J. Franklin Inst., special issue on Time-Frequency Signal Analysis and Applications, vol. 337, no. 4, July 2000, pp. 403–419

Minimax Robust Nonstationary Signal Estimation Based on a p-Point Uncertainty Model*

Gerald Matz and Franz Hlawatsch

Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Vienna, Austria
phone: +43 1 58801 38916, fax: +43 1 58801 38999
e-mail: gmatz@aurora.nt.tuwien.ac.at
web: http://www.nt.tuwien.ac.at/dspgroup/time.html

March 22, 2000

Abstract

We propose a time-varying Wiener filter for nonstationary signal estimation that is robust in a minimax sense. This robust Wiener filter optimizes worst-case performance within novel "p-point" uncertainty classes of nonstationary random processes. Furthermore, it features constant performance within these uncertainty classes and requires less detailed prior knowledge than the ordinary time-varying Wiener filter. We also propose a time-frequency formulation that is intuitively appealing since signal subspaces are replaced by time-frequency regions, and an efficient on-line implementation using local cosine bases. Our theory is illustrated by numerical simulations and a real-data example.

1 Introduction

The estimation of signals corrupted by noise (or other disturbances) is important in many applications. If the signal and noise processes are nonstationary, the linear signal estimator that is optimal with respect to a mean-square error criterion is a time-varying filter (termed *time-varying Wiener filter* in the following). The design of this time-varying Wiener filter uses the second-order statistics of the nonstationary processes involved which, however, are rarely exactly known in practice. Unfortunately, errors in estimating or modeling the second-order statistics can result in severe performance degradation.

In this paper, we propose a *minimax robust* time-varying Wiener filter that is advantageous in two respects. First, it is more practical since the prior knowledge required for its design is less detailed.

*Funding by FWF grant P11904-TEC.

Second, it is less sensitive to deviations from the nominal operating conditions—in fact, it maintains constant performance for all second-order statistics within prescribed uncertainty classes. Our results extend minimax robust time-*in*variant Wiener filters for *stationary* processes [1, 2] to the nonstationary case. For simplicity, we restrict ourselves to a nonstationary version of the so-called *p-point uncertainty model*, although other uncertainty models can be used as well (see [3] for a generalized formulation).

The paper is organized as follows. Based on a novel p-point uncertainty model for nonstationary random processes introduced in Section 2, the minimax robust time-varying Wiener filter is derived in Section 3. Section 4 presents intuitively appealing time-frequency formulations of both the uncertainty model and the corresponding robust Wiener filter, and Section 5 proposes an efficient causal implementation using local cosine bases. Finally, numerical simulations and a real-data (speech enhancement) example are provided in Section 6.

We shall first review some necessary background on ordinary time-varying Wiener filters. We consider the problem of estimating a signal $s(t)$ from a noisy observation $r(t) = s(t) + n(t)$. We assume that both signal $s(t)$ and noise $n(t)$ are nonstationary, zero-mean, real or circularly complex random processes with respective correlation operators[1] $\mathbf{R}_s$ and $\mathbf{R}_n$. Furthermore, we assume that signal and noise are mutually uncorrelated and have finite mean energy, i.e., $\bar{E}_s = \mathrm{E}\{\|s\|^2\} < \infty$ and $\bar{E}_n = \mathrm{E}\{\|n\|^2\} < \infty$ (this implies that the correlation operators $\mathbf{R}_s$ and $\mathbf{R}_n$ are trace-class (nuclear) [4, 5]). A signal estimate $\hat{s}(t)$ is obtained by applying a linear, generally time-varying system $\mathbf{H}$ to the observation $r(t)$, i.e.,

$$\hat{s}(t) = (\mathbf{H}\, r)(t) = \int_{-\infty}^{\infty} h(t,t')\, r(t')\, dt',$$

where $h(t,t')$ denotes the impulse response (kernel) of $\mathbf{H}$. The ordinary Wiener filter, denoted $\mathbf{H}_W$, is the system $\mathbf{H}$ that minimizes the mean-square error (MSE) defined as[2]

$$e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n) \triangleq \mathrm{E}\big\{\|\hat{s} - s\|^2\big\} = \int_{-\infty}^{\infty} \mathrm{E}\big\{|\hat{s}(t) - s(t)|^2\big\}\, dt\,.$$

It is well known [6–8] that, in the absence of causality constraints,[3]

$$\mathbf{H}_W \triangleq \arg\min_{\mathbf{H}}\; e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n) = \mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}\,. \tag{1}$$

The minimal MSE achieved by the Wiener filter $\mathbf{H}_W$ is given by

$$e_{\min}(\mathbf{R}_s, \mathbf{R}_n) \triangleq e(\mathbf{H}_W; \mathbf{R}_s, \mathbf{R}_n) = \mathrm{Tr}\big\{\mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}\mathbf{R}_n\big\}\,, \tag{2}$$

[1] The correlation operator $\mathbf{R}_x$ of a (generally nonstationary) random process $x(t)$ is the positive (semi-)definite linear operator whose kernel equals the correlation $r_x(t,t') = \mathrm{E}\,\{x(t)\, x^*(t')\}$. In a discrete-time setting, $\mathbf{R}_x$ would be a matrix.

[2] While minimization of $e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n)$ is equivalent to minimization of the "instantaneous MSE" $\mathrm{E}\{|\hat{s}(t) - s(t)|^2\}$, it results in a more symmetric formulation and solution of the minimax problem (see Section 3).

[3] If $\mathbf{R}_s + \mathbf{R}_n$ is not invertible, the inverse $(\mathbf{R}_s + \mathbf{R}_n)^{-1}$ has to be replaced by the pseudo-inverse $(\mathbf{R}_s + \mathbf{R}_n)^{\#}$ [8].

where $\mathrm{Tr}\{\cdot\}$ denotes the trace [4, 5]. Note that the design of the time-varying Wiener filter requires complete knowledge of the correlations $\mathbf{R}_s$ and $\mathbf{R}_n$. Errors in estimating or modeling $\mathbf{R}_s$ and $\mathbf{R}_n$ will result in deviations of the actual correlations $\mathbf{R}_s$ and $\mathbf{R}_n$ from the nominal correlations (hereafter denoted by $\mathbf{R}_s^0$ and $\mathbf{R}_n^0$) for which the Wiener filter $\mathbf{H}_W^0 = \mathbf{R}_s^0(\mathbf{R}_s^0 + \mathbf{R}_n^0)^{-1}$ was designed. Unfortunatly, such deviations may cause the filter's performance to degrade significantly.

2 Nonstationary p-Point Uncertainty Model

In order to model the designer's uncertainty about the actual correlations, we next propose an *uncertainty model* for nonstationary random processes [9, 10]. This uncertainty model generalizes the so-called p-point uncertainty model for stationary processes considered in [1, 2] to the nonstationary case. An approximate time-frequency reformulation of this model will be presented in Section 4.

The novel "nonstationary" p-point uncertainty model is based on a partition of the space $L_2(\mathbb{R})$ of square-integrable functions into N mutually orthogonal subspaces $\mathcal{X}_i$ with $i = 1, \ldots, N$:

$$\mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \cdots \oplus \mathcal{X}_N = L_2(\mathbb{R}) \qquad \text{and} \qquad \mathcal{X}_i \perp \mathcal{X}_j \;\; \text{for} \;\; i \neq j\,.$$

Here, the symbol $\oplus$ denotes the direct sum of subspaces. Consequently, the orthogonal projection operators[4] $\mathbf{P}_i$ on the subspaces $\mathcal{X}_i$ form a resolution of the identity,

$$\sum_{i=1}^{N} \mathbf{P}_i = \mathbf{I} \qquad \text{and} \qquad \mathbf{P}_i\mathbf{P}_j = \mathbf{0} \;\; \text{for} \;\; i \neq j\,.$$

The mean energy of a (finite-energy) random process $x(t)$ in a subspace $\mathcal{X}_i$ is given by

$$\mathrm{E}\{\|\mathbf{P}_i\, x\|^2\} = \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_x\mathbf{P}_i\} = \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_x\}\;.$$

We now define the *p-point uncertainty class* $\mathcal{U}$ as the set of all correlations $\mathbf{R}_x$ with prescribed mean subspace energies $x_i \geq 0$,

$$\mathcal{U} \triangleq \left\{\mathbf{R}_x\colon\; \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_x\} = x_i\,,\;\; i = 1, \ldots, N\right\}. \tag{3}$$

Hence, $\mathcal{U}$ is characterized by the space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$ (equivalently, $\{\mathbf{P}_i\}_{i=1,\ldots,N}$) and the subspace energies x_i. Note that for any $\mathbf{R}_x \in \mathcal{U}$, we obtain with $\mathbf{I} = \sum_{i=1}^{N} \mathbf{P}_i$

$$\bar{E}_x = \mathrm{Tr}\{\mathbf{R}_x\} = \mathrm{Tr}\left\{\sum_{i=1}^{N} \mathbf{P}_i\,\mathbf{R}_x\right\} = \sum_{i=1}^{N} \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_x\} = \sum_{i=1}^{N} x_i\,.$$

This shows that the mean energies of all processes $x(t)$ (i.e., correlations $\mathbf{R}_x$) belonging to $\mathcal{U}$ are equal.

[4] We recall that orthogonal projection operators are idempotent, $\mathbf{P}^2 = \mathbf{P}$, and self-adjoint, $\mathbf{P}^+ = \mathbf{P}$ [4, 5].

The p-point uncertainty model corresponds to a prior knowledge that consists of the mean subspace energies x_i relative to a given space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$. In general, this prior knowledge is far less detailed than the full-blown correlation $\mathbf{R}_x$. It is equivalent to $\mathbf{R}_x$ (and, thus, there is no uncertainty) in the extreme case where the $\mathcal{X}_i$ are one-dimensional spaces, each of which is spanned by one of the eigenfunctions of $\mathbf{R}_x$; the subspace energies x_i are then equal to the eigenvalues of $\mathbf{R}_x$. The opposite extreme case is $N = 1$, i.e., $\mathcal{X}_1 = L_2(\mathbb{R})$; here, only the total mean energy $\bar{E}_x$ is determined, corresponding to a minimal amount of prior knowledge and a maximally wide uncertainty class $\mathcal{U}$.

If a nominal correlation $\mathbf{R}_x^0$ is available, the mean subspace energies can be determined as $x_i \triangleq \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_x^0\}$. In most practical situations, however, the prior knowledge must be estimated. Here, the p-point model (3) proves useful since estimation of the N numbers x_i is much simpler and statistically more stable than estimation of the complete correlation $\mathbf{R}_x$. In particular, the subspace energy estimator $\hat{x}_i = \langle \mathbf{P}_i\, x, x\rangle$ is unbiased and its relative variance can be shown to be approximately inversely proportional to the dimension of the corresponding subspace [10, 11], i.e., $\mathrm{var}\{\hat{x}_i\}/x_i^2 \sim 1/\dim\{\mathcal{X}_i\}$. Hence, a cruder space partition (i.e., larger subspace dimensions) yields smaller estimation variance.

The convex combination $\mathbf{R}_\alpha = (1-\alpha)\mathbf{R}_1 + \alpha\mathbf{R}_2$ of two correlations $\mathbf{R}_1 \in \mathcal{U}$, $\mathbf{R}_2 \in \mathcal{U}$ is again in $\mathcal{U}$:

$$\mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_\alpha\} \;=\; (1-\alpha)\,\mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_1\} + \alpha\,\mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_2\} \;=\; (1-\alpha)\,x_i + \alpha\,x_i \;=\; x_i\,.$$

Thus, the p-point uncertainty class $\mathcal{U}$ is a *convex* set. This will be important in what follows.

3 Robust Time-Varying Wiener Filter

Let us return to our estimation problem defined in Section 1. Rather than assuming that $\mathbf{R}_s$ and $\mathbf{R}_n$ are completely known, we hereafter assume that $\mathbf{R}_s$ and $\mathbf{R}_n$ belong to p-point uncertainty classes

$$\mathcal{S} \;\triangleq\; \big\{\mathbf{R}_s\colon\; \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_s\} = s_i\,,\;\; i = 1,\ldots,N\big\}\,, \tag{4a}$$

$$\mathcal{N} \;\triangleq\; \big\{\mathbf{R}_n\colon\; \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_n\} = n_i\,,\;\; i = 1,\ldots,N\big\}\,, \tag{4b}$$

respectively, with a prescribed space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$ (identical for $\mathcal{S}$ and $\mathcal{N}$) and prescribed mean subspace energies $s_i \geq 0$ and $n_i \geq 0$. In general, there will not exist a single system $\mathbf{H}$ that uniformly minimizes the MSE $e(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n)$ for *all* correlations $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}$ simultaneously. Hence, we shall adopt a minimax approach and look for the system $\mathbf{H}$ that has optimal worst-case performance within the uncertainty classes $\mathcal{S}$ and $\mathcal{N}$. We thus define the *robust time-varying Wiener filter* $\mathbf{H}_R$ as[5] [3, 9, 10]

$$\mathbf{H}_R \;\triangleq\; \arg\min_{\mathbf{H}}\, e_{\max}(\mathbf{H})\,, \qquad \text{with } e_{\max}(\mathbf{H}) \;\triangleq\; \max_{\substack{\mathbf{R}_s\in\mathcal{S}\\ \mathbf{R}_n\in\mathcal{N}}} e(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n)\,. \tag{5}$$

[5] While no causality constraint is imposed in this definition, we will show in Section 5 that causality of $\mathbf{H}_R$ can be obtained by appropriate choice of the subspace partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$.

Initially, this optimization problem is difficult to solve since in general no explicit expression for $e_{\max}(\mathbf{H})$ is available. However, an important simplification is possible since—as we will show in a moment—minimization with respect to $\mathbf{H}$ and maximization with respect to $\mathbf{R}_s$, $\mathbf{R}_n$ can be interchanged, i.e.,

$$\min_{\mathbf{H}} \max_{\substack{\mathbf{R}_s \in \mathcal{S} \\ \mathbf{R}_n \in \mathcal{N}}} e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n) = \max_{\substack{\mathbf{R}_s \in \mathcal{S} \\ \mathbf{R}_n \in \mathcal{N}}} \min_{\mathbf{H}} e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n) \equiv \max_{\substack{\mathbf{R}_s \in \mathcal{S} \\ \mathbf{R}_n \in \mathcal{N}}} e_{\min}(\mathbf{R}_s, \mathbf{R}_n)\,, \tag{6}$$

with an explicit expression for $e_{\min}(\mathbf{R}_s, \mathbf{R}_n)$ given by (2). But $e_{\min}(\mathbf{R}_s, \mathbf{R}_n)$ is achieved by the ordinary Wiener filter $\mathbf{H}_W = \mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}$ in (1). Hence, (6) implies that $\mathbf{H}_R$ equals the *ordinary* Wiener filter

$$\mathbf{H}_W^L \triangleq \mathbf{R}_s^L \big(\mathbf{R}_s^L + \mathbf{R}_n^L\big)^{-1}$$

obtained for those correlations $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ that are *least favorable* in the sense that they lead to the maximal $e_{\min}(\mathbf{R}_s, \mathbf{R}_n)$ among all $\mathbf{R}_s \in \mathcal{S}$ and $\mathbf{R}_n \in \mathcal{N}$, i.e.,

$$(\mathbf{R}_s^L, \mathbf{R}_n^L) = \arg \max_{\substack{\mathbf{R}_s \in \mathcal{S} \\ \mathbf{R}_n \in \mathcal{N}}} e_{\min}(\mathbf{R}_s, \mathbf{R}_n)\,. \tag{7}$$

Note that $\mathbf{R}_s^L$ and $\mathbf{R}_n^L$ correspond to those processes within $\mathcal{S}$ and $\mathcal{N}$, respectively that are hardest to separate in a mean-square sense.

We have thus shown that $\mathbf{H}_R = \mathbf{H}_W^L = \mathbf{R}_s^L \big(\mathbf{R}_s^L + \mathbf{R}_n^L\big)^{-1}$ and the original minimax problem (5) has been reduced to the maximization problem (7). This maximization problem is convex since $e_{\min}(\mathbf{R}_s, \mathbf{R}_n)$ is convex in $\mathbf{R}_s$, $\mathbf{R}_n$ [10] and the uncertainty classes $\mathcal{S}$, $\mathcal{N}$ were shown in Section 2 to be convex as well.

Before proceeding with our derivation of $\mathbf{H}_R = \mathbf{H}_W^L$, we shall comment on the pivotal relation (6). It can be shown [12] that (6) is valid if and only if there exist a filter $\mathbf{H}_L$ and correlations $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ forming a *saddle point* of $e(\mathbf{H}; \mathbf{R}_s, \mathbf{R}_n)$ in the sense that

$$e(\mathbf{H}_L; \mathbf{R}_s, \mathbf{R}_n) \;\le\; e(\mathbf{H}_L; \mathbf{R}_s^L, \mathbf{R}_n^L) \;\le\; e(\mathbf{H}; \mathbf{R}_s^L, \mathbf{R}_n^L) \qquad \text{for all } \mathbf{H} \text{ and for all } \mathbf{R}_s \in \mathcal{S},\ \mathbf{R}_n \in \mathcal{N}\,. \tag{8}$$

If $\mathbf{H}_L$ is chosen as $\mathbf{H}_L = \mathbf{H}_W^L = \mathbf{R}_s^L(\mathbf{R}_s^L + \mathbf{R}_n^L)^{-1}$, i.e., as the ordinary Wiener filter for the correlations $\mathbf{R}_s^L$, $\mathbf{R}_n^L$, the right-hand inequality in (8) is trivially true (since $\mathbf{H}_W^L$ minimizes $e(\mathbf{H}; \mathbf{R}_s^L, \mathbf{R}_n^L)$), and it only remains to find correlation operators $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ satisfying the left-hand inequality in (8):

$$e(\mathbf{H}_W^L; \mathbf{R}_s, \mathbf{R}_n) \;\le\; e(\mathbf{H}_W^L; \mathbf{R}_s^L, \mathbf{R}_n^L) \;\equiv\; e_{\min}(\mathbf{R}_s^L, \mathbf{R}_n^L)\,. \tag{9}$$

This inequality means that the filter $\mathbf{H}_W^L$ performs worst for those correlations $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ for which it was designed, i.e., deviations within $\mathcal{S}, \mathcal{N}$ of the true correlations $\mathbf{R}_s$, $\mathbf{R}_n$ from $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ cannot increase the MSE. Using the convexity of the sets $\mathcal{S}$ and $\mathcal{N}$, it can be shown [3, 10] that (9) is satisfied if and only if $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ are chosen as the least favorable correlations defined in (7). Hence, we have found $\mathbf{H}_L$ and $\mathbf{R}_s^L, \mathbf{R}_n^L$ satisfying the saddle-point inequalities (8), and therefore (6) is valid.

Finally, based on the intermediate result $\mathbf{H}_R = \mathbf{H}_W^L = \mathbf{R}_s^L \big(\mathbf{R}_s^L + \mathbf{R}_n^L\big)^{-1}$ with $\mathbf{R}_s^L, \mathbf{R}_n^L$ defined in (7), the following simple explicit expressions for $\mathbf{H}_R$ and $e(\mathbf{H}_R; \mathbf{R}_s, \mathbf{R}_n)$ [9, 10] are proved in the Appendix.

Theorem 1. *For the p-point uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ in* (4)*, the robust time-varying Wiener filter* $\mathbf{H}_R$ *as defined in* (5) *is given by*

$$\mathbf{H}_R = \sum_{i=1}^{N} \frac{s_i}{s_i + n_i} \mathbf{P}_i \,, \tag{10}$$

and the MSE achieved by $\mathbf{H}_R$ *for any* $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}$ *is given by*

$$e(\mathbf{H}_R; \mathbf{R}_s, \mathbf{R}_n) = \sum_{i=1}^{N} \frac{s_i\, n_i}{s_i + n_i} \,.$$

We see that $\mathbf{H}_R$ simply forms a weighted sum of the orthogonal projections of the input signal onto the subspaces $\mathcal{X}_i$. Thus, it treats all signal components lying in a given subspace $\mathcal{X}_i$ alike and it does not exploit cross-correlations between process components in different subspaces $\mathcal{X}_i$. Remarkably, the MSE achieved by $\mathbf{H}_R$ does not depend on the actual correlation operators $\mathbf{R}_s$ and $\mathbf{R}_n$ as long as these lie in $\mathcal{S}$ and $\mathcal{N}$, respectively. Hence, $\mathbf{H}_R$ achieves constant performance within the uncertainty classes. Whereas the ordinary Wiener filter $\mathbf{H}_W$ need not even be a normal operator, the robust Wiener filter $\mathbf{H}_R$ is always self-adjoint and nonnegative definite. Furthermore, its operator norm can be shown to be

$$\|\mathbf{H}_R\|_O \triangleq \sup_{\|x\|=1} \|\mathbf{H}_R\, x\| = \max_i \left\{ \frac{s_i}{s_i + n_i} \right\} \leq 1 \,.$$

If $N = 1$, i.e., only $\bar{E}_s$ and $\bar{E}_n$ are known, $\mathbf{H}_R = \frac{\bar{E}_s}{\bar{E}_s + \bar{E}_n} \mathbf{I}$ which corresponds to a simple gain factor.

4 Time-Frequency Formulation

The minimax robust nonstationary Wiener filter derived above is formulated in terms of operators and signal subspaces, which is rather abstract. Therefore, we shall next develop an approximate but intuitively appealing time-frequency (TF) formulation of the p-point uncertainty classes $\mathcal{S}$, $\mathcal{N}$ and the robust time-varying Wiener filter $\mathbf{H}_R$ [9, 10]. This TF formulation uses the *Weyl symbol* of a linear time-varying system $\mathbf{H}$ [13–15],

$$L_{\mathbf{H}}(t, f) \triangleq \int_{-\infty}^{\infty} h\Big(t + \frac{\tau}{2},\, t - \frac{\tau}{2}\Big)\, e^{-j2\pi f \tau}\, d\tau$$

(here, t and f denote time and frequency, respectively), and the *Wigner-Ville spectrum* (WVS) of a nonstationary random process $x(t)$ [16–19],

$$\overline{W}_x(t, f) \triangleq L_{\mathbf{R}_x}(t, f) = \int_{-\infty}^{\infty} r_x\Big(t + \frac{\tau}{2},\, t - \frac{\tau}{2}\Big)\, e^{-j2\pi f \tau}\, d\tau \,.$$

Note that the WVS is in one-to-one correspondence with the correlation $\mathbf{R}_x$ and thus it constitutes a complete description of the second-order statistics of $x(t)$. Using the relation [13, 14]

$$\mathrm{Tr}\{\mathbf{P}_i \mathbf{R}_x\} = \big\langle L_{\mathbf{P}_i}, \overline{W}_x \big\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L_{\mathbf{P}_i}(t, f)\, \overline{W}_x(t, f)\, dt\, df \,,$$

the uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ in (4) can be reformulated as $\{\overline{W}_s(t,f) : \langle L_{\mathbf{P}_i}, \overline{W}_s\rangle = s_i\,,\ \ i=1,\ldots,N\}$ and $\{\overline{W}_n(t,f) : \langle L_{\mathbf{P}_i}, \overline{W}_n\rangle = n_i\,,\ \ i=1,\ldots,N\}$, respectively. To a *non-sophisticated* [20] subspace $\mathcal{X}_i$, one can associate a TF region $\mathcal{R}_i$ such that $L_{\mathbf{P}_i}(t,f) \approx I_{\mathcal{R}_i}(t,f)$, where $I_{\mathcal{R}_i}(t,f)$ is the indicator function of $\mathcal{R}_i$ [20, 21], and hence there is

$$\langle L_{\mathbf{P}_i}, \overline{W}_x\rangle \;\approx\; \langle I_{\mathcal{R}_i}, \overline{W}_x\rangle \;=\; \iint_{\mathcal{R}_i} \overline{W}_x(t,f)\,dt\,df\,.$$

The TF regions $\mathcal{R}_i$ associated to the subspaces $\mathcal{X}_i$ can be chosen to form a partition of the TF plane,

$$\mathcal{R}_1 \cup \mathcal{R}_2 \cup \cdots \cup \mathcal{R}_N = \mathbb{R}^2 \qquad \text{and} \qquad \mathcal{R}_i \cap \mathcal{R}_j = \emptyset \quad \text{for } i \neq j\,.$$

The TF partition $\{\mathcal{R}_i\}_{i=1,\ldots,N}$ is the TF counterpart of the space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$. Using this TF partition, approximate TF counterparts of the p-point uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ can be defined as[6]

$$\widetilde{\mathcal{S}} \;\triangleq\; \Big\{\overline{W}_s(t,f) : \iint_{\mathcal{R}_i} \overline{W}_s(t,f)\,dt\,df = \tilde{s}_i\,,\ \ i=1,\ldots,N\Big\},$$
$$\widetilde{\mathcal{N}} \;\triangleq\; \Big\{\overline{W}_n(t,f) : \iint_{\mathcal{R}_i} \overline{W}_n(t,f)\,dt\,df = \tilde{n}_i\,,\ \ i=1,\ldots,N\Big\}.$$

The "TF uncertainty classes" $\widetilde{\mathcal{S}}$, $\widetilde{\mathcal{N}}$ comprise all nonstationary processes $s(t)$, $n(t)$ (i.e., WVS $\overline{W}_s(t,f)$, $\overline{W}_n(t,f)$) having prescribed mean energies $\tilde{s}_i$, $\tilde{n}_i$ in given TF regions $\mathcal{R}_i$. This is more intuitive than the definition of the original uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ in terms of subspace energies. Note that $\sum_{i=1}^N \tilde{s}_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{W}_s(t,f)\,dt\,df = \bar{E}_s$ and similarly $\sum_{i=1}^N \tilde{n}_i = \bar{E}_n$. If nominal WVS $\overline{W}_s^0(t,f)$, $\overline{W}_n^0(t,f)$ are given, we can set $\tilde{s}_i \triangleq \iint_{\mathcal{R}_i} \overline{W}_s^0(t,f)\,dt\,df$ and $\tilde{n}_i \triangleq \iint_{\mathcal{R}_i} \overline{W}_n^0(t,f)\,dt\,df$. If at the same time the subspace energies s_i and n_i are defined as $s_i \triangleq \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_s^0\}$ and $n_i \triangleq \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_n^0\}$ (cf. Section 2), then we can expect that $\tilde{s}_i \approx s_i$ and $\tilde{n}_i \approx n_i$ (this will be verified empirically in Section 6). If no nominal second-order statistics are available, the "mean regional energies" $\tilde{s}_i$, $\tilde{n}_i$ have to be estimated from realizations of $s(t)$, $n(t)$. An unbiased estimator of $\tilde{s}_i$ is $\hat{\tilde{s}}_i = \iint_{\mathcal{R}_i} W_s(t,f)\,dt\,df$ (where $W_s(t,f)$ is the Wigner distribution [17, 22, 23] of a realization $s(t)$) and similarly for $\hat{\tilde{n}}_i$. We note that estimation of the numbers $\tilde{s}_i$, $\tilde{n}_i$ is much simpler and more stable than estimation of the complete WVS or correlations.

Next, we present an approximation to the Weyl symbol $L_{\mathbf{H}_R}(t,f)$ of the robust time-varying Wiener filter $\mathbf{H}_R$ in (10). We use the linearity of the Weyl symbol and the approximations $L_{\mathbf{P}_i}(t,f) \approx I_{\mathcal{R}_i}(t,f)$, $s_i \approx \tilde{s}_i$, and $n_i \approx \tilde{n}_i$ to obtain

$$L_{\mathbf{H}_R}(t,f) \;=\; \sum_{i=1}^N \frac{s_i}{s_i+n_i}\, L_{\mathbf{P}_i}(t,f) \;\approx\; \sum_{i=1}^N \frac{\tilde{s}_i}{\tilde{s}_i+\tilde{n}_i}\, I_{\mathcal{R}_i}(t,f)\,. \tag{11}$$

This shows that the Weyl symbol of the robust Wiener filter $\mathbf{H}_R$ is approximately a piecewise constant function, where the weight within a region $\mathcal{R}_i$ is given by $\tilde{s}_i/(\tilde{s}_i+\tilde{n}_i)$. The approximation in (11)

[6] Subsequently, the tilde ˜ is used to denote TF formulations or TF approximations.

motivates the definition of a *robust TF Wiener filter* $\widetilde{\mathbf{H}}_R$ as

$$L_{\widetilde{\mathbf{H}}_R}(t,f) \;\triangleq\; \sum_{i=1}^{N} \frac{\tilde{s}_i}{\tilde{s}_i + \tilde{n}_i}\, I_{\mathcal{R}_i}(t,f)\,.$$

Since $L_{\widetilde{\mathbf{H}}_R}(t,f) \approx L_{\mathbf{H}_R}(t,f)$ according to (11), $\widetilde{\mathbf{H}}_R$ can be expected to perform similarly as $\mathbf{H}_R$. Note that $L_{\widetilde{\mathbf{H}}_R}(t,f)$ is piecewise constant, expressing an equal TF weighting (cf. [15]) of all process components lying in the same TF region $\mathcal{R}_i$. The robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ can be written as

$$\widetilde{\mathbf{H}}_R = \sum_{i=1}^{N} \frac{\tilde{s}_i}{\tilde{s}_i + \tilde{n}_i}\, \tilde{\mathbf{P}}_i\,,$$

with the impulse response $\tilde{p}_i(t,t')$ of $\tilde{\mathbf{P}}_i$ given by the inverse Weyl transform [13, 14] of $I_{\mathcal{R}_i}(t,f)$,

$$\tilde{p}_i(t,t') = \int_{-\infty}^{\infty} I_{\mathcal{R}_i}\Big(\frac{t+t'}{2}, f\Big)\, e^{j2\pi f(t-t')}\, df\,.$$

The inverse Weyl transform can be efficiently computed using FFT methods. The systems $\tilde{\mathbf{P}}_i$ are self-adjoint and orthogonal ($\mathrm{Tr}\{\tilde{\mathbf{P}}_i\tilde{\mathbf{P}}_j\} = 0$ for $i \neq j$), and their sum equals $\mathbf{I}$. However, while $\tilde{\mathbf{P}}_i$ approximates $\mathbf{P}_i$ it is not exactly an orthogonal projection operator [20, 21]. We note that the robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ is related to perfect-reconstruction TF filterbanks as discussed in [20].

The prior knowledge necessary for designing the robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ is given by the $\tilde{s}_i$ and $\tilde{n}_i$, i.e., the mean energies of $s(t)$ and $n(t)$ in the TF regions $\mathcal{R}_i$. This is more intuitive and physically relevant than the mean energies in subspaces. Also, the task of choosing a space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$ has been replaced by the simpler and more intuitive task of choosing a partition $\{\mathcal{R}_i\}_{i=1,\ldots,N}$ of the TF plane. This TF partition may have a regular structure (see Section 5), or the $\mathcal{R}_i$ may correspond to individual components of the processes $s(t)$, $n(t)$ if the TF localization of such components is known.

We finally note that since $\mathbf{H}_R$ is minimax robust for the uncertainty classes $\mathcal{S}$, $\mathcal{N}$, and since furthermore $\widetilde{\mathbf{H}}_R \approx \mathbf{H}_R$ and $\widetilde{\mathcal{S}}$, $\widetilde{\mathcal{N}}$ are approximately equivalent to $\mathcal{S}$, $\mathcal{N}$, the filter $\widetilde{\mathbf{H}}_R$ is approximately minimax robust for the TF uncertainty classes $\widetilde{\mathcal{S}}$, $\widetilde{\mathcal{N}}$ (cf. [3, 10]).

5 Regular Time-Frequency Tilings and On-Line Implementation

The robust Wiener filter $\mathbf{H}_R$ often allows an efficient and causal implementation if the space partition $\{\mathcal{X}_i\}_{i=1,\ldots,N}$ corresponds to a TF partition $\{\mathcal{R}_i\}_{i=1,\ldots,N}$ with a regular structure. As an important example, we shall here discuss a uniform rectangular tiling of the TF plane. Consider rectangular TF regions $\mathcal{R}'_{k,l} = [kT, kT+T] \times [lF, lF+F]$ ($k,l \in \mathbb{Z}$) of duration T and bandwidth F. We assume that the TF regions $\mathcal{R}'_{k,l}$ have area $TF = M$ with $M \in \mathbb{N}$. Combining positive and corresponding negative frequencies, we then define the disjoint TF regions constituting our TF partition as $\mathcal{R}_{k,l} = \mathcal{R}'_{k,l} \cup \mathcal{R}'_{k,-l-1}$ with $k \in \mathbb{Z}$ and $l = 0, 1, \ldots$. Note that the $\mathcal{R}_{k,l}$ are indexed by the double index k,l and that $N = \infty$.

The TF regions $\mathcal{R}_{k,l}$ can be associated to orthogonal signal subspaces $\mathcal{X}_{k,l}$ that are spanned by the orthonormal *local cosine basis functions*[7] [24, 25]

$$x_{k,l}^{(m)}(t) \;=\; w(t-kT)\sqrt{\frac{2}{T}}\,\cos\!\left(\frac{2(lM+m)-1}{2T}\pi(t-kT)\right), \qquad m=1,2,\ldots,M\,, \tag{12}$$

where $w(t)$ is a suitable window function [24, 25]. The M-dimensional subspaces $\mathcal{X}_{k,l}$ form an orthogonal partition of the real space $L_2(\mathbb{R})$. Note that this implies a restriction to real-valued signals.

According to (10), the robust Wiener filter based on our space partition $\{\mathcal{X}_{k,l}\}_{k\in\mathbb{Z},l=0,1,\ldots}$ is given by

$$\mathbf{H}_R = \sum_{k=-\infty}^{\infty}\sum_{l=0}^{\infty} h_{k,l}\,\mathbf{P}_{k,l} \qquad \text{with } \; h_{k,l} = \frac{s_{k,l}}{s_{k,l}+n_{k,l}}\,,$$

where the $s_{k,l}$ and $n_{k,l}$ are the subspace energies that define the uncertainty classes. The kernels of the projection operators $\mathbf{P}_{k,l}$ are $p_{k,l}(t,t') = \sum_{m=1}^{M} x_{k,l}^{(m)}(t)\,x_{k,l}^{(m)}(t')$. Consequently, the signal estimate obtained at the output of $\mathbf{H}_R$ can be written as

$$\hat{s}(t) = (\mathbf{H}_R\,r)(t) = \sum_{k=-\infty}^{\infty}\sum_{l=0}^{\infty} h_{k,l}\left[\sum_{m=1}^{M} c_{k,l}^{(m)} x_{k,l}^{(m)}(t)\right] \qquad \text{with} \quad c_{k,l}^{(m)} = \langle r, x_{k,l}^{(m)}\rangle = \int_{-\infty}^{\infty} r(t)\,x_{k,l}^{(m)}(t)\,dt\,.$$

Note that M orthogonal components $c_{k,l}^{(m)} x_{k,l}^{(m)}(t)$ are summed and the entire sum is then multiplied by the weights $h_{k,l}$. The overall filtering procedure can be split into four stages:

1. *Signal analysis:* Calculation of the expansion coefficients $c_{k,l}^{(m)} = \langle r, x_{k,l}^{(m)}\rangle$ of $r(t)$.

2. *Projection:* Calculation of the orthogonal projections of $r(t)$ onto the subspaces $\mathcal{X}_{k,l}$ as $(\mathbf{P}_{k,l}\,r)(t) = \sum_{m=1}^{M} c_{k,l}^{(m)} x_{k,l}^{(m)}(t)$.

3. *Weighting:* Multiplication of the signals $(\mathbf{P}_{k,l}\,r)(t)$ by the weights $h_{k,l} = s_{k,l}/(s_{k,l}+n_{k,l})$.

4. *Signal synthesis:* Superposition of all $h_{k,l}\,(\mathbf{P}_{k,l}\,r)(t)$ yields the output signal $\hat{s}(t) = (\mathbf{H}_R\,r)(t) = \sum_{k=-\infty}^{\infty}\sum_{l=0}^{\infty} h_{k,l}\,(\mathbf{P}_{k,l}r)(t)$.

The above choice of the orthogonal subspaces $\mathcal{X}_{k,l}$ allows for an on-line implementation that is *causal* with a delay that is determined by the window $w(t)$ in (12). In a discrete-time framework, the underlying orthogonal basis expansion corresponds to a paraunitary cosine-modulated filterbank that can be efficiently implemented using discrete cosine transforms [28].

Under certain conditions, the expansion coefficients $c_{k,l}^{(m)} = \langle r, x_{k,l}^{(m)}\rangle$ that are formed at the analysis stage can be used to estimate the optimal weights $h_{k,l} = s_{k,l}/(s_{k,l}+n_{k,l})$ in a causal, on-line manner.

[7] For simplicity, T and F are assumed fixed even though the general format of a local cosine basis would allow T and F to be time-varying [24, 25]. We note that an alternative to local cosine bases is given by Wilson bases [26, 27].

This leads to a practical adaptive version of the robust Wiener filter $\mathbf{H}_R$ that can be used even if no or very little prior knowledge about the statistics of $s(t)$ and $n(t)$ is available. A problem with such an adaptive filter is that, in general, the signal statistics $s_{k,l}$ and the noise statistics $n_{k,l}$ cannot be estimated individually since only the overall observation $r(t) = s(t) + n(t)$ is available. However, in some cases this problem can be circumvented. We first note that the optimal weights can be written as

$$h_{k,l} \;=\; 1 - \frac{n_{k,l}}{r_{k,l}} \qquad \text{with } r_{k,l} \triangleq s_{k,l} + n_{k,l}\,.$$

Furthermore, there is $r_{k,l} = \mathrm{E}\{\|\mathbf{P}_{k,l}\, r\|^2\} = \sum_{m=1}^{M} \mathrm{E}\{\langle r, x_{k,l}^{(m)}\rangle^2\} = \sum_{m=1}^{M} \mathrm{E}\{(c_{k,l}^{(m)})^2\}$, so that an unbiased estimator of $r_{k,l}$ is given by

$$\hat{r}_{k,l} \;=\; \sum_{m=1}^{M} \left(c_{k,l}^{(m)}\right)^2 .$$

Suppose now that the following two facts are known:

- The mean subspace noise energy $n_{k,l}$ at any given time index k does not depend on the frequency index l, i.e., $n_{k,l} = n_k$. This will be true if $n(t)$ is (nonstationary) white; however, whiteness is not necessary since the power distribution of $n(t)$ *within* a given subspace $\mathcal{X}_{k,l}$ need not be constant.
- For each time index k, there exists at least one "noise-only" subspace, i.e., $s_{k,l'} = 0$ for some l'.

Then, the mean noise energy n_k at time k can be estimated in the noise-only subspace as $\hat{n}_k = \hat{r}_{k,l'} = \sum_{m=1}^{M} \left(c_{k,l'}^{(m)}\right)^2$, and estimates of the optimal filter weights $h_{k,l} = 1 - \frac{n_{k,l}}{r_{k,l}}$ are given by

$$\hat{h}_{k,l} \;=\; 1 - \frac{\hat{n}_k}{\hat{r}_{k,l}} \qquad \text{with } \hat{r}_{k,l} = \sum_{m=1}^{M} \left(c_{k,l}^{(m)}\right)^2 .$$

A similar estimation strategy exists if the mean subspace noise energy $n_{k,l}$ does not depend on the time index k, i.e., $n_{k,l} = n_l$. In particular, this will be true if $n(t)$ is stationary.

The robust Wiener filter discussed above was based on a uniform rectangular tiling of the TF plane. Other tilings can be obtained by grouping the local cosine basis functions in a different manner and/or by allowing T and F to be time-varying. Furthermore, analogous efficient implementations and adaptive versions of $\mathbf{H}_R$ using orthonormal bases or paraunitary filterbanks can be developed for the TF tilings corresponding to wavelet transforms and wavelet packet transforms [29]. We finally note that previously proposed subband-based speech enhancement schemes [30, 31] can be viewed as special cases for $M = 1$.

6 Simulation Results

We now illustrate our results through numerical simulations. In our first example, we designed nominal correlations $\mathbf{R}_s^0$ and $\mathbf{R}_n^0$ using the TF synthesis procedure described in [32]. The corresponding nominal

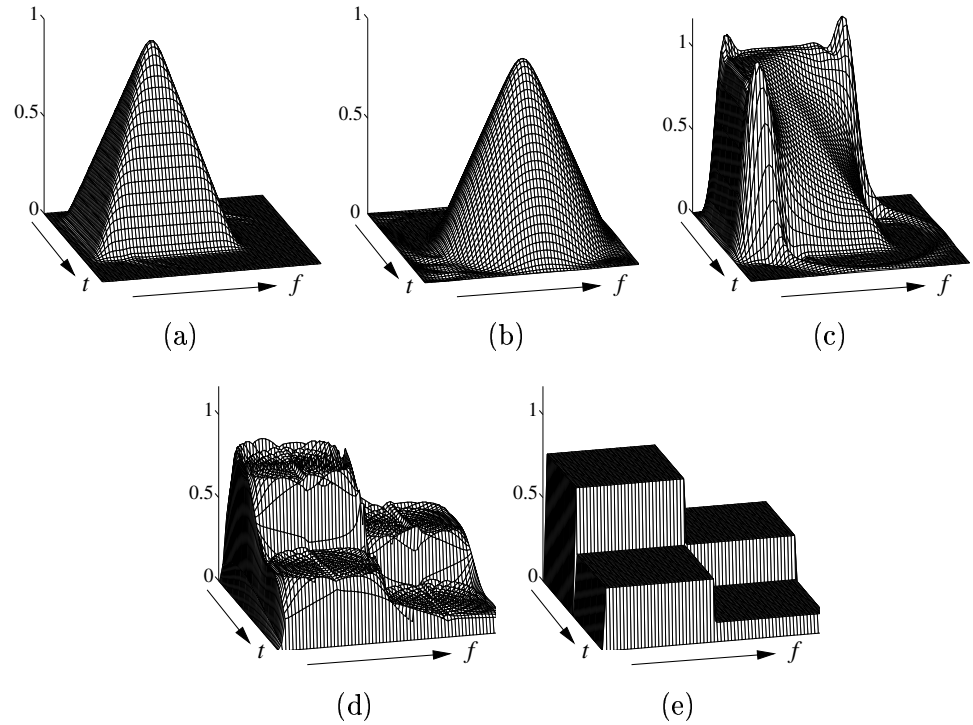

Figure 1: *TF representations of nominal signal and noise statistics and of various Wiener filters: (a) Nominal WVS of* $s(t)$, $\overline{W}_s^0(t,f)$, *(b) nominal WVS of* $n(t)$, $\overline{W}_n^0(t,f)$, *(c) Weyl symbol of Wiener filter* $\mathbf{H}_W^0$, *(d) Weyl symbol of robust Wiener filter* $\mathbf{H}_R$, *and (e) Weyl symbol of robust TF Wiener filter* $\widetilde{\mathbf{H}}_R$.

WVS $\overline{W}_s^0(t,f)$ and $\overline{W}_n^0(t,f)$ are shown in Fig. 1(a),(b), and the Weyl symbol of the ordinary Wiener filter $\mathbf{H}_W^0 = \mathbf{R}_s^0(\mathbf{R}_s^0+\mathbf{R}_n^0)^{-1}$ designed from the nominal correlations is shown in Fig. 1(c). We then constructed TF uncertainty classes $\widetilde{\mathcal{S}}$, $\widetilde{\mathcal{N}}$ with $N=4$ by partitioning the TF plane into four rectangular TF regions $\mathcal{R}_i$ and calculating the mean regional energies as $\tilde{s}_i \triangleq \iint_{\mathcal{R}_i} \overline{W}_s^0(t,f)\,dt\,df$ and $\tilde{n}_i \triangleq \iint_{\mathcal{R}_i} \overline{W}_n^0(t,f)\,dt\,df$ $(i = 1,...,4)$. The weights of the robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ were formed as $\tilde{h}_i = \tilde{s}_i/(\tilde{s}_i+\tilde{n}_i)$. Furthermore, we constructed uncertainty classes $\mathcal{S}$, $\mathcal{N}$ based on an orthogonal space partition $\{\mathcal{X}_i\}_{i=1,...,4}$ that was derived from the TF regions $\mathcal{R}_i$ by means of TF subspace synthesis [20]. The mean subspace energies were calculated as $s_i = \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_s^0\}$ and $n_i = \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_n^0\}$ $(i = 1,...,4)$, and the weights of the robust Wiener filter $\mathbf{H}_R$ were formed as $h_i = s_i/(s_i+n_i)$. The relative errors $\frac{|\tilde{s}_i - s_i|}{s_i}$, $\frac{|\tilde{n}_i - n_i|}{n_i}$, and $\frac{|\tilde{h}_i - h_i|}{h_i}$ were all below 0.022, which shows that the TF approximations are quite good. The Weyl symbols of the robust Wiener filter $\mathbf{H}_R$ and the robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ are shown in Figs. 1(d) and (e), respectively. The rectangular TF tiling underlying the uncertainty classes is clearly visible.

Fig. 2 compares the performance (output SNR $\bar{E}_s/e(\mathbf{H};\mathbf{R}_s,\mathbf{R}_n)$ vs. input SNR[8] $\bar{E}_s/\bar{E}_n$) of the ordinary Wiener filter $\mathbf{H}_W^0 = \mathbf{R}_s^0(\mathbf{R}_s^0+\mathbf{R}_n^0)^{-1}$, the robust Wiener filter $\mathbf{H}_R$, the robust TF Wiener

[8]The input SNR was varied by scaling $\mathbf{R}_s$.

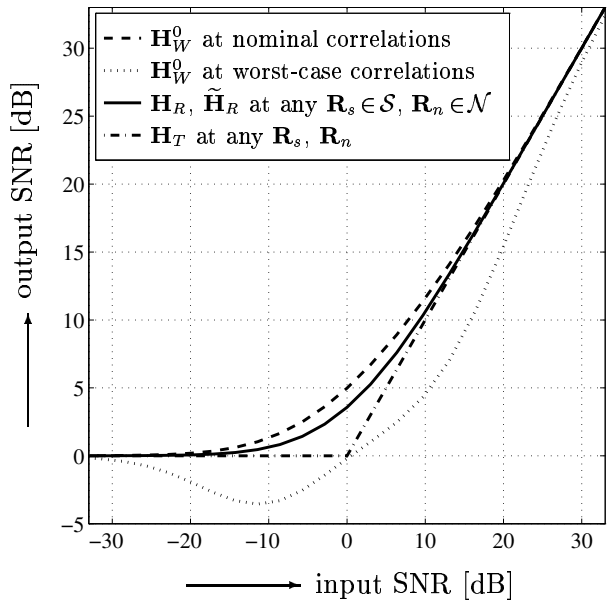

Figure 2: *Performance of the ordinary Wiener filter* $\mathbf{H}_W^0$*, the robust Wiener filters* $\mathbf{H}_R$ *and* $\widetilde{\mathbf{H}}_R$*, and the trivial filter* $\mathbf{H}_T$*. (The performance curve of* $\widetilde{\mathbf{H}}_R$ *coincides with that of* $\mathbf{H}_R$ *for all input SNRs.)*

filter $\widetilde{\mathbf{H}}_R$, and a trivial filter $\mathbf{H}_T$ that suppresses (passes) all signals in the case of negative (positive) SNR. For all input SNRs, the output SNR obtained with $\widetilde{\mathbf{H}}_R$ was observed to be within 0,07 dB of that obtained with $\mathbf{H}_R$, thereby confirming the quality of the TF approximation (11). It is seen that at nominal operating conditions $\mathbf{H}_W^0$ performs only slightly better than $\mathbf{H}_R$ or $\widetilde{\mathbf{H}}_R$ but at its worst-case operating conditions $\mathbf{H}_W^0$ performs much worse than $\mathbf{H}_R$ or $\widetilde{\mathbf{H}}_R$ or even $\mathbf{H}_T$. Hence, in this example, the robust Wiener filters $\mathbf{H}_R$ and $\widetilde{\mathbf{H}}_R$ achieve a drastic performance improvement over $\mathbf{H}_W^0$ at the worst-case operating conditions for $\mathbf{H}_W^0$ with only a slight performance loss at nominal operating conditions.

Fig. 3 compares the performance of the nominal TF Wiener filter[9] $\widetilde{\mathbf{H}}_W^0$ and the robust TF Wiener filter $\widetilde{\mathbf{H}}_R$ for different TF uncertainty models (derived from the same nominal WVS as above) corresponding to regular rectangular TF tilings with $N = K^2$ rectangles $\mathcal{R}_i$. Here, K is the number of time intervals which is equal to the number of frequency bands. The three plots in Fig. 3 show the output SNR versus the number N of TF rectangles for three different input SNRs. As N is increased, the prior knowledge becomes more detailed and thus the uncertainty is reduced. Hence, as N is increased, the worst-case performance of $\widetilde{\mathbf{H}}_W^0$ ('×') tends to improve while the nominal performance of $\widetilde{\mathbf{H}}_W^0$ ('∘') remains constant; furthermore, the performance of $\widetilde{\mathbf{H}}_R$ ('+') approaches the nominal performance of $\widetilde{\mathbf{H}}_W^0$. Again, the performance of $\widetilde{\mathbf{H}}_R$ is much better than the worst-case performance of $\widetilde{\mathbf{H}}_W^0$ and only slightly poorer than the nominal performance of $\widetilde{\mathbf{H}}_W^0$.

[9]The nominal TF Wiener filter $\widetilde{\mathbf{H}}_W^0$ is defined by $L_{\widetilde{\mathbf{H}}_W^0}(t,f) = \overline{W}_s^0(t,f)/\left[\overline{W}_s^0(t,f) + \overline{W}_n^0(t,f)\right]$ [8].

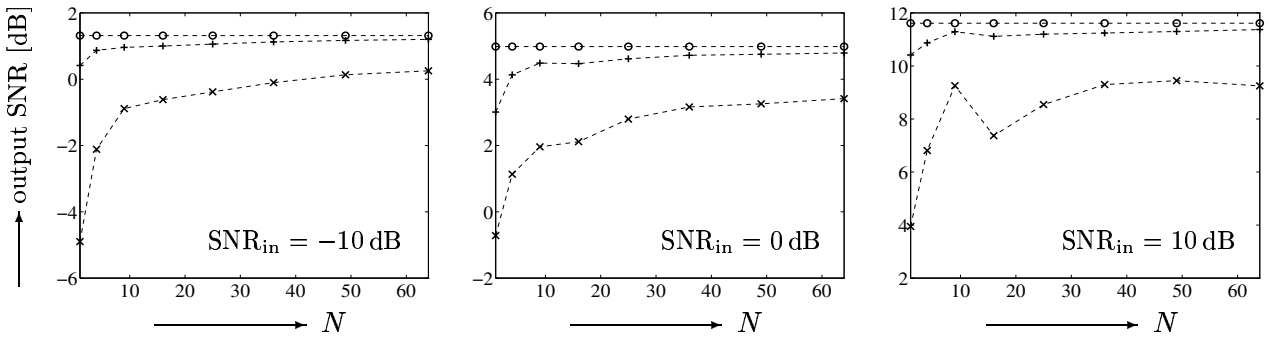

Figure 3: *Output SNR of ordinary TF Wiener filter* $\widetilde{\mathbf{H}}_W^0$ *and robust TF Wiener filter* $\widetilde{\mathbf{H}}_R$ *versus number* N *of TF rectangles, for three different input SNRs. (*$\circ \ldots \widetilde{\mathbf{H}}_W^0$ *at nominal operating conditions;* $\times \ldots \widetilde{\mathbf{H}}_W^0$ *at worst-case operating conditions;* $+ \ldots \widetilde{\mathbf{H}}_R$ *at any operating conditions within* $\widetilde{\mathcal{S}}$, $\widetilde{\mathcal{N}}$*.)*

The previous simulation examples did not consider estimation of the mean subspace energies s_i and n_i. Therefore, our final simulation example demonstrates the application of the adaptive robust Wiener filter using a local cosine basis (see Section 5) to speech enhancement. A segment of speech uttered by a female speaker was corrupted by additive white Gaussian noise. The signal was 3 seconds long and was sampled at 11025 Hz. The input SNR was 3.18 dB. The filter coefficients $h_{k,l}$ were estimated as explained in Section 5, with prior knowledge of the noise level only. We first used a TF tiling with $T = 23.2$ ms and $F = 43.1$ Hz so that $M = TF = 1$, i.e., each subspace $\mathcal{X}_{k,l}$ consisted of a single local cosine basis function. (This is similar to current subband speech enhancement schemes [30, 31].) The SNR improvement achieved with this filter was 3.72 dB. Next, we used the same $T = 23.2$ ms but $F = 1379.3$ Hz, which corresponds to an increased subspace dimension of $M = TF = 32$. Here, the SNR improvement was 7.72 Hz, i.e., 4 dB better than for $M = 1$. We conclude that by trading TF filtering resolution for improved statistical stability of estimation, the adaptive robust Wiener filter is able to achieve noticeable performance gains.

7 Conclusion

Motivated by the sensitivity of ordinary time-varying Wiener filters to deviations from nominal operating conditions, we introduced a p-point uncertainty model for nonstationary random processes and a minimax robust time-varying Wiener filter with constant performance within the uncertainty classes. We also developed intuitively appealing time-frequency formulations of the p-point uncertainty model and the robust Wiener filter. Finally, we discussed an efficient causal implementation of the robust Wiener filter that allows on-line filtering as well as on-line estimation of the prior knowledge required.

Appendix: Proof of Theorem 1

According to Section 3, $\mathbf{H}_R$ equals the ordinary Wiener filter $\mathbf{H}_W^L = \mathbf{R}_s^L\big(\mathbf{R}_s^L + \mathbf{R}_n^L\big)^{-1}$ obtained for least favorable correlations $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ that satisfy the inequality (9). Let us consider the operators

$$\mathbf{R}_s^L \triangleq \sum_{i=1}^N \mathbf{R}_{s,i}\,, \qquad \mathbf{R}_n^L \triangleq \sum_{i=1}^N \mathbf{R}_{n,i}\,,$$

where $\mathbf{R}_{s,i}$ and $\mathbf{R}_{n,i}$ $(i = 1, \ldots, N)$ are positive semidefinite operators with domain and range [4, 5] equal to $\mathcal{X}_i$ and satisfying

$$n_i \mathbf{R}_{s,i} = s_i \mathbf{R}_{n,i}\,. \tag{13}$$

$\mathbf{R}_s^L$ and $\mathbf{R}_n^L$ are positive (semi-)definite and thus valid correlation operators, and furthermore the $\mathbf{R}_{s,i}$, $\mathbf{R}_{n,i}$ can always be normalized such that $\mathbf{R}_s^L \in \mathcal{S}$, $\mathbf{R}_n^L \in \mathcal{N}$. We now show that $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ indeed satisfy (9). The ordinary Wiener filter for $\mathbf{R}_s^L$, $\mathbf{R}_n^L$ is given by

$$\mathbf{H}_W^L \;=\; \mathbf{R}_s^L\big(\mathbf{R}_s^L + \mathbf{R}_n^L\big)^{-1} = \sum_{i=1}^N \mathbf{R}_{s,i}\Big[\sum_{j=1}^N \big(\mathbf{R}_{s,j} + \mathbf{R}_{n,j}\big)\Big]^{-1}.$$

Using the fact that the domain and range of $\mathbf{R}_{s,i}$, $\mathbf{R}_{n,i}$ equal $\mathcal{X}_i$, and denoting the pseudo-inverse [4, 5] of $\mathbf{R}_{s,i} + \mathbf{R}_{n,i}$ by $(\mathbf{R}_{s,i} + \mathbf{R}_{n,i})^{\#}$, we obtain further

$$\mathbf{H}_W^L \;=\; \sum_{i=1}^N \mathbf{R}_{s,i} \sum_{j=1}^N \big(\mathbf{R}_{s,j} + \mathbf{R}_{n,j}\big)^{\#} = \sum_{i=1}^N \mathbf{R}_{s,i}\big(\mathbf{R}_{s,i} + \mathbf{R}_{n,i}\big)^{\#}\,,$$

where the last identity holds since $\mathbf{R}_{s,i}\big(\mathbf{R}_{s,j} + \mathbf{R}_{n,j}\big)^{\#} = \mathbf{0}$ for $i \neq j$. With (13) we finally obtain

$$\mathbf{H}_W^L = \sum_{i=1}^N \mathbf{R}_{s,i}\Big[\Big(1 + \frac{n_i}{s_i}\Big)\mathbf{R}_{s,i}\Big]^{\#} = \sum_{i=1}^N \frac{s_i}{s_i + n_i}\,\mathbf{R}_{s,i}\mathbf{R}_{s,i}^{\#} = \sum_{i=1}^N \frac{s_i}{s_i + n_i}\,\mathbf{P}_i\,. \tag{14}$$

Using $\mathbf{I} - \mathbf{H}_W^L = \sum_{i=1}^N \frac{n_i}{s_i+n_i}\mathbf{P}_i$, the MSE obtained when applying $\mathbf{H}_W^L$ to processes with correlations $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}$ can be developed as

$$\begin{aligned}
e(\mathbf{H}_W^L; \mathbf{R}_s, \mathbf{R}_n) \;&=\; \mathrm{E}\big\{\big\|\mathbf{H}_W^L r - s\big\|^2\big\} \;=\; \mathrm{E}\big\{\big\|\mathbf{H}_W^L s + \mathbf{H}_W^L n - s\big\|^2\big\} \;=\; \mathrm{E}\big\{\big\|\big(\mathbf{I} - \mathbf{H}_W^L\big)s\big\|^2\big\} + \mathrm{E}\big\{\big\|\mathbf{H}_W^L n\big\|^2\big\} \\
&=\; \mathrm{Tr}\Big\{\big(\mathbf{I} - \mathbf{H}_W^L\big)^2\mathbf{R}_s\Big\} + \mathrm{Tr}\Big\{\big(\mathbf{H}_W^L\big)^2\mathbf{R}_n\Big\} \\
&=\; \mathrm{Tr}\Big\{\Big[\sum_{i=1}^N \frac{n_i}{s_i + n_i}\mathbf{P}_i\Big]^2\mathbf{R}_s\Big\} + \mathrm{Tr}\Big\{\Big[\sum_{i=1}^N \frac{s_i}{s_i + n_i}\mathbf{P}_i\Big]^2\mathbf{R}_n\Big\} \\
&=\; \sum_{i=1}^N \Big(\frac{n_i}{s_i + n_i}\Big)^2 \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_s\} \;+ \sum_{i=1}^N \Big(\frac{s_i}{s_i + n_i}\Big)^2 \mathrm{Tr}\{\mathbf{P}_i\mathbf{R}_n\} \\
&=\; \sum_{i=1}^N \frac{n_i^2 s_i + s_i^2 n_i}{(s_i + n_i)^2} = \sum_{i=1}^N \frac{s_i\, n_i}{s_i + n_i}\,,
\end{aligned}$$

which is identical for all $\mathbf{R}_s \in \mathcal{S}$, $\mathbf{R}_n \in \mathcal{N}$. Hence, inequality (9) is satisfied (with equality), and thus the Wiener filter $\mathbf{H}_W^L$ in (14) equals the robust Wiener filter $\mathbf{H}_R$.

References

[1] L. J. Cimini and S. A. Kassam, "Robust and quantized Wiener filters for p-point spectral classes," in *Proc. Conf. Information Sciences and Systems*, (Princeton, NJ), pp. 314–318, March 1980.

[2] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, March 1985.

[3] G. Matz and F. Hlawatsch, "Minimax robust time-frequency filters for nonstationary signal estimation," in *Proc. IEEE ICASSP-99*, (Phoenix, AZ), pp. 1333–1336, March 1999.

[4] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science.* New York: Springer, 2nd ed., 1982.

[5] I. C. Gohberg and M. G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators.* Providence (RI): Amer. Math. Soc., 1969.

[6] T. Kailath, *Lectures on Wiener and Kalman Filtering*, vol. 140 of *CISM Courses and Lectures.* Wien: Springer, 1981.

[7] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory.* New York: Wiley, 1968.

[8] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation," *IEEE Trans. Signal Processing*, vol. 48, May 2000.

[9] G. Matz and F. Hlawatsch, "Robust time-varying Wiener filters: Theory and time-frequency formulation," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Pittsburgh, PA), pp. 401–404, Oct. 1998.

[10] G. Matz and F. Hlawatsch, "Minimax robust time-varying Wiener filters," Tech. Rep. #00-1, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Vienna, 2000.

[11] G. G. Tziritas, "On the distribution of positive-definite Gaussian quadratic forms," *IEEE Trans. Inf. Theory*, vol. 33, pp. 895–906, Nov. 1987.

[12] V. Barbu and T. Precupanu, *Convexity and Optimization in Banach Spaces.* Alphen aan de Rijn (The Netherlands): Sijthoff & Noordhoff, 1978.

[13] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework," *Signal Processing*, vol. 29, pp. 77–92, Oct. 1992.

[14] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, Feb. 1994.

[15] G. Matz and F. Hlawatsch, "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory," *J. Math. Phys., Special Issue on Wavelet and Time-Frequency Analysis*, vol. 39, pp. 4041–4071, Aug. 1998.

[16] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 1461–1470, Dec. 1985.

[17] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego (CA): Academic Press, 1999.

[18] P. Flandrin and W. Martin, "The Wigner-Ville spectrum of nonstationary random signals," in *The Wigner Distribution — Theory and Applications in Signal Processing* (W. Mecklenbräuker and F. Hlawatsch, eds.), pp. 211–267, Amsterdam (The Netherlands): Elsevier, 1997.

[19] G. Matz and F. Hlawatsch, "Time-varying spectra for underspread and overspread nonstationary processes," in *Proc. 32nd Asilomar Conf. Signals, Systems, Computers*, (Pacific Grove, CA), pp. 282–286, Nov. 1998.

[20] F. Hlawatsch, *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation.* Boston: Kluwer, 1998.

[21] F. Hlawatsch and W. Kozek, "Time-frequency projection filters and time-frequency signal expansions," *IEEE Trans. Signal Processing*, vol. 42, pp. 3321–3334, Dec. 1994.

[22] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—A tool for time-frequency signal analysis; Parts I–III," *Philips J. Research*, vol. 35, pp. 217–250, 276–300, and 372–389, 1980.

[23] W. Mecklenbräuker and F. Hlawatsch, eds., *The Wigner Distribution — Theory and Applications in Signal Processing.* Amsterdam (The Netherlands): Elsevier, 1997.

[24] R. R. Coifman and Y. Meyer, "Remarques sur l'analyse de Fourier à fenêtre," *C. R. Acad. Sci.*, vol. 1, no. 312, pp. 259–261, 1991.

[25] P. Auscher, G. Weiss, and M. V. Wickerhauser, "Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets," in *Wavelets: A Tutorial in Theory and Applications* (C. K. Chui, ed.), pp. 237–256, New York: Academic Press, 1992.

[26] I. Daubechies, S. Jaffard, and J. L. Journé, "A simple Wilson orthonormal basis with exponential decay," *SIAM J. Math. Anal.*, vol. 22, pp. 554–572, 1991.

[27] H. Bölcskei, H. G. Feichtinger, K. Gröchenig, and F. Hlawatsch, "Discrete-time Wilson expansions," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Paris, France), pp. 525–528, June 1996.

[28] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs (NJ): Prentice Hall, 1993.

[29] S. G. Mallat, *A Wavelet Tour of Signal Processing.* San Diego: Academic Press, 1998.

[30] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol. 67, pp. 1586–1604, Dec. 1979.

[31] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 32, pp. 1109–1121, Dec. 1984.

[32] F. Hlawatsch and W. Kozek, "Second-order time-frequency synthesis of nonstationary random processes," *IEEE Trans. Inf. Theory*, vol. 41, pp. 255–267, Jan. 1995.