Generalized Wiener Filtering for Time-Variant Linearly Transformed Signals in Noise

William A. Gardner

12 April 2026

Abstract

Second-order Temporal Statistical Analysis of Signals subjected to distorting time-variant linear transformations and additive corruption is conducted. Explicit input/output relations for the time-average mean and spectral density of time-average power are derived and shown to be determined by the corresponding input statistics and the statistics of the time-varying linear transformation, including the mean system function and generalized spectrum of the time-variant linear operator (transformation). The solution is obtained by an application of a simplified modern version of Wiener's concept of generalized harmonic analysis of operator-valued functions of time for the general case of arbitrary but stationary time-variation of the transformation. The input/output relations are used to derive a *Generalized form of Wiener Filter* that combats not only additive corruption (e.g., noise and/or interference) and time-invariant signal dispersion, but also multiplicative corruption, time-variant dispersion, and time warping effects. This work is contrasted with earlier treatments of the similar topic that uses stochastic process models for signals and random but time-invariant convolutions, thereby applying to only populations of convolutions and populations of signals, rather than to individual time-varying transformations and signals.

1. Introduction, Overview, and Results Summary

Norbert Wiener's 1930s treatise [1] introduces the Generalization of *Harmonic Analysis* from real-valued functions of real variables on finite intervals and on the entire real line when the functions die away as time increases in magnitude, especially those whose squared magnitude is integrable over the real line, which renders them Fourier transformable, to *persistent functions* that do not die away and also do not grow without bound over the entire real line. He called such functions *stationary* and gave a mathematical definition of this defining property: the existence of the function's limit temporal autocorrelation function. This is an appropriate model (in common use today) for persistent signals in the field of statistical signal processing and dating back to Wiener's original work in *Generalized Harmonic Analysis (GHA)* [1], cf. [2, Secs.3.2, 10.2-10.6] for a modern treatment. These persistent signals include periodic, almost periodic, and erratic unpredictable functions.

Generalized harmonic analysis gave us the first-ever definition of the *Power Spectral Density* (PSD) for an individual persistent function, which was later translated into the theory of the stationary stochastic process for which the individual function is replaced with a population of functions [2, Secs.3.2], producing a *very different* type of quantity, except in the special case for which the stochastic process model is ergodic in some sense, in which case time averages converge, as averaging time grows without bound, to population averages (expected values) [2, Sec.3.2 & Chap. 8] in some probabilistic sense. But there is nothing stochastic about the functions that Wiener's theory addresses. Despite the use of Wiener's name for the stochastic process counterpart of Wiener's seminal theory for non-stochastic functions, including his Fourier transform relationship between his definition of PSD and his definition of autocorrelation function and his solution for optimum filters in terms of PSDs, Wiener did not originate the stochastic process theory that was slowly introduced in the 1950s and became popular in the 1960s, thirty years after Wiener's introduction of his theory of Generalized harmonic analysis and associated optimum filtering, smoothing, and prediction. Nevertheless, it was back in the 1950s a trivial matter to translate his theory for stationary functions to the analogous theory for wide-sense stationary stochastic processes. So, Wiener's work is seminal.

In this article, PSD means spectral density of *time-averaged* power of an individual function. In the translated theory of the stochastic process (not dealt with in this article), PSD means spectral density of population-averaged power, which is called *expected power*--a quantity that is significantly more abstract than time-averaged power, and generally not directly relevant in a practical sense when signal populations are not of interest [2, Sec.3.2]. This important fact seems to have been lost over time. The substantial abstraction of the stochastic process model essentially everyone today is taught and uses, is both unnecessary and undesirable for empirical problems involving real measurements or observations of individual functions and not hypothesized abstract infinite ensembles or populations of hypothesized functions or realizations of the stochastic process.

In order to gain the maximum understanding and corresponding utility of Wiener's theory and the generalization introduced herein, the reader is encouraged to pretend (s)he knows nothing about stochastic processes. This removes any need for even more abstract notions like ergodicity, which is simply an abstract mathematical patch intended to establish a link between the realizations of a stochastic process and an individual stationary function as defined by Wiener. And this patch doesn't even achieve this link. Rather, it links statistics like the time-average PSD calculated from a single abstract realization of a mathematical stochastic process to the abstract expected-PSD for the process. Nowhere in this patch is there a link between a direct model for an empirical function and a stochastic process model. The properties of the stationary functions in

Wiener's theory are generally not the same as the properties of the sample paths of a stochastic process cf. [3].

All past work on the specific topic of this article, centered on optimum time-invariant filters for signals passing through time-variant linear channels, appears to be based on the more abstract stochastic process model. The earliest work appears to be that of Lotfi Zadeh, dating back to 1950. He treats input/output statistics relations for stochastic time-varying systems with stochastic inputs [4]-[6]. Bello's early 1960s paper presents the stochastic counterpart of the topic addressed in the present paper [7]. A few other papers treat related topics concerning stochastic process models for time-varying filters, such as conditions for separation of signals and noise [8]. A broad comprehensive treatment of time-varying channels in communication systems is provided in the book [9]. It appears that Wiener's relatively abstract treatment of generalized harmonic analysis of operator-valued functions of time, which appears to have not been applied in engineering in the ensuing 95 years, is the only past work that is directly relevant to the approach taken in this article, but Wiener does not appear to have derived input/output relations for time average statistics associated with time-variant operators which, as shown in this article, gives rise to a need for Wiener's concept of generalized harmonic analysis of operators.

In this paper, it is explained how Wiener's concept of generalized harmonic analysis of relatively abstract operator-valued functions of time [1], can be modified and thereby simplified to apply to analysis of 2nd order time and frequency statistics (temporal mean and PSD) for long-term time-variant linear filtering, which we shall call *persistent filtering*. The general result is then simplified by focusing on special cases and specific applications. The contribution of this article is theoretical and should be easily understood by engineering readers properly trained in statistical signal processing, including Wiener's seminal work, which is generalized in this article. Therefore, the applications included herein are among the simplest-to-understand mathematical examples that demonstrate the theory is not vacuous and yields meaningful mathematical solutions for optimum filters. It is shown that the essence of the difference between the classic Wiener filter and the generalization derived herein is the presence, in the explicit solution given in (18)-(20) and (22)-(24), of a single spectral covariance term (defined in terms of time averages, not expected values) for the time-varying channel filter. Consequently, in the interest of elegance, the three examples given focus on this specific term. Anyone familiar with applications of Wiener filter theory to stochastic processes, can trivially formulate various applications of this generalized theory.

This article avoids the level of mathematical abstraction in Wiener's GHA paper [1] by working directly with the time-domain kernels of operators specifically for linear time-varying convolutions, and it also simplifies the generalized spectrum of a linear-operator-valued function of time by focusing on especially tractable examples. One class of examples are those for which the time-variant channel filter is separable in its variables in either of two ways, as illustrated in Case 1 and Case 2 developed below. Another class of examples are those for which long-term filters of interest vary slowing over time, which enables approximation of the relatively complicated general channel-filter input/output relation for PSDs (5)-(6), with a type of convolution of the input PSD with the generalized spectrum of the operator-kernel-valued function of time as shown in (7)-(11).

The results obtained are explicit formulas for 2nd order statistics of the output signal from a linear time-variant channel filter, namely the time-average mean (3) and time-average PSD (5)-(6) and approximation thereof (10)-(11) in terms of these same types of statistics for the input signal and statistics of the filter. The derivation of the Generalized Wiener Filter herein illustrates the use of the derived input/output relations.

The problem of optimum time-invariant filtering of individual signals subjected to linear time-variant distorting transformations and additive corruption is formulated in terms of the optimality criterion of minimum-time-averaged squared-magnitude of error between the desired signal and its estimate produced by the time-*invariant* filter, and an elegant revealing comparison is made between this and the optimum filter for the time-*invariant* time-averaged value of the signal corrupting time-variant transformation.

2. Derivation of Output Statistic for Time-varying Dispersion

Let $g(t,\tau)$ be the time-varying impulse response for the linear operator that is a time-varying convolution in which case the output $y(t)$, produced by this operator, corresponding to the input $x(t)$ is given by

$$y(t) = \int_{-\infty}^{\infty} g(t,\ \tau)\, x(t-\tau)\, d\tau$$

(1)

The variable τ is called the age variable and *t* in the argument of *h* is the time variable whose presence reflects the fact that the nature of the spreading (dispersive) effect on the input depends on the time *t* the output is observed. Both *g* and *x* are assumed to be complex valued in general.

Because of this dependence of *h* on *t*, this operator is typically called a *time-variant filter* in the field of signal processing. In mathematics, by interpreting the input and output to be vectors in the same linear vector space, the filter can be interpreted as a linear operator that maps input vectors to output vectors. In classical liner system theory, the Fourier transform of $g(t, \tau)$ in the age variable τ produces what is called the *system function* and is given

$$G(t,f) = \int_{-\infty}^{\infty} g(t, \tau)\, exp\{-i2\pi f \tau\} d\tau$$

(2)

When the input is a complex sinewave at frequency, f, namely $exp\left\{i2\pi f\tau\right\}$, the output can easily be shown to be given by $G(t,f)\, exp\{i2\pi f t\}$. It is shown in (3) below that the time-average of the system function is useful in the input/output relation of signal means, but in the input/output relation for signal PSDs, it is shown in (5) below that the cross-spectrum of the impulse response function at two distinct values of the age variable is used. In these equations below the angle brackets represent time averages over the entire real line. It is mentioned in passing that for the special case of time-invariant filters, the system function becomes independent of *t* and is identical to the transfer function of the filter.

In the following equation manipulations, it is assumed that the interchange of the convolution integral with the time-average operation is justified by the properties of the particular functions involved, which are assumed to be complex-valued in general, but may be real valued in special cases. Nevertheless, the acceptability of specific conditions under which these assumptions are valid depends on the application. But this interchange is valid for all mathematical models of practical channel filters. The mathematical conditions under which the various operations can be interchanged are various, and the more restrictive they are, the less applicable the obtained results will be. Little is gained here by selecting specific conditions under which specific operation interchanges are valid.

Using (1), the time-average mean of the filter output is given by

$$
\begin{aligned}
\langle y(t) \rangle &= \left\langle \int_{-\infty}^{\infty} g(t,\ \tau)\, x(t-\tau)\, d\tau \right\rangle \\
&= \int_{-\infty}^{\infty} \langle g(t,\ \tau)\, x(t-\tau) \rangle d\tau \\
&= \int_{-\infty}^{\infty} \langle g(t,\ \tau) \rangle \langle x(t-\tau) \rangle d\tau \\
&= \int_{-\infty}^{\infty} \langle g(t,\ \tau) \rangle \langle x(t) \rangle d\tau \\
&= \int_{-\infty}^{\infty} \langle g(t,\ \tau) \rangle d\tau\ \langle x(t) \rangle \\
&= \left\langle \int_{-\infty}^{\infty} g(t,\ \tau)\, d\tau \right\rangle \langle x(t) \rangle \\
&= \langle G(t,0) \rangle \langle x(t) \rangle
\end{aligned}
$$

(3)

The third equality is a result of the assumption that the two functions of time, *t*, in the integrand are statistically independent or at least their cross-correlation is identically zero. This assumption is valid for all practical applications except possibly for those involving signal feedback from filter output to filter input. The last equality follows from (2).

Using (1), the autocorrelation of the filter output is given by

$$
\begin{aligned}
R_y(u) &= \left\langle y(t)\, y^*(t-u) \right\rangle \\
&= \left\langle \int_{-\infty}^{\infty} g(t,\tau)\, x(t-\tau)\, d\tau \int_{-\infty}^{\infty} g^*(t-u,\tau')\, x^*(t-u-\tau')\, d\tau' \right\rangle \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\langle g(t,\tau)\, x(t-\tau)\ \ g^*(t-u,\tau')\, x^*(t-u-\tau') \right\rangle d\tau d\tau' \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\langle g(t,\tau)\, g^*(t-u,\tau') \right\rangle \left\langle x(t-\tau)\, x^*(t-u-\tau') \right\rangle d\tau d\tau' \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_g(u,\ \tau,\tau')\, R_x(u-\tau+\tau')\, d\tau d\tau'
\end{aligned}
$$

(4)

It is obvious from the last two lines of (4) that R_g is the cross-correlation of the impulse response function at two distinct values of the age variable. It follows that its Fourier transform, which appears in (5) below is the corresponding cross-spectrum. This is the one quantity that does not arise in Wiener's classical optimum filtering for time-invariant channel filters. It has been assumed in (4) that the time average of the product of two convolutions is the double convolution of the time average of the product of convolution integrands. This is guaranteed for mathematical models of *all practical filters* and can be more explicitly guaranteed mathematically using any of several standard assumptions about interchanging limits which, in this case, define convolutions and time averages.

Fourier transforming this output autocorrelation to obtain the output PSD yields

$$
\begin{aligned}
S_y(f) &= \mathscr{F}_u\{R_y(u)\} \\
&= \mathscr{F}_u\left\{ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_g(u,\ \tau,\tau')\, R_x(u-\tau+\tau')\, d\tau d\tau' \right\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathscr{F}_u\left\{ R_g(u,\ \tau,\tau')\, R_x(u-\tau+\tau') \right\} d\tau d\tau' \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_g(f,\tau,\tau') \underset{f}{\otimes} \left[S_x(f)\, exp\{-i2\pi f(\tau-\tau')\} \right] d\tau d\tau'
\end{aligned}
$$

(5)

where the convolution theorem for Fourier transform has been used to obtain the last equality. It is assumed that the PSD of the output *y(t)* exists, which requires that the generalized spectrum of the filter operator exists, as in Wiener's work [1]. However, the Fourier transforms of the auto- and cross-correlations in (5) may not exist if they are not continuous at the origin, age-variable equal to zero. This is proven by Wiener in [1] and is the primary reason Wiener works with the integrated Fourier transform. Since all practical models for signals are continuous, I adopt this assumption and thereby avoid the use of the integrated Fourier transform, which is less familiar to engineers. One particularly simple example of a function whose autocorrelation is not continuous is the linear chirp signal that is persistent over the entire real time line, that is $x(t) = \sin(t^2)$.

The quantities that are convolved in the last line of (5) are the linearly phase-shifted input PSD (Fourier transform of the time-shifted autocorrelation of the input) and the *generalized spectrum of the kernel-valued function* of time representing the operator-valued function of time that is the time-variant channel filter, this spectrum being defined to be the Fourier transform, in the variable u, of the autocorrelation $R_g(u, \tau, \tau')$, parameterized by and , of the kernel-valued function of time, $g(t, \tau)$, which is the time-variant impulse response function:

$$S_g(f, \tau, \tau') = \mathscr{F}_u\{R_g(u, \tau, \tau')\}$$

(6)

This generalized spectrum of g is function-valued due to its dependence on the two parameters and , and it can also be interpreted to be the generalized spectrum of the operator represented by the operator kernel g .

Equation (5) is the most general formula we can obtain for the output PSD of a linear time-variant filter. It is quite messy, involving three integrals. However, there are situations of practical interest for which analytical simplifications are possible. One such example is that for which the dispersive part of the transformation and the time varying part are separable, as shown below:

Case 1. Product modulator *a* follows time-invariant dispersion *b*, $g(t, \tau) = a(t)\, b(\tau)$:

$$R_g(u, \tau, \tau') = R_a(u)\, b(\tau)\, b^*(\tau')$$

$$S_g(f, \tau, \tau') = S_a(f)\, b(\tau)\, b^*(\tau')$$

Case 2. Product modulator *a* precedes time-invariant dispersion *b*, $g(t, \tau) = a(t - \tau)\, b(\tau)$

$$R_g(u,\ \tau,\ \tau') = R_a(u-\tau+\tau')\,b(\tau)\,b(\tau')$$
$$S_g(f,\ \tau,\ \tau') = S_a(f)\,exp\{-i2\pi f(\tau-\tau')\}b(\tau)\,b^*(\tau')$$

For applications in which the filter varies with time much more slowly than the input signal varies with time, the input signal autocorrelation function is much narrower in width than the filter cross-correlation function at all pairs of lags $\tau,\ \tau'$, and this enables a useful simplifying approximation. This relationship between narrowness of functions assumes that both functions have approximate widths that are finite, which rules out any finite-strength additive periodic components in either the filter time variation or the input signal. Under this condition, we have the approximation

$$R_g(u,\ \tau,\tau')\,R_x(u-\tau+\tau') \simeq R_g(\tau-\tau',\ \tau,\tau')\,R_x(u-\tau+\tau')$$

(7)

where

$$\begin{aligned} R_g(\tau-\tau',\ \tau,\tau') &= \left\langle g(t,\tau)\,g^*(t-\tau+\tau',\tau')\right\rangle \\ &= \left\langle g(t+\tau,\tau)\,g^*(t+\tau',\ \tau')\right\rangle \end{aligned}$$

(8)

With this approximation, the second line of (5) is approximated by

$$S_y(f) \simeq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathscr{F}_u\{R_g(\tau-\tau',\ \tau,\tau')\,R_x(u-\tau+\tau')\}d\tau d\tau'$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}R_g(\tau-\tau',\ \tau,\tau')\,S_x(f)\,exp\{-i2\pi f(\tau-\tau')\}d\tau d\tau'$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}R_g(\tau-\tau',\ \tau,\tau')\,exp\{-i2\pi f(\tau-\tau')\}d\tau d\tau'\ S_x(f)$$

(9)

As the final step of simplification, the above double Fourier transform is performed to obtain the new spectral quantity

$$\mathcal{S}_g(f) \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_g(\tau-\tau',\ \tau,\tau')\, exp\{-i2\pi f(\tau-\tau')\} d\tau d\tau'$$

(10)

which, when substituted into (9), yields the following desired approximate input/output PSD relation for slowly time-varying filters:

$$S_y(f) \simeq \mathcal{S}_g(f)\, S_x(f)$$

(11)

Various conditions on the memory of the filter characterized by $g(t,\tau)$ (width in τ) can assure existence of the double Fourier transform in (10). For example, finite memory reduces the intervals of integration in (10) to finite intervals, in which case an assumption of continuity of the cross-correlation in (10) is sufficient. This new spectral quantity (10) appearing in (11) has no *direct* relation to the generalized spectrum (6) of the filter, which spectrum arises in (5), because of the approximation (7) used to arrive at a modification of the cross-correlation in (7), leading to (8). Nevertheless, in this special case for which the stationary time-varying filter fluctuates slowly in time relative to the fluctuations in the filter input signal, it is the appropriate quantity for approximating the filter output PSD from the filter input PSD. It is not a genuine approximation to the generalized spectrum (6), which must be convolved with the phase-shifted input PSD to obtain the exact output PSD. Fortunately, the approximation used here produces a more convenient formula for calculation. Nevertheless, its definition (10) defies simple intuitive interpretation beyond its being the double Fourier transform at a single frequency of the somewhat odd cross-correlation (8) of the modified impulse-response function.

3. Derivation of Optimum Filter for Noise Plus Signal distorted by Time-Varying Dispersion

By using the *orthogonal projection theorem*, as demonstrated in [2, Secs.13.2-13.4], it can be shown that the optimum filter, which minimizes the time-averaged squared magnitude of the error between a desired signal *s(t)* and a filtered version of corrupted measurements *z(t)* of *s(t)*, satisfies the necessary and sufficient *orthogonality condition*

$$\left\langle \left[s(t) - \widehat{s}(t)\right] z^*(t-u) \right\rangle = 0$$

(12)

for all real *u* and for which the estimate $\widehat{s}(t)$ is given by

$$\hat{s}(t) = \int_{-\infty}^{\infty} h(t-v)\, z(y)\, \mathrm{d}v$$

(13)

where $\langle . \rangle$ denotes average over all time *t* in the set of real numbers.

It easily follows from (12) that, for the filter impulse response function *h* to be optimum, it is required that

$$R_{sz}(u) = R_{\hat{s}z}(u)$$

(14)

where

$$\begin{aligned} R_{\hat{s}z}(u) &= \left\langle \hat{s}(t)\, z^*(t-u) \right\rangle \\ &= \left\langle \int_{-\infty}^{\infty} h(v)\, z(t-v)\, \mathrm{d}v\, z^*(t-u) \right\rangle \\ &= \int_{-\infty}^{\infty} h(v) \left\langle z(t-v)\, z^*(t-u) \right\rangle \mathrm{d}v \\ &= \int_{-\infty}^{\infty} h(v)\, R_z(u-v)\, \mathrm{d}v \\ &= h(u) \otimes R_z(u) \end{aligned}$$

(15)

The interchange of time integration is justified if the memory of the filter drops off fast enough (which can be confirmed once *h(u)* is solved for) and the autocorrelation of *z(t)* exists. Therefore, using (14), (15), and the convolution theorem for Fourier transforms, we obtain the following necessary and sufficient condition

$$S_{sz}(f) = H(f)\, S_z(f)$$

(16)

where

$$S_{sz}(f) = \int_{-\infty}^{\infty} R_{sz}(u)\, exp\{-i2\pi f u\} \mathrm{d}u$$

(17)

is the cross-spectral density and, similarly, $S_z(f)$ is the PSD, and where $H(f)$ is the transfer function of the filter (Fourier transform of *h*). Consequently, the optimum filter is specified by the following equation, and is known as the *Wiener filter* after its originator Norbert Wiener [10]

$$H(f) = \frac{S_{sz}(f)}{S_z(f)}$$

(18)

The Cauchy-Schwartz inequality can be used to show that the denominator cannot be zero at any frequency *f* unless the numerator also is zero at that *f*.

Let us now consider the following generic model for the corrupted data

$$z(t) = \int_{-\infty}^{\infty} g(t,u)\, s(t-u)\, \mathrm{d}u + n(t)$$

(19)

where

$$\langle s(t-u)\, n^*(t) \rangle = 0$$

(20)

for all real *t* and *u*, and where *g* is the impulse response function for some time-varying linear transformation, which can be considered a time-varying dispersion of the signal *s(t)* except in the cases for which the dependence of *g* on *u* consists of a single Dirac delta representing a time-varying delay $\delta(u - \tau(t))$ or the product of a function of *t* alone and such a direct delta, $a(t)\, \delta(u - \tau(t))$, representing a time-varying delay and time-varying attenuation.

Using (19), we can calculate the cross-correlation

$$
\begin{aligned}
R_{sz}(\tau) &= \langle s(t)\, z^*(t-\tau) \rangle \\
&= \left\langle s(t) \int_{-\infty}^{\infty} g^*(t-\tau, u)\, s^*(t-\tau-u)\, \mathrm{d}u \right\rangle \\
&= \int_{-\infty}^{\infty} \langle g^*(t-\tau, u)\, s(t)\, s^*(t-\tau-u) \rangle \mathrm{d}u \\
&= \int_{-\infty}^{\infty} \langle g^*(t-\tau, u) \rangle \langle s(t)\, s^*(t-\tau-u) \rangle \mathrm{d}u \\
&= \int_{-\infty}^{\infty} \langle g^*(t, u) \rangle R_s(\tau+u)\, \mathrm{d}u \\
&= \langle g^*(t, -\tau) \rangle \otimes \langle R_s(\tau) \rangle
\end{aligned}
$$

(21)

Using the usual assumptions justifying mathematical operations used throughout the earlier part of this article, (21) is valid and Fourier transforming this convolution result yields

$$S_{sz}(f) = \langle G(t,f)^* \rangle S_s(f)$$

where $\bar{G}(f)$ is the *time-averaged system function.*

Substituting this expression into (18) yields

$$H(f) = \frac{\bar{G}(f)^* S_s(f)}{S_z(f)}$$

(22)

Again using (19), we can calculate the PSD in the denominator of (18) by first calculating the autocorrelation function. This calculation is similar to (4) - (6), where *z(t)* here represents *n(t)* plus *y(t)* there and *s* here represents *x* there. The result of this calculation substituted into (22) yields

$$H(f) = \frac{\bar{G}(f)^* S_s(f)}{S_w(f) + S_n(f)}$$

(23)

for which $S_w(f)$ is the PSD of the output of the transformation *g(t, u)* with input *s(t)* and can be calculated as in (5)-(6). No simpler expression for $S_w(f)$ than (5) with *y* replaced by *w* and *x* replaced by *s* can be obtained without considering specific models for the impulse response *g(t, u)* in (19).

As shown in Section 2, if the transformation *g(t, u)* varies with time much more slowly than the input signal varies with time, the input signal autocorrelation function is much narrower in width than the transformation cross-correlation function at all pairs of lags $u = \tau,\ \tau'$, and this enables a useful simplifying approximation to (5) -(6), which is shown in (10)-(11). This relationship between narrowness of correlation functions assumes that both functions have approximate widths that are finite. Nevertheless, this condition under which this simplification of (23) is valid is of limited practical interest because this is the condition under which adaptive filtering can converge to a good approximation to the optimum *time-varying* filter for estimating *s(t)*, although this does require a means for adaptation [9]. The usefulness of a training signal prior to performing filtering for adaptation is minimal for adapting to time-varying corruption because the time variation requires ongoing training. But, when a training signal can be derived from the observed data *z(t)*, then adaptation can be practical. An important example is the class of property-restoral adaptation methods, including the signal properties of constant modulus [11] and cyclostationarity [12]. Algorithms for adaptation are reviewed in [13].

In order to compare the generalized Wiener filter *H(f)* shown in (23) for time-varying linear corruption with the classical optimum filter for the time average of time-varying linear corruption $H_{TI}(f)$, originally derived by Wiener, we can add to and subtract from $S_w(f)$ in (23) the product of the means of $g(t,\tau)$ and $g^*(t,\tau')$ from the cross-correlation of these two functions of time inside the formula for $S_w(f)$ as first defined in (5) with *y* and *x* there replaced with *z* and *s* here. This produces the re-expression of (23) as

$$H(f) = \frac{\overline{G}(f)^* S_s(f)}{\left|\overline{G}(f)\right|^2 S_s(f) + S_{cov}(f) + S_n(f)}$$

(24)

The suboptimum filter $H_{TI}(f)$ for the data containing a time-varying dispersion is simply given by (24) with the contribution $S_{cov}(f)$, from the covariance of $g(t,\tau)$ and $g^*(t,\tau')$ mixed with the PSD $S_s(f)$ as in (5) with the cross-correlation of *g's* replaced with its cross-covariance, removed from the denominator. It can be seen that this $S_{cov}(f)$ term in (24) determined by the cross-covariance of $g(t,\tau)$ and $g^*(t,\tau')$ has an impact on the generalized Wiener filter that is analogous to that of $S_n(f)$ due to the additive noise *n(t)* in the model (19). At frequencies for which the left-most term in the denominator of (24) dominates the middle term, the impact of time variation of the dispersive channel filter is negligible; also, at frequencies for which the sum of both right-most terms in the denominator of (24) are dominated by the left-most term, the Generalized Wiener filter is closely approximated by the reciprocal of the time average of the system function $\overline{G}(f)$ which is simply the equalizer for the average linear corruptive transformation. At all other frequencies, this generalized Wiener filter is substantially affected by the time variations of the dispersive channel filter.

The impact of this cross-covariance $S_{cov}(f)$ on the minimized time-average squared magnitude error also can be revealed using (24) in the time-average analog of the expected-value formulas (13.115), (13.121) - (13.124) in [2, Chap. 13].

Example

Consider a linear time-varying transformation consisting of time-variant attenuation and time-variant wideband Doppler (time dilation). When the time-variations here are not much slower than the time variation of the signal itself, adaptive filtering methods are typically not able to provide accurate tracking. In this case, knowledge of the optimum time-invariant filter and its performance can be useful, particularly if the cross-covariance of the impulse response at one age-variable value with that at another age-variable value, are not too large. The general formula (23) or (24) can be made specific for attenuation and Doppler by using the model $g(t,\tau) = a(t)\,\delta(\tau - \varepsilon(t))$ where is a time-varying delay with can model a time-varying time scaling, which leads to

$$
\begin{aligned}
R_g(u,\tau,\tau') &= \langle g(t,\tau)\, g^*(t-u,\tau') \rangle \\
&= \langle a(t)\, a^*(t-u) \rangle \langle \delta(\tau - \varepsilon(t))\, \delta(\tau' - \varepsilon(t-u)) \rangle \\
&= R_a(u) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_{\varepsilon(t),\,\varepsilon(t-u)}(a,b)\, \delta(\tau - a)\, \delta(\tau' - b)\, \mathrm{d}a\,\mathrm{d}b \\
&= R_a(u)\, p_{\varepsilon(t),\,\varepsilon(t-u)}(\tau,\tau')
\end{aligned}
$$

(25)

where the *p* function is the joint fraction-of-time PDF (13) of $\varepsilon(t)$ and $\varepsilon(t-u)$, which is a function of only *u* and not *t*. The third equality in (25) is simply a direct application of the *Fundamental Theorem of Time Averaging* [14, pp.517-518], [15]. The formula (25) can be substituted into (5)-(6) to calculate either $S_z(f)$ or $S_{cov}(f)$, which can be substituted into (22) or (24), providing an explicit expression for the generalized Wiener filter for a channel with both time-varying attenuation, and time-varying Doppler.

It can be shown by analogy to the derivation of formula (24) that this solution for the Wiener filter transfer function is valid for wide-send stationary (WSS) stochastic processes *s(t)* and *n(t)*, and jointly WSS processes, $g(t,\tau)$, indexed by τ, for all real τ, if we simply replace the time-average operation $\langle . \rangle$ in (3) - (6) with the expectation operation $E\{\cdot\}$. By including both time-average and expectation, this generalizes the Wiener filter derived by Maurer and Franks in [16] from a random *time-invariant* dispersion $g(\tau)$ to a random *time-varying* dispersion.

REFERENCES

[1] Norbert Wiener, Generalized Harmonic Analysis, *Acta Mathematica*, Vol. 55, pp. 117-258, 1930

[2] W. A. Gardner, Introductions to Random Processes, with applications to signals and noise, McGraw Hill, N.Y.,1990

[3] A. Napolitano and W.A. Gardner, Fraction-of-Time Probability: Advancing Beyond the Need for Stationarity and Ergodicity Assumptions. *IEEE Access*, vol. 10, pp. 34,591-34,612, 2022. doi: 10.1109/ACCESS.2022.3162620.

[4] L. A. Zadeh, Time-Varying Networks, I, *Proc. IRE*, pp. 1488-1503, October 1961

[5] L. A. Zadeh, "Correlation functions and power spectra in variable networks," *Proc IRE*, vol. 38, pp. 1342-1345; November 1950.

[6] L.A. Zadeh, "On a class of stochastic operators," *J. Math. and Phys.*, vol. 32, pp. 48-53; April 1953.

[7] P. Bello , Characterization of Randomy Time- Variant Linear Channels, *IEEE Trans. Comm.*, Vol. 11, pp. 360-393, December 1963 https://ieeexplore.ieee.org/abstract/document/1088793

[8] Hopgood, J. R., Rayner, P.J.W., eds. Single Channel Nonstationary Stochastic Signal Separation Using Linear Time-Varying Filters, *IEEE Trans. Sig. Proc.*, Vol. 51, pp. 1739-1752, July 2003: https://ieeexplore.ieee.org/abstract/document/1206684

[9] Hlawatch, F., Matz, M., *Wireless Communications Over Rapidly Time-Varying Channels*, Elsevier, May 2000, https://pdfs.semanticscholar.org/05ce/a2e6f2bf4778115fd4618abc17e6b4b10162.pdf

[10] Norbert Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series: with Engineering Applications, The MIT Press, 1949 (first published in 1942 as a classified document) DOI: https://doi.org/10.7551/mitpress/2946.001.0001

[11] B.G. Agee, The Property Restoral Approach to Blind Adaptive Signal Extraction, PhD Dissertation, Univ. Ca., Davis, 1989, published by UMI

[12] S. V. Schell and W. A. Gardner, Spatio-Temporal Filtering and Equalization for Cyclostationary Signals, Chap. in *Control and Dynamic Systems*, by C. T. Leondes, ed., Academic Press, Vol. 64, pp. 1-85, 1994.

[13] W. A. Gardner, Nonstationary Learning Characteristics of the LMS Algorithm, *IEEE Transactions on Circuits and Systems*, Vol. CAS-34, No. 10, pp. 1199-1207, 1987

[14] W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*, Prentice-Hall, Englewood Cliffs, N.J., 1987

[15] W. A. Gardner, An Introduction to Cyclostationarity, Chap. 1 in *Cyclostationarity in Communications and Signal Processing*, W. A. Gardner, ed., IEEE Press, 1994

[16] Maurer, R., Franks, L.E. Optimal linear processing of randomly distorted signals, *IEEE Transactions on Circuit Theory*, Vol 17, pp. 61-67, 1970. https://doi.org/10.1109/tct.1970.1083054