CM02v250903wed

2025년 9월 3일 수요일 오후 5:28

[J6] Joint Tx and Rx optimization in Additive Cyclostationary Noise

2025년 9월 3일 수요일
13:48

$x > y$. $|x-y|$

= data-like interference.

can only be conveyed via amplitude for the noncoherent memoryless fading channel, information can be carried by the phase as well when the channel has memory.

ACKNOWLEDGMENT

The authors are greatly indebted to Dr. Amos Lapidoth for generously providing the proof of Theorem 5.3, which led to a significantly improved version of the original manuscript. The authors would also like to acknowledge helpful comments from the reviewers.

REFERENCES

[1] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

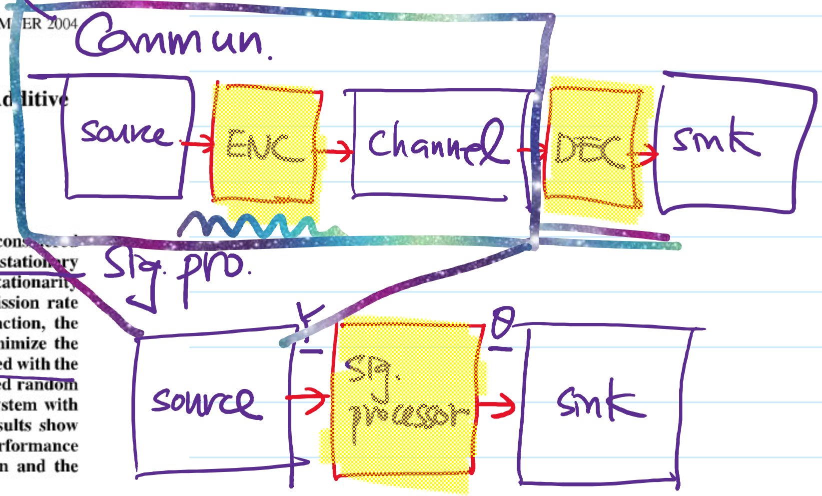
[2] R. R. Chen, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, pp. 1676–1689, Oct. 2003.

Joint Transmitter and Receiver Optimization in Additive Cyclostationary Noise

Joon Ho Cho, *Member, IEEE*

***Abstract*—A joint optimization of transmitter and receiver is considered for strictly band-limited linear modulations in additive cyclostationary noise. Under the assumptions that the period of the cyclostationarity of the noise is the same as the inverse of the symbol transmission rate and that the noise has a positive-definite autocorrelation function, the optimum transmitter and receiver waveforms that jointly minimize the mean-squared error at the output of the linear receiver are derived with the data sequence modeled as a wide-sense stationary (WSS) colored random process and the channel modeled as a linear time-invariant system with a general frequency-selective impulse response. Numerical results show that this joint optimization technique leads to a significant performance improvement over the systems with receiver-only optimization and the systems with no transmitter and receiver optimizations.**

***Index Terms*—Cyclostationary noise, joint transmitter and receiver op-**



Theory, vol. 45, pp. 139–157, Jan. 1999.

[2] R. R. Chen, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, pp. 1676–1689, Oct. 2003.

[3] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective block Rayleigh flat-fading channels," *IEEE. Trans. Inform. Theory*, to be published.

[4] M. Peleg, S. Shamai, and S. Galán, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel," *Proc. Inst. Elec. Eng.–Commun.*, vol. 147, pp. 87–95, Apr. 2000.

[5] R. R. Chen, R. Koetter, and U. Madhow, "Joint noncoherent demodulation and decoding for fast Rayleigh fading channels," in *Proc. 2003 Conf. Information Sciences and Systems (CISS 2003)*, Baltimore, MD, Mar. 2003.

[6] N. Jacobsen and U. Madhow, "Reduced-complexity noncoherent communication with differential QAM and iterative receiver processing," in *Proc. 2003 Conf. Information Sciences and Systems (CISS 2003)*, Baltimore, MD, Mar. 2003.

[7] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2426–2467, Oct. 2003.

[8] G. Taricco and M. Elia, "Capacity of fading channels with no side information," *Electron. Lett.*, vol. 33, pp. 1368–1370, July 1997.

[9] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, May 2001.

[10] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, CA: Holden-Day, 1964. Translation into English by A. Feinstein.

[11] A. Lapidoth and S. Shamai (Shitz), "Fading channels: How perfect need "perfect side information" be?," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1118–1133, May 2002.

[12] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. 18, pp. 725–730, 1972.

that this joint optimization technique leads to a significant performance improvement over the systems with receiver-only optimization and the systems with no transmitter and receiver optimizations.

***Index Terms*—Cyclostationary noise, joint transmitter and receiver optimization, minimum mean-squared error (MMSE).**

h(t)

I. Introduction

One of the classical problems in information theory is the problem of jointly optimizing the transmitter and the receiver of a communication system. Among many different approaches to this fundamental problem is the joint optimization of transmitter and receiver waveforms to minimize the mean-squared error (MSE) at the output of a linear receiver. In [1], Berger and Tufts solved the joint optimization problem for pulse amplitude modulation with a wide-sense stationary (WSS) data symbol sequence, a frequency-selective channel, and a WSS additive colored noise. One of many unintuitive results obtained in their work and the work followed [2]–[4] is that the spectrum of the optimum waveform pair has nonzero values only on a generalized Nyquist interval [3]. Later in [5], Graef extended this work by replacing the WSS data symbol sequence with a wide-sense cyclostationary (WSCS) sequence.

After these problems with a single-input/single-output (SISO) channel were studied, more general joint optimization with a multiple-input/multiple-output (MIMO) channel attracted a lot of interest. It is worth noting that, unlike an SISO channel, there are two types of MIMO channels because we can either allow the coordination of transmitters or not. In [6], Yang and Roy solved the joint optimization problem with a fully coordinated MIMO system. By allowing the coordination of transmitters, they were able to show that the MIMO problem can be divided into separate SISO joint optimization problems. In [7] and [8], Golden *et al.* tackled the other type of MIMO problems with noncoordinating transmitters. They successfully derived the optimum zero-forcing solution for symmetrically cross-coupled two-user channels. However, it was shown that the optimum minimum mean-squared error (MMSE) solution is not analytically obtainable

Manuscript received July 30, 2003; revised August 23, 2004. This work was supported in part by the Ministry of Information and Communication, Korea, under the ITRC program supervised by the IITA. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Yokohama, Japan, June 2003.

The author was with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA USA. He is now with the Department of Electronic and Electrical Engineering, Pohang University of Science and Technology (POSTECH), Kyungbuk 790-784, Korea (e-mail: jcho@postech.ac.kr).

Communicated by V. V. Veeravalli, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2004.838342

Opt. single-objective/criterion

multi-objective/critera

~~min an objective ft. of x~~

$$\min_{x} f_0(x)$$

subject to (s.t.) $x \in \Omega$

the constraint set

$d_n \in \{+1,-1\}$ $p(t)$ t

$V(t)$

$\sum_{n=-\infty}^{\infty} d_n p_n(t)$

$p_n(t) = p_0(t-nT)$

$d_n \in \{+1,-1\}$ $p(t)$

$X(t)$

BPSK mod. signal $\mathrm{Re}\{X(t)e^{j(2\pi f_c t+\theta)}\}$

QAM, PAM $X(t)=\sum_{n=-\infty}^{\infty} d_n p(t-nT)$ $\sum_{n=-\infty} d_n p_n(t)$

input to the ch.

even with such a simple scenario.[1] Later in [9], Kumar and Roy further investigated this problem for symmetrically cross-coupled N-user channels, and again showed that a set of coupled, nonlinear equations needs to be solved to obtain jointly optimum transmitter and receiver waveforms. A comprehensive treatment on this topic can be found in [22].

In this correspondence, we consider a joint transmitter and receiver optimization for an SISO system with an additive WSCS noise in contrast to the previous works with additive WSS noises. Note that cochannel interference, crosstalk, adjacent-cell interference, and many other types of data-like interference of digital communication systems are best modeled as WSCS noises rather than WSS noises. It is assumed that the period of the cyclostationarity of the noise is the same as the inverse of the symbol transmission rate and that the noise has a positive-definite autocorrelation function. The data sequence is modeled as a proper-complex wide-sense stationary colored random process and the channel is modeled as a linear time-invariant system with a general frequency-selective impulse response. After a weighted water-filling is performed to find the optimum transmission energy allotment, the optimum transmitter and receiver waveforms that jointly minimize the MSE at the output of a linear receiver are obtained in terms of the vectorized Fourier transforms of waveforms and the matrix-valued power spectral density of the cyclostationary noise.

The solution of this variational problem is applicable in designing the optimum transmit filter when a new system employing a linear MMSE receiver is overlaid on an existing system having a linear modulation with the same symbol transmission rate. It is also applicable in canceling data-like jamming signals. It is also worth noting that this joint optimization is the locally optimum action taken by each pair of a transmitter and a receiver in a noncooperative game, where each pair competes with other pairs sharing the same frequency band to achieve a better quality of service in a wireless *ad hoc* network. Note that in all these cases, the interference is accurately modeled as a WSCS noise with the aforementioned characteristic.

The rest of this correspondence is organized as follows. The signal and system models are described in Section II. The frequency-domain expression of the MSE is derived in Section III, and the jointly optimum transmitter and receiver waveforms are derived in Section IV. Numerical results are provided in Section V. Finally, concluding remarks are offered in Section VI.

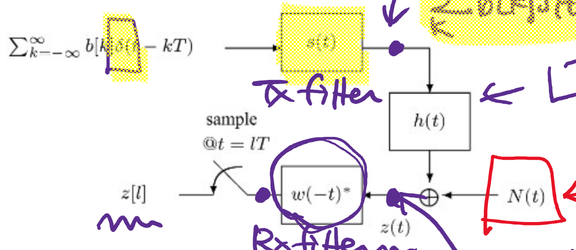


Fig. 1. System block diagram.

impulse response $h(t)$ having bandwidth B (Hz). Without loss of generality, we assume that the carrier frequency used for the complex baseband representation is the center frequency of the channel. The output of the channel is corrupted by a zero-mean proper-complex[2] noise $N(t)$. Hence, the received signal can be written as

$$z(t) = \sum_{k=-\infty}^{\infty} b[k]p(t-kT) + N(t) \tag{2}$$

where $p(t) \triangleq h(t) \otimes s(t)$ is the response of the channel to the input $s(t)$ with $\otimes$ denoting convolution.

It is assumed that the noise process is band-limited with the same bandwidth B as the transmitted signal, positive definite, and having the period of cyclostationarity T.[3] Related definitions are as follows.

Definition 1: [11, p. 7] A second-order proper-complex random process $N(t)$ is band-limited with bandwidth B if it is unaffected by passing it through an ideal low-pass filter with cutoff frequency B.

Definition 2: A second-order proper-complex random process $N(t)$ with bandwidth B and the autocorrelation function

$$r_N(t,s) \triangleq \mathrm{E}\{N(t)N(s)^*\}$$

is called positive definite if

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t)^* r_N(t,s) w(s)\,dt\,ds > 0 \tag{3}$$

for all nonzero $w(t) \in \mathrm{L}^2(B)$.[4]

Definition 3: A second-order random process $N(t)$ is called WSCS

$w^*(-t)$

$\underline{w}^H(-t)$

$\underline{w}(-t)^H$

$w(t) \circledast z(t) = \int_{-\infty}^{\infty} w^*(\tau-t)\, z(\tau)\,d\tau$

$\otimes$ \otimes $w^H z$

Tx pulse/wavefm $s(t)$

Rx pulse/ $w(t)$

transmitter and receiver waveforms are derived in Section IV. Numerical results are provided in Section V. Finally, concluding remarks are offered in Section VI.

II. Signal and System Model

The system block diagram in complex baseband is shown in Fig. 1, where and in what follows the superscript * denotes complex conjugation. Modification of this signal and system models to the real-valued baseband case is straightforward. The transmitter employs a linear modulation scheme with a transmitter waveform $s(t)$ and a symbol transmission rate $1/T$ (symbols/s). The data sequence $(b[k])_k$ is modeled as a discrete-time WSS random process with the auto-correlation function

$$m[k] \triangleq \mathbb{E}\{b[k+l]b[l]^*\}$$

and the power spectral density

$$M(f) \triangleq \sum_{k=-\infty}^{\infty} m[k]e^{-j2\pi fk}. \quad (1)$$

The transmitted signal is passed through a strictly band-limited channel, which is modeled as a linear time-invariant system with the

[1] This is because there exists no adequate mathematical tools for this type of a complicated nonlinear functional optimization problem [7].

for all nonzero $w(t) \in \mathbb{L}^2(B)$.[4]

Definition 3: A second-order random process $N(t)$ is called WSCS with mean zero and period T if

$$\mu_N(t) \triangleq \mathbb{E}\{N(t)\} = 0, \quad \forall t$$

and

$$r_N(t,s) = r_N(t+kT, s+kT), \quad \text{for all } t,s \in \mathbb{R} \text{ and } k \in \mathbb{Z}.$$

The simplest example is the case where an additive white Gaussian noise (AWGN) corrupts the channel output. Since the out-of-band signal is irrelevant [12], only the in-band portion of the noise can be regarded as $N(t)$. This band-limited white Gaussian noise is certainly positive definite because $\int_{-\infty}^{\infty} |w(t)|^2 dt > 0$ for all nonzero $w(t) \in \mathbb{L}^2(B)$ by definition. From this example, we see that the solution in [1] is a special case of what is considered in this correspondence.

[2] Roughly speaking, a WSS complex-valued random process $N(t)$ satisfying $\mathbb{E}\{N(t)N(s)\} = 0, \forall t, s$, is called proper. For details about proper-complex random processes, see [10].

[3] For the cases where the frequency band of $h(t)$ being apparently contained in the frequency band of $N(t)$, we just need to redefine the frequency band of $h(t)$ to fully include all relevant out-of-band signals that have nonzero spectral correlation with the in-band signals.

[4] $\mathbb{L}^2(B)$ is the set of all finite energy, band-limited signals with bandwidth B.

Rx pulse/waveform w(t) ←

$E\{|z[l]-b[l]|^2\}$

MSE b/w b[l] & z[l]

$(s(t), w(t))$

~~$\int_{-\infty}^{\infty} |s(t)|^2 dt =$~~

const.

$A[E\{|x(t)|^2\}] = AT$

The receiver consists of a linear time-invariant filter with the impulse response $w(-t)^*$ and a sampler with the sampling rate $1/T$ (samples/s). The decision statistic $z[l]$ for the detection of the lth symbol $b[l]$ is defined as

$$\begin{aligned} z[l] &\triangleq w(-t)^* \boxed{\otimes} z(t)\Big|_{t=lT} = \int_{-\infty}^{\infty} w(t-lT)^* z(t)dt \\ &= \sum_{k=-\infty}^{\infty} b[k]q((l-k)T) + n[l] \quad (4) \end{aligned}$$

where the waveform $q(t)$ is defined as

$$q(t) \triangleq w(-t)^* \otimes p(t) = \int_{-\infty}^{\infty} w(\tau)^* p(t+\tau)d\tau \quad (5)$$

Proof: By definition, the inverse discrete-time Fourier transform of (1) is given by

$$m[k] = \int_{-\frac{1}{2}}^{\frac{1}{2}} M(f)e^{j2\pi fk}df.$$

Since $m[0] = \mathbb{E}\{|b[k]|^2\}$, we obtain

$$\mathbb{E}\{|b[k]|^2\} = \int_{-\frac{1}{2}}^{\frac{1}{2}} M(f)df.$$

Now, a simple change of variable leads to the conclusion. □

It is worth noting that the power spectral density $M(f)$ is periodic

$k=-\infty$

where the waveform $q(t)$ is defined as

$$q(t) \triangleq w(-t)^* \otimes p(t) = \int_{-\infty}^{\infty} w(\tau)^* p(t+\tau) d\tau \tag{5}$$

and the noise component $n[l]$ is defined as

$$n[l] \triangleq \int_{-\infty}^{\infty} w(t - lT)^* N(t) dt. \tag{6}$$

The noise variance can be written as

$$\begin{aligned}\mathbb{E}\{|n[0]|^2\} &= \mathbb{E}\left\{\left(\int_{-\infty}^{\infty} w(t)^* N(t) dt\right)\left(\int_{-\infty}^{\infty} w(s)^* N(s) ds\right)^*\right\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(t)^* r_N(t,s) w(s) dt ds. \end{aligned} \tag{7}$$

Hence, the condition (3) leads to nonzero noise in the observation after filtering.

The performance metric used in this correspondence is the MSE defined as

$$\begin{aligned}\varepsilon &\triangleq \mathbb{E}\{|b[k] - z[k]|^2\} && (8)\\ &= \mathbb{E}\{|b[k]|^2 - 2\Re(b[k]^* z[k]) + |z[k]|^2\} \\ &= \mathbb{E}\{|b[k]|^2\} && (9a)\\ &\quad - 2\Re(\mathbb{E}\{b[k]^* z[k]\}) && (9b)\\ &\quad + \mathbb{E}\{|z[k] - n[k]|^2\} && (9c)\\ &\quad + \mathbb{E}\{|n[k]|^2\} && (9d)\end{aligned}$$

where $\Re$ denotes the real part, (9a) is the reference component, (9b) is the cross-correlation component between the reference signal and the output, (9c) is the signal component, and (9d) is the noise component.

We are going to find the optimum transmitter and receiver waveforms $s_{\mathrm{opt}}(t)$ and $w_{\mathrm{opt}}(t)$ that jointly minimize the MSE defined as above. Similar to [1], the MSE can be made arbitrarily small by spending more and more power in transmitting the signal. So, it is reasonable to impose a constraint on the average power of the transmitted signal. Hence, the problem is formally stated as

$$\begin{aligned}&\underset{s(t), w(t)}{\text{minimize}} \quad \varepsilon \\ &\text{subject to } \mathbb{E}\left\{\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} \left|\sum_{k=-\infty}^{\infty} b[k] s(t - kT)\right|^2 dt\right\} = A \end{aligned} \tag{10}$$

where $A > 0$ is the average transmitted power.

This problem is inherently a variational problem, where the MSE is a complicated objective functional of two time functions $s(t)$ and $w(t)$. In the next section, we convert this time-domain expression to a frequency-domain expression term by term, to solve the problem in the frequency domain.

III. Frequency-Domain Representation of MSE

A. Reference Component (9a)

Proposition 1: The reference component (9a) in the MSE can be rewritten as

$$\mathbb{E}\{|b[0]|^2\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} T M(fT) df. \tag{11}$$

$\mathbb{E}\{|b[k]|^2\} = \int_{-\frac{1}{2}} M(f) df.$

Now, a simple change of variable leads to the conclusion. □

It is worth noting that the power spectral density $M(f)$ is periodic with period 1 because it is the Fourier transform of a discrete-time signal $m[k]$. Consequently, $M(fT) = M((f - l/T)T), \forall l \in \mathbb{Z}$. We will use this property in the next two subsections.

B. Cross Term (9b)

Before we derive the expression for the cross term, we first define the excess bandwidth and the vectorized Fourier transform. The excess bandwidth is a parameter that is used to quantify the extra bandwidth, over the Nyquist bandwidth, spent in pulse shaping. It will be seen that the excess bandwidth determines the degrees of freedom in our optimal waveform design. The vectorized Fourier transform is a representation as a vector-valued function of the ordinary Fourier transform that is a scalar-valued function. This alternate expression of the Fourier transform facilitates our optimal system design in the frequency domain.

Definition 4: The excess bandwidth β is defined by the relation

$$BT = \frac{1+\beta}{2} \tag{12}$$

where B and $1/T$ are, respectively, the bandwidth and the transmission rate of the symbol waveform $s(t)$.

Definition 5: The vectorized Fourier transform $\boldsymbol{x}(f)$ with sampling period $1/T$ of the time function $x(t)$ having bandwidth B is defined as

$$\boldsymbol{x}(f) \triangleq \begin{bmatrix} X\left(-\frac{L}{T} + f\right) \\ X\left(-\frac{L-1}{T} + f\right) \\ \vdots \\ X\left(\frac{L}{T} + f\right) \end{bmatrix} \tag{13}$$

for $f \in [-1/(2T), 1/(2T))$ with $L \triangleq \lceil \beta/2 \rceil$, where $X(\xi)$ is the Fourier transform[5] of $x(t)$ and β is given by the relation (12).

As shown in Fig. 2, the vectorized Fourier transform requires the sampling of an ordinary Fourier transform in the frequency domain. Note that the sampling rate is T (samples/Hz) and that the argument f serves as the offset of the uniform samples from the integer multiples of $1/T$. In the vectorization, the excess bandwidth determines the number of samples to be taken. As will be later seen, the dimension $2L+1$ is an upper bound on the degrees of freedom in the waveform design.

Proposition 2: The cross term (9b) in the MSE can be rewritten as

$$-2\Re(\mathbb{E}\{b[k]^* z[k]\}) = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} -2\Re\Big(\boldsymbol{w}(f)^H (M(fT)\boldsymbol{p}(f))\Big) df \tag{14}$$

where $\boldsymbol{w}(f)$ and $\boldsymbol{p}(f)$ are, respectively, the vectorized Fourier transform of the receiver waveform $w(t)$ and that of the channel response $p(t)$ of the transmitter waveform.

[5]Throughout this correspondence, a capitalized frequency function represents the Fourier transform of the corresponding time function.

$$\mathsf{E}\{|b[0]|^2\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} T M(fT) df. \qquad (11)$$

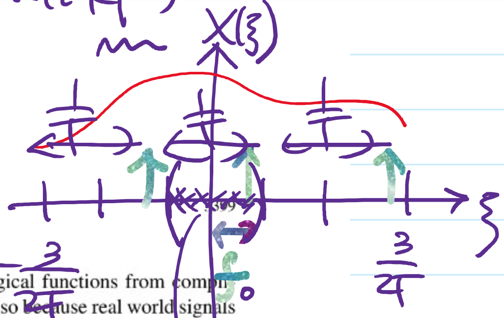
[5]Throughout this correspondence, a capitalized frequency function represents the Fourier transform of the corresponding time function.

$x(t) \overset{\mathcal{F}}{\longleftrightarrow} \cancel{X(f)}$

$X(\xi) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi\xi t} d\xi$

$\mathbf{x}(f)$

Q. $\sum b[k] s(t-kT)$



$$\mathbf{x}(f_0) = \begin{bmatrix} X(f_0 - \frac{1}{T}) \\ X(f_0) \\ X(f_0 + \frac{1}{T}) \end{bmatrix}$$

$|f| < \frac{1}{2T}$.

Nyquist zone

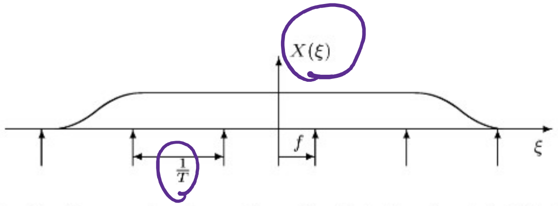


Fig. 2. Frequency-domain sampling with offset f and period $1/T$ to construct the vectorized Fourier transform $\boldsymbol{x}(f)$ of $x(t)$ from the Fourier transform $X(\xi)$.

Proof: Using Parseval's relation and the definition of $M(f)$, we obtain

$$\mathsf{E}\{b[0]^* z[0]\} = \mathsf{E}\left\{b[0]^* \int_{-\infty}^{\infty} w(t)^* z(t) dt\right\} \qquad (15a)$$

$$= \sum_{k=-\infty}^{\infty} m[k] \int_{-\infty}^{\infty} w(t)^* p(t-kT) dt \qquad (15b)$$

$$= \sum_{k=-\infty}^{\infty} m[k] \int_{-\infty}^{\infty} W(\xi)^* P(\xi) e^{-j2\pi\xi kT} d\xi \qquad (15c)$$

$$= \int_{-\infty}^{\infty} W(\xi)^* P(\xi) M(\xi T) d\xi. \qquad (15d)$$

By dividing the integration interval into length-$1/T$ subintervals and properly rearranging terms, we obtain

$$\mathsf{E}\{b[0]^* z[0]\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \boldsymbol{w}(f)^H \big(M(fT)\boldsymbol{p}(f)\big) df \qquad (16)$$

where we used the periodicity of $M(\xi T)$ with period $1/T$. From (16), the conclusion follows immediately. □

C. Signal Component (9c)

Proposition 3: The signal component (9c) in the MSE can be rewritten as

$$\mathsf{E}\{|z[0] - n[0]|^2\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \boldsymbol{w}(f)^H \left(\frac{M(fT)}{T} \boldsymbol{p}(f)\boldsymbol{p}(f)^H\right) \boldsymbol{w}(f) df.$$

not only because it prevents all pathological functions from complicating the derivation of the solution but also because real world signals are accurately modeled as signals having piecewise smooth spectra. The piecewise or sectional smoothness is also a standard assumption in the calculus of variations [14]. We follow this convention and, hence, all frequency functions in this correspondence are assumed piecewise smooth.

D. Noise Component (9d)

An autocorrelation function of a cyclostationary random process is a two-dimensional signal. In order to handle this signal in the frequency domain, we need the Fourier transform and the inverse Fourier transform defined for general two-dimensional signals.

Definition 6: The double Fourier transform of $r(t, s)$ is defined as

$$R(\xi, \tilde{\xi}) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(t, s)\, e^{-j2\pi(\xi t - \tilde{\xi} s)} dt ds$$

and the inverse transform as

$$r(t, s) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(\xi, \tilde{\xi})\, e^{+j2\pi(\xi t - \tilde{\xi} s)}\, d\xi d\tilde{\xi}.$$

Notice that, unlike the conventional definition of the two-dimensional Fourier transform, we correlate $r(t, s)$ with $e^{-j2\pi(\xi t - \tilde{\xi} s)}$ instead of $e^{-j2\pi(\xi t + \tilde{\xi} s)}$ to obtain $R(\xi, \tilde{\xi})$.

By combining this definition and the definition of Fourier transform pair for one-dimensional signals, we obtain the following lemma, which is a simple extension of Parseval's relation.

Lemma 1:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(t)^* r(t, s) v(s) dt ds \qquad (19a)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\xi)^* R(\xi, \tilde{\xi}) V(\tilde{\xi}) d\xi d\tilde{\xi}. \qquad (19b)$$

Proof: Substituting the inverse Fourier transforms of $u(t)^*$ and $v(s)$ into (19a) and then using the definition of double Fourier transform, we obtain (19b). □

Proposition 3: The signal component (9c) in the MSE can be rewritten as

$$\mathsf{E}\{|z[0]-n[0]|^2\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \boldsymbol{w}(f)^H \left(\frac{M(fT)}{T}\boldsymbol{p}(f)\boldsymbol{p}(f)^H\right)\boldsymbol{w}(f)df. \tag{17}$$

Proof: Note that

$$\begin{aligned}
&\mathsf{E}\{|z[0]-n[0]|^2\} \\
&\quad= \mathsf{E}\left\{\left|\sum_{l=-\infty}^{\infty} b[l]q(-lT)\right|^2\right\} &(18a)\\
&\quad= \sum_{k=-\infty}^{\infty} m[-k]\left(\sum_{l=-\infty}^{\infty} q(lT)q((l-k)T)^*\right) &(18b)\\
&\quad= \sum_{k=-\infty}^{\infty} m[-k]\Bigg(\frac{1}{T}\sum_{l=-\infty}^{\infty}\int_{-\infty}^{\infty} Q(\xi)Q\left(\xi-\frac{l}{T}\right)^* \\
&\qquad\cdot e^{j2\pi(\xi-\frac{l}{T})kT}d\xi\Bigg) &(18c)\\
&\quad= \frac{1}{T}\sum_{l=-\infty}^{\infty}\int_{-\infty}^{\infty} Q(\xi)Q\left(\xi-\frac{l}{T}\right)^* M(\xi T)d\xi &(18d)
\end{aligned}$$

where the Poisson summation formula [21, p. 577] leads to (18c) with $Q(\xi)$ being the Fourier transform of $q(t)$ defined in (5). After dividing the integration interval into length-$1/T$ subintervals and properly rearranging the terms, we obtain (17). □

A sufficient condition on $q(t)$ for the applicability of the Poisson summation formula is the piecewise smoothness of $Q(\xi)$.[6] Since the piecewise smoothness of $W(\xi)$, $H(\xi)$, and $S(\xi)$ leads to the piecewise smoothness of $Q(\xi)$, we assume as such. This assumption is reasonable

[6]For details, see [13, Proposition 1 and the related footnote].

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U(\xi)^* R(\xi,\tilde{\xi})V(\tilde{\xi})d\xi d\tilde{\xi}. \tag{19b}$$

Proof: Substituting the inverse Fourier transforms of $u(t)^*$ and $v(s)$ into (19a) and then using the definition of double Fourier transform, we obtain (19b). □

By using Lemma 1, we can rewrite the noise component (7) as

$$\mathsf{E}\{|n[0]|^2\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(\xi)^* R_N(\xi,\tilde{\xi})W(\tilde{\xi})d\xi d\tilde{\xi} \tag{20}$$

which we need to convert to a single integration with a finite integration interval, in order to conform to the frequency-domain expressions for other components in the MSE.

We start by observing the following lemma from [15].

Lemma 2: [15, p. 242] If $r(t,s)$ satisfies

$$r(t,s) = r(t+kT, s+kT), \qquad \forall k \in \mathbb{Z}, \tag{21a}$$

there exists $r^{(k)}(\tau)$, $k \in \mathbb{Z}$ such that

$$r(t,s) = \sum_{k=-\infty}^{\infty} r^{(k)}(t-s)e^{j2\pi\frac{kt}{T}}. \tag{21b}$$

Proof: Let $\tilde{r}(t,\tau) = r(t, t-\tau)$, then $\tilde{r}(t,\tau)$ is periodic in t with period T. If we expand $\tilde{r}(t,\tau)$ using the Fourier series, we obtain (21b). □

Remark 1: If $r(t,s)$ satisfies $r(t,s) = r(t-s,0)$, every $r^{(k)}(t-s)$ vanishes except $r^{(0)}(t-s) = r(t-s)$ because

$$\begin{aligned}
r(t-s,0) &= \sum_{k=-\infty}^{\infty} r^{(k)}(t-s)e^{j2\pi\frac{kt}{T}} = r(0, s-t)\\
&= \sum_{k=-\infty}^{\infty} r^{(k)}(t-s)
\end{aligned}$$

for all t and s.

Note that the double Fourier transform of $r(t,s)$ satisfying (21b) is given as

$$R(\xi,\tilde{\xi}) = \sum_{k=-\infty}^{\infty} R^{(k)}\left(\xi-\frac{k}{T}\right)\delta\left(\xi-\tilde{\xi}-\frac{k}{T}\right) \tag{22}$$

where $R^{(k)}(\xi)$ is the Fourier transform of $r^{(k)}(\tau)$. In this case, the

where the matrix-valued power spectral density $\boldsymbol{R}_N(f)$ of the cyclostationary noise $N(t)$ is defined as

$$[\boldsymbol{R}_N(f)]_{k,l} \triangleq \begin{cases} R_N^{(k-l)}\left(f+\frac{k-L-1}{T}\right), & \text{for } |k-l| \le 2L\\ 0, & \text{elsewhere}\end{cases} \tag{28}$$

with $k, l = 1, 2, \ldots, 2L+1$.

Proof: Using (26), we can rewrite (23) as

$$R(\xi, \tilde{\xi}) = \sum_{k=-\infty}^{\infty} R^{(k)}\left(\xi - \frac{k}{T}\right)\delta\left(\xi - \tilde{\xi} - \frac{k}{T}\right) \tag{22}$$

where $R^{(k)}(\xi)$ is the Fourier transform of $r^{(k)}(\tau)$. In this case, the double Fourier transform of $r(t, s)$ satisfying (21) consists of impulse fences on lines parallel to the $\xi = \tilde{\xi}$ diagonal and separated by $1/T$ [11]. Now, by applying this result to the double integration (20), we obtain an intermediate result as

$$\mathbb{E}\{|n[0]|^2\} = \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} W(\xi)^* R_N^{(k)}\left(\xi - \frac{k}{T}\right) W\left(\xi - \frac{k}{T}\right) d\xi \tag{23}$$

where $R_N^{(k)}(\xi)$, $k \in \mathbb{Z}$, are the Fourier transforms of $r_N^{(k)}(\tau)$, $k \in \mathbb{Z}$, such that

$$r_N(t, s) = \sum_{k=-\infty}^{\infty} r_N^{(k)}(t - s)e^{j2\pi \frac{kt}{T}}. \tag{24}$$

Remark 2: If $N(t)$ is wide-sense stationary, i.e., $r_N(t, s) = r_N(t - s, 0)$, then (23) becomes $\int_{-\infty}^{\infty} R_N^{(0)}(\xi)|W(\xi)|^2 d\xi$, which is a well-known result from the filtering theory of random processes. In other words, the double Fourier transform of the two-dimensional autocorrelation function consists of a single impulse fence on the line $\xi = \tilde{\xi}$, and the power spectral density (PSD) corresponds to the magnitude of the fence.

The expression (23) can be further simplified when the WSCS process $N(t)$ is band-limited, which is our case. Recall Definition 1 of a band-limited process. This definition implies that the autocorrelation function $r_N(t, s)$ is the same as the autocorrelation of the ideal low-pass filtered version of $N(t)$. Hence, using Lemma 1, we obtain

$$\begin{aligned} &\mathbb{E}\{N(t)N(s)^*\} = \mathbb{E}\{L_B[N(t)]L_B[N(s)]^*\} \Longrightarrow \\ &r_N(t, s) = \int_{-B}^{B}\int_{-B}^{B} R_N(\xi, \tilde{\xi})e^{+j2\pi(\xi t - \tilde{\xi} s)}d\xi d\tilde{\xi} \end{aligned} \tag{25}$$

where $L_B[x(t)]$ denotes the response of the ideal low-pass filter with bandwidth B to the input $x(t)$. From the definition of the inverse Fourier transform of $R_N(\xi, \tilde{\xi})$ combined with (25), we find that $R_N(\xi, \tilde{\xi}) = 0$ outside the square area defined as

$$\{(\xi, \tilde{\xi})\| \xi| \le B, |\tilde{\xi}| \le B)\}.$$

Hence, the infinite sum (22) has at most $4L + 1$ nonzero terms as

$$R_N(\xi, \tilde{\xi}) = \sum_{k=-2L}^{2L} R_N^{(k)}\left(\xi - \frac{k}{T}\right)\delta\left(\xi - \tilde{\xi} - \frac{k}{T}\right) \tag{26}$$

where $R_N^{(k)}(\xi)$ has an additional restriction that it can be nonzero only on a finite interval

$$\xi \in \left[\max\left(-B, -B - \frac{k}{T}\right), \min\left(B, B - \frac{k}{T}\right)\right]$$

for $k = -2L, -2L + 1, \ldots, 2L$, i.e., $R_N^{(k)}(\xi - \frac{k}{T}) = 0$ outside the square area.

Proposition 4: The noise component in the MSE can be rewritten

$$[\boldsymbol{R}_N(f)]_{k,l} = \begin{cases} \ldots \\ 0, & \text{elsewhere} \end{cases} \tag{28}$$

with $k, l = 1, 2, \ldots, 2L + 1$.

Proof: Using (26), we can rewrite (23) as

$$\mathbb{E}\{|n[0]|^2\} = \int_{-B}^{B} \sum_{k=-2L}^{2L} W(\xi)^* R_N^{(k)}(\xi - \frac{k}{T})W(\xi - \frac{k}{T})d\xi.$$

After dividing the integration interval into length-$1/T$ subintervals and properly rearranging terms, we obtain (27). □

E. MSE

Now, we present the final result of this section.

Theorem 1: The MSE (8) can be expressed as

$$\begin{aligned} \varepsilon = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} & TM(fT) + \boldsymbol{w}(f)^H \boldsymbol{R}(f)\boldsymbol{w}(f) \\ & - 2\Re\Big(\boldsymbol{w}(f)^H (M(fT)\boldsymbol{p}(f))\Big)df \end{aligned} \tag{29}$$

where

$$\boldsymbol{R}(f) \triangleq \boldsymbol{R}_N(f) + \frac{1}{T}M(fT)\boldsymbol{p}(f)\boldsymbol{p}(f)^H. \tag{30}$$

Proof: By combining Propositions 1 to 4, we obtain (29). □

Since $p(t) = h(t) \otimes s(t)$, we can replace $\boldsymbol{p}(f)$ in (30) with $\boldsymbol{H}(f)\boldsymbol{s}(f)$, where $\boldsymbol{s}(f)$ is the vectorized Fourier transform of $s(t)$ and $\boldsymbol{H}(f)$ is a $(2L + 1)$-dimensional, diagonal, matrix-valued frequency function defined as

$$\boldsymbol{H}(f) \triangleq \text{diag}\left\{H\left(-\frac{L}{T} + f\right), \ldots, H\left(\frac{L}{T} + f\right)\right\} \tag{31}$$

for $f \in [-1/(2T), 1/(2T))$.

IV. Optimization of Waveforms

In this section, we find the jointly optimum transmitter and receiver waveforms using the objective functional (29) derived in the previous section. In order to conform to the frequency-domain expression of the objective functional, we rewrite the average power constraint in (10) as

$$\begin{aligned} A &= \mathbb{E}\left\{\frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} \left|\sum_{k=-\infty}^{\infty} b[k]s(t - kT)\right|^2 dt\right\} \\ &= \frac{1}{T}\int_{-\infty}^{\infty} M(\xi T)|S(\xi)|^2 d\xi \\ &= \frac{1}{T}\int_{-\frac{1}{2T}}^{\frac{1}{2T}} M(fT)\boldsymbol{s}(f)^H \boldsymbol{s}(f)df. \end{aligned} \tag{32}$$

Our approach is to divide the problem into a cascade of individual receiver and transmitter optimizations as

$$\begin{aligned} &\underset{\boldsymbol{s}(f)}{\text{minimize}} \quad \underset{\boldsymbol{w}(f)}{\text{minimize}}\ \varepsilon \\ &\text{subject to} \quad \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \frac{M(fT)}{T} a(f)df = A \end{aligned} \tag{33}$$

for $k = -2L, -2L+1, \ldots, 2L$, i.e., $R_N^{(k)}(\xi - \frac{k}{T}) = 0$ outside the square area.

Proposition 4: The noise component in the MSE can be rewritten as

$$\mathsf{E}\{|n[0]|^2\} = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \boldsymbol{w}(f)^H \boldsymbol{R}_N(f)\boldsymbol{w}(f)df \tag{27}$$

$$\underset{\boldsymbol{s}(f)}{\text{minimize}} \quad \underset{\boldsymbol{w}(f)}{\text{minimize}}$$

$$\text{subject to} \quad \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \frac{M(fT)}{T} a(f)df = A \tag{33}$$

where the energy density $a(f)$ of $\boldsymbol{s}(f)$ is defined as

$$a(f) \triangleq \boldsymbol{s}(f)^H \boldsymbol{s}(f) \tag{34}$$

for $f \in [-1/(2T), 1/(2T))$.