Causal Restriction and Its Generalization for the Wiener Filter

Yun-xiu Yang¹ · Qin Shu² · Fei Yuan¹ ·
Shijie Deng¹

Received: 27 March 2016 / Revised: 3 June 2017 / Accepted: 6 June 2017 /

Published online: 29 June 2017

© The Author(s) 2017. This article is an open access publication

Abstract The classical principles of Wiener filter design consider only the causality constraint. However, additional constraints are often involved in practice. This paper provides a comprehensive theoretical approach for designing Wiener filters with constraint conditions beyond causality. Using integration, we first derive an extension of the Wiener–Hopf equation with an additional term that captures the effect of the constraint conditions on the resulting Wiener filters. Next, we analytically solve the extended Wiener–Hopf equation and develop a new iterative algorithm to obtain its optimum value. We prove the convergence of our iterative algorithm and provide several simulations to illustrate the applicability of our theoretical approach.

Keywords Wiener filter · Wiener–Hopf equation · MMSE · General restriction · Iterative algorithm

1 Introduction

The Wiener filter [15] is a well-known tool in signal processing and communications [9]. It presents a classical approach to determining the optimum filter by providing

✉ Qin Shu
shuqin@scu.edu.cn

Yun-xiu Yang
yangyang_judy@126.com

Fei Yuan
yuanfei312@126.com

Shijie Deng
dsjie520@163.com

¹ South-West Institute of Technical Physics, Chengdu, Sichuan, China

² School of Electrical Engineering and Information, Sichuan University, Chengdu, Sichuan, China

Fig. 1 Diagram of the Wiener filter

the minimum mean-squared error estimate of a desired process $d(t)$, from a noisy random input signal $x(t)$, as illustrated in Fig. 1.

In Fig. 1, $x(t)$ and $d(t)$ are assumed to be jointly wide-sense stationary random processes with known autocorrelations, $R_x(\tau) = E\{x(t+\tau)x^*(t)\}$ and $R_d(\tau) = E\{d(t+\tau)d^*(t)\}$, and a known cross-correlation, $R_{dx}(\tau) = E\{d(t+\tau)x^*(t)\}$, in which $E\{\bullet\}$ denotes the mathematical expectation and the asterisk indicates the complex conjugation operation. The output of the filter is $\hat{d}(t)$, which denotes the estimate of $d(t)$ and is given by

$$\hat{d}(t) = \int_{-\infty}^{\infty} x(t-\tau)w(\tau)d\tau \quad (1)$$

The estimated error is

$$e(t) = d(t) - \hat{d}(t) \quad (2)$$

The mean-squared error is

$$\begin{aligned} E\{|e(t)|^2\} &= E\left\{\left(d(t) - \int_{-\infty}^{\infty} x(t-\tau)w(\tau)d\tau\right)\left(d^*(t) - \int_{-\infty}^{\infty} x^*(t-\theta)w^*(\theta)d\theta\right)\right\} \\ &= R_d(0) - \int_{-\infty}^{\infty} R_{dx}^*(\tau)w(\tau)d\tau - \int_{-\infty}^{\infty} R_{dx}(\theta)w^*(\theta)d\theta \\ &\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(\theta-\tau)w(\tau)w^*(\theta)d\tau d\theta \end{aligned} \quad (3)$$

The optimum filter, denoted by $w_{\text{opt}}(t)$, is determined by minimizing $E\{|e(t)|^2\}$.

In the classical principle of the Wiener filter design, there is only one causality restriction on the filter to be considered. However, in practice, to meet the needs of different applications, the filter may be subject to more broad constraints. For example, in signal prediction problems and adaptive control applications, the filter is generally required to be a time-limited system to realize finite-step prediction or finite-order control [10]. In past years, numerous researchers have studied the design of the Wiener filter with finite support in the time domain [4, 6, 14]. Other common restrictions on the Wiener filter are in the frequency domain. For example, the Wiener filter may be required to be band-limited, band-pass, or low-pass or to satisfy other bandwidth limitations to guarantee that the frequency support of the filter can match the bandwidth of the desired signal and to avoid situations in which the filter works in noise-only periods [7, 8, 12]. For instance, in a speech enhancement system, to reduce the effects of

environmental noise, the Wiener filter usually operates in the audible frequency range. Much research has been reported on constrained Wiener filters in binaural hearing aids and speech enhancement [1, 2, 4, 5, 12]. In addition to time-domain and frequency-domain constraints, linear constraints, which require the Wiener filter to satisfy a set of linear equations, are also widely applied in many fields, including adaptive beamforming [13], spectral analysis, and function interpolation [11]. Moreover, in filter design, when the power spectrum of the input is close to zero at some frequency points, the corresponding frequency response of the filter at these frequency points will become unbounded. A filter with a boundless frequency response is undesirable because it will be sensitive to noise and data errors. In this situation, the finite-energy constraint is imposed to smooth the peak value of the frequency response of the filter. The performance of the Wiener filter with the finite-energy constraint is better than that of the unconstrained Wiener filter [3]. In addition to the common constraints described above, studies on wider classes of constraints for the Wiener filter can be found in the literature [7, 8], including the output-spectrum constraint, the estimated signal-power constraint, the filter-energy constraint, and the filter-stability constraint. Moreover, the closed-form solutions to these constrained design problems have been derived.

The work presented in this paper makes three contributions. First, this is the first work to incorporate the different constraint conditions into a single integration equation. Moreover, since the functions in the integral can be chosen according to different application needs, single integration allows more flexibility and applicability in processing various restrictions on the Wiener filter. Second, we develop a closed formula to represent the optimal solution of the Wiener filter under the constrained integration equation. Because the closed formula has one additional term compared to the Wiener–Hopf equation, the formula is called the extended Wiener–Hopf equation in this paper. Third, we develop a uniform analytical method along with an iterative algorithm to solve the extended Wiener–Hopf equation. Finally, some simulation examples are given to further illustrate the solution steps of our algorithm.

2 Constraint Condition

Assume that $w(t)$ is subject to the following constraint:

$$\int_{-\infty}^{\infty} (w(t) - g(t)) \lambda^*(t) dt = \beta \quad (4)$$

where $g(t)$, $\lambda(t)$, and β may be fully known or some of their characteristics may be known.

In particular, Eq. (4) is a more general expression of the following three constraints.

2.1 Time-Domain Constraints

In this case, $w(t)$ is equal to a given function $g(t)$ within the time subset T , i.e.,

$$w(t) = g(t), t \in T \quad (5)$$

More specifically, if $g(t) = 0$ and $T = (-\infty, 0)$, then $w(t)$ is obviously required to be causal; if $g(t) = 0$ and $T = (-\infty, 0) \cup [t_0, \infty)$, where $t_0 > 0$, then $w(t)$ is a causal time-limited system.

For this case, if we take $\lambda(t) = 0$, $t \notin T$ and $\beta = 0$, then Eq. (5) can be described by Eq. (4).

2.2 Frequency-Domain Constraints

Here, $W(f)$ is limited by

$$W(f) = G(f), f \in \Omega \quad (6)$$

where $W(f)$ is the Fourier transformation of the filter $w(t)$, $G(f)$ is a known frequency-domain function, and Ω indicates a subset of the frequency domain. If $G(f) = 0$ and $\Omega = (-\infty, -\sigma) \cup (\sigma, \infty)$, in which σ is a positive number, $w(t)$ must be a low-pass filter. If $G(f) = 0$ and $\Omega = (-\infty, 0)$, then $w(t)$ is required to be an analytic function. If $G(f) = 0$ and $\Omega = (-\infty, -\sigma_2) \cup (-\sigma_1, \sigma_1) \cup (\sigma_2, \infty)$, where $\sigma_2 > \sigma_1 > 0$, then $w(t)$ must be a band-pass system.

For this state, we take $\Lambda(f) = 0$, $f \notin \Omega$, and $\beta = 0$. Thus:

$$\int_{-\infty}^{\infty} (W(f) - G(f)) \Lambda^*(f) df = \beta \quad (7)$$

According to Parseval's theorem [9], Eq. (7) can be regarded as the expression of Eq. (4) in the frequency domain.

2.3 Orthogonal Constraints

In this case, $w(t)$ is assumed to be orthogonal to the signal space H , i.e.,

$$\int_{-\infty}^{\infty} w(t) \lambda^*(t) dt = 0, \quad \forall \lambda(t) \in H \quad (8)$$

Equation (8) is a special case of Eq. (4), in which $g(t) = 0$ and $\beta = 0$. In practice, the orthogonal space of the filter $w(t)$ is usually viewed as noisy space or undesirable signal space, which needs to be depressed by introducing the constraint defined in Eq. (8).

According to these three particular cases, we can construct the Wiener filter with more general constraints to meet the requirements of different practical applications because there are many possibilities for $g(t)$ or $\lambda(t)$.

3 Principle

To consider the constraints of Eq. (4), we first multiply the two sides of Eq. (4) by a Lagrange factor μ . As a result, we have $\mu^* (\int_{-\infty}^{\infty} (w(t) - g(t)) \lambda^*(t) dt - \beta) = 0$. Then, we add it and its conjugate to Eq. (3) as follows:

$$\begin{aligned}
L(w, \mu) &= E \left\{ |e(t)|^2 \right\} + \mu^* \left(\int_{-\infty}^{\infty} (g(t) - w(t)) \lambda^*(t) dt - \beta \right) \\
&\quad + \mu \left(\int_{-\infty}^{\infty} (g(t) - w(t))^* \lambda(t) dt - \beta^* \right) \\
&= R_d(0) - \int_{-\infty}^{\infty} R_{dx}^*(\tau) w(\tau) d\tau - \int_{-\infty}^{\infty} R_{dx}(\theta) w^*(\theta) d\theta \\
&\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(\theta - \tau) w^*(\theta) w(\tau) d\tau d\theta \\
&\quad + \int_{-\infty}^{\infty} (w(t) - g(t)) \mu^* \lambda^*(t) dt - \mu^* \beta \\
&\quad + \int_{-\infty}^{\infty} (w(t) - g(t))^* \mu \lambda(t) dt - \mu \beta^*
\end{aligned} \tag{9}$$

Let $S_x(f)$, $S_d(f)$, $S_{dx}(f)$, $W(f)$, $G(f)$, and $\Lambda(f)$ denote the Fourier transformations of $R_x(t)$, $R_d(t)$, $R_{dx}(t)$, $w(t)$, $g(t)$, and $\lambda(t)$, respectively. Note that the Fourier transformation of $\int_{-\infty}^{\infty} R_x(\theta - \tau) w(\tau) d\tau$ is $S_x(f)W(f)$, where $S_x(f)$ is nonnegative. We assume $S_x(f) > 0$. Therefore, according to Parseval's equation, we have

$$\begin{aligned}
L(w, \mu) &= \int_{-\infty}^{\infty} \left(S_d(f) - S_{dx}^*(f)W(f) - S_{dx}(f)W^*(f) + S_x(f)|W(f)|^2 \right) df \\
&\quad + \int_{-\infty}^{\infty} (W(f) - G(f)) \mu^* \Lambda^*(f) df + \int_{-\infty}^{\infty} (W^*(f) - G^*(f)) \mu \Lambda(f) df - \mu^* \beta - \mu \beta^* \\
&= \int_{-\infty}^{\infty} S_d(f) df + \int_{-\infty}^{\infty} \left(S_x(f)|W(f)|^2 - \frac{(S_{dx}(f) - \mu \Lambda(f))^*}{\sqrt{S_x(f)}} \sqrt{S_x(f)} W(f) \right. \\
&\quad \left. - \frac{(S_{dx}(f) - \mu \Lambda(f))}{\sqrt{S_x(f)}} \sqrt{S_x(f)} W^*(f) \right) df \\
&\quad + \int_{-\infty}^{\infty} \left(\frac{|(S_{dx}(f) - \mu \Lambda(f))|^2}{S_x(f)} - \frac{|(S_{dx}(f) - \mu \Lambda(f))|^2}{S_x(f)} \right) df \\
&\quad - \int_{-\infty}^{\infty} (G(f) \mu^* \Lambda^*(f) + G^*(f) \mu \Lambda(f)) df - \mu^* \beta - \mu \beta^* \\
&= \int_{-\infty}^{\infty} S_d(f) df + \int_{-\infty}^{\infty} \left(S_x(f) \left| W(f) - \frac{(S_{dx}(f) - \mu \Lambda(f))}{S_x(f)} \right|^2 - \frac{|(S_{dx}(f) - \mu \Lambda(f))|^2}{S_x(f)} \right) df \\
&\quad - \int_{-\infty}^{\infty} (G(f) \mu^* \Lambda^*(f) + G^*(f) \mu \Lambda(f)) df - \mu^* \beta - \mu \beta^*
\end{aligned} \tag{10}$$

Consequently, if $W(f)$ is taken to be equal to $\frac{S_{dx}(f) - \mu \Lambda(f)}{S_x(f)}$, then $L(w, \mu)$ achieves the minimum of $E \left\{ |e(t)|^2 \right\}$, denoted by $L_{\min}(w_{\text{opt}}, \mu)$, and we have

$$\begin{aligned}
L_{\min}(w_{\text{opt}}, \mu) &= \int_{-\infty}^{\infty} \left(S_x(f) - \frac{|(S_{dx}(f) - \mu \Lambda(f))|^2}{S_x(f)} \right) df \\
&\quad - \int_{-\infty}^{\infty} (G(f) \mu^* \Lambda^*(f) + G^*(f) \mu \Lambda(f)) df - \mu^* \beta - \mu \beta^*
\end{aligned} \tag{11}$$

and

$$W_{\text{opt}}(f) = \frac{S_{dx}(f) - \mu\Lambda(f)}{S_x(f)} \quad (12)$$

or, equivalently,

$$S_x(f)W_{\text{opt}}(f) = S_{dx}(f) - \mu\Lambda(f) \quad (13)$$

Correspondingly, by the inverse Fourier transform, in the time domain, Eq. (13) can be converted into a convolution integral of the following form,

$$\int_{-\infty}^{\infty} R_x(\tau - \theta) w_{\text{opt}}(\theta) d\theta = R_{dx}(\tau) - \mu\lambda(\tau) \quad (14)$$

Equation (14) has one more term, $-\mu\lambda(\tau)$, than the Wiener–Hopf equation [9], i.e., $\int_{-\infty}^{\infty} R_x(\tau - \theta) w_{\text{opt}}(\theta) d\theta = R_{dx}(\tau)$, which is easily derived from the unconstrained Wiener filter design problem [9] or from Eq. (9) or (10) by ignoring the integral of the constraint. Since $\mu\lambda(\tau)$ is the function of the constraint on the design result and since $\mu\lambda(\tau) = 0$ if and only if there is no constraint on the Wiener filter, Eq. (14) can be called the extended Wiener–Hopf equation. Finally, we take our familiar causal constraint as an example to further illustrate the role of $\mu\lambda(\tau)$ in the extended Wiener–Hopf equation. As stated in the discussion in Eq. (5), requiring the filter to be causal is equivalent to requiring that $w(t) = g(t) = 0$ for $t \in (-\infty, 0)$ and $\lambda(t) = 0$ for $t \in [0, \infty)$ in Eq. (4). Consequently, in the corresponding extended Wiener–Hopf equation under the causal constraint, $\mu\lambda(\tau)$ is an anti-causal function whose purpose is to offset the non-causal component of the Wiener–Hopf equation and to guarantee that the optimal filter is a causal system.

In addition, we develop the other computational formulas of the minimum mean-squared error of the Wiener filter under the integral constraint. From Eqs. (3), (12), (13), and (9), we can obtain

$$\begin{aligned} L(w, \mu) &= \int_{-\infty}^{\infty} \left(S_d(f) - S_{dx}^*(f)W_{\text{opt}}(f) - S_{dx}(f)W_{\text{opt}}^*(f) + S_x(f)|W_{\text{opt}}(f)|^2 \right) df \\ &= \int_{-\infty}^{\infty} S_d(f)df - \int_{-\infty}^{\infty} S_x(f)|W_{\text{opt}}(f)|^2 df \\ &\quad - \int_{-\infty}^{\infty} (G(f)\mu^*\Lambda^*(f) + G^*(f)\mu\Lambda(f)) df - \mu^*\beta - \mu\beta^* \\ &= \int_{-\infty}^{\infty} S_d(f)df - \int_{-\infty}^{\infty} (S_{dx}(f) - \mu\Lambda(f))W_{\text{opt}}^*(f)df \\ &\quad - \int_{-\infty}^{\infty} (G(f)\mu^*\Lambda^*(f) + G^*(f)\mu\Lambda(f)) df - \mu^*\beta - \mu\beta^* \\ &= \int_{-\infty}^{\infty} S_d(f)df - \int_{-\infty}^{\infty} S_{dx}(f)W_{\text{opt}}^*(f)df - \int_{-\infty}^{\infty} G(f)\mu^*\Lambda^*(f)df - \mu^*\beta \end{aligned} \quad (15)$$

The first equality is immediately derived from the Fourier transformation of Eq. (3), after $w(t)$ is replaced by $w_{\text{opt}}(t)$. The second equality is obtained by replacing $S_{dx}(f) -$

$\mu \Lambda(f)$ with $S_x(f)W_{\text{opt}}(f)$ in Eq. (11), since $S_x(f)W_{\text{opt}}(f) = S_{dx}(f) - \mu \Lambda(f)$. The third equality is obtained by replacing $S_x(f) |W_{\text{opt}}(f)|^2$ with $S_{dx}(f)W_{\text{opt}}^*(f) - \mu \Lambda(f)W_{\text{opt}}^*(f)$ in the second equality. The last equality is a simplification of the third equality since $\mu \left(\int_{-\infty}^{\infty} (W_{\text{opt}}(f) - G(f))^* \Lambda(f) df - \beta^* \right) = 0$.

4 Algorithms

In this section, we discuss how $W_{\text{opt}}(f)$ can be obtained from Eq. (12). First, we give the solution of Eq. (12) for the constraint condition in which the filter is required to have finite frequency support. Then, we develop an analytical approach and an iterative technique to obtain $W_{\text{opt}}(f)$ from Eq. (13).

In practice, the filter may be required to be low-pass, band-pass, high-pass, or band-limited. These restrictions can be summarized as the finite frequency-support constraint, which is explained as follows:

If the filter is required to satisfy

$$W(f) = 0, f \notin \Omega \quad (16)$$

where Ω denotes a non-empty subset of the frequency domain, then the filter has finite frequency support. In this case, $\Lambda(f)$ should be chosen such that

$$\Lambda(f) = 0, f \in \Omega \quad (17)$$

Then, we have the corresponding constrained integral $\int_{-\infty}^{\infty} W(f)\Lambda^*(f)df = 0$.

Let

$$\Gamma(f) = \begin{cases} 1, & f \in \Omega \\ 0, & f \notin \Omega \end{cases} \quad (18)$$

From Eqs. (17) and (18), we have $W(f)\Gamma(f) = W(f)$ and $\Lambda(f)\Gamma(f) = 0$. Then, we multiply both sides of Eq. (12) by $\Gamma(f)$. As a result, we obtain the solution of the Wiener filter with finite frequency support:

$$\begin{aligned} W_{\text{opt}}(f) &= W_{\text{opt}}(f)\Gamma(f) = \frac{S_{dx}(f) - \mu \Lambda(f)}{S_x(f)} \Gamma(f) \\ &= \frac{S_{dx}(f)}{S_x(f)} \Gamma(f) = \begin{cases} \frac{S_{dx}(f)}{S_x(f)}, & f \in \Omega \\ 0, & f \notin \Omega \end{cases} \end{aligned} \quad (19)$$

Next, we introduce the analytical approach. By inserting $\frac{S_{dx}(f) - \mu \Lambda(f)}{S_x(f)}$ into the frequency form of Eq. (4), we have $\int_{-\infty}^{\infty} \left(\frac{S_{dx}(f) - \mu \Lambda(f)}{S_x(f)} - G(f) \right) \Lambda^*(f) df = \beta$. Then, solving this integration equation, we obtain the Lagrange factor μ as

$$\mu = \left(\int_{-\infty}^{\infty} \left(\frac{S_{dx}(f)}{S_x(f)} - G(f) \right) \Lambda^*(f) df - \beta \right) / \int_{-\infty}^{\infty} \frac{|\Lambda(f)|^2}{S_x(f)} df \quad (20)$$

As a result, $W_{\text{opt}}(f)$ can be analytically calculated from Eqs. (20) and (12) if the analytical form of $\Lambda(f)$ is known. However, this approach is invalid when $\Lambda(f)$ is unknown or when only its characteristics are known without any information about its form.

Now, we introduce an iterative technique, which is valid regardless of whether the analytical form of $\Lambda(f)$ is known. The iterative formulas are given in the frequency domain by Eq. (21):

$$W'_{i+1}(f) = W_i(f) + \eta (S_{dx}(f) - S_x(f) W_i(f)) \quad (21a)$$

$$\int_{-\infty}^{\infty} (W'_{i+1}(f) - G(f)) \Lambda^*(f) df = \beta_{i+1} \quad (21b)$$

$$\rho_{i+1} = (\beta_{i+1} - \beta) / \int_{-\infty}^{\infty} |\Lambda(f)|^2 df \quad (21c)$$

$$W_{i+1}(f) = W'_{i+1}(f) - \rho_i \Lambda(f) \quad (21d)$$

$$d(W_{i+1}, W_i) = \int_{-\infty}^{\infty} |W_{i+1}(f) - W_i(f)|^2 df \quad (21e)$$

where η is the convergence factor and the subscript (i) denotes the iteration number. $W'_{i+1}(f)$ represents the result of iteration ($i+1$), and $W_{i+1}(f)$ is the largest constrained component of $W'_{i+1}(f)$, which will be defined in the next paragraph. Because $W'_{i+1}(f)$ generally does not meet the constraint defined in Eq. (4), for the sake of the restriction condition in the next iterative operation and fast convergence, we need to extract the largest constrained component of $W'_{i+1}(f)$ to satisfy Eq. (4). The iterative procedure can be summarized as follows:

- (1) Set the convergence factor $0 < \eta < \frac{2}{\max S_x(f)}$ and a positive number ε , which is the iteration termination threshold. Then, arbitrarily choose an initial value $W_0(f)$ and insert it into Eq. (21a) to obtain $W'_1(f)$.
- (2) Using Eq. (21b) through Eq. (21d), or by immediately applying the constraint condition, extract the largest constrained component of $W'_1(f)$, denoted by $W_1(f)$.
- (3) Insert $W_1(f)$ and $W_0(f)$ into (21e) to obtain $d(W_1, W_0)$. If $d(W_1, W_0) < \varepsilon$, the iteration cycle is terminated and $W_1(f)$ is taken as $W_{\text{opt}}(f)$; otherwise, $W_1(f)$ is carried into Eq. (21a) to another iteration cycle.

Before discussing the convergence, we define the largest constrained component.

Definition 1 The largest constrained component of $W'_i(f)$, denoted by $W_i(f)$, along with its residual $\Delta W_i(f)$, is defined as follows:

- (a) $W'_i(f) = W_i(f) + \Delta W_i(f)$, $i > 0$
- (b) $W_i(f)$ meets the constraint defined in Eq. (4)
- (c) The sum of $W_1(f)$ and any component of $\Delta W_i(f)$ does not satisfy Eq. (4).

To show that the largest constrained component is well defined, one immediate task is to show the existence and uniqueness of $W_i(f)$. First, we analyze the existence of $W_i(f)$ as follows:

If $W'_i(f)$ naturally satisfies Eq. (4), then $W_i(f)$ is simply the largest constrained component; in this case, $W_i(f) = W'_i(f)$. If $W'_i(f)$ does not satisfy Eq. (4), i.e.,

$$\int_{-\infty}^{\infty} (W'_i(f) - G(f)) \Lambda^*(f) df = \beta_i \neq \beta \quad (22)$$

Then let $\Delta W_i(f) = \rho_i \Lambda(f)$, in which $\rho_i = \frac{(\beta_i - \beta)}{\int_{-\infty}^{\infty} |\Lambda(f)|^2 df}$. We will show that $W'_i(f) - \Delta W_i(f)$ satisfies the above definition and thus is the largest constrained component. We insert $W'_i(f) - \Delta W_i(f)$ into the frequency-domain form of Eq. (4). By applying Eq. (22), we obtain

$$\begin{aligned} & \int_{-\infty}^{\infty} (W'_i(f) - \Delta W_i(f) - G(f)) \Lambda^*(f) df \\ &= \int_{-\infty}^{\infty} (W'_i(f) - G(f)) \Lambda^*(f) df - \int_{-\infty}^{\infty} \Delta W_i(f) \Lambda^*(f) df \\ &= \beta_i - \int_{-\infty}^{\infty} \rho_i \Lambda(f) \Lambda^*(f) df = \beta_i - (\beta_i - \beta) \\ &= \beta \end{aligned}$$

Therefore, $W'_i(f) - \Delta W_i(f)$ satisfies item (b) of the above definition. We call this term a constrained component of $W'_i(f)$, and $\Delta W_i(f)$ is the corresponding residual. Let us assume that there is a nonzero component of $\Delta W_i(f)$, denoted $\partial W_i(f)$, such that $W'_i(f) - \Delta W_i(f) + \partial W_i(f)$ also satisfies Eq. (4). We insert $W'_i(f) - \Delta W_i(f) + \partial W_i(f)$ into Eq. (4) and can easily obtain $\int_{-\infty}^{\infty} \partial W_i(f) \Lambda^*(f) df = 0$. Since $\Delta W_i(f) = \rho_i \Lambda(f)$, $\Delta W_i(f)$ is orthogonal to its own component $\partial W_i(f)$. However, this orthogonality is possible only if $\partial W_i(f)$ is equal to zero. As a consequence, $W'_i(f) - \Delta W_i(f)$ satisfies item (c) of the above definition and is thus the largest constrained component of $W'_i(f)$. This completes the proof of existence.

We can rewrite the calculation formula for the largest component of $W'_i(f)$, i.e., $W_i(f)$, as follows:

$$W_i(f) = W'_i(f) - \Delta W_i(f) \quad (23)$$

where $\Delta W_i(f) = \rho_i \Lambda(f)$, $\rho_i = (\beta_i - \beta) / \int_{-\infty}^{\infty} |\Lambda(f)|^2 df$. In other words, using the result of the i th iteration of $W'_i(f)$ and Eqs. (22) and (23), we can easily obtain $W_i(f)$. Then, the next iteration will be implemented.

Finally, it should be mentioned that because $\Delta W_i(f)$ is dependent on $\Lambda(f)$, when $\Lambda(f)$ is known, $W_i(f)$ can be immediately derived from Eq. (23). If only some characteristics of $W(f)$ or $\Lambda(f)$ are known, we can make use of these characteristics to extract $W_i(f)$ from $W'_i(f)$. In these cases, $W_i(f)$ is sometimes easily extracted; in Sect. 5, we will present a simulation example for further illustration.

Now we discuss the uniqueness. Let us assume that the largest constrained component of $W'_i(f)$ is not unique, namely both $W_{i,0}(f)$ and $W_{i,1}(f)$ satisfy the above

definition, along with their residuals $\Delta W_{i,0}(f)$ and $\Delta W_{i,1}(f)$, respectively. It is easy to show that $W_{i,0}(f) - W_{i,1}(f)$ is orthogonal to $\Lambda(f)$ since they both satisfy Eq. (4). Because $W_{i,0}(f) + \Delta W_{i,0}(f) = W_{i,1}(f) + \Delta W_{i,1}(f) = W'(f)$, $\Delta W_{i,1}(f) - \Delta W_{i,0}(f)$ is also orthogonal to $\Lambda(f)$. Let $\partial \Lambda(f)$ denote $\Delta W_{i,1}(f) - \Delta W_{i,0}(f)$. Without loss of generality, let $W_{i,0}(f) = W'_i(f) - \rho_i \Lambda(f)$. Then $\Delta W_{i,0}(f) = \rho_i \Lambda(f)$ and $\Delta W_{i,1}(f) = \rho_i \Lambda(f) + \partial \Lambda(f)$. Because $\partial \Lambda(f)$ is orthogonal to $\Lambda(f)$, $\partial \Lambda(f)$ can be regarded as a component of $\Delta W_{i,1}(f)$. For the orthogonality between $\Lambda(f)$ and $\partial \Lambda(f)$, $W_{i,1}(f) + \partial \Lambda(f)$ satisfies Eq. (4), which means that $W_{i,1}(f)$ cannot be the largest component if $\partial \Lambda(f) \neq 0$, which contradicts the conjecture. Therefore, $\partial \Lambda(f)$ is equal to zero. As a result, $\Delta W_{i,1}(f) = \Delta W_{i,0}(f)$ and $W_{i,0}(f) = W_{i,1}(f)$. This completes the proof of the uniqueness.

Now we prove that $W_i(f) \xrightarrow{i \rightarrow \infty} W_{\text{opt}}(f)$ in the sense of the l_2 norm when $0 < \eta < \frac{2}{\max_x S_x(f)}$.
i.e., if:

$$|1 - \eta S_x(f)| < 1 \quad (24)$$

then

$$\lim_{i \rightarrow \infty} \int_{-\infty}^{\infty} |W_i(f) - W_{\text{opt}}(f)|^2 df = 0 \quad (25)$$

Proof According to the definition, $W_{i+1}(f)$ is the largest constrained component of $W'_{i+1}(f)$ and satisfies the restriction condition. Thus, we have

$$\int_{-\infty}^{\infty} (W_{i+1}(f) - G(f)) \Lambda^*(f) df = \beta \quad (26)$$

Note that

$$\int_{-\infty}^{\infty} (W_{\text{opt}}(f) - G(f)) \Lambda^*(f) df = \beta \quad (27)$$

Hence, by subtracting Eq. (26) from (27), we obtain

$$\int_{-\infty}^{\infty} (W_{\text{opt}}(f) - W_{i+1}(f)) \Lambda^*(f) df = 0 \quad (28)$$

In other words, $W_{\text{opt}}(f) - W_{i+1}(f)$ and $\Lambda(f)$ are mutually orthogonal. Since $\Delta W_{i+1}(f) = \rho_{i+1} \Lambda(f)$, $W_{\text{opt}} - W_{i+1}(f)$ is also orthogonal to $\Delta W_{i+1}(f)$.

After finishing the orthogonality analysis, we will prove Eq. (25) in the following discussion.

By subtracting $W_{\text{opt}}(f)$ from both sides of Eq. (21) and substituting $S_x(f)W_{\text{opt}}(f) + \mu \Lambda(f)$ for $S_{dx}(f)$ in Eq. (21), we have

$$W'_{i+1}(f) - W_{\text{opt}}(f) = (1 - \eta S_x(f)) (W_i(f) - W_{\text{opt}}(f)) + \eta \mu \Lambda(f) \quad (29)$$

By replacing $W'_{i+1}(f)$ with $W_{i+1}(f) + \rho_{i+1} \Lambda(f)$ in Eq. (29), we obtain

$$(1 - \eta S_{dx}(f)) (W_i(f) - W_{\text{opt}}(f)) = W_{i+1}(f) - W_{\text{opt}}(f) + (\rho_{i+1} - \eta \mu) \Lambda(f)$$

Since $W_{i+1}(f) - W_{\text{opt}}(f)$ and $\Lambda(f)$ are mutually orthogonal, we have

$$\begin{aligned} & \int_{-\infty}^{\infty} \left| (1 - \eta S_{dx}(f)) (W_i(f) - W_{\text{opt}}(f)) \right|^2 df \\ &= \int_{-\infty}^{\infty} \left| (W_{i+1}(f) - W_{\text{opt}}(f)) + (\rho_{i+1} - \eta\mu) \Lambda(f) \right|^2 df \\ &= \int_{-\infty}^{\infty} |W_{i+1}(f) - W_{\text{opt}}(f)|^2 df + |\rho_{i+1} - \eta\mu|^2 \int_{-\infty}^{\infty} |\Lambda(f)|^2 df \\ &\geq \int_{-\infty}^{\infty} |W_{i+1}(f) - W_{\text{opt}}(f)|^2 df \end{aligned}$$

Then, if Eq. (24) is satisfied, i.e., $|1 - \eta S_x(f)| < 1$, we obtain

$$\begin{aligned} \int_{-\infty}^{\infty} |W_i(f) - W_{\text{opt}}(f)|^2 df &> \int_{-\infty}^{\infty} |(1 - \eta S_{dx}(f)) (W_i(f) - W_{\text{opt}}(f))|^2 df \\ &\geq \int_{-\infty}^{\infty} |W_{i+1}(f) - W_{\text{opt}}(f)|^2 df \end{aligned} \quad (30)$$

As a result, $W_{i+1}(f)$ converges to $W_{\text{opt}}(f)$ in the l_2 norm. Expanding Eq. (24), we have

$$0 < \eta < \frac{2}{\max S_x(f)} \quad (31)$$

This completes the proof of convergence for the iterative algorithm.

When $W_{i+1}(f)$ converges to $W_{\text{opt}}(f)$, $W_{i+1}(f) - W_i(f)$ will converge to zero in the same norm. Hence, we can take $d(W_{i+1}, W_i) < \varepsilon$ as the criterion to terminate the iterative procedure. In addition, we define the approximation error of $W_{\text{opt}}(f)$ by $W_{i+1}(f)$ as

$$\delta(W_{\text{opt}}(f), W_{i+1}(f)) = \frac{d(W_{\text{opt}}(f), W_{i+1}(f))}{d(W_{\text{opt}}(f), 0)} \times 100\% \quad (32)$$

It is clear that $\delta(W_{\text{opt}}(f), W_{i+1}(f))$ is positively correlated with ε . Therefore, to obtain a high approximation degree, ε should be as small as possible, but this increases the number of iterations that must be computed.

Finally, it should be mentioned that although the iterative technique is developed in the frequency domain, it may be possible to implement this technique in the time domain by making use of the Fourier transformation.

5 Examples

In this section, three simulation examples are presented to illustrate the applicability of our methods. In all three examples, we take the same noisy random input $x(t)$ and the same desired output $d(t)$, but impose different restrictions. We assume that $R_x(\tau) =$

Fig. 2 The wave form of $W_{\text{opt}}(f)$ with a low-pass constraint

$$e^{-0.01|\tau|} + 0.01^2\delta(\tau), R_{dx}(\tau) = R_d(\tau) = e^{-0.01|\tau|}, S_x(f) = \frac{0.02}{0.01^2 + (2\pi f)^2} + 0.01^2,$$

$$\text{and } S_{dx}(f) = S_d(f) = \frac{0.02}{0.01^2 + (2\pi f)^2}.$$

Case 1 The Wiener filter is required to be a low-pass system with a 40-Hz bandwidth, i.e.,

$$W(f) = 0, f \notin (-20, 20) \quad (33)$$

Obviously, this is a finite frequency-support constraint. From Eqs. (16), (18), and (19), we obtain

$$W_{\text{opt}}(f) = \begin{cases} \frac{S_{dx}(f)}{S_x(f)} = \frac{0.02}{0.02 + 0.01^4 + (0.02\pi f)^2}, & f \in (-20, 20) \\ 0, & f \notin (-20, 20) \end{cases}$$

The wave form of $W_{\text{opt}}(f)$ is shown in Fig. 2.

Case 2 The Wiener filter is required to satisfy the following integration equation:

$$\int_{-200}^{200} W(f) df = 2 \quad (34)$$

It is straightforward to show that this integral is equivalent to Eq. (4) when $G(f) = 0$ and

$$\Lambda(f) = \begin{cases} 1, & f \in (-200, 200) \\ 0, & f \notin (-200, 200) \end{cases} \quad (35)$$

Fig. 3 The wave form of $W_{\text{opt}}(f)$ from the analytical method for case 2

The computational procedure of the analytical method is very simple. First, using Eq. (20), we obtain $\mu = 1.36 \times 10^{-5}$. Then, from Eq. (12), we obtain

$$\begin{aligned} W_{\text{opt}}(f) &= \frac{S_{dx}(f) - \mu \Lambda(f)}{S_x(f)} = \frac{0.02 - 1.36 \times 10^{-5} \times (0.01^2 + (2\pi f)^2) \Lambda(f)}{0.02 + 0.01^4 + (0.02\pi f)^2} \\ &\approx \frac{0.02 - 0.536 \times 10^{-3} f^2 \Lambda(f)}{0.02 + 0.01^4 + (0.02\pi f)^2} \end{aligned}$$

The wave form of $W_{\text{opt}}(f)$ is shown in Fig. 3.

Next, we use the iterative algorithm to solve this problem again. The initial parameters of this iteration is chosen as $W_0(f) = \frac{S_{dx}(f)}{S_x(f)} = \frac{0.02}{0.02 + 0.01^4 + (0.02\pi f)^2}$, $\varepsilon = 10^{-6}$, and $\eta = 0.0099 < \frac{2}{\max S_x(f)} \approx 0.01$. In the above iterative algorithm, after 171 iterations, the criterion to end the iterative procedure is satisfied. As a result, we obtain $W_{171}(f)$ as shown in Fig. 4.

From Eq. (32), we obtain the approximation error, $\delta(W_{\text{opt}}(f), W_{i+1}(f)) = 1.37\%$. We can further decrease $\delta(W_{\text{opt}}(f), W_{i+1}(f))$ by reducing ε , but the computation time will increase. In this instance, if we take $\varepsilon = 10^{-13}$ and hold the other initial parameters constant, then, after 162,834 iterations, the iterative procedure is terminated, and we obtain $\delta(W_{\text{opt}}(f), W_{162834}(f)) = 0.33\%$.

Case 3 The Wiener filter is required to satisfy the following constraint condition:

$$w(t) = \begin{cases} g(t), & t \notin [0, 500) \\ \text{unknown}, & t \in [0, 500) \end{cases} \quad (36)$$

Fig. 4 The wave form of $W_{171}(f)$ from the iterative method for case 2

where

$$g(t) = \begin{cases} 0.03e^{-10t} \sin(40\pi t), & t \in [500, 1000) \\ 0, & t \notin [500, 1000) \end{cases} \quad (37)$$

For this restriction, we take

$$\lambda(t) = \begin{cases} \text{unknown}, & t \notin [0, 500) \\ 0, & t \in [0, 500) \end{cases} \quad (38)$$

Then we can convert the restriction condition to the integration equation

$$\int_{-\infty}^{\infty} (w(t) - g(t)) \lambda^*(t) dt = 0$$

Because $\lambda(t)$ is unknown, we can use only the iterative algorithm.

We implement the iterative procedure in the time domain. We set $w_0(t) = IFT \left\{ \frac{S_{dx}(f)}{S_x(f)} \right\}$, $\varepsilon = 10^{-25}$, and $\eta = 0.0099$, where $IFT \{\bullet\}$ denotes the inverse Fourier transform. In this instance, we cannot use Eq. (23) to extract the largest component of $w'_i(t)$, because $\lambda(t)$ is unknown. However, according to the constraint given in Eq. (36), we can immediately obtain

$$w_i(t) = \begin{cases} g(t) & , t \notin [0, 500) \\ w'_i(t) & , t \in [0, 500) \end{cases}$$

Fig. 5 The wave form of $w_4(t)$ from the iterative algorithm for case 3

For this instance, in the fourth iteration, the criterion for terminating the iterative procedure is satisfied. We obtain $w_4(t)$ as shown in Fig. 5.

6 Conclusions

This paper presents a general theoretical method for designing Wiener filters with additional constraint conditions. By introducing an integral expression to represent the constraints, we derive and analytically solve the formula for a more constrained Wiener filter. Compared to the classical design method, our new approach not only incorporates broader constraints but also relaxes the requirement of rational power spectra. The classical method of designing Wiener filters is appropriate only for cases of causality, and the causal solution is obtained only if the input self-power spectrum and the cross-power spectrum of the input and the output are both rational functions. Finally, this paper demonstrates several practical applications of our techniques using simulations.

Acknowledgements This work is supported by Science and Technology Supporting Program Project of Sichuan Province (2014GZ0083).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. N. Cazi and T.V. Sreenivas, Enhancement of binaural speech using codebook constrained iterative binaural Wiener filter, in *Interspeech, Conference of the International Speech Communication Association*, Brighton, UK, pp. 1335–1338 (2009)

2. S. Chehresa and M.H. Savoji, Improved codebook constrained Wiener filter speech enhancement, in *2010 5th International Symposium on Telecommunications (IST'2010)*
3. V.P. Dimri, The multichannel constrained Wiener filter, chapter 33, in *Geophysical Data Inversion Methods and Applications, Proceedings of the 7th International Mathematical Geophysics Seminar Held at the Free University of Berlin*, February 8–11 (1989). ISBN: 978-3-528-06396-2 (Print) 978-3-322-89416-8 (online)
4. V.P. Dimri, L.S. Resende et al., A linearly-constrained approach for filtered-X Wiener filtering, in *20th European Signal Processing Conference*, pp. 16–20, Bucharest, Romania, Aug. 27–31 (2012)
5. J.N. Laneman and G.W. Wornell, Energy-efficient antenna sharing and relaying for wireless networks, in *Wireless Communications and Networking Conference (WCNC 2000)*, vol. 1 (2000)
6. D. Marquardt, E. Hadad et al., Theoretical analysis of linearly constrained multi-channel wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids. *IEEE Trans. Audio Speech Lang. Process.* **23**(12), 2384–2397 (2015)
7. T. Michaeli and Y.C. Eldar, Constrained linear minimum MSEE estimation, in *CCIT Report 641*, Electrical Engineering Department, Technion Israel Institute of Technology (2007)
8. T. Michaeli and Y.C. Eldar, Constrained nonlinear minimum MSE estimation. in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3681–3684 Las Vegas, NV, (2008)
9. A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1977)
10. B. Picinbono, M. Bouvet, Constrained Wiener filtering. *IEEE Trans. Inf. Theory* **33**(1), 160–166 (1987)
11. B. Picinbono, A geometrical interpretation of signal detection and estimation. *IEEE Trans. Inf. Theory* **26**, 493–497 (1980)
12. V.R. Rao, R. Murthy, K.S. Rao, Speech enhancement using cross-correlation compensated multi-band Wiener filter combined with harmonic regeneration. *J. Signal Inf. Process.* **2**, 117–124 (2011)
13. R. Sharma, B.D. Van Veen, Large modular structures for adaptive. Beamforming and the Gram–Schmidt preprocessor. *IEEE Trans. Signal Process.* **42**(2), 448–451 (1994)
14. W.D. Stirling, Least squares subject to linear constraints. *J Roy Stat Soc C-App* **30**, 204–212 (1981)
15. N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (The Technology Press, New York, 1950)