SINGLE CHANNEL SIGNAL SEPARATION USING LINEAR TIME VARYING FILTERS: SEPARABILITY OF NON-STATIONARY STOCHASTIC SIGNALS

James R. Hopgood and Peter J. W. Rayner

Signal Processing Laboratory, Department of Engineering,
University of Cambridge, England CB2 1PZ
jrh1008@eng.cam.ac.uk, pjwr@eng.cam.ac.uk

ABSTRACT

Separability of signal mixtures given only one mixture observation is defined as the identification of the accuracy to which the signals can be separated. The paper shows that when signals are separated using the generalised Wiener filter, the degree of separability can be deduced from the filter structure. To identify this structure, the processes are represented on an arbitrary spectral domain, and a sufficient solution to the Wiener filter is obtained. The filter is composed of a term independent of the signal values, corresponding to regions in the spectral domain where the desired signal components are not distorted by interfering noise components, and a term dependent on the signal correlations, corresponding to the region where components overlap. An example of determining perfect separability of modulated random signals is given.

1. INTRODUCTION

This paper investigates the *separability* of signal mixtures given, at each time instance, only **one** observation of the mixture in the temporal domain. In this context, separability refers to identifying whether the mixture of two signals can be separated to a given degree of accuracy. The problem of actually separating the signals is then referred to as *signal separation*. Formally, this paper deals with the following problem:[1]

Definition 1 (Separability). Suppose a desired signal $\mathbf{d}(t)$ is corrupted by an additive noise signal $\mathbf{n}(t)$, such that the observation, $\mathbf{x}(t)$ of the desired signal, is given by $\mathbf{x}(t) = \mathbf{d}(t)+\mathbf{n}(t)$, $\forall t \in T$. The separability problem is to determine conditions on $\mathbf{d}(t)$ and $\mathbf{n}(t)$ such that, given $\mathbf{x}(t)$, an estimate of the desired signal, $\hat{\mathbf{d}}(t)$, can be obtained to a given degree of accuracy.

This will be answered assuming that signal separation is possible using linear time-varying (LTV) filters. It is well known that stationary signals are "perfectly" separable if their Power Spectrums do not overlap in the Fourier domain. Separation can only then be achieved when the frequency bands of the Power Spectrums are known *a-priori*. Furthermore, non-stationary signals that are non-overlapping in the time domain can only be separated when the times at which the signals "are switched on" are known. Thus, although signals may be separable, signal separation can only be performed by exploiting *a-priori* known signal features, since the separation problem is inherently under-constrained; at each time instant, there are two unknowns, and one equation.

As finite bandwidth signals in the Fourier domain can be separated by a linear time invariant (LTI) (bandpass) filter, and finite duration signals can be separated by a LTV filter (a switch), this raises the question as to whether there exists some other signal domain where the representation of two classes of signals are disjoint, such that the signals can be recovered using a generalised bandpass filter; the filter may neither be as simple as a switch or conventional bandpass filter, but somewhere "in-between".

Definition 2 (Perfect Signal Separation). "Perfect Signal Separation" is achieved when the mean squared error (MSE),

$$\sigma^2(t) = E\left[\boldsymbol{\varepsilon}^2(t)\right] \stackrel{\text{def}}{=} E\left[\left|\hat{\mathbf{d}}(t) - \mathbf{d}(t)\right|^2\right] \tag{1}$$

is zero at each desired time instance; $\sigma^2(t) = 0, \forall t \in T$. Bode and Shannon [1] provide a stimulating discussion of the problems and consequences of using MSE as an error criterion, but regardless of these problems, this criterion will be taken here as the definition of separation accuracy.

Definition 3 (The Autocorrelation Function). The *autocorrelation function* (ACF) of a stochastic process at $\{t, \tau\}$ is defined as[2]

$$R_{xx}\left(t, \tau\right) = E\left[\mathbf{x}(t)\, \mathbf{x}(\tau)\right] \tag{2}$$

where $E\left[\mathbf{z}(t)\right]$ denotes expectation with respect to $\mathbf{z}(t)$.

Theorem 1 (Wiener-Hopf Filter (WHF)). *If signal separation is achieved by LTV filtering of the observed signal, then*

$$\hat{\mathbf{d}}(t) = \int_{\mathcal{T}} h(t, \alpha)\, \mathbf{x}(\alpha)\, d\alpha, \quad \forall t \in T \tag{3}$$

where $h(t, \tau)$, is the response at time t given an impulse occurred at the filter input at time τ, and is chosen such that $\sigma^2(t)$ is minimised. This leads to the WHF, $h(t, \tau)$ given by the solution of

$$R_{dx}\left(t, \tau\right) = \int_{\mathcal{T}} h(t, \alpha)\, R_{xx}\left(\alpha, \tau\right)\, d\alpha, \; \forall t \in T, \forall \tau \in \mathcal{T} \tag{4}$$

with the MSE, $\forall t \in T$, given by

$$\sigma^2(t) = R_{dd}\left(t, t\right) - \int_{\mathcal{T}} h(t, \alpha)\, R_{dx}\left(t, \alpha\right)\, d\alpha \tag{5}$$

For "perfect separation",

$$R_{dd}\left(t, t\right) = \int_{\mathcal{T}} h(t, \alpha)\, R_{dx}\left(t, \alpha\right)\, d\alpha, \quad \forall t \in T \tag{6}$$

This work was supported by the Schiff Foundation, Cambridge Univ.

[1] Note that bold symbols represent scalar stochastic processes, whereas normal type symbols represent scalar deterministic processes.

[2] for simplicity, all processes are assumed to be real and have zero mean.

It is desirable that the required prior knowledge for signal separation will be common to a "class" of process; eg., the prior knowledge required to separate stationary signals is the Fourier frequency range over which the signals have non-zero spectral components. It is not necessary to know, however, the phase and magnitude of these components to achieve separation. Thus, it is desirable that the WHF should not depend on the exact form of the ACFs, but some distinguishing feature of a "class" of stochastic signals. The problem of estimating these distinguishing features is left as further work; here, only the *form* of these features is required to determine separability.

2. GENERALISED POWER SPECTRUM

In signal theory, spectra are often associated with Fourier transforms, and this idea may be extended by using a general integral transform to represent a particular realisation of a stochastic process as the superposition or integral of some given basis functions, with stochastic coefficients. This concept is a generalisation of the Fourier transform of a stochastic process discussed in [2], and a further widely known special case is the Karhunen-Loève (KL) Transform [2].

The stochastic spectral representation of a **continuous** time stochastic process $\mathbf{x}(t)$ on an arbitrary **continuous** spectral domain λ is $\mathbf{X}(\lambda)$, defined by

$$\mathbf{X}(\lambda) = \int_T \mathbf{x}(t)\, K(t,\lambda)\, dt\ , \quad \forall \lambda \in \Lambda \tag{7}$$

where this integral is interpreted in a Mean Square (MS) limit, and Λ is the region in the signal space in which $\mathbf{X}(\lambda)$ lies. The function $K(t,\lambda)$ is called the *direct transform basis kernel*. Conversely, $\mathbf{X}(\lambda)$ may be represented on the time domain t as $\hat{\mathbf{x}}(t)$;

$$\hat{\mathbf{x}}(t) = \int_\Lambda \mathbf{X}(\lambda)\, k(\lambda,t)\, d\lambda\ , \quad \forall t \in T \tag{8}$$

where $k(\lambda,t)$ is called the *inverse transform* or *reciprocal basis kernel* of $K(t,\lambda)$. It may be shown that the representation $\hat{\mathbf{x}}(t)$ equals $\mathbf{x}(t)$ in the MS sense:

$$E\left[|\mathbf{x}(t) - \hat{\mathbf{x}}(t)|^2\right] = 0 \tag{9}$$

(The proof follows a similar line to the proof in [2], given for the stochastic Fourier series). It is assumed that the transformation is isomorphic, so for a given $\mathbf{x}(t)$, there exists a unique $\mathbf{X}(\lambda)$ and conversely, for a given $\mathbf{X}(\lambda)$ there exists a unique $\hat{\mathbf{x}}(t)$. This imposes the constraints [3]

$$\begin{aligned} \delta(t-\tau) &= \int_\Lambda k(\lambda,t)\, K(\tau,\lambda)\, d\lambda\ ,\ \forall t,\tau \in T \\ \delta(\lambda-\hat{\lambda}) &= \int_T k(\lambda,t)\, K(t,\hat{\lambda})\, dt\ ,\ \forall \lambda,\hat{\lambda} \in \Lambda \end{aligned} \tag{10}$$

where $\delta(t)$ is the Dirac delta function. Any pair of functions $k(\lambda,t)$ and $K(t,\lambda)$ which satisfy (10) are called a transform pair.

The spectral decomposition for the finite **discrete** time, finite **discrete** spectral case can be viewed as a change in basis vectors.[3]

[3]Given the discrete transform, often results obtained in this paper for the continuous case can be converted into the finite discrete case by replacing integrals with sums.

The representation of a discrete signal $\mathbf{x}(n),\ n \in \mathbb{Z}$, on an arbitrary finite discrete spectral domain $p,\ p \in \mathbb{Z}$, is $\mathbf{X}(p)$, where

$$\mathbf{X}(p) = \sum_{n\in\mathcal{N}} \mathbf{x}(n)\, K(n,p), \quad \hat{\mathbf{x}}(n) = \sum_{p\in\mathcal{P}} \mathbf{X}(p)\, k(p,n) \tag{11}$$

The support of each domain is finite, eg., $\mathcal{N} = \{0, N-1\}$, and $\mathcal{P} = \{0, P-1\}$. The kernels must satisfy

$$\mathrm{k}\,\mathrm{K} = \mathrm{K}\,\mathrm{k} = \mathrm{I} \tag{12}$$

where $\left[\mathrm{k}\right]_{pm} = k(p,m),\ \left[\mathrm{K}\right]_{mp} = K(m,p),\ m \in \{0, N-1\},\ p \in \{0, P-1\}$, and I is the identity martix.

Since $R_{xx}\,(t,\tau)$ is a well defined 2D deterministic function, then, as a natural extension to the concept of the stationary power spectrum [2], it may be expressed on an arbitrary spectral domain using a 2D integral transform. It will be shown in a forthcoming paper that, by considering the form of the ACF of $\mathbf{X}(\lambda)$ in (7), and the innovations representation of a stochastic process [2], [4], the kernel of this 2D transform should be *separable*.

Definition 4 (Generalised Power Spectrum (GPS). The GPS of the process $\mathbf{x}(t)$ and the inverse relationship are defined as

$$\mathcal{P}_{xx}(\lambda,\hat{\lambda}) = \iint\limits_{T\,T} R_{xx}\,(t,\tau)\ K(t,\lambda)\, K^*(\tau,\hat{\lambda})\, dt\, d\tau \tag{13}$$

$$R_{xx}\,(t,\tau) = \iint\limits_{\Lambda\,\Lambda} \mathcal{P}_{xx}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},\tau)\, d\lambda\, d\hat{\lambda} \tag{14}$$

where $K(t,\lambda)$ and $k(\lambda,t)$ are related by (10). Some basic relations are; $\mathcal{P}^*_{xx}(\lambda,\hat{\lambda}) = \mathcal{P}_{xx}(\hat{\lambda},\lambda)$, $\mathcal{P}_{\hat{x}\hat{x}}(\lambda_t,\lambda_\tau) = \mathcal{P}_{xx}(\lambda_t,\lambda_\tau)$, and $R_{\hat{x}\hat{x}}\,(t,\tau) = R_{xx}\,(t,\tau)$. Special cases of the GPS include Priestley's Evolutionary Spectrum [5] and Loèves Harmonizable processes [4], where the power spectrum is the symplectic 2D Fourier transform of the ACF.

Definition 5. The cross-correlation of the processes $\mathbf{y}(t)$ and $\mathbf{x}(t)$ is $R_{yx}\,(t,\tau) = E\left[\mathbf{y}(t)\,\mathbf{x}^*(\tau)\right]$, and the Generalised Cross Power Spectrum (GCPS) of the processes $\mathbf{y}(t)$ and $\mathbf{x}(t)$ is defined as $\mathcal{P}_{yx}(\lambda,\hat{\lambda}) = E[\mathbf{Y}(\lambda)\,\mathbf{X}^*(\hat{\lambda})]$, where $\forall \lambda,\hat{\lambda} \in \Lambda_0$,

$$\mathcal{P}_{yx}(\lambda,\hat{\lambda}) = \iint\limits_{T\,T} R_{yx}\,(t,\tau)\ K(t,\lambda)\, K^*(\tau,\hat{\lambda})\, dt\, d\tau \tag{15}$$

and the inverse relationship is of the form

$$R_{yx}\,(t,\tau) = \iint\limits_{\Lambda_0\,\Lambda_0} \mathcal{P}_{yx}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},\tau)\, d\lambda\, d\hat{\lambda} \tag{16}$$

where Λ_0 is the region over which the stochastic spectra of $\mathbf{x}(t)$ and $\mathbf{y}(t)$ overlap. Note that $\mathcal{P}^*_{yx}(\lambda,\hat{\lambda}) = \mathcal{P}_{xy}(\hat{\lambda},\lambda)$.

3. SOLUTIONS OF THE WIENER FILTER

Solutions of equation (4) are not easily found, although it is easy to show [6] that the solution can be reduced to the factorisation of the ACFs; the methods of Zadeh and Miller [6], Shinbrot [7], [8], and Darlington [9] are either explicitly based on the factorisation of $R_{xx}\,(t,\tau)$ and $R_{dx}\,(t,\tau)$, or make implicit use of it. A sufficient solution in this paper relies on the factorisation of the ACFs into the GPS, although a necessary condition for a solution has not yet been found. The case when the input to the Wiener filter $\mathbf{x}(t) = \mathbf{d}(t) + \mathbf{n}(t)$, is called the *additive case*.

Theorem 2. *Suppose $R_{dd}(t,\tau)$, $R_{nn}(t,\tau)$ and $R_{dn}(t,\tau)$ can be written as the generalised spectral decompositions*

$$R_{dd}(t,\tau) = \iint_{\hat{\Lambda}_d\,\hat{\Lambda}_d} \mathcal{P}_{dd}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},\tau)\, d\lambda\, d\hat{\lambda} \qquad (17)$$

$$R_{nn}(t,\tau) = \iint_{\hat{\Lambda}_n\,\hat{\Lambda}_n} \mathcal{P}_{nn}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},\tau)\, d\lambda\, d\hat{\lambda} \qquad (18)$$

$$R_{dn}(t,\tau) = \iint_{\Lambda_0\,\Lambda_0} \mathcal{P}_{dn}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},\tau)\, d\lambda\, d\hat{\lambda} \qquad (19)$$

where $\hat{\Lambda}_d \equiv \Lambda_d \oplus \Lambda_0$ and $\hat{\Lambda}_n \equiv \Lambda_n \oplus \Lambda_0$.[4] *Here, Λ_d and Λ_n are the regions of the Λ space over which the spectral components of $\mathbf{d}(t)$ and $\mathbf{n}(t)$ respectively, do not overlap, and Λ_0 is the region over which spectral components of $\mathbf{d}(t)$ and $\mathbf{n}(t)$ do overlap. Hence, $\Lambda_d \cap \Lambda_n = \{\emptyset\}$, $\Lambda_d \cap \Lambda_0 = \{\emptyset\}$ and $\Lambda_n \cap \Lambda_0 = \{\emptyset\}$.*

A sufficient solution $h(t,\tau)$, to the WHF equation (4) *for the additive case, when $\{\mathcal{T}\} = \{T\}$, is*

$$\begin{aligned} h(t,\tau) = &\int_{\Lambda_d} k(\lambda,t)\, K(\tau,\lambda)\, d\lambda \\ &+ \iint_{\Lambda_0\,\Lambda_0} H_0(\lambda,\hat{\lambda})\, k(\lambda,t)\, K(\tau,\hat{\lambda})\, d\lambda\, d\hat{\lambda} \end{aligned} \qquad (20)$$

where $H_0(\lambda_t,\lambda_\tau)$ is the solution of

$$\mathcal{P}_{dx}(\lambda,\hat{\lambda}) = \int_{\Lambda_0} H_0(\lambda,\bar{\lambda})\, \mathcal{P}_{xx}(\bar{\lambda},\hat{\lambda})\, d\bar{\lambda},\ \forall \lambda,\hat{\lambda} \in \Lambda_0 \qquad (21)$$

where $\mathcal{P}_{xx}(\cdot) = \mathcal{P}_{dx}(\cdot) + \mathcal{P}_{nx}(\cdot)$, $\mathcal{P}_{dx}(\cdot) = \mathcal{P}_{dd}(\cdot) + \mathcal{P}_{dn}(\cdot)$ and $\mathcal{P}_{nx}(\cdot) = \mathcal{P}_{nd}(\cdot) + \mathcal{P}_{nn}(\cdot)$. The resulting MSE is

$$\sigma^2(t) = \iint_{\Lambda_0\,\Lambda_0} \mathcal{P}_{\sigma\sigma}(\lambda,\hat{\lambda})\, k(\lambda,t)\, k^*(\hat{\lambda},t)\, d\lambda\, d\hat{\lambda} \qquad (22)$$

where the spectrum of $\sigma^2(t)$ is given by

$$\begin{aligned} \mathcal{P}_{\sigma\sigma}(\lambda,\hat{\lambda}) &= \mathcal{P}_{dd}(\lambda,\hat{\lambda}) - \int_{\Lambda_0} H_0(\lambda,\bar{\lambda})\, \mathcal{P}_{dx}(\bar{\lambda},\hat{\lambda})\, d\bar{\lambda} \\ &= \int_{\Lambda_0} H_0(\lambda,\bar{\lambda}) \left\{ \mathcal{P}_{nd}(\bar{\lambda},\hat{\lambda}) + \mathcal{P}_{nn}(\bar{\lambda},\hat{\lambda}) \right\} d\bar{\lambda} \end{aligned}$$

It is assumed that the initial state of the filter $h(t,\tau)$ is at rest.

The proof of this theorem is essentially by substitution, and will be published in a forthcoming paper. The first term in equation (20) is *independent of the signal ACFs*, unlike the second term. This first term is the expression for an ideal filter, first proposed by Zadeh in 1952 [10], [11] and [12], defined on any arbitrary domain as a filter which passes without distortion all *generalised frequency* components falling in a certain range and rejecting all others. This term corresponds to the perfect separation of the non-overlapping signal components. Hence, immediately, a condition for perfect signal separation is obtained:

[4] where $\oplus$ denotes the *orthogonal direct sum*.

Theorem 3. *Perfect single channel signal separation is only possible if there exists some domain where the generalised spectral representations of the desired and noise signals are disjoint. Thus, the desired and noise signals must be uncorrelated processes.*

Since the ideal filter is independent of the signals *value*, the filter will separate the class of all non-stationary stochastic signals which are disjoint in the filters domain, provided the components of the desired signal lie in the passband of the filter. The only *prior knowledge* required for separation is *the domain in which the signals are disjoint, and the spectral regions over which the spectral components lie*; full knowledge of the ACF is not required.

The second term of (20) is signal dependent, given by the solution of (21); as expected, the filtered power spectrum of each process in the overlapping spectral region sums to the power spectrum of the desired signal. The MSE is the energy of the filtered noise process contained in the overlapping spectra.

4. SEPARATING MODULATED SIGNALS

This section will determine separability constraints for "filtered modulated" signals of the form

$$\left.\begin{aligned} \mathbf{d}(t) &= \int_T h_d(t,\tau)\, \mathbf{a}(\tau)\, d\tau \\ \mathbf{n}(t) &= \int_T h_n(t,\tau)\, \mathbf{b}(\tau)\, d\tau \end{aligned}\right\}\ t \in T \qquad (23)$$

where, since some prior knowledge is required to determine the structure of the signals $\mathbf{d}(t)$ and $\mathbf{n}(t)$, it is assumed that $\mathbf{d}(t)$ and $\mathbf{n}(t)$ overlap in the time and Fourier spectral domains (to avoid trivial cases), $\mathbf{a}(t)$ and $\mathbf{b}(t)$ are bandlimited to $\pm\omega_c$ (but otherwise unknown), and that $h_d(t,\tau)$ and $h_n(t,\tau)$ are known deterministic signals. A special case of filtered modulation is "uniform modulation" [5], examples of which include quadrature and spread spectrum modulation schemes. This problem will be solved by "concatenating power spectras" of each process.

Noting that $\mathbf{a}(t)$ and $\mathbf{b}(t)$ are bandlimited in the Fourier domain, they admit the representations

$$\mathbf{a}(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \mathbf{A}(\omega)\, e^{j\omega t}\, d\omega,\quad \mathbf{b}(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \mathbf{B}(\omega)\, e^{j\omega t}\, d\omega$$

and therefore, after substitution into (23), and some slight rearrangement, $\mathbf{d}(t)$ and $\mathbf{n}(t)$ admit the representation

$$\mathbf{d}(t) = \int_0^{2\omega_c} \mathbf{D}(\omega)\, k(\omega,t)\, d\omega,\quad \mathbf{n}(t) = \int_{-2\omega_c}^{0} \mathbf{N}(\omega)\, k(\omega,t)\, d\omega$$

where the kernel $k(\omega,t)$ is defined as

$$k(\omega,t) = \begin{cases} \dfrac{1}{2\pi} \displaystyle\int_T h_d(t,\tau)\, e^{j(\omega-\omega_c)\tau}\, d\tau & \text{for } \omega \in \{0, 2\omega_c\} \\ \dfrac{1}{2\pi} \displaystyle\int_T h_n(t,\tau)\, e^{j(\omega+\omega_c)\tau}\, d\tau & \text{for } \omega \in \{-2\omega_c, 0\} \\ \dfrac{1}{2\pi} \displaystyle\int_T h_v(t,\tau)\, e^{j\omega\tau}\, d\tau & \text{for } \omega \notin \{-2\omega_c, 2\omega_c\} \end{cases}$$

and $\mathbf{D}(\omega) = \mathbf{A}(\omega-\omega_c)$, $\mathbf{N}(\omega) = \mathbf{B}(\omega+\omega_c)$. Notice that the form of $k_v(\omega,t)$ is arbitrary and has been chosen for symmetry. It follows that a convenient form of expressing the inverse kernel is,

$$K(t,\omega)=\begin{cases}\int_T g_d(t,\tau)\,e^{-j(\omega-\omega_c)\tau}\,d\tau & \text{for } \omega\in\{0,2\omega_c\}\\ \int_T g_n(t,\tau)\,e^{-j(\omega+\omega_c)\tau}\,d\tau & \text{for } \omega\in\{-2\omega_c,0\}\\ \int_T g_v(t,\tau)\,e^{-j\omega\tau}\,d\tau & \text{for } \omega\notin\{-2\omega_c,2\omega_c\}\end{cases}$$

If $\mathbf{d}(t)$ and $\mathbf{n}(t)$ are separable, then $k(\omega,t)$ and $K(t,\omega)$ must satisfy (10); substituion of these expressions into (10) creates constraints on the filters $h_d(t,\tau)$ and $h_n(t,\tau)$. When possible, $h_v(t,\tau)$ will be chosen to complete a basis function set. However, in the continuous case, finding general constraints on the filters $h_{(\cdot)}(t,\tau)$ is quite difficult, and no intuitive results have yet been found. Fortunately, the discrete case leads to more tractable results. In the discrete – time, discrete spectrum case these expressions become

$$\mathbf{a}(m)=\frac{1}{N}\sum_{-p_c}^{p_c}\mathbf{A}(p)\,e^{jmp\frac{2\pi}{N}},\quad \mathbf{b}(m)=\frac{1}{N}\sum_{-p_c}^{p_c}\mathbf{B}(p)\,e^{jmp\frac{2\pi}{N}}$$

where, for clarity, assume $N=P$, $\mathbf{a}(m)$, $\mathbf{b}(m)$, $\mathbf{d}(m)$, $\mathbf{n}(m)\in\mathbb{K}^N$, and $\mathbf{a}(m)$ and $\mathbf{b}(m)$ are bandlimited to p_c. Hence, $\mathbf{d}(m)$ and $\mathbf{n}(m)$ admit the representations,

$$\mathbf{d}(m)=\sum_{p\in\mathcal{P}}\mathbf{D}(p)\,k(p,m),\quad \mathbf{n}(m)=\sum_{p\in\mathcal{P}}\mathbf{N}(p)\,k(p,m)$$

where $\mathbf{D}(p)=\mathbf{A}(p-p_c)$, $\mathbf{N}(p)=\mathbf{B}(p+p_c)$, and

$$k(p,m)=\begin{cases}\frac{1}{N}\sum_{\hat m\in\mathcal{N}}h_d(m,\hat m)\,e^{j\hat m(p-p_c)\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_D\\ \frac{1}{N}\sum_{\hat m\in\mathcal{N}}h_n(m,\hat m)\,e^{j\hat m(p+p_c)\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_N\\ \frac{1}{N}\sum_{\hat m\in\mathcal{N}}h_v(m,\hat m)\,e^{j\hat m p\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_V\end{cases}$$

$$K(m,p)=\begin{cases}\sum_{\hat m\in\mathcal{N}}g_d(m,\hat m)\,e^{-j\hat m(p-p_c)\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_D\\ \sum_{\hat m\in\mathcal{N}}g_n(m,\hat m)\,e^{-j\hat m(p+p_c)\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_N\\ \sum_{\hat m\in\mathcal{N}}g_v(m,\hat m)\,e^{-j\hat m p\frac{2\pi}{N}} & \text{for } p\in\mathcal{P}_V\end{cases}$$

with $m\in\mathcal{N}$, $\mathcal{P}_D=\{0,2p_c-1\}$, $\mathcal{P}_N=\{P-2p_c,P-1\}$ and $\mathcal{P}_v=\{2p_c,P-2p_c-1\}$. If separation is possible, then these kernels must be transform pairs; thus the square "kernel matrix" k must be invertible (have full rank). Defining $W_N^{mp}=\frac{1}{N}e^{jmp\frac{2\pi}{N}}$, and the matrices

$$\left[\hat{\mathrm{k}}_d\right]_{pm}=W_N^{m(p-p_c)},\ p\in\mathcal{P}_D,\ \left[\hat{\mathrm{k}}_v\right]_{pm}=W_N^{mp},\ p\in\mathcal{P}_V,$$
$$\left[\hat{\mathrm{k}}_n\right]_{pm}=W_N^{m(p+p_c)},\ p\in\mathcal{P}_N,\ \left[\mathrm{H}_{(\cdot)}\right]_{m\hat m}=h_{(\cdot)}(m,\hat m)$$

where $m,\ \hat m\in\mathcal{N}$, and $\mathrm{H}_d,\ \mathrm{H}_n\in\mathbb{K}^{N\times N}$; k can be partitioned as

$$\mathrm{k}^T=\left[\mathrm{H}_d\,\hat{\mathrm{k}}_d^T\ \middle|\ \mathrm{H}_n\,\hat{\mathrm{k}}_n^T\ \middle|\ \mathrm{H}_v\,\hat{\mathrm{k}}_v^T\right] \quad (24)$$

where $\mathrm{H}_v\,\hat{\mathrm{k}}_v$ are the unused basis vectors.[5] Since the Fourier kernel is an orthonormal basis, then $\hat{\mathrm{k}}_{(\cdot)}$ have full rank;

$$\text{rank}\,[\hat{\mathrm{k}}_d]=\text{rank}\,[\hat{\mathrm{k}}_n]=2p_c \ \text{ and } \ \text{rank}\,[\hat{\mathrm{k}}_v]=P-4p_c \quad (25)$$

[5]Note $\hat{\mathrm{k}}_d$ and $\hat{\mathrm{k}}_n$ are identical; their explicit form is shown to emphasis that for separability, $\mathrm{H}_{(\cdot)}\,\hat{\mathrm{k}}_{(\cdot)}$ span different regions of the p-domain.

Theorem 4. *The signals $\mathbf{d}(m)$ and $\mathbf{n}(m)$ are separable if the fully known matrix k of equation* (24) *is of full rank.*

One physical way of interpreting the rôle of H_d and H_n in equation (24), is to consider H_d and H_n as linear transformations of $\mathbf{a}(m)$ and $\mathbf{b}(m)$, mapping them to *different subspaces*; $\mathrm{d}(m)$ and $\mathrm{n}(m)$ are not separable when these linear transformations do not map $\mathbf{a}(m)$ and $\mathbf{b}(m)$ to *disjoint subspaces*.

5. APPLICATIONS

Filtered and uniformly modulated signals that overlap in the Fourier domain are often assumed to be inseparable. If the modulating (bandlimited) functions $\mathbf{a}(t)$ and $\mathbf{b}(t)$ are unknown, but the modulated functions $h_{(\cdot)}(t)$ are known, then provided (24) is satisfied, the resulting processes are in fact separable. This result has applications where modulated signals share the same single channel and separation is required, for example when taking multiple measurements in seismology, and, of course, transmission schemes in communications. A forthcoming paper will discuss further separable classes of nonstationary signals across a single channel.

6. REFERENCES

[1] H. W. Bode and C. E. Shannon, "A Simplified Derivation of Linear Least Square Smoothing and Prediction Thoery", *Proc. I.R.E.*, Vol. 38, pp. 417-425, April 1950

[2] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Third Edition, 1991

[3] H. D'Angelo, *Linear Time-Varying Systems : Analysis and Synthesis*, Allyn and Bacon, Boston, 1970

[4] H. Cramér, "On Some Classes of Nonstationary Stochastic Processes", in *Proc. 4th Berkeley Symp. Math, Stat. Prob.*, Vol. 2, pp. 57-78, 1961

[5] M. B. Priestly, *Non-linear and Non-stationary Time Series Analysis*, Academic Press, 1988

[6] K. S. Miller and L. A. Zadeh, "Solution of an Integral Equation Occuring in the Theories of Prediction and Detection", *I.R.E. Transactions on Information Theory*, Vol. IT-2, pp. 72-72, June 1956

[7] M. Shinbrot, "On the Integral Equation Occuring in Optimization Theory With Nonstationary Inputs", *Journal of Mathematics and Physics*, Vol. 36, pp. 121-129, July 1957

[8] M. Shinbrot, "A Generalization of a Method for the Solution of the Integral Equation Arising in Optimization of Time-Varying Linear Systems with Nonstationary Inputs", *I.R.E. Transactions on Information Theory*, Vol. IT-3, pp. 220-224, December 1957

[9] S. Darlington, "Nonstationary Smoothing and Prediction using Network Theory Concepts", *I.R.E. Transactions on Circuit Theory (special suppl.)*, Vol. CT-6, pp. 1-11, May 1959

[10] L. A. Zaheh, "A general Theory of Linear Signal Transmission Systems", *Journal Franklin Institute*, Vol. 353 pp. 293-312, 1952

[11] L. A. Zaheh and K. S. Miller, "Generalised Ideal Filters", *Jour. Appl. Physics*, Vol. 23 pp. 223-228, 1952

[12] L. A. Zaheh, "On the Theory of Filtration of Signals", *Zeitschrift fur Angenwante Mathematix und Physick*, Vol. 111, pp. 149-156, 1952