RESULTS ON PRINCIPAL COMPONENT FILTER BANKS: COLORED NOISE SUPPRESSION AND EXISTENCE ISSUES[1]

Sony Akkarakaran, Student member, IEEE and P.P.Vaidyanathan, Fellow, IEEE

Department of Electrical Engineering 136-93

California Institute of Technology

Pasadena, CA 91125 USA.

Phone: (626) 395-4681 Fax: (626) 795-8649

Contact Author: P.P.Vaidyanathan, Email: ppvnath@sys.caltech.edu

Submission to IEEE Transactions on Information Theory (accepted as Regular Paper)

Abstract. We have recently made explicit the precise connection between the optimization of orthonormal filter banks (FB's) and the *principal component* property: The principal component filter bank (PCFB) is optimal whenever the minimization objective is a concave function of the subband variances of the FB. This explains PCFB optimality for compression, progressive transmission and various hitherto unnoticed white–noise suppression applications such as subband Wiener filtering. The present work examines the nature of the FB optimization problems for such schemes when PCFB's do not exist. Using the geometry of the optimization search–spaces, we explain exactly why these problems are usually analytically intractable. We show the relation between compaction filter design (i.e., variance maximization) and optimum FB's: A sequential maximization of subband variances produces a PCFB if one exists, but is otherwise suboptimal for several concave objectives. We then study PCFB optimality for *colored* noise suppression. Unlike case when the noise is white, here the minimization objective is a function of both the signal and the noise subband variances. We show that for the *transform coder* class, if a *common* signal and noise PCFB (KLT) exists, it is optimal for a large class of concave objectives. Common PCFB's for general FB classes have a considerably more restricted optimality, as we show using the class of unconstrained orthonormal FB's. For this class, we also show how to find an optimum FB when the signal and noise spectra are both *piecewise constant* with all discontinuities at rational multiples of π.

Keywords: filter bank, majorization, noise suppression, optimal basis, principal component.

[1] Work supported in part by the National Science Foundation under Grant MIP 0703755.

1 Introduction

Filter bank optimization has been a problem of considerable interest in recent literature, with many notable results and open problems. In a companion paper [1], we have presented a number of results on the optimality of *principal component* filter banks (PCFB's) for several signal processing schemes, especially involving suppression of additive white noise. The present paper aims to extend and generalize these results in several directions. We first examine the nature of the filter bank optimization when PCFB's do not exist, and explain why many of these problems become analytically intractable. We then consider the problem of *colored* noise suppression, and show the optimality of simultaneous signal and noise PCFB's in certain situations.

A generic signal processing scheme using an M-channel uniform perfect reconstruction filter bank (FB) is shown in Fig. 1. The FB is said to be **orthonormal** [1] if the $(M \times M)$ analysis polyphase matrix $\mathbf{E}(e^{j\omega})$

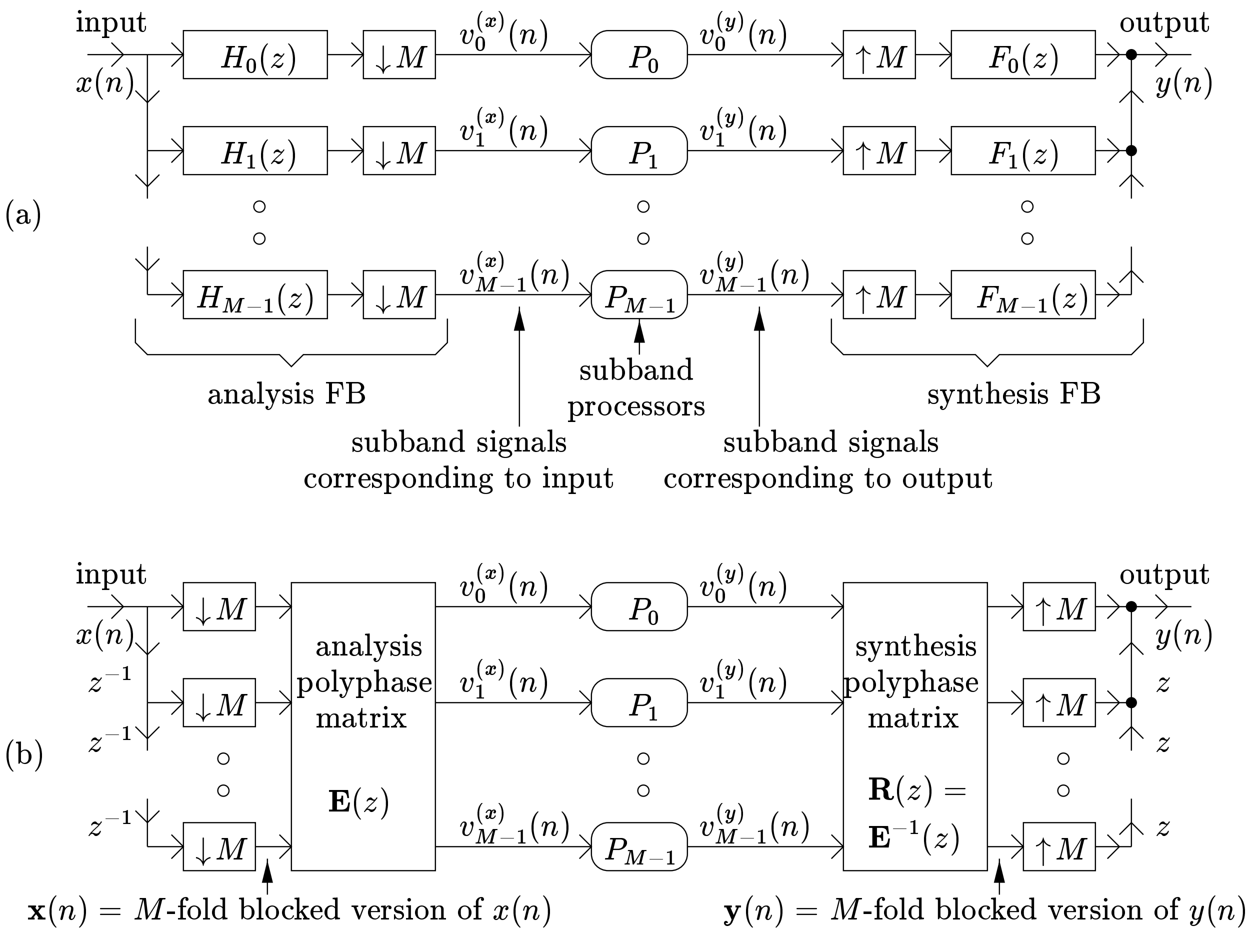


Figure 1: **Generic FB based signal processing scheme.**
(a) Showing analysis and synthesis filters, (b) Polyphase representation.

is unitary for all ω. The input vector $\mathbf{x}(n)$ is the M-fold blocked version of the scalar input $x(n)$. We assume that $\mathbf{x}(n)$ is a zero mean wide sense stationary (WSS) random process with a given power spectral density (psd) matrix $\mathbf{S_{xx}}(e^{j\omega})$. We are also given a class $\mathcal{C}$ of **orthonormal uniform M-channel FB's**. Examples are the class of FB's in which all filters are FIR with a given bound on their order, or the class of unconstrained FB's (in which there are no constraints on the filters besides those imposed by orthonormality). The FB optimization problem is that of **finding the best FB from** $\mathcal{C}$ for the given input statistics $\mathbf{S_{xx}}(e^{j\omega})$, for use in the system of Fig. 1. The term 'best FB' here means one that minimizes a well-defined objective function over the class $\mathcal{C}$. To formulate this objective, we need to describe the purpose or application of the FB in Fig. 1, and the nature of the subband processors P_i.

1.1 Relevant earlier work

Consider for example the case when the P_i are quantizers for signal compression. A commonly used quantizer model [10] replaces each P_i by an additive white noise of variance $f_i(b_i)\sigma_i^2$. Here b_i is the number of bits alloted to the quantizer, σ_i^2 is its input variance and f_i is the normalized quantizer function which is assumed not to depend on the input statistics. If the quantization noise processes are uncorrelated to each other, the overall mean square reconstruction error can be shown to be $g = \sum_{i=0}^{M-1} \frac{1}{M} f_i(b_i)\sigma_i^2$. The minimization objective here is this error g. It is shown in [10] that for any given set of bits b_i, the best FB for this problem is a **principal component filter bank (PCFB)** for the class $\mathcal{C}$ for the given input spectrum $\mathbf{S_{xx}}(e^{j\omega})$.

The notion of a PCFB will be reviewed in Section 2.3. PCFB's for certain classes of FB's have been studied earlier. For example, consider the orthogonal transform coder class $\mathcal{C}^t$ having all FB's as in Fig. 1 where $\mathbf{E}(z)$ is a constant unitary matrix $\mathbf{T}$. The Karhunen Loeve transform (KLT), which diagonalizes the autocorrelation matrix of the blocked input $\mathbf{x}(n)$ of Fig. 1, is the PCFB for $\mathcal{C}^t$ [8]. For the class $\mathcal{C}^u$ of *all* FB's with no constraints on filters (besides those imposed by orthonormality), PCFB construction has been studied in [20],[17]. Some optimality properties of PCFB's have been studied in [18]. A more recent work [1] by the authors has made explicit the precise connection between FB optimization and the principal component property: We have shown that the PCFB is optimal whenever the minimization objective is a concave function of the subband variances of the FB. This result explains the known optimality properties of PCFB's for compression and progressive transmission. It also shows the hitherto unnoticed optimality of PCFB's for various noise suppression schemes [1]. Suppose the FB input $x(n)$ in Fig. 1 is a signal corrupted by uncorrelated additive white noise, and the subband processors P_i are aimed at rejecting the noise components in their inputs. If the P_i are chosen as any combination of constant multipliers, zeroth order Wiener filters and hard thresholds (explained in Section 2.2), the PCFB is the optimum FB.

1.2 Paper outline

Section 2 describes the general structure of the FB optimization problems, with specific signal processing situations resulting in such problems. It also reviews the definition, properties and optimality of PCFB's [1].

Section 3 studies the nature of these FB optimizations when a PCFB does not exist. It shows that a PCFB exists iff there is a single FB that simultaneously minimizes all concave functions of the subband variances. By studying the structure of a certain convex set associated with the optimizations, we show exactly why they are usually analytically intractable in absence of a PCFB. We explain the relation between compaction filters and FB optimization. We show that a sequential maximization of subband variances always yields a PCFB if it exists, but is suboptimum for large classes of concave objectives if a PCFB does not exist.

Section 4 studies PCFB optimality for *colored* noise suppression. With white noise, the minimization objective g is a function of only the signal subband variances. The signal PCFB is optimal if g is concave [1]. With colored noise however, the objective depends on both the signal and noise subband variances. So the results of [1] no longer hold. We show that for the *transform coder* class, if a *common* signal and noise PCFB (KLT) exists, it minimizes a large class of concave objectives. Common PCFB's for a general FB class do not have such optimality, as we show using the unconstrained FB class $\mathcal{C}^u$. We show how to find the optimum FB in $\mathcal{C}^u$ for certain piecewise constant input spectra. We conclude with some open problems, especially on biorthogonal FB's and PCFB existence. Some of our results appear in preliminary form in [2],[3].

1.3 Notations

Superscripts $(^*)$ and $(^T)$ denote the complex conjugate and matrix (or vector) transpose respectively, while superscript dagger $(^\dagger)$ denotes the conjugate transpose. Boldface letters are used for matrices and vectors. Lower case letters are used for discrete sequences while upper case letters are used for Fourier transforms. $\mathcal{R}^M$ denotes the set of M-tuples of real numbers, and $\mathcal{R}^M_+$ denotes that of M-tuples of non-negative real numbers. We denote by $\text{diag}(\mathbf{A})$ the column vector consisting of the diagonal entries of the square matrix $\mathbf{A}$. The convex hull of a set D is denoted by $\text{co}(D)$. The Cartesian product of two sets A, B is denoted by $A \times B$.

2 Review of PCFB optimality

2.1 FB optimization problem formulation

We are given a class $\mathcal{C}$ of **orthonormal uniform M-channel FB's**. Recall that a FB is fully specified by its analysis polyphase matrix $\mathbf{E}(z)$, or alternatively by the ordered set of analysis and synthesis filter pairs $(H_k(z), F_k(z))$, $k = 0, 1, \ldots, M-1$ (see Fig. 1). We are also given subband processors P_i, $i = 0, 1, \ldots, M-1$, where P_i is the processor acting on the i-th subband. Each P_i is simply a function mapping input sequences to output sequences. The specification of this function may or may not depend on the input statistics.

The system of Fig. 1 is built using a FB in $\mathcal{C}$ and the processors P_i. Its aim is to produce a certain **desired signal** $d(n)$ at the FB output. For example, for signal compression, the P_i are quantizers and the desired output is the input, i.e., $d(n) = x(n)$. For noise reduction, the input $x(n) = s(n) + \mu(n)$ where $\mu(n)$ is additive noise, the desired output $d(n) = s(n)$, the pure signal; and the P_i could for instance be

Wiener filters. The FB optimization problem is to find a FB in $\mathcal{C}$ minimizing some measure of the error signal $e(n) \stackrel{\triangle}{=} d(n) - y(n)$, where $y(n)$ is the true FB output. To formulate the error measure, we impose random process models on $x(n)$ and $d(n)$. We assume that the blocked input $\mathbf{x}(n)$ (see Fig. 1) is a WSS vector process with given psd matrix $\mathbf{S_{xx}}(e^{j\omega})$. Equivalently, $x(n)$ is CWSS(M), i.e., wide sense cyclostationary with period M (in particular, it could be WSS). All processes are assumed zero mean unless otherwise stated. In all our problems, the $d(n)$ and the P_i are such that the error $e(n)$ is also a zero mean CWSS(M) random process. Thus we choose as **error measure** ε, the variance of $e(n)$ averaged over its period of cyclostationarity M.

As shown in Fig. 1, we denote by $v_i^{(x)}(n)$ the i-th subband signal produced when the FB input is the scalar signal $x(n)$. If the error $e(n)$ is CWSS(M), the signals $v_i^{(e)}(n)$, $i = 0, 1, \ldots, M-1$ are jointly WSS; and orthonormality of the FB can be used to show that the error measure becomes

$$\varepsilon = \frac{1}{M}\sum_{i=0}^{M-1} E[|v_i^{(e)}(n)|^2], \quad \text{where} \quad v_i^{(e)}(n) = v_i^{(d)}(n) - v_i^{(y)}(n). \tag{1}$$

Thus the processor P_i must try to produce output 'as close to' $v_i^{(d)}(n)$ as possible, i.e., to minimize $E[|v_i^{(e)}(n)|^2]$.

2.2 General form of optimization objective

In many signal processing problems [1], the processors P_i are such that the subband error variance is

$$E[|v_i^{(e)}(n)|^2] = h_i(\sigma_i^2). \tag{2}$$

Here $\sigma_i^2 = E[|v_i^{(x)}(n)|^2]$ is the variance of $v_i^{(x)}(n)$; and h_i is some function whose specification depends only on the processor P_i and not on the choice of FB. Thus, the minimization objective over the class $\mathcal{C}$ of FB's is

$$g(\sigma_0^2, \sigma_1^2, \ldots, \sigma_{M-1}^2) = \frac{1}{M}\sum_{i=0}^{M-1} h_i(\sigma_i^2). \tag{3}$$

We now summarize several such examples:

Compression: Here the desired FB output $d(n)$ equals the FB input $x(n)$, and the subband processors P_i are quantizers. We model the quantizers as a set of jointly stationary additive noise sources which thus represent the subband error signals $v_i^{(e)}(n)$. The noise variance is proportional to the variance σ_i^2 of the quantizer input. The proportionality constant $f_i(b_i)$ is called the normalized quantizer function, and depends on the number of bits b_i alloted to the quantizer. Thus, the error variance has the form of (2), i.e.,

$$E[|v_i^{(e)}(n)|^2] = f_i(b_i)\sigma_i^2. \qquad \text{[as in (2) with } h_i(x) = f_i(b_i)x\text{]} \tag{4}$$

The standard high bitrate quantizer model [20] chooses $f_i(b_i) = 2^{-2b_i}$. Here the arithmetic mean - geometric mean (AM-GM) inequality can be used to explicitly compute the optimal allocation of bits b_i (minimizing the overall error $\frac{1}{M}\sum_{i=0}^{M-1} E[|v_i^{(e)}(n)|^2]$ subject to a total bitrate constraint $\sum_{i=0}^{M-1} b_i = B$). With this optimal allocation, minimizing the overall error is equivalent to minimization of the GM of the variances σ_i^2, or equivalently of its logarithm. Thus we can rewrite the objective as in (3) with $h_i(x) = \log(x)$ for all i.

Discrete multitone (DMT) communications: This is the 'dual' of the compression problem. The system setup differs somewhat from that in Fig. 1: The analysis and synthesis banks exchange positions, and the 'subbands' are now 'subchannels' (or 'tones') carrying digitally modulated data. These are multiplexed by the FB into a single stream and sent on the 'full band' communications channel which is represented by a transfer function $C(e^{j\omega})$ followed by WSS additive noise with psd $S_e(e^{j\omega})$. Each subchannel is associated with a bitrate, power requirement and probability of error. We refer the reader to [9],[22] for details on the setup. Here we merely mention that given the error probabilities, both the total power for given bitrates (to be minimized) and the total bitrate for given subchannel powers (to be maximized) are objectives of the form (3); where the σ_i^2 are the subband variances of the FB when its input has psd $S_e(e^{j\omega})/\left|C(e^{j\omega})\right|^2$ [22].

Noise suppression: Suppose the input $v_i^{(x)}(n)$ to each subband processor P_i is the sum of a signal component $v_i^{(s)}(n)$ and a zero mean noise component $v_i^{(\mu)}(n)$. The P_i aims at rejecting the noise component, i.e., its desired output is $v_i^{(d)}(n) = v_i^{(s)}(n)$. We study the case when all the P_i are memoryless multipliers k_i. The subband error $v_i^{(e)}(n)$ of (1) thus equals $v_i^{(s)}(n) - k_i v_i^{(x)}(n)$. We assume that the $2M$ signals $v_i^{(s)}(n), v_i^{(\mu)}(n); i = 0, \ldots, M-1$ are jointly WSS; hence so are the M errors $v_i^{(e)}(n)$. Let σ_i^2, η_i^2 be the variances of $v_i^{(s)}(n), v_i^{(\mu)}(n)$ respectively, and let each $v_i^{(s)}(n)$ be uncorrelated to $v_i^{(\mu)}(n)$. We consider three different schemes to choose the k_i, each yielding an error variance $E[|v_i^{(e)}(n)|^2]$ of the form (2) for appropriate functions h_i:

- **Constants**, independent of choice of FB. Here

$$E[|v_i^{(e)}(n)|^2] = |1-k_i|^2\,\sigma_i^2 + |k_i|^2\,\eta_i^2, \quad \text{i.e.,} \quad h_i(x) = |1-k_i|^2\,x + |k_i|^2\,\eta_i^2. \tag{5}$$

- **Zeroth order Wiener filters**, $k_i = \frac{\sigma_i^2}{\sigma_i^2+\eta_i^2}$. These minimize the error variance $E[|v_i^{(e)}(n)|^2]$. Here

$$E[|v_i^{(e)}(n)|^2] = \frac{\sigma_i^2\eta_i^2}{\sigma_i^2+\eta_i^2}, \quad \text{i.e.,} \quad h_i(x) = \frac{x\eta_i^2}{x+\eta_i^2}. \tag{6}$$

- **Hard-thresholds**, that keep the subband if its signal component is stronger than its noise component, and drop it otherwise. Similar thresholds, applied separately to each element in the subband sequences (rather than to the entire subband), have been extensively studied by Donoho et al [5]. In this case,

$$k_i = \begin{cases} 1 & \text{if } \sigma_i^2 \geq \eta_i^2 \\ 0 & \text{otherwise} \end{cases}, \quad \text{and} \tag{7}$$

$$E[|v_i^{(e)}(n)|^2] = \min(\sigma_i^2, \eta_i^2), \quad \text{i.e.,} \quad h_i(x) = \min(x, \eta_i^2). \tag{8}$$

The above noise suppression problem could arise in many ways:

1. When the FB input is $x(n) = s(n)$, and the corresponding subband signals $v_i^{(s)}(n)$ are transmitted on separate communication lines. Here $v_i^{(\mu)}(n)$ represents the noise in the i-th line, and its variance η_i^2 (which affects the function h_i) is assumed to be independent of choice of FB.

2. When the FB input is $x(n) = s(n) + \mu(n)$ and the desired output is $s(n)$. Here $\mu(n)$ is zero mean additive *white* noise with variance η^2, and is uncorrelated to the pure signal $s(n)$. Thus $v_i^{(\mu)}(n)$ is the i-th subband signal corresponding to $\mu(n)$, with variance $\eta_i^2 = \eta^2$ independent of the FB. Here a minor difference from (2) is that the variances σ_i^2 now correspond not to the FB input $x(n)$ but to $s(n)$.

3. In item 1 above, if the line noise variance η_i^2 depends on choice of FB, Equation (2) no longer holds, as the function h_i then depends on the FB. However, in some cases this can be accounted for by a modified h_i that is independent of the FB. For example, suppose the noise $v_i^{(\mu)}(n)$ arises due to quantization of the signal $v_i^{(s)}(n)$. By the usual quantizer model (4), $\eta_i^2 = f_i(b_i)\sigma_i^2$, which depends on the FB only through the subband signal variance σ_i^2. Substituting this in the expressions (5),(6),(8) for the error variances $E[|v_i^{(e)}(n)|^2]$ shows that they still have the form of (2), with the modified h_i given by

$$h_i(x) = \begin{cases} \left[|1-k_i|^2 + |k_i|^2 f_i(b_i)\right] x & \text{for constant multiplier } k_i \\ \left[\frac{f_i(b_i)}{1+f_i(b_i)}\right] x & \text{for 0-th order Wiener filter } k_i \\ \left[\min(1, f_i(b_i))\right] x & \text{for hard threshold } k_i \end{cases} \tag{9}$$

4. Likewise, if the input noise $\mu(n)$ in item 2 is *colored*, the subband noise variances η_i^2 depend on the FB (just as the signal variances σ_i^2 do), rendering (2) invalid in general. Section 4 is devoted to analyzing this situation. Again as in item 3, modifying the h_i solves the problem in some very restricted cases; e.g. the h_i can be modified exactly as in item 3 if $\eta_i^2 = c\sigma_i^2$ for some c independent of FB. This happens in the special case when $\mathbf{S}_{\mu\mu}(e^{j\omega}) = c\mathbf{S}_{ss}(e^{j\omega})$, where $\mathbf{S}_{ss}(e^{j\omega}), \mathbf{S}_{\mu\mu}(e^{j\omega})$ are the psd matrices of the M-fold blocked versions of $s(n)$ and $\mu(n)$ respectively. If all FB's in the given class $\mathcal{C}$ have memoryless polyphase matrices, it suffices that the respective autocorrelation matrices $\mathbf{R}_{ss}, \mathbf{R}_{\mu\mu}$ satisfy $\mathbf{R}_{\mu\mu} = c\mathbf{R}_{ss}$.

Due to the general form (3) of the objective, the following definitions and observations become important:

The subband variance vector. For each FB in the given class $\mathcal{C}$, the subband variance vector associated with the input process $x(n)$ is defined as the vector $\mathbf{v} = (\sigma_0^2, \sigma_1^2, \ldots, \sigma_{M-1}^2)^T$, where σ_i^2 is the variance of the process $v_i^{(x)}(n)$. Here $v_i^{(x)}(n)$ is the i-th subband signal generated by feeding $x(n)$ as input to the FB.

The FB optimization objective g of (3) is purely a function of the subband variance vector of the FB; this function is fully specified given the description of the subband processors P_i. Given the FB analysis polyphase matrix $\mathbf{E}(z)$ and the psd matrix $\mathbf{S}_{\mathbf{xx}}(e^{j\omega})$ of the vector input $\mathbf{x}(n)$ in Fig. 1, the vector process $(v_0^{(x)}(n), v_1^{(x)}(n), \ldots, v_{M-1}^{(x)}(n))^T$ has psd matrix $\mathbf{E}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{E}^\dagger(e^{j\omega})$. Thus the subband variance vector is

$$\mathbf{v} = \frac{1}{2\pi}\int_0^{2\pi} \operatorname{diag}\left(\mathbf{E}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{E}^\dagger(e^{j\omega})\right)\, d\omega. \tag{10}$$

The search–space. The optimization search–space is defined as the set $\mathcal{S}$ of all subband variance vectors corresponding to all FB's in the given class $\mathcal{C}$. So $\mathcal{S}$ is fully specified given the class $\mathcal{C}$ and the input statistics $\mathbf{S_{xx}}(e^{j\omega})$. With objectives of the form (3), the FB optimization problem is reduced to that of finding the minima of the real-valued function g on the set $\mathcal{S}$; hence the importance of $\mathcal{S}$. As we will see in Section 2.4, the optimality of PCFB's is mainly due to a very special structure that $\mathcal{S}$ has whenever a PCFB exists.

Some general properties of $\mathcal{S}$. We have $\mathcal{S} \subset \mathcal{R}_+^M \subset \mathcal{R}^M$ (by definition). Also, $\mathcal{S}$ is *bounded* and *lies entirely on an* $M-1$ *dimensional hyperplane in* $\mathcal{R}^M$. (This follows from (10) using the fact that $\mathbf{E}(e^{j\omega})$ is unitary for all ω, i.e., that the FB is orthonormal [1].) Lastly, $\mathcal{S}$ has a *permutation symmetry* property: Any permutation of any vector $\mathbf{v} \in \mathcal{S}$ is also in $\mathcal{S}$. To understand this, recall that a FB is defined by an *ordered* set of analysis and synthesis filters. Changing this ordering technically produces a new FB, which we call a *permutation* of the original FB. Different permutations of a FB have subband variance vectors that are corresponding permutations of the variance vector of the original FB, and could thus achieve different values for the optimization objective. However, they are all essentially the same FB in the sense of being equally easy to implement. Hence we assume that any permutation of any FB in the given class $\mathcal{C}$ is also in $\mathcal{C}$. This most reasonable assumption about $\mathcal{C}$ causes the permutation symmetry of $\mathcal{S}$.

2.3 PCFB's: Definition and properties

Definition: Majorization. Let $A = \{a_0, a_1, \ldots, a_{M-1}\}$ and $B = \{b_0, b_1, \ldots, b_{M-1}\}$ be two sets each having M real numbers (not necessarily distinct). The set A is defined to **majorize** the set B if the elements of these sets, ordered so that $a_0 \geq a_1 \geq \ldots \geq a_{M-1}$ and $b_0 \geq b_1 \geq \ldots \geq b_{M-1}$, obey the property that

$$\sum_{i=0}^{P} a_i \geq \sum_{i=0}^{P} b_i \qquad \text{for all } P = 0, 1, \ldots, M-1, \text{ with equality holding when } P = M-1. \tag{11}$$

Given two vectors $\mathbf{v}_1, \mathbf{v}_2$ in $\mathcal{R}^M$, we will say that $\mathbf{v}_1$ majorizes $\mathbf{v}_2$ when the set of entries of $\mathbf{v}_1$ majorizes that of $\mathbf{v}_2$. Evidently in this case any permutation of $\mathbf{v}_1$ majorizes any permutation of $\mathbf{v}_2$.

Definition: PCFB's. Let $\mathcal{C}$ be the given class of orthonormal uniform M-channel FB's, and let $\mathbf{S_{xx}}(e^{j\omega})$ be the power-spectrum matrix of the vector process input $\mathbf{x}(n)$ (shown in Fig. 1). A FB in $\mathcal{C}$ is said to be a **principal component filter bank (PCFB)** for the class $\mathcal{C}$ for the input psd $\mathbf{S_{xx}}(e^{j\omega})$, if its subband variance vector (defined in Section 2.2) majorizes the subband variance vector of every FB in the class $\mathcal{C}$.

PCFB optimality for progressive transmission. In Fig. 1, suppose the FB has subbands numbered in decreasing order of their variances σ_i^2, i.e., $\sigma_0^2 \geq \sigma_1^2 \geq \ldots \geq \sigma_{M-1}^2$, and the P_i are constant multipliers m_i :

$$m_i = \begin{cases} 1 & \text{for } 0 \leq i \leq P-1 \\ 0 & \text{for } P \leq i \leq M-1 \end{cases} \tag{12}$$

for a fixed integer P $(0 \leq P \leq M)$. This system keeps the P strongest (largest variance) subbands, discarding the others. Due to FB orthonormality, the expected mean square error between the output and input is then

$$\frac{1}{M}\sum_{i=P}^{M-1}\sigma_i^2 = \frac{1}{M}(c-\sum_{i=0}^{P-1}\sigma_i^2), \quad \text{where} \quad c=\sum_{i=0}^{M-1}\sigma_i^2,$$

and c is the same for all orthonormal FB's. Thus by definition, the PCFB minimizes this error for all P.

Existence of PCFB. Given the class $\mathcal{C}$ of FB's and the input psd $\mathbf{S_{xx}}(e^{j\omega})$, a PCFB for $\mathcal{C}$ may not always exist. The PCFB and its existence depends on both $\mathcal{C}$ and $\mathbf{S_{xx}}(e^{j\omega})$. For a white input (i.e., when $\mathbf{S_{xx}}(e^{j\omega})$ is the identity matrix), PCFB's always exist: In fact, all FB's in $\mathcal{C}$ are PCFB's, no matter what $\mathcal{C}$ is. There are three important classes $\mathcal{C}$ for which PCFB's exist irrespective of the input psd [1]. These are

1. Any class of *two-channel* orthonormal FB's,
2. The orthogonal transform coder class $\mathcal{C}^t$, which has all FB's as in Fig. 1 where the analysis polyphase matrix $\mathbf{E}(z)$ is a constant unitary matrix $\mathbf{T}$,
3. The unconstrained class $\mathcal{C}^u$, which has *all* orthonormal M-channel FB's with no constraints on the filters besides those that orthonormality imposes (thus $\mathcal{C}^u$ contains FB's having ideal brickwall filters).

The DFT and cosine-modulated FB classes do not have PCFB's for large sets of input psd's, as shown in [1].

Construction of PCFB. For classes of two-channel FB's, a PCFB is simply a FB that maximizes its larger subband variance (thus minimizing the smaller one) over all FB's in the class. A FB in the transform coder class $\mathcal{C}^t$ is a PCFB for $\mathcal{C}^t$ for the given input psd iff it *decorrelates* the input, i.e., its analysis polyphase matrix $\mathbf{T}$ *diagonalizes* the autocorrelation matrix $\mathbf{R_{xx}}$ of the vector input $\mathbf{x}(n)$ of Fig. 1 (i.e., $\mathbf{TR_{xx}T}^\dagger$ is diagonal) [8]. A FB in the unconstrained class $\mathcal{C}^u$ is a PCFB for $\mathcal{C}^u$ iff its subband processes satisfy two properties [20]: Total decorrelation, i.e., the polyphase matrix $\mathbf{E}(e^{j\omega})$ diagonalizes the input psd $\mathbf{S_{xx}}(e^{j\omega})$ for all ω; and spectral majorization, i.e., if the subbands are numbered in decreasing order of variances, their spectra $S_i(e^{j\omega})$ satisfy $S_0(e^{j\omega}) \geq S_1(e^{j\omega}) \geq \ldots \geq S_{M-1}(e^{j\omega})$ for all ω.

Uniqueness of PCFB. From the definition of majorization we see that any permutation of a PCFB is also a PCFB. Further it is also possible that two different FB's which are not permutations of each other are both PCFB's, i.e., the PCFB need not be unique. However, all PCFB's must produce the same subband variance vector upto permutation. This is because **two sets majorizing each other must be identical** (directly from the definition of majorization). Since all our FB optimization problems involve not the actual FB but only its subband variance vector, we will often speak of *the* PCFB even though it may not be unique.

2.4 Principal components, convex polytopes and PCFB optimality

Let $\mathcal{C}$ be the given class of orthonormal uniform M-channel FB's, and $\mathbf{S_{xx}}(e^{j\omega})$ the psd matrix of the vector input $\mathbf{x}(n)$ (shown in Fig. 1). The search–space $\mathcal{S}$ is the set of all subband variance vectors of all FB's in $\mathcal{C}$ for the input $\mathbf{x}(n)$. The fundamental explanation of PCFB optimality (the core result of [1]) involves connections between convexity and majorization theory, and is summarized by the following theorems:[2]

Theorem 1: Optimality of extreme points of polytopes. Let f be a function whose domain P is a convex polytope. If f is concave on P, at least one of the extreme points of P attains the minimum of f over P. Further if f is strictly concave, its minimum over P *has to be* at an extreme point of P. (See Fig. 2.) ■

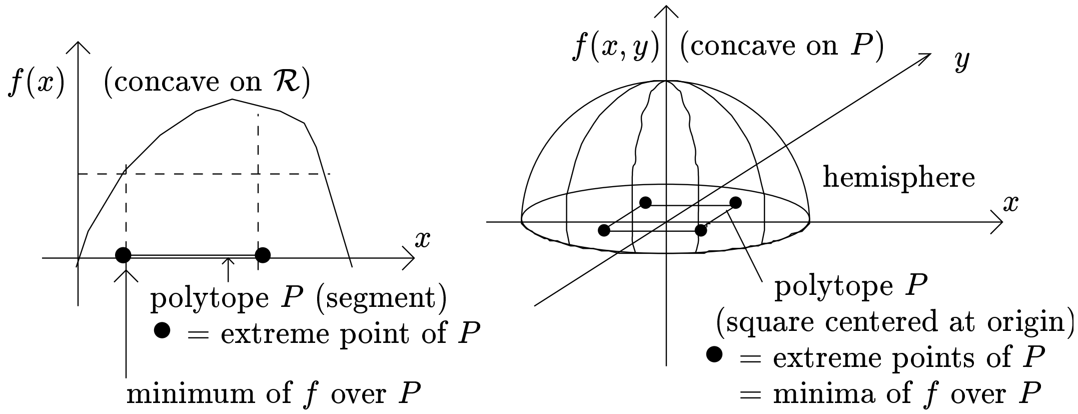


Figure 2: **Optimality of extreme points of polytopes (Theorem 1).**

Theorem 2: PCFB's and convex polytopes. A PCFB for the class $\mathcal{C}$ for input psd $\mathbf{S}_{\mathbf{xx}}(e^{j\omega})$ exists if and only if the convex hull co($\mathcal{S}$) is a polytope whose extreme points consist of all permutations of a single vector $\mathbf{v}_*$. Under this condition, $\mathbf{v}_*$ is the subband variance vector produced by the PCFB. ■

Theorem 3: Optimality of PCFB's. The PCFB for the class $\mathcal{C}$ (if it exists) is the optimum FB in $\mathcal{C}$ whenever the minimization objective is a concave function on the domain co($\mathcal{S}$). Further if this function is strictly concave, the optimum FB is necessarily a PCFB. ■

Theorem 3 shows PCFB optimality for several signal processing problems. In Section 2.2 we saw many situations in which the minimization objective g had the general form of (3). It can further be verified [1] that in all these cases, the functions h_i in (3) are concave on the non-negative real line. This implies that g is concave on the domain co($\mathcal{S}$), and hence that a PCFB is optimal for all these problems (Theorem 3).

Theorem 3 evidently follows directly from Theorems 1 and 2. The FB optimization involves finding the best vector from $\mathcal{S}$, but Theorem 1 is used here to find the best one from co($\mathcal{S}$) $\supset \mathcal{S}$. However, Theorem 2 shows that the best vector from co($\mathcal{S}$) is in fact in $\mathcal{S}$, hence it must be optimum over $\mathcal{S}$. Also note that all permutations of a PCFB are PCFB's, and the above results do not specify which of them is the optimum. In general they are not all equally good, but as they are finitely many, it is easy to pick the best one. For

[2]We refer the reader to [1] for detailed explanations of terms such as concavity and polytopes, and for proofs of the theorems.

the objective (3), if all h_i are identical then all permutations are equally good, while if $h_i(x) = k_i x$ for all i then we assign the largest σ_i^2 to the least k_i and so on. More generally, finding the best permutation of the PCFB is an instance of the *assignment problem*, well studied in operations research literature [4].

3 What if there is no PCFB?

When a PCFB exists, the search–space $\mathcal{S}$ consisting of all realizable subband variance vectors has a very special structure: Its convex hull co($\mathcal{S}$) is a polytope whose extreme points are all permutations of the PCFB subband variance vector (Theorem 2). The optimality of PCFB's under concave objectives (Theorem 3) follows from this structure and the optimality of extreme points of polytopes (Theorem 1). If a PCFB does not exist, $\mathcal{S}$ does not have this structure. Thus co($\mathcal{S}$) is a general convex set. For such sets too, there is a notion of extreme points, which coincides with the usual definition when the convex sets are polytopes, and further allows the following generalization of Theorem 1: If a function f is concave over a compact convex domain D, at least one extreme point of D is a minimum of f over D. Thus in this case, to minimize f over D it suffices to minimize f over the extreme points of D. Polytopes are exactly the compact convex sets having *finitely many* extreme points.

This section uses these observations to study the effect of nonexistence of PCFB's on the FB optimizations. When a PCFB exists, all the (finitely many) extreme points of the set co($\mathcal{S}$) correspond to the PCFB. So the PCFB is always optimal for *all* concave minimization objectives. On the other hand if a PCFB does not exist, co($\mathcal{S}$) could in general have infinitely many extreme points. This explains the analytical intractability of many FB optimizations when PCFB's do not exist. Finally, we explain the relation between PCFB's and 'compaction filters' that maximize their output variance among certain classes of filters.

3.1 Arbitrary convex sets: Extreme points and their optimality

Definition [7]. For a convex set $B \subset \mathcal{R}^M$, a point $\mathbf{z} \in B$ is said to be an *extreme point*, or a *corner* of B if

$$\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \quad \text{with} \quad \alpha \in (0,1),\ \mathbf{x},\mathbf{y} \in B \quad \text{implies} \quad \mathbf{x} = \mathbf{y}\,(= \mathbf{z}).$$

Geometrically, no line-segment passing *through* $\mathbf{z}$ (i.e., containing $\mathbf{z}$ but not as an endpoint) can lie wholly in the set B. The interior of B cannot have any extreme points, since around each point in the interior there is a ball lying wholly in B. So all extreme points lie on the boundary, though all boundary points need not be extreme points. If B is a polytope, the above definition can be verified to coincide with the usual definition of extreme points of a polytope. Figure 3 illustrates these facts, showing the extreme points of some closed and bounded (or compact) convex sets. It is not hard to show that every (nonempty) compact convex set is the convex hull of its boundary, and that it has at least one extreme point. A stronger result is true:

Krein–Milman theorem (Internal representation of convex sets) [7, 16]: Every compact convex set D is

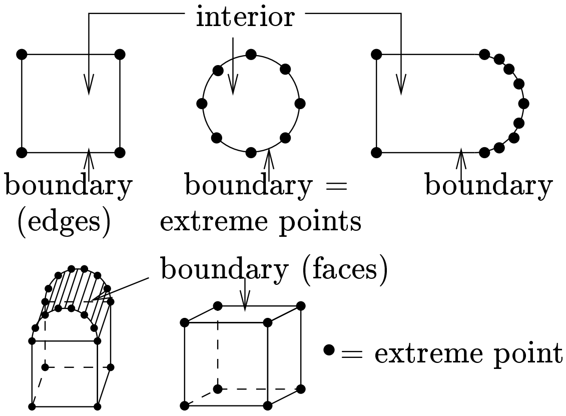


Figure 3: **Extreme points of compact convex sets.**

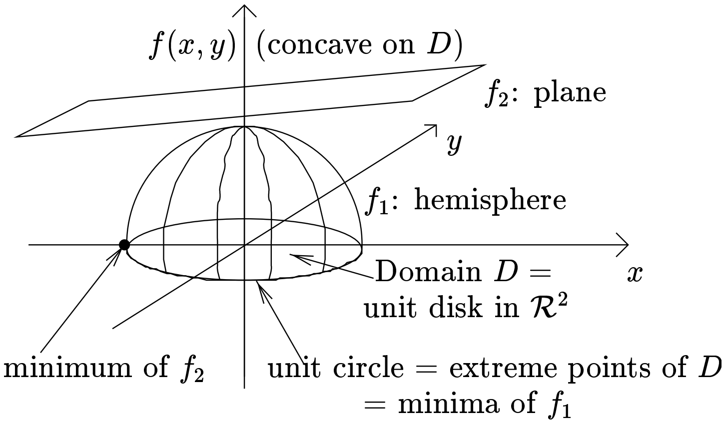


Figure 4: **Optimality of extreme points of compact convex sets (Theorem 4).**

the convex hull of its extreme points. Hence the set of extreme points of D is the minimal subset of D having D as its convex hull. This fact can serve as an equivalent definition of extreme points of compact convex sets.

This result evidently holds for polytopes, and is verifiable in the examples of Fig. 3. Thus it is intuitive, though its formal proof [16] may not be trivial. It is important as it immediately proves the following:

Theorem 4: Optimality of extreme points. If a function g is concave on a compact convex set D, at least one of the extreme points of D is a minimum of g over D. Further if g is *strictly* concave, its minimum *has to be* at an extreme point of D. ■

This result reduces to Theorem 1 if D is a polytope, and is illustrated in Fig. 4 for a compact convex D that is *not* a polytope.

Proof : Let $\mathbf{v}_{opt}$ minimize g over D. (Existence of $\mathbf{v}_{opt}$ is either assumed or follows if g is assumed continuous.) By the Krein–Milman theorem, $\mathbf{v}_{opt}$ is a convex combination of some extreme points of D, i.e.,

$$\mathbf{v}_{opt} = \sum_{j=1}^{J} \beta_j \mathbf{z}_j \quad \text{where} \quad 0 \leq \beta_j \leq 1, \quad \sum_{j=1}^{J} \beta_j = 1 \tag{13}$$

for some distinct extreme points $\mathbf{z}_j$ of D. If none of these $\mathbf{z}_j$ minimizes g over D, $g(\mathbf{z}_j) > g(\mathbf{v}_{opt})$ for all j, so

$$g(\mathbf{v}_{opt}) = g(\sum_{j=1}^{J} \beta_j \mathbf{z}_j) \geq \sum_{j=1}^{J} \beta_j g(\mathbf{z}_j) > \sum_{j=1}^{J} \beta_j g(\mathbf{v}_{opt}) = g(\mathbf{v}_{opt}), \tag{14}$$

i.e., $g(\mathbf{v}_{opt}) > g(\mathbf{v}_{opt})$, a contradiction. Hence at least one extreme point of D is a minimum of g over D. If g is strictly concave, the first inequality above (Jensen's inequality) is strict unless $\beta_j = 1$ for some j. So in this case $\mathbf{v}_{opt} = \mathbf{z}_j$, i.e., the minimum is necessarily at an extreme point of D. $\bigtriangledown\bigtriangledown\bigtriangledown$

3.2 Filter bank optimization and extreme points of convex sets

In our FB optimizations, the objective is concave on the set co($\mathcal{S}$) where $\mathcal{S}$ is the search–space. We seek its minima over $\mathcal{S}$. We assume from now on that $\mathcal{S}$ (and hence co($\mathcal{S}$)) is compact. This is true for most input power spectra and practical FB classes (Appendix A), and allows use of Theorem 4. Let E be the set of extreme points of co($\mathcal{S}$). From Theorem 4, for any concave objective over co($\mathcal{S}$), at least one of its minima lies in E (and all them do if the concavity is strict). From the definition of extreme points, we can show that $E \subset \mathcal{S}$. So the minima over co($\mathcal{S}$) $\supset \mathcal{S}$ found by minimizing over E in fact lie in $\mathcal{S}$, and are hence minima over $\mathcal{S}$ too. Thus, minimization over $\mathcal{S}$ has been reduced to one over the set E of extreme points of co($\mathcal{S}$). Now for 'almost every' extreme point $\mathbf{z}$ in E there is a concave (in fact linear) function that is minimized over co($\mathcal{S}$) uniquely by $\mathbf{z}$.[3] So without further information on the specific concave objective, nothing can be said about its minima over $\mathcal{S}$ apart from the fact that a search over E will yield at least one of them.

When a PCFB exists, *all* points in E correspond to it. This explains the remarkable optimality of PCFB's for *all* concave objectives. If there is no PCFB, E has at least two points that are not permutations of each other, i.e., that correspond to essentially different FB's. Thus, no single FB can be simultaneously optimal for all concave objectives f. If E is finite, the optimal FB for any given concave f can still be found by a finite exhaustive search over E. Unfortunately in general there is no reason to expect E to be finite, hence a numerical search is required. Any derivation of analytical results on the optimum FB will have to take into account the specific nature of both the concave objective at hand and the set E (which depends on the FB class $\mathcal{C}$ and input psd at hand). This explains why these optimizations are usually analytically intractable.

3.3 The sequential compaction algorithm

This is an algorithm that has sometimes been proposed [20],[13] to find a 'good' FB in classes $\mathcal{C}$ that may not have PCFB's. We first state the algorithm in a precise manner that holds for any general class $\mathcal{C}$. We

[3]This is because for any compact convex set D, the set of extreme points is the closure of the set of *exposed* points [16], which by definition are points $\mathbf{v} \in D$ for which there is a linear function minimized (or maximized) over D uniquely by $\mathbf{v}$.

then show that it produces FB's for which the corresponding subband variance vector *is an extreme point of* co($\mathcal{S}$). We examine the optimality of the algorithm in this light.

Let $\mathcal{C}$ be the given class of FB's, and $\mathcal{S}$ the corresponding optimization search–space. The algorithm involves rearranging all vectors in $\mathcal{S}$ in decreasing order of their entries, and then picking from these the vector $\mathbf{v}_\alpha \in \mathcal{S}$ defined as the greatest one in the 'dictionary ordering' on $\mathcal{R}^M$. This means that the greatest (first) entry of $\mathbf{v}_\alpha$ is greater than or equal to the greatest entry of any of the other vectors. Among vectors for which equality prevails, the second greatest entry of $\mathbf{v}_\alpha$ is greater than or equal to the second greatest entry of the other vectors, and so on. The output of the algorithm is any FB with subband variance vector $\mathbf{v}_\alpha$ (or any of its permutations). The vector $\mathbf{v}_\alpha$ is well-defined, and finding it involves a sequential maximization of subband variances giving the algorithm its name. (Existence of the maxima follows from compactness of $\mathcal{S}$.)

Relation to compaction filters. The *ideal compaction filter* [20] for an input process is defined as the filter maximizing its output variance among all filters $H(e^{j\omega})$ whose magnitude squared $\left|H(e^{j\omega})\right|^2$ is Nyquist(M). The Nyquist(M) constraint is imposed because these filters are used to build an orthonormal M-channel FB, and any filter in such a FB obeys this constraint [19]. For WSS inputs, a procedure from [20] finds the compaction filter given the input psd. It always yields a 'brickwall' filter, i.e., one with constant magnitude on its support. If such a filter is to be an analysis filter in an orthonormal FB, its support cannot overlap with that of any other analysis filter. Thus the FB can be built by a sequential design of compaction filters: The next filter maximizes its output variance among all filters that have a Nyquist(M) magnitude squared and a support that does not overlap with the supports of the previously designed filters.[4]

This FB design method from [20] is exactly the sequential algorithm described above, applied to the unconstrained FB class $\mathcal{C}^u$ when the input is WSS (as distinct from CWSS(M)). The variance maximization in the algorithm corresponds to an ideal compaction filter design. This connection has motivated the study and design of *FIR compaction filters* [12]. These are defined as filters maximizing their output variance among all filters of order not exceeding N whose magnitude squared is Nyquist(M). It was believed that such filters would play a role in PCFB design for the class $\mathcal{C}^{fir}$ of all M-channel orthonormal FB's in which all filters are FIR with order not exceeding N ($N \geq M$). Indeed, it may seem that the first step in the sequential algorithm for the class $\mathcal{C}^{fir}$ is to design a FIR compaction filter. However this is not true for a general M and input psd, as there may not even be a FB in $\mathcal{C}^{fir}$ having the FIR compaction filter as one of its filters. The correct first step in the sequential algorithm for a general FB class $\mathcal{C}$ is to design a filter maximizing its output variance *among all filters belonging to FB's in* $\mathcal{C}$. It seems quite infeasible to propose any variant of the sequential algorithm or the class $\mathcal{C}^{fir}$ in which FIR compaction filters will play any serious role. The only notable exception is when $M = 2$, where the FB is fully determined by any one

[4]Equivalently, it is an ideal compaction filter for the psd that is obtained by setting to zero the bands of the original input psd falling within the supports of the previously designed filters.

of its filters. Thus, a clear relation between the sequential algorithm and compaction filters exists only for the unconstrained class $\mathcal{C}^u$ when the input is WSS (as opposed to CWSS(M)).

3.4 Is the sequential algorithm optimal?

The optimality properties of the sequential algorithm of Section 3.3 follow easily from the following result:

Assertion 1. The subband variance vector $\mathbf{v}_\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{M-1})^T \in \mathcal{S}$ (with $\alpha_0 \geq \ldots \geq \alpha_{M-1}$) produced by the sequential algorithm **is an extreme point of** co($\mathcal{S}$).

Proof. Let $\mathbf{v}_\alpha = \gamma\mathbf{x}+(1-\gamma)\mathbf{y}$ for $\gamma \in (0,1)$ and $\mathbf{x}, \mathbf{y} \in$ co($\mathcal{S}$). By definition of an extreme point (Section 3.1), showing that $\mathbf{x} = \mathbf{y} = \mathbf{v}_\alpha$ will complete the proof. Now by definition of the convex hull co($\mathcal{S}$), we see that $\mathbf{x}, \mathbf{y}$ and hence $\mathbf{v}_\alpha$ can be written as convex combinations of elements of $\mathcal{S}$, i.e., $\mathbf{v}_\alpha = \sum_{j=1}^{J} \beta_j \mathbf{v}^j$ for some $\mathbf{v}^j = (v_0^j, v_1^j, \ldots, v_{M-1}^j) \in \mathcal{S}$ and $\beta_j \in (0,1]$ satisfying $\sum_{j=1}^{J} \beta_j = 1$. We now show $\mathbf{x} = \mathbf{y} = \mathbf{v}_\alpha$ by showing $\mathbf{v}^j = \mathbf{v}_\alpha$ for all $j = 1, 2, \ldots, J$. To this end, since $\mathbf{v}_\alpha$ exceeds (or equals) all the $\mathbf{v}^j$ in the dictionary ordering on $\mathcal{R}^M$, we have $\alpha_0 \geq v_0^j$, but α_0 is a convex combination of the v_0^j. Hence $\alpha_0 = v_0^j$ for all j. This in turn leads to $\alpha_1 \geq v_1^j$, and hence to $\alpha_1 = v_1^j$ and so on; until finally $\mathbf{v}_\alpha = \mathbf{v}^j$ for all $j = 1, 2, \ldots, J$. $\bigtriangledown\bigtriangledown\bigtriangledown$

When the class $\mathcal{C}$ has a PCFB, all extreme points of co($\mathcal{S}$) correspond to the PCFB. Hence the sequential algorithm always yields the PCFB and is thus optimal for many problems (Section 2). The subband variance vector $\mathbf{v}_\alpha$ produced by the algorithm here has an additional property: If its entries are arranged in *increasing order*, then in fact it becomes the *least* vector in $\mathcal{S}$ in the dictionary ordering.[5] On the other hand, if a PCFB does not exist, then there will be at least two extreme points that do not correspond to essentially the same FB, i.e., whose coordinates are not permutations of each other. The algorithm of Section 3.3 produces one extreme point, but the minimum could easily be at another one. Thus the algorithm could be suboptimum.

The following hypothetical example with $M = 4$ channels illustrates this point: Let co($\mathcal{S}$) = co(E) where E is the set of all permutations of $\mathbf{v}_1 = (25, 10, 10, 2)^T$ and $\mathbf{v}_2 = (24, 17, 3, 3)^T$. This would happen for a WSS input with psd shown in Fig. 5, when the class $\mathcal{C}$ has exactly the two FB's in the figure. As E is finite, co($\mathcal{S}$) is a polytope whose extreme points lie in E. In fact all points in E are extreme points of co($\mathcal{S}$) as neither of $\mathbf{v}_1, \mathbf{v}_2$ *majorizes* the other. A PCFB does not exist, as $\mathbf{v}_1$ is not a permutation of $\mathbf{v}_2$. Now consider the high bit-rate coding problem of [20]. Here the objective to be minimized over $\mathcal{S}$ is $\pi(\mathbf{v})$, the geometric mean of the entries of $\mathbf{v} \in \mathcal{S}$. (As noted in Section 2.2, this is equivalent to minimizing an objective that is concave on co($\mathcal{S}$).) Now $[\pi(\mathbf{v}_1)]^4 = 5000 > [\pi(\mathbf{v}_2)]^4 = 3672$, so $\mathbf{v}_2$ is the minimum. However, the algorithm of Section 3.3 yields $\mathbf{v}_\alpha = \mathbf{v}_1$, and is thus suboptimum. Further it remains so even if it is run to sequentially minimize rather than maximize variances (again, giving in general some extreme point of co($\mathcal{S}$), in this case $\mathbf{v}_1$).

In fact one can even create a family of (necessarily artificial) concave objectives that the algorithm actually

[5] However, the fact that $\mathbf{v}_\alpha$ has this property does not imply that a PCFB exists, unless the number of channels is $M \leq 3$. *Majorization* is a stronger requirement. For example, $\mathbf{v}_1 = (25, 10, 10, 2)$ exceeds $\mathbf{v}_2 = (24, 17, 3, 3)$ and its permutations, and also becomes less than them if its entries are rearranged in increasing order; but still $\mathbf{v}_1$ does not majorize $\mathbf{v}_2$.

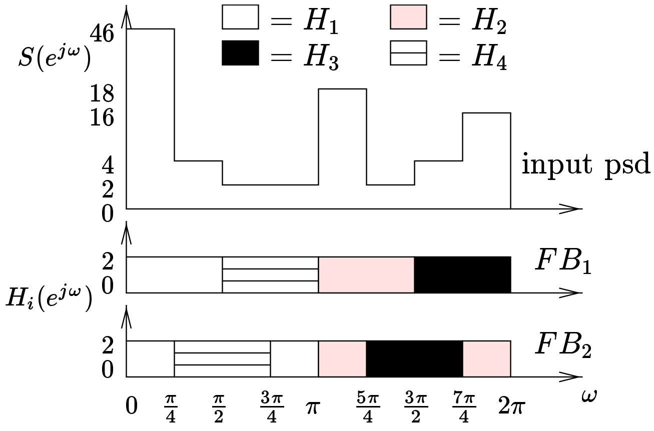


Figure 5: **Suboptimality of sequential compaction algorithm.**

maximizes instead of minimizing. Let $P \subseteq \mathrm{co}(\mathcal{S})$ be the polytope with extreme points as permutations of the vector $\mathbf{v}_\alpha$ output by the algorithm, so $P = \mathrm{co}(\mathcal{S})$ iff a PCFB exists. Let $f(\mathbf{v}) = -d(\mathbf{v}, P)$, where $d(\mathbf{v}, P) = \min\{\|\mathbf{v} - \mathbf{x}\| : \mathbf{x} \in P\}$ is the minimum distance from $\mathbf{v}$ to P (well–defined since P is compact) using any valid norm $\|\cdot\|$ on $\mathcal{R}^M$. Now f is continuous and concave on $\mathcal{R}^M$ (Appendix B). Its definition shows that f is constant (zero) on P, and that if a PCFB does not exist, P is actually the set of *maxima* of f over $\mathrm{co}(\mathcal{S})$. Thus FB's with subband variance vector $\mathbf{v}_\alpha \in P$ or its permutations (output by the sequential algorithm) perform the worst. Even if these examples may seem artificial, they should convince the reader of the total absence of intrinsic connection between FB optimality and variance maximization / compaction filter design *except if a PCFB exists*, in which case the sequential algorithm yields exactly the PCFB's.

4 Optimum FB's for colored noise suppression

This section considers the case when the FB input in Fig. 1 is $x(n) = s(n) + \mu(n)$, where $s(n)$ is the pure signal desired at the FB output, and $\mu(n)$ is zero mean additive noise uncorrelated to $s(n)$. Both $s(n)$ and $\mu(n)$ are assumed CWSS(M) random processes. This problem has been described in Section 2.2. Using the same notation, let σ_i^2, η_i^2 be the variances of the i-th subband signals $v_i^{(s)}(n)$ and $v_i^{(\mu)}(n)$ corresponding to $s(n)$ and $\mu(n)$ respectively. The subband processors are memoryless multipliers k_i which could be constants, zeroth order Wiener filters or hard-thresholds (Section 2.2). We have already seen that the mean square error between the true and desired FB output is

$$f(\mathbf{v}_\sigma, \mathbf{v}_\eta) = \frac{1}{M} \sum_{i=0}^{M-1} f_i(\sigma_i^2, \eta_i^2) \tag{15}$$

where $\mathbf{v}_\sigma = (\sigma_0^2, \sigma_1^2, \ldots, \sigma_{M-1}^2)^T$, $\mathbf{v}_\eta = (\eta_0^2, \eta_1^2, \ldots, \eta_{M-1}^2)^T$ are respectively the signal and noise subband variance vectors, and

$$f_i(x,y) = \begin{cases} |1-k_i|^2 x + |k_i|^2 y & \text{for constant multiplier } k_i \\ \frac{xy}{x+y} & \text{for 0-th order Wiener filter } k_i \\ \min(x,y) & \text{for hard threshold } k_i \end{cases} \tag{16}$$

In Section 2.2, $f_i(\sigma_i^2, \eta_i^2)$ was denoted by $h_i(\sigma_i^2)$ with the dependence of the function h_i on η_i^2 being implicit. PCFB optimality was shown only if all the h_i were concave and *independent of choice of FB*. This happens for example if the noise $\mu(n)$ is white, in which case for all FB's we have $\eta_i^2 = \eta^2$, the variance of $\mu(n)$. It also happens if $\eta_i^2 = c\sigma_i^2$ for constant c independent of the FB, which would happen for a very restricted class of input signal and noise psd's. As explained in Section 2.2, in this case the dependence of h_i on η_i^2 can be accounted for by a modified h_i that is independent of η_i^2. In both these cases, the results of Section 2.4 can be used to show that a PCFB for the input signal $s(n)$ is optimum in the sense of minimizing (15).

For general input signal and noise psd's, the minimization objective f of (15) indeed depends on both the signal and noise subband variances. So the problem is outside the scope of Section 2, and is the topic of the present one. Possible questions that arise are: Is the PCFB for $s(n)$ still optimal? Or is it the PCFB for $\mu(n)$ (or for $x(n) = s(n) + \mu(n)$) that is now optimal? Here we answer these questions. At the outset, re-examine the two special cases noted above where a PCFB for $s(n)$ is still optimal. When $\mu(n)$ is white, since any FB is a PCFB for a white input, the optimum FB is actually a *common* PCFB for the signal $s(n)$ and the noise $\mu(n)$ (and is in fact also the PCFB for $x(n) = s(n) + \mu(n)$). When $\eta_i^2 = c\sigma_i^2$, this relation itself ensures (by definition of PCFB's) that a PCFB for any one of the signals $s(n), \mu(n)$ and $x(n)$ is also a PCFB for the others. So in both these cases, the optimum FB is a *common* signal and noise PCFB. **Is this true in greater generality?** We answer this question in detail. We show that *for the transform coder class* $\mathcal{C}^t$, the common PCFB (if it exists) is indeed optimal; while the same is not always true for other classes of FB's, specifically for the unconstrained FB class $\mathcal{C}^u$. We also show how to find the optimum FB in $\mathcal{C}^u$ when the input signal and noise spectra are both *piecewise constant* with all discontinuities at rational multiples of π.

4.1 Notation and study of search–spaces

To study the issues mentioned above, we need notations for certain sets associated with the optimization problem. We now introduce these notations, which will hold throughout Section 4.

1. *Signal and noise variance spaces* S_σ, S_η. The set of all realizable subband signal variance vectors $\mathbf{v}_\sigma$ is denoted by S_σ. Similarly, the set of all realizable subband noise variance vectors $\mathbf{v}_\eta$ is denoted by S_η.

2. *Optimization search space* S_v. We denote by S_v the set of all realizable *pairs* of signal and noise variance vectors $\begin{pmatrix} \mathbf{v}_\sigma \\ \mathbf{v}_\eta \end{pmatrix}$. The minimization objectives for the problems studied here have the form (15), i.e., they are real-valued functions on S_v. Thus S_v is the 'search–space' for these problems, just as S_σ is for those

of Section 2. As both $\mathbf{v}_\sigma$ and $\mathbf{v}_\eta$ have entries whose sum is independent of the FB, the set S_v is bounded and lies on a $2M-2$ dimensional hyperplane in $\mathcal{R}_+^{2M}$. It also has a permutation symmetry, slightly different from that of S_σ but arising from the same reason (see end of Section 2.2). It is expressed as $\begin{pmatrix}\mathbf{v}_\sigma \\ \mathbf{v}_\eta\end{pmatrix} \in S_v \Rightarrow \begin{pmatrix}\mathbf{P}\mathbf{v}_\sigma \\ \mathbf{P}\mathbf{v}_\eta\end{pmatrix} \in S_v$ for any permutation matrix $\mathbf{P}$. Also, $\mathbf{v}_\sigma \in S_\sigma, \mathbf{v}_\eta \in S_\eta$ does not always imply $\begin{pmatrix}\mathbf{v}_\sigma \\ \mathbf{v}_\eta\end{pmatrix} \in S_v$; i.e., S_v is some subset of the Cartesian product $S_\sigma \times S_\eta$, usually a proper subset.[6] We also assume S_v (and hence $\text{co}(S_v)$) to be compact, for similar reasons as in Section 3.2.

3. *Objective function domain T.* We will consider general minimization objectives concave over the set

$$T \triangleq \text{co}(S_\sigma) \times \text{co}(S_\eta) = \text{co}(S_\sigma \times S_\eta) \supset \text{co}(S_v). \tag{17}$$

(We have used above, and will freely use the set identity $\text{co}(A \times B) = \text{co}(A) \times \text{co}(B)$.) Note that if all the f_i in (15) are concave on $\mathcal{R}_+^2$, the objective f of (15) is concave on $\mathcal{R}_+^{2M}$ and hence on T. Also, the f_i of (16) arising for the noise suppression problems above are indeed concave on $\mathcal{R}_+^2$ (Appendix B). We know that minimizing a concave function over S_v is reducible to minimizing it over the set of extreme points of $\text{co}(S_v)$ (Section 3). So we will try to study the structure of this set of extreme points.

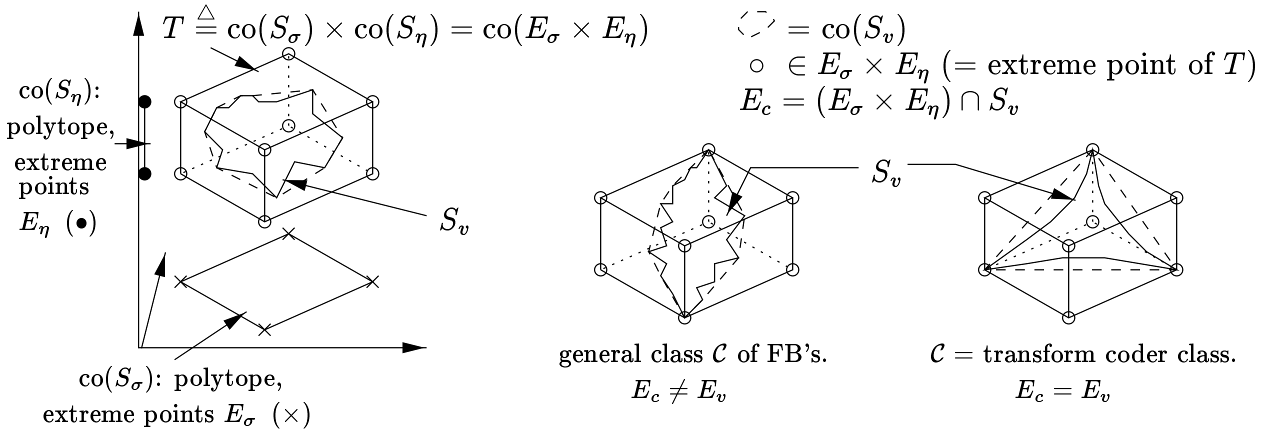
4. *Extreme point sets E_σ, E_η, E_v.* We denote by $E_\sigma \subset S_\sigma$, $E_\eta \subset S_\eta$, $E_v \subset S_v$ the sets of extreme points of $\text{co}(S_\sigma), \text{co}(S_\eta), \text{co}(S_v)$ respectively. (Extreme points of $\text{co}(A)$ always lie in A.) From definitions it is easily shown that $E_\sigma \times E_\eta$ is the set of extreme points of the set T of (17). In all problems in this section we assume that separate PCFB's for the signal and noise psd's always exist (otherwise most optimizations are analytically intractable for similar reasons as explained in Section 3). Thus E_σ, E_η are both finite sets, each one being the set of all permutations of a single vector that corresponds to the relevant PCFB. Also T is a polytope, as its set of extreme points $E_\sigma \times E_\eta$ is also finite.

5. *Common PCFB point set E_c.* We denote by E_c the set of all points in S_v that correspond to a *common* signal and noise PCFB for the given FB class $\mathcal{C}$. (E_c is empty iff there is no such PCFB.) From earlier discussions, a FB in $\mathcal{C}$ will be such a common PCFB iff its corresponding point in the search space S_v lies in the finite set $E_\sigma \times E_\eta$. However, even when a common PCFB exists, in general all points of $E_\sigma \times E_\eta$ will not correspond to such PCFB's: In fact usually many of them will be unrealizable, i.e., outside the search space S_v. Thus, $E_c = (E_\sigma \times E_\eta) \cap S_v$, i.e., E_c consists of the extreme points of the polytope T that lie in $S_v \subset T$. Points in E_c are hence also extreme points of $\text{co}(S_v)$, i.e., $E_c \subset E_v$.

From the above definitions and discussions, the optimum FB for minimizing functions that are concave on the domain T of (17) can be found by a search over the FB's corresponding to points in $E_v \subset S_v$. On the other hand, common signal and noise PCFB's correspond to points in the finite set $E_c \subset E_v$. Now as noted

[6] $S_v = S_\sigma \times S_\eta$ only in artificial / degenerate cases, e.g. if $\mu(n)$ (or $s(n)$) is white. (For white $\mu(n)$, S_η has only one element.)

in Section 3.2, for almost every $\mathbf{z} \in E_v$ there is a concave objective minimized over S_v uniquely by $\mathbf{z}$. Thus, **the common signal and noise PCFB will minimize all concave objectives over T if and only if** $E_c = E_v$. For the transform coder class $\mathcal{C}^t$, it turns out that indeed $E_c = E_v$ whenever a common signal and noise PCFB (KLT) exists. For the unconstrained class $\mathcal{C}^u$ on the other hand, even when a common PCFB exists (i.e., E_c is nonempty), $E_c \neq E_v$ in general, except for some very restricted input spectra (for example with *constant* signal and noise psd matrices, in which case the PCFB's are the corresponding KLT's). We formally state results on PCFB optimality for colored noise suppression in the next section; their proofs follow from the above comments on the relation between E_c and E_v which will be proved later. Figure 6 shows the various geometries of S_v as a subset of T arising in the different situations discussed above. (The figure only serves as illustration: Actually T lies in $\mathcal{R}^{2M}$ and not $\mathcal{R}^3$ as the figure shows.)



separate but no common signal and noise PCFB's exist: E_c empty

common signal, noise PCFB exists: E_c nonempty
$E_c \subset E_v \triangleq$ set of extreme points of $\mathrm{co}(S_v)$

Figure 6: **Colored noise suppression: Geometry of search–space.**

4.2 Statement and discussion of results

Theorem 5. Optimality of common KLT. Consider any minimization objective that is concave on the set T of (17). The common signal and noise PCFB for the transform coder class $\mathcal{C}^t$ (i.e., the common KLT) if it exists, is the optimum FB in $\mathcal{C}^t$ for all these problems. Thus, it is optimum in the mean square sense for the noise-suppression system using any combination of constant multipliers, zeroth order Wiener filters and hard-thresholds (Section 2.2) in the subbands. ■

Theorem 6. Suboptimality of common PCFB. The optimality of the common signal and noise PCFB *for the transform coder class* $\mathcal{C}^t$ (Theorem 5) *does not hold* for all classes of FB's. In particular it is violated for large families of input signal and noise spectra for the class $\mathcal{C}^u$ of unconstrained 2-channel FB's. ■

Theorem 7. Optimality for a restricted class of concave objectives. For any FB class $\mathcal{C}$, the common signal and noise PCFB if it exists is always optimal for a certain well-defined subset of the minimization objectives that are concave over the domain T of (17). There is a finite procedure to identify whether or not a given concave objective falls in this subset. ■

Theorem 7 is easily proved: As long as separate PCFB's exist for the signal and noise, the set T of (17) is a polytope, and a search over the finite set $E_\sigma \times E_\eta$ of its extreme points will yield a minimum $\mathbf{z}_f$ of any concave objective f over T. If $\mathbf{z}_f$ *lies in the true search–space* $S_v \subset T$, then it also minimizes f over S_v, and is in E_c, i.e., corresponds to a common signal and noise PCFB. In general $\mathbf{z}_f$ does not lie in S_v, but the common PCFB minimizes all concave objectives f for which it does, thus proving Theorem 7.

As explained in Section 4.1, we will complete the proof of Theorem 5 (in Section 4.3) by showing that if a common signal and noise KLT exists, $E_c = E_v$ for the class $\mathcal{C}^t$. Section 4.3 also proves Theorem 6, using a specific illustration of PCFB suboptimality. We may also note here another speciality of the class $\mathcal{C}^t$ besides that shown by Theorems 5 and 6: The common signal and noise PCFB for $\mathcal{C}^t$ (i.e., common KLT) is also the PCFB (i.e., KLT) for the noisy FB input $x(n) = s(n) + \mu(n)$. This need not be true for a general FB class $\mathcal{C}$ (for example, for the unconstrained class $\mathcal{C}^u$). For the noise suppression problems, we have already shown in Section 2 and [1], the restriction of Theorem 5 to the case when the noise is white. Even with colored noise, if all subbands use zeroth order Wiener filters, a stronger result is shown in [1]:

Theorem 8. Optimality over biorthogonal class. Consider the (colored) noise suppression system using zeroth order Wiener filters in all subbands. For this scheme, the common signal and noise KLT if it exists, is the mean-square sense optimum memoryless *biorthogonal* transform (i.e., FB as in Fig. 1 with a constant but not necessarily unitary polyphase matrix). ■

However with other subband operations (e.g. constant multipliers or hard-thresholds), it is an open problem as to whether this optimality over the biorthogonal class still holds (even if the noise is white).

The above results show that PCFB optimality for noise suppression is considerably restricted when the noise is colored rather than white. If the PCFB is not optimal, can we find the true optimal FB? We know that searching the extreme point set E_v suffices to find an optimal FB, but in general E_v may be infinite, making analytical solutions difficult. However for one special case involving unconstrained FB's and *piecewise constant spectra*, E_v is finite and easily characterized, as shown by the next result (proved in Section 4.4):

Theorem 9: Optimum unconstrained FB for piecewise constant spectra. Consider the problem of finding within the unconstrained M-channel orthonormal FB class $\mathcal{C}^u$, a FB minimizing an objective function that is concave on the set T of (17). From Section 4.1, this is reducible to a minimization of the objective over the set E_v of extreme points of the convex hull $\mathrm{co}(S_v)$ (where S_v is the search space, defined in Section 4.1). Suppose the input signal and noise are WSS with psd's that are constant on all intervals

$(\frac{2\pi k}{MN}, \frac{2\pi(k+1)}{MN})$ for all integers k for some fixed positive integer N. Then,

1. S_v is a polytope, i.e., $S_v = \mathrm{co}(S_v)$ and E_v is finite. Further, let $\mathcal{F}$ be the set of all *brickwall* FB's in $\mathcal{C}^u$ having all filter band edges at integer multiples of $\frac{2\pi}{MN}$. Then $\mathcal{F}$ has $|\mathcal{F}| = (M!)^N$ FB's, and for each point of E_v there is a FB in $\mathcal{F}$ corresponding to it.

2. For fixed M, though the size of $\mathcal{F}$ is exponential in N, the number of FB's in $\mathcal{F}$ that actually correspond to points in E_v is polynomial: $|E_v| \leq K_M N^{2M-3}$, where $K_M = 4(2M-3)(M!(M!-1)/2)^{2M-3}$. These FB's can be extracted from $\mathcal{F}$ in $C_1 N^{2M-2}(M!)^{4M-5}$ arithmetic operations if $M > 2$ and in $C_2 N$ operations if $M = 2$ (where constants C_1, C_2 are independent of M, N), again polynomial in N. ■

Discussion on Theorem 9

1. *On brickwall orthonormal M channel FB's* [20, 19]: In these FB's, all filters have piecewise constant responses $H_i(e^{j\omega}) \in \{0, \sqrt{M}\}$ for all ω. Their supports are nonoverlapping and aliasfree(M), i.e., for any ω, exactly one of the M numbers $H_i(e^{j(\omega+\frac{2\pi k}{M})})$, $k = 0, 1, \ldots, M-1$ is nonzero. If further all filter band edges (i.e., points of discontinuity of $H_i(e^{j\omega})$) are integer multiples of $\frac{2\pi}{MN}$, the number of such FB's is evidently finite and not hard to compute; our proof (Section 4.4) gives a way to compute it.

2. *Result appeals but is not obvious:* The theorem shows that the optimum FB can always be chosen to lie in $\mathcal{F}$, i.e., to be brickwall with nonoverlapping filter responses having shapes similar to the input spectra (i.e., piecewise constant with the same allowed discontinuities). While intuitively very appealing, this is certainly not obvious; e.g. it is in general not true without the concavity of the objective.

3. *Bounds on $|E_v|$:* Items (1) and (2) of the theorem statement give two different bounds $(M!)^N (= |\mathcal{F}|)$ and $K_M N^{2M-3}$ respectively on the size of E_v. The latter bound is stronger when $N \gg M$, while the former is when $M \gg N$. There are no bounds that are polynomial in both M and N.

4. *Common PCFB's and the case of $N = 1$:* Theorem 9 holds whether or not a common signal and noise PCFB for $\mathcal{C}^u$ exists for the given spectra. If such a PCFB exists, it also corresponds to points of E_v (often it is also in $\mathcal{F}$). However it need not always be optimal (Theorem 6), as E_v could in general have other points as well. In the special case when $N = 1$ however, $|\mathcal{F}| = M!$, and all elements of $\mathcal{F}$ are permutations of the same FB, namely the usual contiguous–stacked brickwall FB, which is hence always optimal. This FB is a common signal and noise PCFB in this case: It produces *white* and totally decorrelated signal and noise subband processes. The comments after the proof of Theorem 5 in Section 4.3 provide an independent proof of the optimality of FB's producing such subband processes.

5. *Approximating optimum FB's for arbitrary spectra:* Most spectra can be approximated by the piecewise constant ones in the premise of Theorem 9, to any desired accuracy by sufficiently increasing M

and/or N. Thus Theorem 9 in principle allows approximation of the optimum FB in $\mathcal{C}^u$ for any input spectra to any desired accuracy. However the complexity of the algorithm for this is polynomial in N but super–exponential in M. Thus, we have good algorithms for low M (especially $M = 2$, where the complexity is linear in N). For sufficiently large M, we get good enough approximations of the true spectra by taking $N = 1$. The earlier remark then gives at no cost, the optimum FB in $\mathcal{C}^u$, i.e., the usual contiguous–stacked brickwall FB. There are no good algorithms if both M and N are large.

4.3 Proof and comments on Theorems 5,6

Proof of Theorem 5. Using the notations and discussion of Section 4.1, we need to show that for the transform coder class $\mathcal{C}^t$, $E_c = E_v$ whenever a common signal and noise PCFB (KLT) exists. Let $\mathbf{R}_\sigma, \mathbf{R}_\eta$ be the autocorrelation matrices of the M-fold blocked versions of the signal $s(n)$ and noise $\mu(n)$ respectively. Let the unitary $\mathbf{K}$ be a common KLT. Its subband signal and noise variance vectors are thus $\mathbf{z}_\sigma = \text{diag}(\mathbf{\Lambda}_\sigma)$ and $\mathbf{z}_\eta = \text{diag}(\mathbf{\Lambda}_\eta)$ respectively, where $\mathbf{\Lambda}_\sigma = \mathbf{K}\mathbf{R}_\sigma\mathbf{K}^\dagger$, $\mathbf{\Lambda}_\eta = \mathbf{K}\mathbf{R}_\eta\mathbf{K}^\dagger$ are both diagonal. The set of points in S_v corresponding to the KLT $\mathbf{K}$ and its permutations is thus

$$E'_c = \left\{ \begin{bmatrix} \mathbf{P}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_j \end{bmatrix} \begin{bmatrix} \mathbf{z}_\sigma \\ \mathbf{z}_\eta \end{bmatrix} \quad : j = 1, 2, \ldots, M! \right\} \tag{18}$$

where $\mathbf{P}_j$ are the $M \times M$ permutation matrices. Now E_c is the set of points in S_v corresponding to any common KLT, so $E'_c \subset E_c$. (It will turn out that $E'_c = E_c$, but this needs proof due to possible nonuniqueness of the KLT $\mathbf{K}$.) We now compute S_v: Note that $(\mathbf{v}_\sigma^T, \mathbf{v}_\eta^T)^T \in S_v$ iff there is a transform coder producing $\mathbf{v}_\sigma, \mathbf{v}_\eta$ as signal and noise subband variance vectors respectively, i.e., iff there is a unitary matrix $\mathbf{T}$ such that $\text{diag}(\mathbf{T}\mathbf{R}_\sigma\mathbf{T}^\dagger) = \mathbf{v}_\sigma = \text{diag}(\mathbf{T}\mathbf{K}^\dagger\mathbf{\Lambda}_\sigma\mathbf{K}\mathbf{T}^\dagger)$ and $\text{diag}(\mathbf{T}\mathbf{R}_\eta\mathbf{T}^\dagger) = \mathbf{v}_\eta = \text{diag}(\mathbf{T}\mathbf{K}^\dagger\mathbf{\Lambda}_\eta\mathbf{K}\mathbf{T}^\dagger)$. Let $\mathbf{Q}$ be the orthostochastic matrix [1],[7] corresponding to $\mathbf{T}\mathbf{K}^\dagger$, i.e., the doubly stochastic matrix formed by replacing each entry of the unitary $\mathbf{T}\mathbf{K}^\dagger$ by the square of its absolute value. Then $\mathbf{v}_\sigma = \mathbf{Q}\mathbf{z}_\sigma$ and $\mathbf{v}_\eta = \mathbf{Q}\mathbf{z}_\eta$. Thus

$$S_v = \left\{ \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathbf{z}_\sigma \\ \mathbf{z}_\eta \end{bmatrix} : \mathbf{Q} \text{ orthostochastic} \right\} \subseteq \left\{ \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \mathbf{z}_\sigma \\ \mathbf{z}_\eta \end{bmatrix} : \mathbf{Q} \text{ doubly stochastic} \right\} \triangleq A. \tag{19}$$

By Birkhoff's theorem [1],[7] we can express $\mathbf{Q}$ above as a convex combination of permutation matrices, thus obtaining $A = \text{co}(E'_c)$. Since $E'_c \subset S_v \subseteq A = \text{co}(E'_c)$, we have $\text{co}(E'_c) \subseteq \text{co}(S_v) \subseteq \text{co}(A) = \text{co}(E'_c)$, i.e., $\text{co}(S_v) = \text{co}(E'_c)$, which is thus a polytope whose extreme points lie in E'_c. But E_v is by definition the set of these extreme points, so $E_v \subset E'_c$. Together with $E'_c \subset E_c \subset E_v$, this gives $E_c = E_v$ as desired. $\nabla\nabla\nabla$

We may note here that the set S_σ of realizable subband signal variance vectors $\mathbf{v}_\sigma$ is convex [1], and that

$$S_\sigma = \{\mathbf{Q}\mathbf{z}_\sigma : \mathbf{Q} \text{ orthostochastic}\} = \{\mathbf{Q}\mathbf{z}_\sigma : \mathbf{Q} \text{ doubly stochastic}\} = \text{co}(S_\sigma). \tag{20}$$

Is S_v convex too? For dimension $M \leq 2$, every doubly stochastic matrix is orthostochastic [1]. So from (19), $S_v = A = \text{co}(S_v)$, i.e., S_v is indeed convex, as we also verify in Section 4.5 by explicitly computing S_v. Even

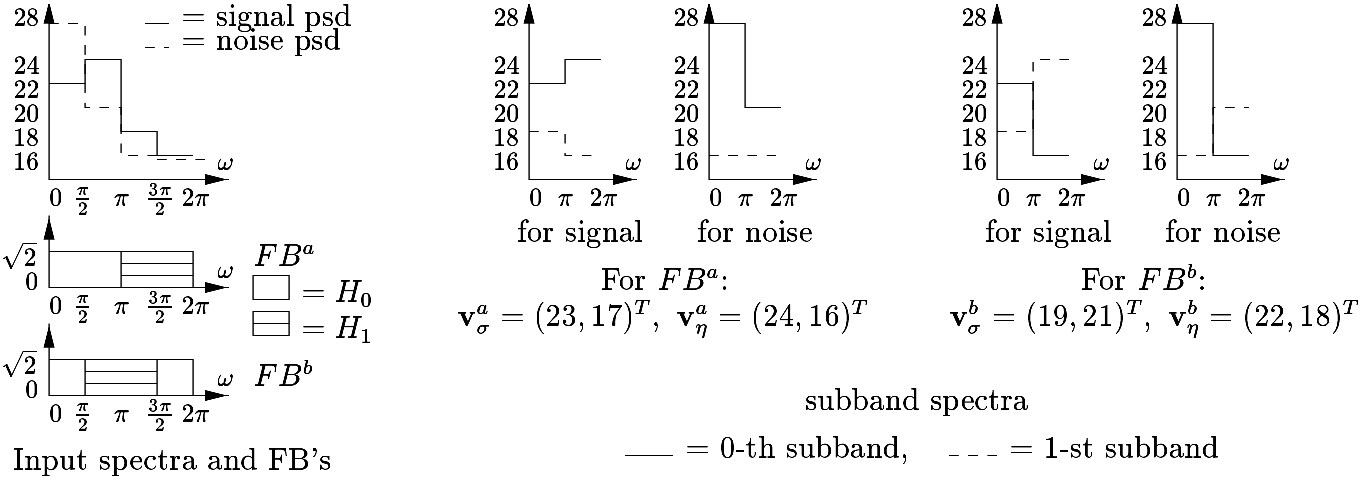


Figure 7: **Suboptimality of common unconstrained signal and noise PCFB.**

for general M, the same argument that proves convexity of S_σ also shows that S_v is convex in two very special cases: (1) if all entries of $\mathbf{z}_\eta$ (or $\mathbf{z}_\sigma$) are equal, i.e., $\mathbf{R}_\eta$ (respectively, $\mathbf{R}_\sigma$) is the identity matrix upto scale – the 'white noise' case, and (2) if $\mathbf{z}_\eta = c\mathbf{z}_\sigma$ (i.e., $\mathbf{R}_\eta = c\mathbf{R}_\sigma$). However if $M > 2$, S_v is *not* convex for several pairs of values of $\mathbf{z}_\sigma, \mathbf{z}_\eta$ (some shown in Appendix C).

We can try to modify the above proof to show that $E_c = E_v$ for the class $\mathcal{C}^u$ too. To do this we replace the autocorrelation matrices $\mathbf{R}_\sigma, \mathbf{R}_\eta$ with psd matrices $\mathbf{S}_\sigma(e^{j\omega}), \mathbf{S}_\eta(e^{j\omega})$ and try to use the earlier arguments at each ω. We cannot complete the proof for *all* psd matrices, for else a common signal and noise PCFB would always be optimal for the class $\mathcal{C}^u$ too, contradicting Theorem 6. However, we can in fact complete the proof for some restricted classes of psd's: (1) If $\mathbf{S}_\eta(e^{j\omega})$ (or $\mathbf{S}_\sigma(e^{j\omega})$) is the identity matrix upto scale – the 'white noise' case, (2) if $\mathbf{S}_\eta(e^{j\omega}) = c\mathbf{S}_\sigma(e^{j\omega})$, and (3) if the diagonalized versions of $\mathbf{S}_\sigma(e^{j\omega}), \mathbf{S}_\eta(e^{j\omega})$ are both *constant* (independent of ω). We have seen cases (1) and (2) earlier, as situations where a signal PCFB is automatically also a noise PCFB and minimizes all concave objectives of the form (15). In case (3), the common PCFB for $\mathcal{C}^u$ has *white* and uncorrelated subband signal and noise components. Examples of this case are (a) if $\mathbf{S}_\sigma(e^{j\omega}), \mathbf{S}_\eta(e^{j\omega})$ are themselves independent of ω– the PCFB's for $\mathcal{C}^u$ are then the corresponding KLT's, and (b) if $N = 1$ in Theorem 9 – the common PCFB for $\mathcal{C}^u$ is then the usual contiguous–stacked brickwall FB.

Proof of Theorem 6. We provide a specific example of PCFB suboptimality. For the class $\mathcal{C}^u$ of unconstrained 2-channel FB's, consider the input signal and noise spectra and the two FB's from $\mathcal{C}^u$ shown in Fig. 7. The figure also shows the resulting subband spectra and signal and noise variance vectors. As the analysis filters are non-overlapping, the subbands are totally decorrelated. From Fig. 7, the subbands of FB^a also obey spectral majorization (Section 2.3), while those of FB^b do not. Thus FB^a is a common signal and noise PCFB while FB^b is neither a signal PCFB nor a noise PCFB for the class $\mathcal{C}^u$. However, consider the concave objectives of the noise suppression problem with either zeroth order Wiener filters or hard-thresholds in *both* subbands (see (15),(16)). By evaluation using the subband variances in Fig. 7, FB^b achieves a lower value than FB^a for these objectives. Thus the common PCFB is not always optimal. More examples of PCFB suboptimality can be created by slight perturbations of the spectra of Fig. 7. $\bigtriangledown\bigtriangledown\bigtriangledown$

The spectra in Fig. 7 are piecewise constant and Theorem 9 can be applied to them (with $M = N = 2$). This shows that one of the two FB's (FB^a and FB^b) in Fig. 7 is optimum in $\mathcal{C}^u$ for any concave minimization objective. Thus, in the example in the proof of Theorem 6, not only is FB^b better than the common signal and noise PCFB (FB^a), but it is in fact the best possible two–channel (unconstrained orthonormal) FB.

4.4 Proof of Theorem 9

Let $H_i(e^{j\omega}), i = 0, 1, \ldots, M-1$ be the analysis filters of a M-channel orthonormal FB (i.e., a FB from $\mathcal{C}^u$). For $i = 0, 1, \ldots, M-1$ and $k = 0, 1, \ldots, MN-1$, define

$$f_{ik} \triangleq \frac{N}{2\pi} \int_{\frac{2\pi k}{MN}}^{\frac{2\pi(k+1)}{MN}} \left|H_i(e^{j\omega})\right|^2 d\omega. \tag{21}$$

Let the constant values of the input signal and noise psd's $D_\sigma(e^{j\omega})$ and $D_\eta(e^{j\omega})$ on the interval $(\frac{2\pi k}{MN}, \frac{2\pi(k+1)}{MN})$ be a_k, b_k respectively. Let σ_i^2, η_i^2 respectively be the signal and noise variances in the i-th subband. Then

$$\sigma_i^2 = \frac{1}{2\pi} \int_0^{2\pi} \left|H_i(e^{j\omega})\right|^2 D_\sigma(e^{j\omega})\, d\omega = \frac{1}{N} \sum_{k=0}^{MN-1} f_{ik} a_k. \tag{22}$$

Similarly $N\eta_i^2 = \sum_{k=0}^{MN-1} f_{ik} b_k$. Thus all subband variances are linear functions of the f_{ik}. So the search space S_v (Section 4.1) is the image under a linear transformation of the set of all possible arrays f_{ik} corresponding to all FB's in $\mathcal{C}^u$. Hence we now proceed to study this set. By FB orthonormality, from [19],

$$\sum_{i=0}^{M-1} \left|H_i(e^{j\omega})\right|^2 \;=\; M \quad \text{(power complementarity), and} \tag{23}$$

$$\sum_{k=0}^{M-1} \left|H_i(e^{j(\omega + \frac{2\pi k}{M})})\right|^2 \;=\; M \quad \text{for } i = 0, 1, \ldots, M-1 \text{ (Nyquist}(M)\text{ constraint), hence} \tag{24}$$

$$0 \le f_{ik} \le 1 \qquad \text{for all } i, k \text{ (for which } f_{ik} \text{ is defined)}, \tag{25}$$

$$\sum_{i=0}^{M-1} f_{ik} = 1 \qquad \text{for all } k, \tag{26}$$

$$\sum_{k=0}^{M-1} f_{i(l+Nk)} = 1 \qquad \text{for all } i, \text{ for each } l = 0, 1, \ldots, N-1. \tag{27}$$

Here (25) follows from $0 \le \left|H_i(e^{j\omega})\right|^2 \le M$ for all i, ω (which follows from (23) or (24)), while (26),(27) follow from (23),(24) respectively. Define the $M \times M$ matrix $\mathbf{G}^{(l)}$ to have entries $g_{ik}^{(l)} = f_{i(l+Nk)}$ (where $i, k \in \{0, 1, \ldots, M-1\}$), for $l = 0, 1, \ldots, N-1$. Then (25),(26),(27) are equivalent to the following:

$$\mathbf{G}^{(l)} \text{ is } \textbf{doubly stochastic} \text{ for all } l = 0, 1, \ldots, N-1. \tag{28}$$

Let $\mathcal{G}$ be the collection of all ordered sets $(\mathbf{G}^{(0)}, \mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(N-1)})$ corresponding to all FB's in $\mathcal{C}^u$. Instead of studying the set of all arrays f_{ik}, we can study $\mathcal{G}$ (as S_v is also the image of $\mathcal{G}$ under a linear transform). Let $\mathcal{Q}, \mathcal{P}$ respectively denote the sets of all $M \times M$ doubly stochastic matrices and permutation matrices. From (28), $\mathcal{G} \subset \mathcal{Q}^N$ $(= \mathcal{Q} \times \mathcal{Q} \times \ldots \times \mathcal{Q})$.

Claim: $\mathcal{G} = \mathcal{Q}^N$, which (by Birkhoff's theorem [1, 7]) is a polytope with $\mathcal{P}^N$ as its set of extreme points. Also, FB's in the set $\mathcal{F}$ (defined in stating Theorem 9) correspond directly (one–to–one) with points in $\mathcal{P}^N$.

Showing this claim will prove item (1) in the statement of Theorem 9. Recall that S_v is the image of $\mathcal{G}$ under a linear transform $\mathcal{L}$. So if $\mathcal{G}$ is a polytope, so is S_v; further all its extreme points are images of some extreme points of $\mathcal{G}$ under $\mathcal{L}$. The claim above thus means that there is a FB in $\mathcal{F}$ for every extreme point of S_v. The correspondence between $\mathcal{F}$ and $\mathcal{P}^N$ also means that $\mathcal{F}$ has $|\mathcal{F}| = |\mathcal{P}^N| = |\mathcal{P}|^N = (M!)^N$ FB's (counting separately all permutations of each FB in $\mathcal{F}$ – else we must divide the number by $M!$).

Proof of Claim: We show that $\mathcal{G} = \mathcal{Q}^N$ by building a brickwall FB in $\mathcal{C}^u$ corresponding to any given $\underline{\mathbf{G}} = (\mathbf{G}^{(0)}, \mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(N-1)}) \in \mathcal{Q}^N$. To do this, let σ_m, $m = 0, 1, \ldots, M! - 1$ be the $M!$ permutation functions on the set $\{0, 1, \ldots, M-1\}$. Now there is a one–to–one correspondence between brickwall FB's and functions ϕ mapping each $\omega \in [0, \frac{2\pi}{M})$ to one of the σ_m. This is described by the following construction of the analysis filters $H_i(e^{j\omega}), i = 0, 1, \ldots, M-1$ of the FB given the function ϕ: Let $\omega \in [0, \frac{2\pi}{M})$ and $\phi(\omega) = \sigma_m$. Then $H_{\sigma_m(k)}(e^{j(\omega + \frac{2\pi k}{M})}) = \sqrt{M}$ for $k = 0, 1, \ldots, M-1$. In other words, the permutation $\sigma_m = \phi(\omega)$ decides which of the M filter responses is nonzero at the M frequencies $\omega + \frac{2\pi k}{M}$. The construction ensures nonoverlapping aliasfree(M) filter responses resulting in a valid FB in $\mathcal{C}^u$. Now for each $l = 0, 1, \ldots, N-1$, let $x_m^{(l)}$ be the fraction of length of the interval $[\frac{2\pi l}{MN}, \frac{2\pi(l+1)}{MN})$ that is mapped by ϕ to σ_m, for $m = 0, 1, \ldots, M!-1$. For a brickwall FB, f_{ik} of (21) is the fraction of length of the interval $[\frac{2\pi k}{MN}, \frac{2\pi(k+1)}{MN})$ on which $H_i(e^{j\omega})$ is nonzero (i.e., $= \sqrt{M}$). Thus, the chosen ϕ yields a FB corresponding to the given $\underline{\mathbf{G}} \in \mathcal{Q}^N$ (i.e., given set of f_{ik} obeying (25),(26),(27)) iff for $i, k = 0, 1, \ldots, M-1$ and $l = 0, 1, \ldots, N-1$ we have

$$\sum_{\text{all } m \text{ obeying } \sigma_m(k) = i} x_m^{(l)} \quad = \quad f_{i(l+Nk)} \ \ (\stackrel{\triangle}{=} g_{ik}^{(l)}) \tag{29}$$

Thus, given $\underline{\mathbf{G}}$, we must find $x_m^{(l)} \in [0, 1]$ obeying (29). This is easy if $\underline{\mathbf{G}} \in \mathcal{P}^N$: Here for each l, $\mathbf{G}^{(l)}$ (with entries $g_{ik}^{(l)}$) is a permutation matrix, i.e., there is a $m_*(l)$ such that $g_{ik}^{(l)}$ is 1 if $\sigma_{(m_*(l))}(k) = i$ and is 0 otherwise. We then simply set $x_m^{(l)}$ to be 1 for $m = m_*(l)$ and 0 for all other m. Note that this yields a FB in the set $\mathcal{F}$ defined in stating Theorem 9. For a general $\underline{\mathbf{G}} \in \mathcal{Q}^N$, we use Birkhoff's theorem to write $\mathbf{G}^{(l)} \in \mathcal{Q}$ as a convex combination of elements of $\mathcal{P}$. The same convex combination of the solution vectors $(x_0^{(l)}, x_1^{(l)}, \ldots, x_{M!-1}^{(l)})$ corresponding to each element of $\mathcal{P}$ yields the corresponding solution vector for $\mathbf{G}^{(l)}$. Repeating the process for $l = 0, 1, \ldots, N-1$ completes the solution. This shows that $\mathcal{G} = \mathcal{Q}^N$, a polytope with $\mathcal{P}^N$ as its set of extreme points. The proof has also associated to each of these extreme points a unique FB in $\mathcal{F}$. Conversely for any FB in $\mathcal{F}$, the f_{ik} of (21), and hence all entries of the doubly stochastic matrices $\mathbf{G}^{(l)}$, are either 0 or 1. Hence the corresponding point in $\mathcal{G}$ is in $\mathcal{P}^N$. This proves the one–to–one correspondence between $\mathcal{F}$ and $\mathcal{P}^N$. $\bigtriangledown\bigtriangledown\bigtriangledown$

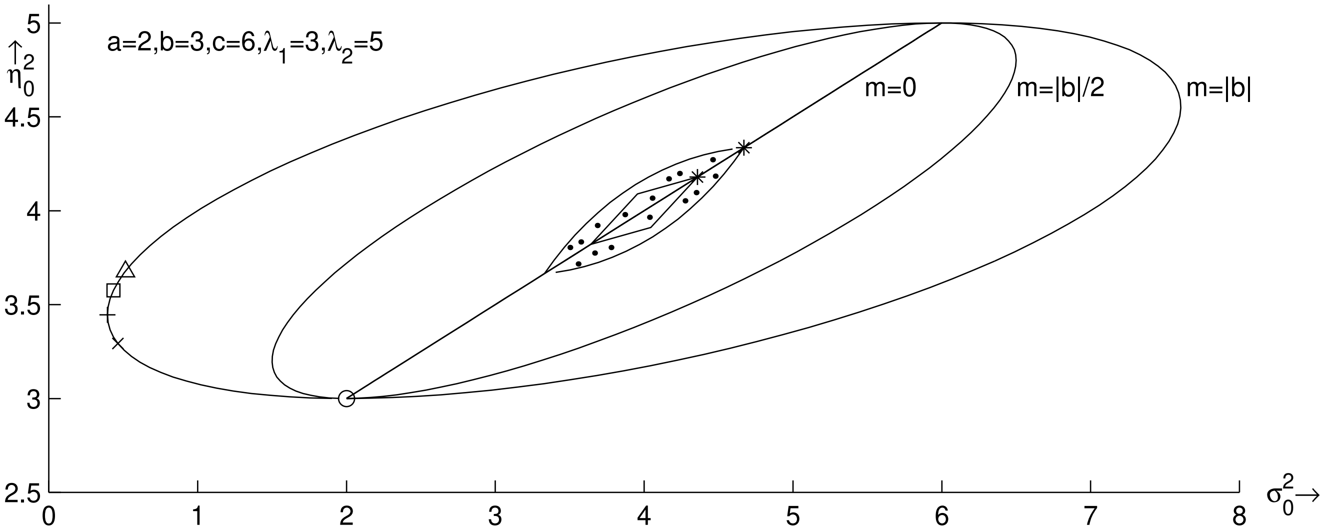
Proof of item (2) of Theorem 9 statement: By (22), the map $\mathcal{L}$ from $\underline{\mathbf{G}} = (\mathbf{G}^{(0)}, \mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(N-1)}) \in \mathcal{G}$ to the associated point in S_v is given by $\mathcal{L}(\underline{\mathbf{G}}) = \sum_{l=0}^{N-1} \begin{pmatrix} \mathbf{G}^{(l)}\mathbf{a}^{(l)} \\ \mathbf{G}^{(l)}\mathbf{b}^{(l)} \end{pmatrix}$, where $\mathbf{a}^{(l)} = (a_l, a_{l+N}, \ldots, a_{l+(M-1)N})^T/N$, $\mathbf{b}^{(l)} = (b_l, b_{l+N}, \ldots, b_{l+(M-1)N})^T/N$. For any fixed $l = 0, 1, \ldots, N-1$, as $\mathbf{G}^{(l)}$ can be any element of $\mathcal{Q}$, the set of possible values of $\begin{pmatrix} \mathbf{G}^{(l)}\mathbf{a}^{(l)} \\ \mathbf{G}^{(l)}\mathbf{b}^{(l)} \end{pmatrix}$ is itself a polytope $T^{(l)}$. It lies on a $2M-2$ dimensional hyperplane in $\mathcal{R}^{2M}$, and its extreme points correspond to the $M!$ possible choices of $\mathbf{G}^{(l)} \in \mathcal{P}$. Thus, $S_v = \left\{\sum_l \mathbf{x}^{(l)} \mid \mathbf{x}^{(l)} \in T^{(l)}\right\}$, which is known as the *Minkowski sum* of the polytopes $T^{(l)}$. Minkowski sums have been well studied in computational geometry [6],[15], e.g. in context of robot motion planning algorithms in 2 and 3 dimensions [15]. Theorem 2.1.10 and Corollary 2.1.11 of [6] bound the number of extreme points of the Minkowski sum of k polytopes of dimension d with not more than p extreme points each. Theorem 2.3.7′ and Proposition 2.3.9 of [6], with their proofs, outline algorithms to find the extreme points of this Minkowski sum, thus bounding the number of arithmetic operations needed for the same. Applying these bounds with $k = N$, $d = 2M - 2$ and $p = M!$ yields item (2) of the statement of Theorem 9. $\bigtriangledown\bigtriangledown\bigtriangledown$

Note that like S_v, the set S_σ of realizable signal subband variance vectors is also the image of $\mathcal{G}$ under a linear map $\mathcal{L}_\sigma$ given by $\mathcal{L}_\sigma(\underline{\mathbf{G}}) = \sum_{l=0}^{N-1} \mathbf{G}^{(l)}\mathbf{a}^{(l)}$. However, while $\mathcal{G}$ has $(M!)^N$ extreme points (i.e., points in $\mathcal{P}^N$), and S_v has $|E_v| \leq K_M N^{2M-3}$ of them, we know from [1] that S_σ has at most $M!$ of them – its extreme points are the permutations of the signal PCFB subband variance vector $\mathbf{v}_*$. Indeed, here $\mathbf{v}_* = \mathcal{L}_\sigma(\underline{\mathbf{G}})$ when each $\mathbf{G}^{(l)}$ is a permutation matrix rearranging the entries of $\mathbf{a}^{(l)}$ in decreasing order. It is not hard to see (by definition of majorization) that $\mathbf{v}_*$ majorizes all points $\mathcal{L}_\sigma(\underline{\mathbf{G}})$ for $\underline{\mathbf{G}} \in \mathcal{P}^N$ (i.e., for all choices of $\underline{\mathbf{G}}$ as extreme points of $\mathcal{G}$). Hence [1] all these points are some convex combinations of the permutations of $\mathbf{v}_*$. Thus these permutations are the only extreme points of S_σ. Note that S_σ too is expressible as a Minkowski sum of polytopes $T_\sigma^{(l)}$, where $T_\sigma^{(l)}$ is the set of all permutations of $\mathbf{a}^{(l)}$. Using [6] to bound the number of extreme points of S_σ gives a bound that grows with both M and N, whereas the true number is independent of N. The tightening of the bound of [6] has been achieved by exploiting the special structure of the summand polytopes $T_\sigma^{(l)}$. The summands $T^{(l)}$ of S_v also have a special structure (different from that of the $T_\sigma^{(l)}$), but it is not clear whether the bound on $|E_v|$ can be similarly tightened using this structure.

4.5 Study of S_v for two-channel transform coders

This section explicitly computes for the class $\mathcal{C}^t$ of two-channel transform coders, the set $S_v \subset \mathcal{R}^4$ of all realizable pairs of subband signal and noise variance vectors $(\sigma_0^2, \sigma_1^2)^T$ and $(\eta_0^2, \eta_1^2)^T$ respectively. Since

$$\sigma_0^2 + \sigma_1^2 = k_\sigma \quad \text{and} \quad \eta_0^2 + \eta_1^2 = k_\eta \tag{30}$$



Δ = optimum for hard thresholding, $\square$ = optimum for Wiener filtering.
$+$ = signal KLT, $\times$ = signal + noise KLT, $\circ$ = noise KLT.
= search space for which common PCFB exists ($*$), but is not always optimal.

Figure 8: **Search–space $S_{\sigma\eta}$ for 2-channel transform coders.**

are constants (independent of choice of FB from $\mathcal{C}^t$), it is enough to study instead the set $S_{\sigma\eta} \subset \mathcal{R}^2$ of all realizable pairs $(\sigma_0^2, \eta_0^2)^T$. Unlike S_v, the set $S_{\sigma\eta}$ can be plotted, and (30) directly relates the properties of S_v and $S_{\sigma\eta}$. For example S_v is convex iff $S_{\sigma\eta}$ is convex, there is an obvious correspondence between the extreme points of the sets $\text{co}(S_v)$ and $\text{co}(S_{\sigma\eta})$, and the permutation symmetry of S_v is equivalently restated as $(\sigma_0^2, \eta_0^2)^T \in S_{\sigma\eta} \Rightarrow (k_\sigma - \sigma_0^2, k_\eta - \eta_0^2)^T \in S_{\sigma\eta}$. The result of computation of $S_{\sigma\eta}$ is summarized as follows:

Theorem 10. Search space for two-channel transform coders. Consider the class $\mathcal{C}^t$ of two-channel transform coders, and the associated set $S_{\sigma\eta} \subset \mathcal{R}^2$ defined above. If a common signal and noise PCFB (KLT) exists for $\mathcal{C}^t$, then $S_{\sigma\eta}$ is a **line-segment** whose endpoints correspond to the common PCFB. Otherwise, it is an **elliptical disk**. ■

Discussion. When a common KLT exists, $\text{co}(S_{\sigma\eta})$, and hence $\text{co}(S_v)$, is a polytope ($\text{co}(S_{\sigma\eta})$ is a 1–dimensional polytope, i.e., a line-segment). Further the extreme points of the polytope are precisely the points corresponding to the common KLT. This corroborates for two-channel FB's, the result $E_c = E_v$ proved in Section 4.3 for any transform coder class with a common signal and noise KLT. Recall (from Section 4.1) that this result was the key to the optimality of the common KLT (Theorem 5). Also note that $S_{\sigma\eta} = \text{co}(S_{\sigma\eta})$, i.e., $S_{\sigma\eta}$ is convex, and hence so is S_v. This also was independently proved earlier for *two-channel* transform coders, though it does not always hold with more than two channels (Section 4.3).

If there is no common KLT, $S_{\sigma\eta}$ is an elliptical disk – a compact convex set whose extreme points are the points on its elliptical boundary. Thus S_v is a compact convex set with infinitely many extreme points. The minima over S_v of different concave objectives are at different extreme points. Figure 8 shows a plot of $S_{\sigma\eta}$; the parameters a, b, c, etc. are constants depending on the input spectra (defined shortly). The individual signal and noise KLT's are extreme points of $S_{\sigma\eta}$ – respectively the points at which the vertical and horizontal tangents to the disk $S_{\sigma\eta}$ touch it. This verifies a general fact: The individual signal PCFB for any class of FB's corresponds to boundary points of $\text{co}(S_v)$, and further if it uniquely defines the subband noise variance vector, it actually corresponds to an extreme point of $\text{co}(S_v)$. However, the individual signal and noise KLT's need not be optimum: Figure 8 shows that different concave objectives yield different minima, all lying on the disk boundary. The figure also shows contrived examples of FB classes for which common signal and noise PCFB's exist but do not minimize all concave objectives. The classes are defined as sets of all FB's in $\mathcal{C}^t$ whose variance pairs $(\sigma_0^2, \eta_0^2)^T$ lie in well–chosen subsets of $S_{\sigma\eta}$, marked as dotted areas in the figure. Note that these subsets obey the required permutation symmetry property. These examples are artificial due to the choice of these subsets, and also because the FB class definition depends on the input spectra.

Proof of Theorem 10. Let the input signal and noise autocorrelation matrices $\mathbf{R}_\sigma, \mathbf{R}_\eta$, and a general element $\mathbf{T} \in \mathcal{C}^t$ (i.e., a general 2×2 unitary matrix), respectively be given by

$$\mathbf{R}_\sigma = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix}, \quad \mathbf{R}_\eta = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} \cos(\theta) & \sin(\theta)e^{-j\phi} \\ \sin(\theta)e^{j\phi} & -\cos(\theta) \end{bmatrix}. \tag{31}$$

Here $a, c, \lambda_1, \lambda_2 \geq 0$ and $ac \geq |b|^2$ as $\mathbf{R}_\sigma, \mathbf{R}_\eta$ are positive semidefinite. By initially passing the noisy input through the KLT for the noise, $\mathbf{R}_\eta$ can be assumed diagonal without loss of generality. A common signal and noise KLT exists iff one (or both) of the following hold: (1) $\mathbf{R}_\sigma$ is diagonal too, i.e., $b = 0$, or (2) $\mathbf{R}_\eta$ is the identity matrix upto scale (so that any unitary matrix diagonalizes it), i.e., $\lambda_1 = \lambda_2$ (e.g. this happens with white input noise). Also in (31), $\theta, \phi \in [0, 2\pi)$, and the unitary $\mathbf{T}$ is fully general upto multiplication by a *diagonal* unitary matrix, which does not affect its subband variances. By direct computation, the subband signal and noise variance vectors $(\sigma_0^2, \sigma_1^2)^T = \text{diag}(\mathbf{T}\mathbf{R}_\sigma\mathbf{T}^\dagger)$ and $(\eta_0^2, \eta_1^2)^T = \text{diag}(\mathbf{T}\mathbf{R}_\eta\mathbf{T}^\dagger)$ respectively are:

$$\begin{pmatrix} \sigma_0^2 \\ \sigma_1^2 \end{pmatrix} = \begin{pmatrix} a\cos^2(\theta) + c\sin^2(\theta) + \sin(2\theta)\text{Re}(be^{j\phi}) \\ a\sin^2(\theta) + c\cos^2(\theta) - \sin(2\theta)\text{Re}(be^{j\phi}) \end{pmatrix}, \quad \begin{pmatrix} \eta_0^2 \\ \eta_1^2 \end{pmatrix} = \begin{pmatrix} \lambda_1\cos^2(\theta) + \lambda_2\sin^2(\theta) \\ \lambda_1\sin^2(\theta) + \lambda_2\cos^2(\theta) \end{pmatrix}. \tag{32}$$

Note that $\mathbf{T}$ is the signal KLT iff $be^{j\phi}$ is real (i.e., $|\text{Re}(be^{j\phi})|$ is maximized) and the choice of θ then maximizes (or minimizes) σ_0^2. Of course $\mathbf{T}$ is the noise KLT iff it is diagonal or antidiagonal, i.e., iff $\cos(\theta)\sin(\theta) = 0$.

From (32), $S_{\sigma\eta}$ is the set of all $(\sigma_0^2, \eta_0^2)^T$ satisfying for some θ, ϕ, the equation

$$\begin{pmatrix} \sigma_0^2 \\ \eta_0^2 \end{pmatrix} - \mathbf{e} = \mathbf{A}_\phi \begin{pmatrix} \cos(2\theta) \\ \sin(2\theta) \end{pmatrix}, \quad \text{where} \quad \mathbf{A}_\phi = \frac{1}{2}\begin{bmatrix} a - c & 2\text{Re}(be^{j\phi}) \\ \lambda_1 - \lambda_2 & 0 \end{bmatrix}, \quad \mathbf{e} = \frac{1}{2}\begin{pmatrix} a + c \\ \lambda_1 + \lambda_2 \end{pmatrix}. \tag{33}$$

For each fixed ϕ, let $\mathcal{I}_\phi$ be the set of vectors in $\mathcal{R}^2$ given by the right side as θ varies. Then $S_{\sigma\eta}$ is the union of these sets $\mathcal{I}_\phi$ as ϕ varies, with origin shifted to $\mathbf{e}$. As $\mathbf{e}$ is constant, it suffices to prove Theorem 10 replacing $S_{\sigma\eta}$ by the union $\bigcup_\phi \mathcal{I}_\phi$. From (33), $\mathcal{I}_\phi$ is the image of the unit circle under a linear map $\mathbf{A}_\phi$. So $\mathcal{I}_\phi$ is a line segment with midpoint at the origin if $\mathbf{A}_\phi$ is singular, and an ellipse centered at the origin otherwise. Suppose a common signal and noise KLT exists, i.e., $\lambda_1 = \lambda_2$ or $b = 0$ (or both). Then $\mathbf{A}_\phi$ is singular for all ϕ. If $\lambda_1 = \lambda_2$, $\mathcal{I}_\phi$ is horizontal, while if $b = 0$, it lies along the line $(\lambda_1 - \lambda_2)x = (a-c)y$, for all ϕ. So in either case, $\bigcup_\phi \mathcal{I}_\phi$ is a line-segment with midpoint at the origin. Its endpoints correspond to extremum (maximum or minimum) values of both σ_0^2 and η_0^2, i.e., to the common KLT.

Now suppose a common signal and noise KLT does not exist. Then $\mathcal{I}_\phi$ is an ellipse centered at the origin for general ϕ. It degenerates into a line segment for exactly two values of ϕ in $[0, 2\pi)$ at which $\mathrm{Re}(be^{j\phi}) = 0$, i.e., $\mathbf{A}_\phi$ is singular. To compute $\bigcup_\phi \mathcal{I}_\phi$, we write (using (33)) the nonparametric equation of the ellipse $\mathcal{I}_\phi$:

$$\left(\frac{x - (\frac{a-c}{\lambda_1-\lambda_2})y}{m}\right)^2 + \left(\frac{2y}{\lambda_1 - \lambda_2}\right)^2 = 1, \quad \text{where} \quad m = \mathrm{Re}(be^{j\phi}) \ \text{and} \ \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sigma_0^2 \\ \eta_0^2 \end{pmatrix} - \mathbf{e}. \tag{34}$$

This shows that (1) the ellipses for $+m$ and $-m$ are the same, (2) the ellipse for $m_1 \geq 0$ lies inside that for $m_2 > m_1$, and (3) for every point in the interior of the ellipse for $m_2 > 0$ there is another ellipse for $m_1 \geq 0$ (for some $m_1 < m_2$) passing through it (the 'ellipse' for $m = 0$ is the line segment $x = \frac{a-c}{\lambda_1-\lambda_2}y$ with endpoints having $y = \pm|\lambda_1 - \lambda_2|/2$). Since the range of values of m is $[-|b|, |b|]$, we conclude that $\bigcup_\phi \mathcal{I}_\phi$ is an elliptical disk whose boundary is the ellipse corresponding to $m = |b|$. $\triangledown\triangledown\triangledown$

For the present example, with certain concave objectives of the form (15), it is easy to explicitly compute the optimum FB $\mathbf{T}$ of (31), by inserting the variances of (32) into the objective and analytically optimizing θ and ϕ. For example it can be done for noise suppression using constant multipliers in both subbands (see (16)). This will verify that the optimum FB indeed corresponds to a boundary point of $S_{\sigma\eta}$, and further that the common signal and noise KLT if it exists is indeed optimum.

5 Conclusion

We have extended the study of principal component filter banks [1] in many ways. A central theme in our analysis is to study the geometry of the relevant search spaces of realizable subband variances, and to exploit concavity of the minimization objective on these spaces. However, many interesting issues are still unresolved.

An important question is whether there are any useful classes of FB's for which PCFB's exist for all (or large families of) input spectra. Indeed it seems possible that the two-channel, the transform coder and the unconstrained classes may be the only such classes (ruling out contrived situations where the class definition depends on the input spectrum). However this has not been proved. Analytical study of PCFB existence and FB optimization for classes of FIR FB's has proven to be very complicated. The problem stated in most general form is as follows: Given a class of orthonormal FB's, find all input spectra for which a PCFB exists.

Regarding the FIR classes, we could reach a partial solution by solving the following problem: Find a family of input spectra for which there is no PCFB for some general FIR class, say that of all FB's with a given bound N on the McMillan degree or order of the polyphase matrix. At present, a few such results are known for specific low values of the bound N, for isolated input spectra [13],[11]. Even in these cases, the proofs of PCFB nonexistence need numerical optimizations. Further one of these, from [11], is suspect due to the assumption that the FB maximizing its largest subband variance must contain a FIR compaction filter. Some insight may possibly be obtained by analytical computation of the search-spaces for simple examples of these classes (e.g. the class of all 3-channel FIR FB's with polyphase matrices of McMillan degree unity).

Another area of open issues involves biorthogonal FB's. The compression and noise reduction systems of this paper remain well-defined if the FB used is biorthogonal rather than orthonormal; however the FB optimization objective no longer depends purely on the subband variances. We have seen certain cases where the best orthonormal FB is also the best biorthogonal one. For example, the KLT is not only the best orthogonal transform but also the best memoryless biorthogonal one for both the high bitrate coding problem with optimal bit allocation and for noise reduction with Wiener filters in all subbands. However it is not known whether this is true with other subband operations, e.g. low bitrate coding and noise reduction by hard-thresholds. For the unconstrained biorthogonal FB class, even for the high bitrate coding problem the best FB was known only in certain cases [21] until recently when [14] has claimed a full solution.

With regard to noise suppression, we have considered only Wiener filters of order $N = 0$ in the subbands. If $N > 0$, the objective depends not only on the subband variances but also on other coefficients in the autocorrelation sequences of the subband processes. In this case, analytical results on the optimum FB are not known. The performance gain due to increasing the order of the subband Wiener filters could instead be obtained by using a FB with more channels, however the exact nature of this tradeoff is not known.

Appendix A: Compactness of search–space

Here we justify the assumption of Section 3.2 that the search–space $\mathcal{S}$ is compact, i.e., closed and bounded. (In fact we already know [1] that it is bounded.) Many FB classes $\mathcal{C}$ are parameterized by a vector of real numbers, that is free to take any values in a set P which may be called the parameter space. It often happens that P is compact, and that for any bounded non-impulsive input spectrum, there is a continuous function mapping parameter vectors (from P) to the subband variance vectors (in $\mathcal{S}$) produced by the corresponding FB. Thus $\mathcal{S}$ is the continuous image of the compact set P, and is hence compact. This reasoning works for example when $\mathcal{C}$ is the set of all FIR orthonormal M-channel FB's with a given McMillan degree – Here $\mathcal{C}$ is parameterized by a finite set of unit norm vectors in $\mathcal{R}^M$ and a unitary matrix [19]. Thus P is compact, being the Cartesian product of finitely many sphere-surfaces in $\mathcal{R}^M$ and the set of $M \times M$ unitary matrices.

Appendix B: Concavity proofs for some functions in the paper

$f(\mathbf{v}) = -d(\mathbf{v}, P)$ *(Section 3.4)*: Continuity of f follows from that of the norm. To show concavity of f, we must show that $d(\mathbf{z}, P) \leq \alpha d(\mathbf{x}, P) + (1-\alpha)d(\mathbf{y}, P)$, where $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, for $0 \leq \alpha \leq 1$, $\mathbf{x}, \mathbf{y} \in \mathcal{R}^M$. Let $\mathbf{a}, \mathbf{b}$ be the points in P that are closest to $\mathbf{x}, \mathbf{y}$ respectively. (They exist because P is compact.) Thus

$$\alpha d(\mathbf{x}, P) + (1-\alpha)d(\mathbf{y}, P) = \alpha \|\mathbf{x} - \mathbf{a}\| + (1-\alpha)\|\mathbf{y} - \mathbf{b}\| \geq \|\alpha(\mathbf{x} - \mathbf{a}) + (1-\alpha)(\mathbf{y} - \mathbf{b})\| = \|\mathbf{z} - \mathbf{c}\|$$

where $\mathbf{c} = \alpha\mathbf{a} + (1-\alpha)\mathbf{b} \in P$ since P is convex. Thus $\|\mathbf{z} - \mathbf{c}\| \geq d(\mathbf{z}, P)$, which completes the proof.

Functions of (16): Linear functions and the minimum of concave functions are concave [16], so $f_i(x, y) = |1-k_i|^2 x + |k_i|^2 y$ and $f_i(x, y) = \min(x, y)$ are concave on $\mathcal{R}^2$. For $f_i(x, y) = \frac{xy}{x+y}$, we need to show that

$$\frac{[\alpha x + (1-\alpha)a]\,[\alpha y + (1-\alpha)b]}{\alpha x + (1-\alpha)a + \alpha y + (1-\alpha)b} \geq \alpha\left(\frac{xy}{x+y}\right) + (1-\alpha)\left(\frac{ab}{a+b}\right) \quad \text{when} \quad 0 \leq \alpha \leq 1. \tag{35}$$

If $x, y, a, b \geq 0$, by cross-multiplying and defining $L = \alpha^2 xy + (1-\alpha)^2 ab$, this is equivalent to proving that

$$[L + \alpha(1-\alpha)(xb + ya)]\,(x+y)(a+b) \geq [\alpha xy(a+b) + (1-\alpha)ab(x+y)]\,[\alpha(x+y) + (1-\alpha)(a+b)]\,.$$

The right side is $L(x+y)(a+b) + \alpha(1-\alpha)\left[ab(x+y)^2 + xy(a+b)^2\right]$, hence as $\alpha \in [0, 1]$, it suffices to show that $(xb+ya)(x+y)(a+b) \geq ab(x+y)^2 + xy(a+b)^2$, i.e., (expanding and simplifying) that $x^2b^2 + y^2a^2 - 2xyab = (xb - ya)^2 \geq 0$, which is true. Thus $f_i(x, y) = \frac{xy}{x+y}$ arising in colored noise reduction (see (16)) is concave on $\mathcal{R}_+^2$. However it is not *strictly* concave as equality holds in (35) when $xb = ya$. Note that fixing x (or y) in $f_i(x, y)$ yields univariate functions that appear in white noise reduction and are *strictly* concave on $\mathcal{R}_+$ [1].

Appendix C: Nonconvexity of search–space S_v (Section 4.3)

For $M = 3$, let $\mathbf{Q}_* = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}$ which is doubly stochastic but not orthostochastic [1]. Let $\mathbf{z}_\sigma = (a, b, c)^T, \mathbf{z}_\eta = (k, 0, 0)^T$ where $k(b-c) \neq 0$. By (19), $\mathbf{v}_* \triangleq \begin{bmatrix} \mathbf{Q}_* & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_* \end{bmatrix} \begin{bmatrix} \mathbf{z}_\sigma \\ \mathbf{z}_\eta \end{bmatrix} = \begin{bmatrix} \mathbf{v}_\sigma \\ \mathbf{v}_\eta \end{bmatrix} \in A = \mathrm{co}(S_v)$, where $\mathbf{v}_\sigma = 0.5(a+b, a+c, b+c)^T, \mathbf{v}_\eta = k(0.5, 0.5, 0)^T$. Now $\mathbf{v}_* \in S_v$ iff $\mathbf{Q}\mathbf{z}_\sigma = \mathbf{v}_\sigma$ and $\mathbf{Q}\mathbf{z}_\eta = \mathbf{v}_\eta$ for some orthostochastic $\mathbf{Q}$; but it can be verified that a doubly stochastic $\mathbf{Q}$ satisfies these equations iff $\mathbf{Q} = \mathbf{Q}_*$ which is not orthostochastic. So $\mathbf{v}_* \notin S_v$, proving that S_v is not convex. A somewhat more restricted class of pairs $\mathbf{z}_\sigma = (a, 0, \ldots, 0)^T$, $\mathbf{z}_\eta = (0, b, 0, \ldots, 0)^T$ (with $ab \neq 0$) also produces nonconvex S_v *for any* $M > 2$. To show this we use the earlier argument replacing $\mathbf{Q}_*$ by $\mathbf{Q}_{**} = \begin{bmatrix} \mathbf{Q}_* & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$, where $\mathbf{I}$ is the identity and the $\mathbf{0}$'s are zero matrices of suitable sizes. Here a doubly stochastic $\mathbf{Q}$ satisfying $\mathbf{Q}\mathbf{z}_\sigma = \mathbf{v}_\sigma$ and $\mathbf{Q}\mathbf{z}_\eta = \mathbf{v}_\eta$ need not be $\mathbf{Q}_{**}$, but must agree with it in the first two columns. This already prevents it from being orthostochastic.

References

[1] Sony Akkarakaran and P.P.Vaidyanathan, "Filter Bank Optimization with Convex Objectives, and the Optimality of Principal Component Forms," *IEEE Trans. Sig. Proc.,* vol. 49, pp. 100-114, Jan. 2001.

[2] ——, "The best basis problem, compaction problem and PCFB design problems," in *Proc. IEEE ISCAS,* Orlando, FL, June 1999.

[3] ——, "Optimized Orthonormal Transforms for SNR Improvement by Subband Processing," in *Proc. IEEE Wkshop SP Adv. Wireless Comm.,* Annapolis, MD, May 1999.

[4] D.P.Bertsekas, "The Auction Algorithm for Assignment and other Network Flow Problems: A Tutorial," *Interfaces,* vol. 20, no. 4, pp. 133-149, July–August 1990.

[5] D.L.Donoho and I.M.Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika,* vol. 81, no. 3, pp. 425-455, 1994.

[6] P.Gritzmann and B.Sturmfels, "Minkowski Addition of Polytopes: Computational Complexity and Applications to Grobner Bases," *SIAM J. Discrete Math,* vol. 6, no. 2, pp. 246-269, May 1993.

[7] R.A.Horn and C.R.Johnson, *Matrix Analysis.* Cambridge University Press, 1985.

[8] Y.Huang and P.M.Schultheiss, "Block quantization of correlated Gaussian random variables," *IEEE Trans. Commun. Syst.,* vol. C-10, pp. 289-296, Sept. 1963.

[9] I.Kalet, "The multitone channel," *IEEE Trans. Comm.,* vol. 37, no. 2, pp. 119-124, Feb. 1989.

[10] A.Kirac and P.P.Vaidyanathan, "Optimality of Orthonormal Transforms for Subband Coding," in *Proc. IEEE DSP Workshop,* Utah, August 1998.

[11] ——, "On Existence of FIR Principal Component Filter Banks," in *Proc. ICASSP,* Seattle, May 1998.

[12] ——, "Theory and Design of Optimum FIR Compaction Filters," *IEEE Trans. Signal Processing,* vol. 46, no. 4, pp. 903-919, April 1998.

[13] P.Moulin and M.K.Mihcak, "Theory and Design of Signal-Adapted FIR Paraunitary Filter Banks," *IEEE Trans. Signal Processing,* vol. 46, no. 4, pp. 920-929, April 1998.

[14] P.Moulin, M.Anitescu, and K.Ramchandran, "Theory of Rate-Distortion Optimal, Constrained Filter Banks – Application to IIR and FIR Biorthogonal Designs," *IEEE Trans. Signal Processing,* vol. 48, no. 4, pp. 1120-1132, April 2000.

[15] J.O'Rourke, *Computational Geometry in C.* Cambridge University Press, 1998.

[16] R.T.Rockafellar, *Convex Analysis.* Princeton University Press, 1970.

[17] M.K.Tsatsanis and G.B.Giannakis, "Principal Component Filter Banks for Optimal Multiresolution Analysis," *IEEE Trans. Signal Processing,* vol. 43, no. 8, pp. 1766-1777, Aug. 1995.

[18] M.Unser, "An extension of the KLT for wavelets and perfect reconstruction filter banks," in *Proc. SPIE no. 2034, Wavelet Appl. Signal Image Processing,* San Diego, CA, 1993, pp.45-56.

[19] P.P.Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[20] ——, "Theory of Optimal Orthonormal Subband Coders," *IEEE Trans. Signal Processing,* vol. 46, no. 6, pp. 1528-1543, June 1998.

[21] P.P.Vaidyanathan and A.Kirac, "Results on Optimal Biorthogonal Filter Banks," *IEEE Trans. Circuits and Systems II,* vol. 45, no. 8, pp. 932-947, August 1998.

[22] P.P.Vaidyanathan, Y-P.Lin, S.Akkarakaran and S-M. Phoong, "Optimality of principal component filter banks for discrete multitone communication systems," in *Proc. IEEE ISCAS,* Geneva, May 2000.

LIST OF FIGURES

FOOTNOTES

1. Manuscript received ____________________

2. Work supported in part by the National Science Foundation under Grant MIP 0703755.

3. The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

Biography of Sony Akkarakaran

Sony Akkarakaran was born in Thrissur, India in 1975. He received the B.Tech degree from the Indian Institute of Technology, Bombay, in 1996, and the M.S. degree from the California Institute of Technology in 1997, both in Electrical Engineering. He is currently working towards the Ph.D. degree in the field of digital signal processing at the California Institute of Technology. His research interests are multirate systems, wavelets, and their communications applications.

Biography of P. P. Vaidyanathan

P. P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India on Oct. 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, all from the University of Calcutta, India, in 1974, 1977 and 1979, respectively, and the Ph.D degree in electrical and computer engineering from the University of California at Santa Barbara in 1982. He was a post doctoral fellow at the University of California, Santa Barbara from Sept. 1982 to March 1983. In March 1983 he joined the electrical engineering department of the Calfornia Institute of Technology as an Assistant Professor, and since 1993 has been Professor of electrical engineering there. His main research interests are in digital signal processing, multirate systems, wavelet transforms and adaptive filtering.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International symposium on Circuits and Systems. He was an Associate editor for the IEEE Transactions on Circuits and Systems for the period 1985-1987, and is currently an associate editor for the journal *IEEE Signal Processing letters,* and a consulting editor for the journal *Applied and computational harmonic analysis.* He has been a guest editor in 1998 for special issues of the IEEE Trans. on Signal Processing and the IEEE Trans. on Circuits and Systems II, on the topics of filter banks, wavelets and subband coders. Dr. Vaidyanathan has authored a number of papers in IEEE journals, and is the author of the book *Multirate systems and filter banks.* He has written several chapters for various signal processing handbooks. He was a recepient of the Award for excellence in teaching at the California Institute of Technology for the years 1983-1984, 1992-93 and 1993-94. He also received the NSF's Presidential Young Investigator award in 1986. In 1989 he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990 he was recepient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the IETE journal. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE SP Transactions, for which the first author (Truong Nguyen) received the *Young outstanding author* award in 1993. Dr. Vaidyanathan was elected Fellow of the IEEE in 1991. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education, especially the book *Multirate systems and filter banks* published by Prentice Hall in 1993. He has given several plenary talks including at the Eusipco'98, Asimolar'88 and SPCOM'95 conferences on signal processing. He has been chosen a distinguished lecturer for the IEEE Signal Processing Society for the year 1996-97. In 1999 he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal.