c-a} + \int_b^c \left(\tfrac{x-a}{c-a}\right)^{k-\nu-1} \tfrac{dx}{c-a} \\ &= \tfrac{1}{k} r^k + \tfrac{1}{k-\nu} \left(1 - r^{k-\nu}\right). \end{aligned} \tag{5}$$

Based on n samplings, parametric estimators of a, b, and c for each potential change-point k are $\widehat{a} = \min_{j \leq n} x_j$,

$\widehat{b} = \max_{j\leq k} x_j$, and $\widehat{c} = \max_{j\leq n} x_j$, and the CUSUM process is constructed as in (3), only using probabilities (5) with estimated parameters $\widehat{a}$, $\widehat{b}$, and $\widehat{c}$.

In conclusion, we note that more than one change-point may occur over an observation of a signal. The algorithm for detecting multiple change-points is detailed, for example, in [3]. After detecting the first change-point, we adapt the bandwidth accordingly and continue sampling from the signal. Further samplings allow us to refine the estimator of a change-point location. Then we drop observations that occurred before the change-point and use the post-change samplings to detect the next change-point, and so on.

III. Adaptive Sampling via Projection

The *sequential change-point detection algorithm* can provide a near real-time method to determine a change in bandwidth. To engineer a complete system, a procedure of signal splitting and a delay in one signal would be needed. Given that, how can we design a signal sampling method to adjust to the change in bandwidth? We propose a system that first windows the signal in the analog domain, and then uses the Malvar-Coifman-Meyer basis folding system to construct the system with ON basis elements tailored to the signal (following the seminal work of Malvar [14] and Coifman and Meyer [9]). We refer to this as the *Projection Method*. The windows decompose the signal into a basis via a continuous-time inner product operation, computing the basis coefficients in parallel, preserving orthogonality of any orthonormal system between adjacent blocks. Moreover, the windows have *variable* partitioning length, roll-off and smoothness, and are designed to preserve orthogonality of any orthonormal system between adjacent blocks. We use these to create a basis system, computing the basis coefficients in parallel. We show how one can use statistical change point detection to detect changing bandwidth, and how changing one parameter – window size – can change the system to optimally adapt to the changing bandwidth. Moreover, the systems are designed with flexible parameters that can be tailored to optimize the efficiency of the processing of a given signal. We develop a system of time-domain windows, separating a function into parts $f_k = \mathbb{W}_k \cdot f$, which maintain orthogonality between signal blocks.

Definition 1 (ON Window System): An **ON Window System** is a set of functions $\{\mathbb{W}_k(t)\}$ such that for all $k \in \mathbb{Z}$,

$$\begin{aligned} &(i.) \quad \operatorname{supp}(\mathbb{W}_k(t)) \subseteq [kT - r, (k+1)T + r]\,, \\ &(ii.) \quad \mathbb{W}_k(t) \equiv 1 \text{ for } t \in [kT + r, (k+1)T - r]\,, \\ &(iii.) \quad \mathbb{W}_k \text{ is symmetric about its midpoint}\,, \\ &(iv.) \quad \sum [\mathbb{W}_k(t)]^2 \equiv 1\,, \\ &(v.) \quad \{\widehat{\mathbb{W}_k{}^{\circ}}[n]\} \in l^1\,. \end{aligned} \tag{6}$$

Conditions $(i.)$ and $(ii.)$ are partition properties, in that they give an exact snapshot of the input function f on $[kT + r, (k+1)T - r]$ with smooth roll-off at the edges. Conditions $(iii.)$ and $(iv.)$ are needed to preserve orthogonality between adjacent blocks. Condition $(v.)$ for the periodization $\bullet^{\circ}$ is needed for the computation of Fourier coefficients, and expresses a certain smoothness of the window. Indeed, let $I = T + 2r$ and let $\mathbb{PW}_{\Omega}$ denote the Paley-Wiener space for bandlimit Ω. Let $f \in \mathbb{PW}_{\Omega}$ and let $\{\mathbb{W}_k(t)\}$ be an ON window system with generating window $\mathbb{W}_I$. Then $\frac{1}{I}\int_{-T/2-r}^{T/2-r} [f \cdot \mathbb{W}_I]^{\circ}(t) \exp(-2\pi i n t/[I])\, dt = \widehat{f} * \widehat{\mathbb{W}_I}[n]\,.$

We generate our systems by translations and dilations of a given window $\mathbb{W}_I$, where $\operatorname{supp}(\mathbb{W}_I) = [-\frac{T}{2} - r, \frac{T}{2} + r]$. Our general window function $\mathbb{W}_I$ is m-times differentiable, has $\operatorname{supp}(\mathbb{W}_I) = [-\frac{T}{2} - r, \frac{T}{2} + r]$, and has values

$$\mathbb{W}_I = \begin{cases} 0 & |t| \geq T/2 + r\,, \\ 1 & |t| \leq T/2 - r\,, \\ \rho(\pm t) & T/2 - r < |t| < T/2 + r\,. \end{cases} \tag{7}$$

We solve for $\rho(t)$ by solving the Hermite interpolation problem

$$\begin{cases} (a.) & \rho(T/2 - r) = 1\,, \\ (b.) & \rho^{(n)}(T/2 - r) = 0\,,\ n = 1, 2, \ldots, m\,, \\ (c.) & \rho^{(n)}(T/2 + r) = 0\,,\ n = 0, 1, 2, \ldots, m\,. \end{cases}$$

with the conditions that $\rho \in C^m$ and

$$[\rho(t)]^2 + [\rho(-t)]^2 = 1 \text{ for } t \in [\pm(\tfrac{T}{2} - r), \pm(\tfrac{T}{2} + r)]\,. \tag{8}$$

The C^m solution for ρ is given by a theorem of Schoenberg (see [20], pp. 7-8). The spline $S(t)$ for the Hermite problem with endpoints -1 and 1 such that $S(1) = 1$, $S^{(n)}(1) = 0$, $n = 1, 2, \ldots, m$, and $S^{(n)}(-1) = 0$, $n = 0, 1, \ldots, m$, is given by the integral of the function $M(t) = (-1)^m \sum_{j=0}^{m} \frac{\Psi(t - t_j)}{\phi'(t_j)}$, where Ψ is the $m+1$ convolution of characteristic functions, the knot points are $t_j = -\cos(\frac{\pi j}{m})$, $j = 0, 1, \ldots, m$, and $\phi(t) = \prod_{j=0}^{k}(t - t_j)$. Given these knots, we have to choose α to fit the knot points. If m is even, the midpoint occurs at the $m/2$ knot point. If m is odd, the midpoint occurs at the midpoint between the $m/2$ and $(m+1)/2$ knot points. Let $\xi = l(t) = \frac{r}{2}(t - 1)$, and let $\alpha(\xi) = S \circ l(\pm\xi)\,, |\xi| \leq r$. Let $A = \int_{-r}^{r} \alpha(\zeta)\, d\zeta$. Now, normalize α by letting $\beta(\xi) = \frac{\pi}{2A}\alpha(\xi)\,,$ and let $\Theta(\tau) = \int_{-r}^{\tau} \beta(\xi)\, d\xi\,, |\tau| \leq r$. Define $\rho_{\text{up}}(\tau) = \sin(\Theta(\tau))\,, \rho_{\text{down}}(\tau) = \cos(\Theta(\tau))$. We define our C^m window $\mathbb{W}_I(t) = \mathbb{ON}_{C^m}(t)$ as follows:

$$\begin{cases} 0 & |t| \geq T/2 + r\,, \\ 1 & |t| \leq T/2 - r\,, \\ \rho_{\text{up}}(t + (T/2 + r)) & -T/2 - r < t < -T/2 + r\,, \\ \rho_{\text{down}}(t - (T/2 - r)) & T/2 - r < t < T/2 + r\,. \end{cases} \tag{9}$$

We translate the window as needed. The resultant windowing system has variable partitioning length, variable roll-off, and variable smoothness. With each degree of smoothness, we get an additional degree of decay in frequency.

We designed the ON windows $\{\mathbb{W}_k(t)\}$ so that they preserve orthogonality of basis elements of overlapping blocks, using the techniques of Malvar, Meyer, and Coiffman. Because of the partition properties of these systems, we need only check the orthogonality of adjacent overlapping blocks. We can show the following ([7]):

Theorem 1: Let φ_j be an orthononormal (ON) basis for $L^2(\mathbb{R})$. Then $\{\Psi_{k,j}\} = \{\mathbb{W}_k \widetilde{\varphi_j}\}$ is also an ON basis.

Given characteristics of the class of input signals, the choice of basis functions used can be tailored to optimal representation of the signal or a desired characteristic in the signal. A direct consequence is ([7]):

Theorem 2: Let $\{\mathbb{W}_k(t)\}$ be ON windows, and let $\{\Psi_{k,n}\} = \{\mathbb{W}_k \widetilde{\varphi_n}\}$ be an ON basis that preserves orthogonality between adjacent windows. Let $f \in \mathbb{PW}_\Omega$ and $N = N(T, \Omega)$ be such that $\langle f, \Psi_{k,n} \rangle = 0$ for all $n > N$ and all k. Then, $f(t) \approx f_{\mathcal{P}}(t)$, where

$$f_{\mathcal{P}}(t) = \sum_{k \in \mathbb{Z}} \left[\sum_{n=-N}^{N} \langle f, \Psi_{k,n} \rangle \Psi_{k,n}(t) \right]. \tag{10}$$

Using the original Ω band-limit gives us a lower bound on the number of non-zero Fourier coefficients $\mathbb{W}_k \widetilde{\varphi_n}$ as follows. We have $\frac{n}{T} \leq \Omega$, i.e. , $n \leq T \cdot \Omega$. So, choose $N = \lceil T \cdot \Omega \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function. For this fixed value of N, if bandwidth increases or decreases, simply adjust the width T of the time window.

We finish by discussing an ON basis that is tailored to work extremely well is the AFB system – the Modified Gegenbauer basis. The Gegenbauer polynomials are the symmetric specialization of the Jacobi polynomials ([16, Chapter 18]). They are used in a AFB communication system to construct pulses with narrow widths. The Gegenbauer waveform is used to modulate data, and has demonstrated superior performance to classic waveforms, e.g., Gaussian waveforms and the Hermite systems. We develop a modified Gegenbauer system, and use it to construct a windowed basis system for AFB signals with exponentially small error on each block.

We define an ON basis for $L^2[-\frac{T}{2}, \frac{T}{2}]$ using modified Gegenbauer functions, constructed from Gegenbauer polynomials. The Gegenbauer polynomials are modified so that they zero out at the endpoints and normalized to create an ON system. This then allows AFB signals to be expanded in the folding method using the modified Gegenbauer system.

The Gegenbauer polynomials $C_n^\nu : \mathbb{C} \to \mathbb{C}$ are orthogonal over $(-1, 1)$ with orthogonality relation given by [16, Table 18.3.1] $\int_{-1}^{1} C_n^\nu(x) C_m^\nu(x) w(x;\nu) dx = h_n^\nu \delta_{n,m}$, for $\nu \in (-\frac{1}{2}, \infty) \setminus \{0\}$, where $w(x;\nu) := (1-x^2)^{\nu - 1/2}$, $h_n^\nu := \frac{2^{1-2\nu}\pi\Gamma(2\nu+n)}{(\nu+n)\Gamma^2(\nu)n!}$. Gottlieb and Shu [11] give a detailed analytic argument showing how Gegenbauers minimize the Gibbs phenomenon. They have two parameters – truncation q_T and regularization q_R – to minimize the Gibbs jump. We let $q = \max\{q_T, q_R\}$, and note that $q < 1$.

The modified Gegenbauer function $\mathcal{C}_n^\nu : [-\frac{T}{2}, \frac{T}{2}] \times (0, \infty) \to \mathbb{R}$ are defined by

$$\mathcal{C}_n^\nu(t;T) := \sqrt{\frac{2\, w\left(\frac{2t}{T};\nu\right)}{T h_n^\nu}} C_n^\nu\left(\frac{2t}{T}\right). \tag{11}$$

These functions form an ON basis for $L^2[-\frac{T}{2}, \frac{T}{2}]$ with $\nu \in (\frac{1}{2}, \infty)$, namely $\int_{-T/2}^{T/2} \mathcal{C}_n^\nu(t;T)\mathcal{C}_m^\nu(t;T)dt = \delta_{m,n}$. Note that we exclude the parameters $\nu \in (-\frac{1}{2}, \frac{1}{2}]$ in order to keep the endpoints $\pm\frac{T}{2}$ in the domain of integration. By using $w(x;\nu)$ and h_n^ν one has

$$\mathcal{C}_n^\nu(t;T) = \frac{2^{2\nu-1/2}\Gamma(\nu)}{T^\nu} \sqrt{\frac{(n+\nu)n!}{\pi\Gamma(2\nu+n)}} \times \left(\left(\tfrac{T}{2}\right)^2 - t^2 \right)^{\nu/2-1/4} C_n^\nu\left(\frac{2t}{T}\right). \tag{12}$$

The modified Gegenbauer system zeros out at the endpoints, which allows us to use it to create the windowed ON basis $\{\Psi_{k,n}\} = \{\mathbb{W}_k\, \mathcal{C}_n^\nu(t;T)^\curlywedge\}$, where we window with $\mathbb{ON}_{C^m}$, and define the *folded basis elements* $\mathcal{C}_n^\nu(t;T)^\curlywedge$ by

$$\begin{array}{cl} 0 & |t| \geq T/2 + r \\ \mathcal{C}_n^\nu(t;T) & |t| \leq T/2 \\ -\mathcal{C}_n^\nu(-T-t;T) & -T/2 - r < t < -T/2 \\ \mathcal{C}_n^\nu(T-t;T) & T/2 < t < T/2 + r\,. \end{array} \tag{13}$$

We close by computing the error in each window $\mathcal{E}_{k_{\mathcal{P}}}$ in terms of the modified Gegenbauer system. This system minimizes the Gibbs phenomenon, giving the point values of a piecewise smooth signal with essentially the same accuracy as a smooth approximation. Moreover, in [11], we also get a decay parameter $q < 1$. Let $\sigma \in \mathbb{N}$ be the smoothness parameter, and assume $\mathbb{W}_k$ is C^σ, and so $\widehat{\mathbb{W}_k}(\omega) = \mathcal{O}(1/(\omega)^{\sigma+2})$. Now approximate the signal f with the windowed ON basis $\{\Psi_{k,n}\} = \{\mathbb{W}_k \mathcal{C}_n^\nu(t;T)^\curlywedge\}$, where we window with $\mathbb{ON}_{C^\sigma}$. Let $q < 1$ be the decay parameter. Then, the error $\mathcal{E}_{k_{\mathcal{P}}}$ on a given block is

$$\sup\left| (f(t)\cdot\mathbb{W}_k) - \left[\sum_{n=-N}^{N} \langle f, \Psi_{n,k} \rangle \Psi_{n,k}(t) \right] \mathbb{W}_k(t) \right| \tag{14}$$

$$\leq \sup\left[\sum_{|n|>N} \left| \langle f, \Psi_{n,k} \rangle \Psi_{n,k}(t) \right| \right] \mathbb{W}_k(t) \leq \sum_{|n|>N} \frac{e^{\log(q)N}}{n^{\sigma+2}}\,.$$

Since $q < 1$, $e^{\log(q)N}$ decays exponentially as N increases.

Acknowledgments: Second author's research was partially supported by U. S. Army Research Office Scientific Services program, administered by Battelle (TCN 06150, Contract DAAD19-02-D-0001) and U. S. Air Force Office of Scientific Research Grant Number FA9550-12-1-0430.

References

[1] M. Ahsanullah, A. L. Rukhin, and B. Sinha. *Applied Change Point Problems in Statistics*. Nova Science Publishers, New York, 1995.

[2] M. Baron. Nonparametric adaptive change-point estimation and on-line detection. *Sequential Analysis*, 19(12):1–23, 2000.

[3] M. Baron. Sequential methods for multistate processes. In N. Mukhopadhyay, S. Datta, and S. Chattopadhyay, editors, *Applied Sequential Methodologies*, pages 55–73. Dekker, New York, 2004.

[4] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. PTR Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.

[5] J. J. Benedetto, *Harmonic Analysis and Applications*, CRC Press, Boca Raton, FL (1997).

[6] L. Borup and M. Neilsen, "Frame Decomposition of Decomposition Spaces" *Jour. Fourier Anal. Appl.* **13** (1), 39–70 (2007).

[7] S. D. Casey, "Adaptive Signal Processing," Chapter 11 in *Excursions in Harmonic Analysis, Volume 4*, Springer-Birkhäuser, New York, 261–290 (2015).

[8] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*. Birkhäuser, Boston, MA, 2012.

[9] R. Coifman and Y. Meyer, "Remarques sur l'analyse de Fourier a fenetre," *CR Acad. Sci. Paris* **312**, 259–261 (1991).

[10] H. G. Feichtinger, T. Strohmer *Gabor Analysis and Algorithms - Theory and Applications* Birkhäuser Boston (1998).

[11] D. Gottlieb and C-W. Shu, "On the Gibbs phenomenon and its resolution," *SIAM Rev.* **39** (4), 644–668 (1997).

[12] B. Jawerth and W. Sweldens, "Biorthogonal smooth local trigonometric bases," *Jour. Fourier Anal. Appl.*, **2** (2), 109–133, (1995).

[13] G. Lorden. Procedures for reacting to a change in distribution. *Ann. Math. Statist.*, 42:1897–1908, 1971.

[14] H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Process.*, **46** (4) 1043-1053 (1998).

[15] G. V. Moustakides. Optimal stopping times for detecting a change in distribution. *Ann. Statist.*, 14:1379–1388, 1986.

[16] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST handbook of mathematical functions*, Cambridge University Press, Cambridge (2010).

[17] E. S. Page. Continuous inspection schemes. *Biomterika*, 41:100–115, 1954.

[18] H. V. Poor and O. Hadjiliadis. *Quickest Detection*. Cambridge University Press, Cambridge, UK, 2009.

[19] Y. Ritov. Decision theoretic optimality of CUSUM procedure. *Ann. Statist.*, 18:1464–1469, 1990.

[20] I. J. Schoenberg, *Cardinal Spline Interpolation* (CBMS–NSF Conference Series in Applied Mathematics, 12), SIAM, Philadelphia, PA (1973).

[21] A. G. Tartakovsky, I. V. Nikiforov, and M. Basseville. *Sequential Analysis: Hypothesis Testing and Change-Point Detection*. Chapman & Hall/CRC, 2014.