CM12v251020mon

2025년 9월 18일 목요일 09:51

[J18] Capacity_o...

Capacity of Second-Order Cyclostationary Improper-Complex Gaussian Noise Channels – Part I :

Byung Wook Han and Joon Ho Cho, *Member, IEEE*

***Abstract*—In this paper, we derive the capacity of a continuous-time, single-input single-output (SISO), frequency-selective, band-limited, linear time-invariant (LTI) channel, whose output is corrupted by a second-order cyclostationary (SOCS) complex Gaussian noise. By using a pair of invertible, linear-conjugate linear time-varying operators called a properizing FREquency SHift (p-FRESH) vectorizer and a p-FRESH scalarizer, it is shown that, whether the complex noise is proper or improper, the SISO channel can always be converted to an equivalent multiple-input multiple-output (MIMO) LTI channel whose output is now corrupted by a proper-complex vector wide-sense stationary noise. A variational problem is then formulated in the frequency domain to find the optimal input distribution that maximizes the throughput of the equivalent MIMO channel. It turns out that the optimal input obtained through a procedure similar to the water filling, is an SOCS complex Gaussian random process with the same cycle period as the noise. It is shown that this procedure, named cyclic water filling, significantly outperforms ordinary water filling by effectively utilizing the spectral correlation of the cyclostationary noise.**

***Index Terms*—Channel capacity, cyclostationarity, multi-input multi-output channel, water filling.**

I. Introduction

Signals transmitted through wireless channels are susceptible to various types of noise and interference [1]. In many cases, the interference is well modeled by a wide-sense stationary (WSS) random process. However, man-made interference such as co-channel and adjacent channel interference is often much accurately modeled by a wide-sense cyclostationary (WSCS) random process [2]. For example, pulse amplitude modulation (PAM) or quadrature amplitude modulation (QAM) of a WSS data sequence results in a WSCS signal in general [3, Ch. 4].

As well documented in [4]–[6], cyclostationarity has been noticed as an important classifier for random processes that are encountered in communications and signal processing [7], [8]. One of the seminal features of a cyclostationary random process is that it exhibits spectral correlation as well as temporal correlation. To efficiently characterize such structural properties, a scheme called the *harmonic series representation (HSR)* of a cyclostationary random process is proposed in [9]. The HSR scheme has been intensively used, either implicitly or explicitly, in signal design and parameter estimation. This includes the generalized Nyquist criterion to design a set of band-limited orthogonal pulses [10], [11, Ch. 6], the optimal transmit pulse design for PAM, QAM, or direct-sequence spread-spectrum (DS/SS) modulated systems with linear receivers [12]–[16], the cyclic Wiener filtering of a cyclostationary random process [17], and the sparse representation of shift-invariant signals [18]. This scheme is also used in [19] and [20] for joint transmitter and receiver optimizations in additive WSCS noise. However, only the minimization of the mean squared error (MSE) at the linear receiver output is considered to design the optimal transmit and receive waveforms of the QAM signaling format with a WSS proper-complex data sequence.

For a continuous-time, band-limited, linear time-invariant channel corrupted only by an additive stationary Gaussian noise, its capacity and capacity-achieving input distribution are both well known [21]. The optimal input distribution that maximizes the throughput turns out to be WSS and Gaussian as the additive noise. Depending on the noise power spectral density (PSD) and the channel gain, the optimal solution allocates the transmit power in the frequency domain through a geometric procedure called water filling. For the same single-input single-output (SISO) channel, however, no capacity result is available if the Gaussian noise is no longer stationary but cyclostationary. Neither the channel capacity nor the optimal input distribution and associated power allocation scheme is discovered yet.

In this paper, we derive the capacity of a continuous-time, SISO, frequency-selective, band-limited, linear time-invariant (LTI) channel, whose output is corrupted by a *second-order cyclostationary (SOCS)* complex Gaussian noise. An SOCS complex random process is defined as a proper- or improper-complex random process whose auto-covariance and complementary auto-covariance (a.k.a. pseudo-covariance) functions both exhibit periodicity with a common period.[1] It can be easily shown that, under certain conditions, PAM, QAM, and their variants such as single sideband (SSB) PAM, vestigial sideband (VSB) PAM, minimum shift keying (MSK), staggered quaternary phase-shift keying (SQPSK), DS/SS modulation, and some forms of orthogonal frequency-division multiplexing (OFDM) all result in SOCS complex envelopes.

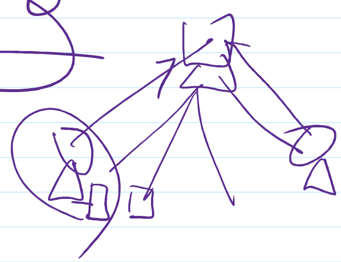
Paper approved by M. R. Beuhrer, the Editor for Cognitive Radio and UWBA of the IEEE Communications Society. Manuscript received July 14, 2010; revised April 18, 2011 and August 25, 2011.

This work was supported in part by the Ministry of Knowledge Economy, Korea, under grant NIPA-2011-C1090-1111-0011 for the BrOMA-ITRC@POSTECH supervised by the NIPA, and in part by the Ministry of Education, Science and Technology, Korea, under NRF Basic Science Research Program grant 2009-0088483. The material in this paper was presented in part at the 2009 Information Theory and Applications Workshop (ITA 2009), San Diego, CA, 8–13, Feb. 2009, and in part at the IEEE Military Communications Conference (MILCOM), Baltimore, MD, 7–10, Nov. 2011.

The authors are with the Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Gyeongbuk 790-784, Korea (e-mail: {hanbw, jcho}@postech.ac.kr).

Digital Object Identifier 10.1109/TCOMM.2011.110711.100414

[1] See *Definitions 1, 2,* and *3* for the formal definitions of the SOCS random processes and related terms.



□ Lemma. If proper-complex RP $N(t)$ is WSS then

$$r_{NN}(t,s) \triangleq E\{N(t)N^*(s)\} = r_{NN}(t+T, s+T), \quad \forall (t,s) \in \mathbb{R}^2.$$

proof) WSS ⇒ $r_{NN}(t,s) = r_{NN}(t-s, 0), \ \forall (t,s)$

$$r_{NN}(t+T, s+T) = r_{NN}((t+T)-(s+T), 0) = r_{NN}(t-s, 0) = r_{NN}(t,s).$$

□ Def. A proper complex RP is WSS if.

(i) $E\{N(t)\} = \mu_N, \ \forall t$

(ii) $r_{NN}(t,s) = r_{NN}(t-s, 0), \ \forall (t,s)$

(iii) $\tilde{r}_{NN}(t,s) = \tilde{r}_{NN}(t-s, 0), \ \forall (t,s)$ $=0$

Initially, $N_A(t)$, $N_c(t)$ both proper-complex Gaussian RPs.

$\underline{X}(t) \rightarrow \boxed{\mathrm{diag}\{\underline{h}(t)\}} \rightarrow \oplus \rightarrow \oplus \rightarrow Y(t)$

Digital Object Identifier 10.1109/TCOMM.2011.110711.100414

processes and related terms.

0090-6778/12$31.00 © 2012 IEEE

The Gaussian assumption not only makes the problem mathematically tractable but also is well justified if the additive noise consists of an ambient white Gaussian noise and a worst-case jamming signal. It is well known [22] that, as the Gaussian noise minimizes the capacity of a discrete-time memoryless channel given a noise power [23], the worst-case jammer chooses to transmit a Gaussian jamming signal given second-order statistics of the interference. It can be shown that the Gaussianity of the worst-case jamming signal also applies to our case with an SOCS interference. Of course, this assumption does not hold in general when the interference is generated by a co-channel or an adjacent channel communicator with a finite symbol constellation of small cardinality. However, the channel capacity and the capacity-achieving scheme derived in this paper can still serve well as a lower bound and a reference system, respectively, for general non-Gaussian SOCS noise channels.

Motivated by the HSR that leads to the equivalence [9] between a scalar-WSCS process and a vector-WSS process, we convert the SISO channel to an equivalent multiple-input multiple-output (MIMO) channel by placing a linear-conjugate linear time-varying (LCL-TV) operator and its inverse at the output and the input of the channel, respectively. Unlike the conventional HSR-based linear time-varying operator such as that used in [24], this operator always generates a proper-complex vector-WSS random process regardless of the propriety of the SOCS signal input.

A variational problem is then formulated in the frequency domain to find the capacity of the MIMO channel. Using the capacity results on MIMO communications [25], [26], the optimal input distribution to the equivalent MIMO channel is obtained through the water filling procedure for parallel Gaussian channels. It turns out that the corresponding optimal input to the SISO channel is SOCS and Gaussian with the same cycle period as the noise. It is shown that the proposed water filling significantly improves the spectral efficiency by exploiting the spectral correlation of the noise.

The rest of this paper is organized as follows. In Section II, the channel model is described and the problem is formulated. In Section III, the LCL-TV operator and its inverse are described and an equivalent MIMO channel is derived. In Section IV, the optimal power distribution is obtained and the channel capacity is derived. Discussions and numerical results are provided in Section V, and concluding remarks are offered in Section VI.

II. Channel Model and Problem Formulation

In this section, the continuous-time band-limited additive SOCS complex Gaussian noise channel is described and the problem is formulated. We consider communicating over a real passband channel. Fig. 1 shows its equivalent channel model in complex baseband. The channel input $X(t)$ passes through a frequency-selective channel modeled by an LTI filter with impulse response $h(t)$, and is received in the presence of an additive noise $N(t)$. Thus, the channel output $Y(t)$ is given by

$$Y(t) = X(t) * h(t) + N(t), \tag{1}$$

where the operator $*$ denotes the convolution integral.

BL Ch.

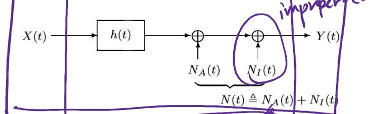


Fig. 1. Complex baseband equivalent model of additive SOCS Gaussian noise channel.

To proceed, we define the following terms.

Definition 1: The mean, the auto-covariance, and the complementary auto-covariance functions of a complex random process $X(t)$ are defined, respectively, as

$$\mu_X(t) \triangleq \mathbb{E}\{X(t)\}, \tag{2}$$
$$c_X(t,s) \triangleq \mathbb{E}\{(X(t)-\mu_X(t))(X(s)-\mu_X(s))^*\}, \text{and} \tag{3}$$
$$\tilde{c}_X(t,s) \triangleq \mathbb{E}\{(X(t)-\mu_X(t))(X(s)-\mu_X(s))\}, \tag{4}$$

where the operator $\mathbb{E}\{\cdot\}$ denotes the expectation and the superscript * denotes the complex conjugation.[2]

Definition 2: [27, *Definition 2*] A complex random process $X(t)$ is called proper if its complementary auto-covariance function vanishes, i.e., $\tilde{c}_X(t,s) = 0, \forall t, \forall s$, and is called improper otherwise.

Definition 3: A complex random process $X(t)$ is SOCS with cycle period $T > 0$ if, $\forall t, \forall s$,

$$\mu_X(t) = \mu_X(t+T), \text{and} \tag{5}$$
$$c_X(t,s) = c_X(t+T, s+T), \text{and} \tag{6}$$
$$\tilde{c}_X(t,s) = \tilde{c}_X(t+T, s+T). \tag{7}$$

The second-order cyclostationarity is motivated by the fact that in [28] as $\mu_X(t) = \mu_X(t+\tau), c_X(t,s) = c_X(t+\tau, s+\tau)$, and $\tilde{c}_X(t,s) = \tilde{c}_X(t+\tau, s+\tau), \forall t, \forall s, \forall \tau$. Note that, like the second-order stationarity, the second-order cyclostationarity is defined in terms of the complementary auto-covariance function as well as the auto-covariance function.

It is assumed that the additive noise consists of an ambient noise and an interfering signal, i.e.,

$$N(t) \triangleq N_A(t) + N_I(t), \tag{8}$$

where the ambient noise $N_A(t)$ is modeled as a proper-complex white Gaussian noise. The interfering signal $N_I(t)$ is modeled as a zero-mean SOCS proper- or improper-complex Gaussian random process with cycle period $T_0 > 0$.

Such an SOCS interference can be observed when there is a co-channel or an adjacent channel user that transmits a second-order stationary proper- or improper-complex data sequence using a linear modulation scheme such as double sideband (DSB) PAM, SSB PAM, VSB PAM, QAM, MSK, SQPSK, etc. Note that these linear modulation schemes are among the most popular digital modulation schemes and that their complex envelopes are all in the form of

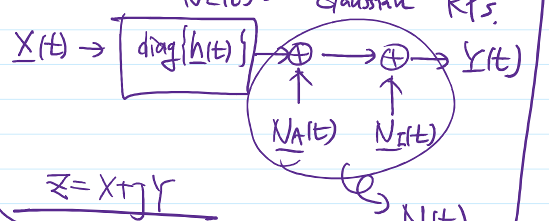
$$N_I(t) = \sum_{n=-\infty}^{\infty} \left\{ c_1[n] p_1(t - nT_0) + jc_2[n] p_2(t - nT_0) \right\}, \tag{9}$$

[2] There should be no confusion from the operator $*$ that denotes the convolution integral.

where j denotes $\sqrt{-1}$, $(c_1[n])_n$ and $(c_2[n])_n$ are jointly WSS real-valued random data sequences, and $p_1(t)$ and $p_2(t)$ are waveforms that includes the effect of the transmit waveforms

$m \in \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of all integers. In other words, there exist functions $\{R_N^{(k)}(f)\}_{k\in\mathbb{Z}}$ such that

Gaussian RPs.



□ Def $Z = X + jY$

where X & Y are jointly random variable.

real-Random

□ Def. $\mu_Z = E\{Z\} = E\{X\} + jE\{Y\}$ existence of μ_Z assumed!

$Var\{Z\} \triangleq E\{|Z-\mu_Z|^2\} = Var\{X\} + Var\{Y\}$

$\widetilde{Var}\{Z\} \triangleq E\{(Z-\mu_Z)^2\} = \boxed{Var\{X\} - Var\{Y\}} - 2jCov(X,Y)$

pseudo-complementary

□ Def. Z is proper if $\boxed{\widetilde{Var}\{Z\}=0}$

Lemma Z is proper iff
(i) $Var\{X\} = Var\{Y\}$
(ii) $Cov\{X,Y\} = 0$.

Z is improper otherwise.

complex RVs

proper | improper

□ z □ Z.

□ z(t)

□ Z(t)

/ex/ $\begin{bmatrix}X\\Y\end{bmatrix} \sim N\left(\begin{bmatrix}\mu_X\\\mu_Y\end{bmatrix}, \frac{\sigma^2}{2}\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)$

$E[Z] = E[X+jY] = \mu_X + j\mu_Y = \mu_Z$

$Var[Z] = \cancel{2\sigma^2}\ \sigma^2$

$\widetilde{Var}[Z] = 0$

$f_Z(z) = f_{X+jY}(x,y)$
$= f_X(x) f_Y(y)$
$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu_X)^2}{2\sigma^2}\right)$
$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y-\mu_Y)^2}{2\sigma^2}\right)$
$= \frac{1}{\pi(2\sigma^2)} \exp\left(-\frac{|z-\mu_z|^2}{(2\sigma^2)}\right)$

where j denotes $\sqrt{-1}$, $(c_1[n])_n$ and $(c_2[n])_n$ are jointly WSS real-valued random data sequences, and $p_1(t)$ and $p_2(t)$ are waveforms that includes the effect of the transmit waveforms and the channel. Then, it is straightforward to show that the joint wide-sense stationarity of $(c_1[n])_n$ and $(c_2[n])_n$ is equivalent to the second-order stationarity of $(c_1[n]+\mathrm{j}c_2[n])_n$, and that, by *Definition 3*, $N_I(t)$ in (9) is SOCS with cycle period T_0.

As mentioned in the previous section, the Gaussian assumption is fully justified when the interference is the worst-case jamming signal. The outline of the proof is provided in Section V. Although this is not the case in general, the capacity and the capacity-achieving scheme obtained under this assumption can be used, respectively, as a lower bound on the channel capacity and a reference system, of which performance any well-designed system for a non-Gaussian SOCS interference must exceed.

Since the overall additive noise $N(t)$ is now a zero-mean SOCS Gaussian random process with cycle period T_0, the second-order statistics of $N(t)$ satisfy

$$\mu_N(t) \triangleq \mathbb{E}\{N(t)\} = 0, \tag{10a}$$

$$r_N(t,s) \triangleq \mathbb{E}\{N(t)N(s)^*\} = r_N(t+T_0, s+T_0), \text{ and} \tag{10b}$$

$$\tilde{r}_N(t,s) \triangleq \mathbb{E}\{N(t)N(s)\} = \tilde{r}_N(t+T_0, s+T_0), \forall t, \forall s, \tag{10c}$$

where $r_N(t,s)$ and $\tilde{r}_N(t,s)$ are called the auto-correlation and the complementary auto-correlation functions, respectively.

It is also assumed that the channel and the noise are band-limited to $f \in [-B_0, B_0)$, where and in what follows the left-end is included but the right-end is not included in a frequency interval just for convenience and consistency. Let $H(f)$, $R_N(f,f')$, and $\tilde{R}_N(f,f')$ be the Fourier transform of $h(t)$, the double Fourier transform of $r_N(t,s)$, and the double Fourier transform of $\tilde{r}_N(t,s)$, respectively, i.e.,

$$H(f) \triangleq \mathscr{F}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-\mathrm{j}2\pi f t}dt, \tag{11a}$$

$$\begin{aligned} R_N(f,f') &\triangleq \mathscr{F}^2\{r_N(t,s)\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r_N(t,s)e^{-\mathrm{j}2\pi(ft-f's)}dtds, \text{ and} \end{aligned} \tag{11b}$$

$$\begin{aligned} \tilde{R}_N(f,f') &\triangleq \mathscr{F}^2\{\tilde{r}_N(t,s)\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{r}_N(t,s)e^{-\mathrm{j}2\pi(ft-f's)}dtds, \end{aligned} \tag{11c}$$

where the operators $\mathscr{F}\{\cdot\}$ and $\mathscr{F}^2\{\cdot\}$ denote the Fourier transform and the double Fourier transform, respectively. Then, this assumption can be written as

$$H(f) = 0, \forall f \notin [-B_0, B_0), \text{ and} \tag{12a}$$

$$\begin{aligned} R_N(f,f') &= \tilde{R}_N(f,f') = 0, \\ &\forall (f,f') \notin [-B_0,B_0) \times [-B_0,B_0), \end{aligned} \tag{12b}$$

where the set operator $\times$ denotes the Cartesian product. In what follows, we alternately call $R_N(f,f')$ and $\tilde{R}_N(f,f')$ the two-dimensional (2-D) PSD and the 2-D complementary PSD of $N(t)$, respectively.

It is well known [9] that the 2-D PSD $R_N(f,f')$ of a cyclostationary random process $N(t)$ with cycle period T_0 consists of impulse fences on the lines $f = f' - m/T_0$, for $m \in \mathbb{Z}$, where $\mathbb{Z}$ denotes the set of all integers. In other words, there exist functions $\{R_N^{(k)}(f)\}_{k\in\mathbb{Z}}$ such that

$$R_N(f,f') = \sum_{k=-\infty}^{\infty} R_N^{(k)}\left(f - \frac{k}{T_0}\right)\delta\left(f - f' - \frac{k}{T_0}\right), \tag{13a}$$

where $\delta(\cdot)$ denotes the Dirac delta function. Similarly, the 2-D complementary PSD $\tilde{R}_N(f,f')$ can be rewritten as

$$\tilde{R}_N(f,f') = \sum_{k=-\infty}^{\infty} \tilde{R}_N^{(k)}\left(f - \frac{k}{T_0}\right)\delta\left(f - f' - \frac{k}{T_0}\right). \tag{13b}$$

Note that the non-zero self-coherence (10b) and conjugate self-coherence (10c) in the form of the time-domain periodicity in t and s correspond to the non-zero impulse fences that are spaced the integer multiples of the cycle frequency $1/T_0$ apart in the frequency domain. These results can be obtained by double-Fourier transforming the Fourier series representations of $r_N(t,s)$ and $\tilde{r}_N(t,s)$, where the functions $\{R_N^{(k)}(f)\}_{k\in\mathbb{Z}}$ and $\{\tilde{R}_N^{(k)}(f)\}_{k\in\mathbb{Z}}$ represent the heights of the impulse fences. For a measure-theoretic treatment of these 2-D PSDs, see [29]. To avoid considering pathological functions, it is simply assumed in this paper that all frequency-domain functions such as $H(f)$, $\{R_N^{(k)}(f)\}_{k\in\mathbb{Z}}$, and $\{\tilde{R}_N^{(k)}(f)\}_{k\in\mathbb{Z}}$ are piecewise smooth.

Given the channel model described above, the objective of this paper is to find the channel capacity and the capacity-achieving distribution of the complex-valued channel input $X(t)$, subject to the average input power constraint

$$\lim_{\Delta\to\infty} \frac{1}{2\Delta}\int_{-\Delta}^{\Delta} \mathbb{E}\{|X(t)|^2\}dt = P. \tag{14}$$

III. Derivation of Equivalent MIMO Channel Using p-FRESH Vectorizer

In this section, we introduce two LCL-TV operators that convert back and forth between a scalar-valued signal and a vector-valued one. These operators are then used to find an equivalent MIMO channel model, which facilitates the derivation of the channel capacity in the next section. This approach is strongly motivated by the past work on co-channel interference excision, e.g., [30]–[32], that converts a SISO channel into a MIMO channel in order to better exploit the spectral correlation in a cyclostationary interference. Before proceeding, we define Nyquist zones and half-Nyquist zones.

Definition 4: Given a pair $(B, 1/T)$ of a reference bandwidth and a reference rate, the lth Nyquist zone $\mathcal{F}_l$, its center frequency f_l, the half-Nyquist zone $\mathcal{F}^+$, and the lth half-Nyquist zone $\mathcal{F}_l^+$ are defined, respectively, as

$$\mathcal{F}_l \triangleq \left\{f : f_l - \frac{1}{2T} \le f < f_l + \frac{1}{2T}\right\}, \tag{15a}$$

$$f_l \triangleq \frac{l - \lfloor L/2 \rfloor - 1}{T}, \tag{15b}$$

$$\mathcal{F}^+ \triangleq \left\{f : 0 \le f < \frac{1}{2T}\right\}, \text{ and} \tag{15c}$$

$$\mathcal{F}_l^+ \triangleq \mathcal{F}^+ + f_l = \left\{f : f_l \le f < f_l + \frac{1}{2T}\right\}, \tag{15d}$$

for $l = 1, 2, ..., L$, where the integer L is given by

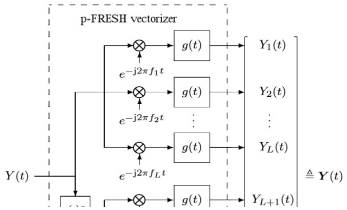
$$L \triangleq \lceil 2BT \rceil. \tag{15e}$$

Note that a half-Nyquist zone is just the right-half of the corresponding Nyquist zone. Also note that L is chosen as the number of length-$1/(2T)$ intervals that can fully cover the one-sided frequency band $f \in [0, B)$ of interest.

First, we define an LCL-TV operator that converts a scalar-valued signal to a vector-valued one.

Definition 5: Given a reference pair $(B, 1/T)$ and an input $Y(t)$, the *properizing FREquency SHift (p-FRESH)*[3] vectorizer is defined as a single-input multiple-output (SIMO) LCL-TV system, whose output $\boldsymbol{Y}(t)$ is given by

$$\left[\; Y_1(t) \;\right]$$



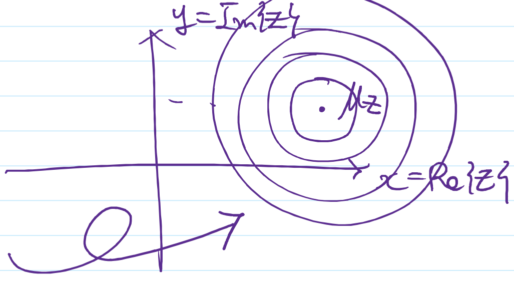
□ $z(t)$

□ $\underline{z}(t)$

□ $c(t,\tau)$

$$= \frac{1}{\pi(2\sigma^2/2)} \exp\left(-\frac{|z-\mu_z|^2}{(2\sigma^2/2)}\right)$$

$$f_z(z) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|z-\mu_z|^2}{\sigma^2}\right)$$



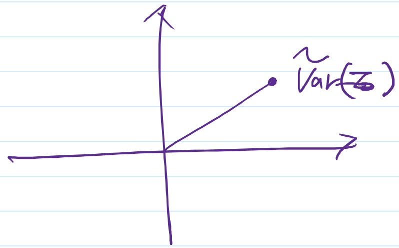
Lemma. If $\underline{z} = \underline{z}_0 + \mu_z$

where $f_{\underline{z}_0}(z_0)$ is circular w.r.t. $z_0 = 0$

$(z-\mu_z)e^{j\theta}$

?, then $\underline{z}$ is proper.

proof. $\widetilde{\mathrm{Var}}(\underline{z}) = E\{\underline{z}_0^2\} =$



$\underline{z}_1 \triangleq \underline{z}_0 e^{j\theta}$

$f_{\underline{z}_1}(z_0) = f_{\underline{z}_0}(z_0), \forall z_0 \in \mathbb{C}$.

$\boxed{\widetilde{\mathrm{Var}}(\underline{z}_1)} = E\{(\underline{z}_0 e^{j\theta})^2\} = e^{j2\theta} E\{\underline{z}_0^2\} = e^{j2\theta} \boxed{\widetilde{\mathrm{Var}}(\underline{z}_0)}, \forall\theta$

$= 0.$

$\theta = 0$
$\theta = \frac{\pi}{2}$

□ $\underline{z} = \underline{x} + j\underline{y}$

$E\{\underline{z}\} = \mu_z = \mu_x + j\mu_y$

$\mathrm{Cov}\{\underline{z}\} = E\{(\underline{z}-\mu_z)(\underline{z}-\mu_z)^H\} =$ ← $\mathrm{Cov}(\underline{x})$, $\mathrm{Cov}(\underline{x},\underline{y})$, $\mathrm{Cov}(\underline{y})$

$\widetilde{\mathrm{Cov}}\{\underline{z}\} = E\{(\underline{z}-\mu_z)(\underline{z}-\mu_z)^T\} =$

valued signal to a vector-valued one.

Definition 5: Given a reference pair $(B, 1/T)$ and an input $Y(t)$, the *properizing FREquency SHift (p-FRESH)*[3] vectorizer is defined as a single-input multiple-output (SIMO) LCL-TV system, whose output $\boldsymbol{Y}(t)$ is given by

$$\boldsymbol{Y}(t) \triangleq \begin{bmatrix} Y_1(t) \\ Y_2(t) \\ \vdots \\ Y_{2L}(t) \end{bmatrix}, \tag{16a}$$

where the lth entry $Y_l(t)$, for $l = 1, 2, ..., 2L$, is given by

$$Y_l(t) \triangleq \begin{cases} \left(Y(t)e^{-j2\pi f_l t}\right) * g(t), & \text{for } l \le L, \\ \left(Y(t)^* e^{-j2\pi f_{l-L} t}\right) * g(t), & \text{for } l > L, \end{cases} \tag{16b}$$

with $g(t)$ being the impulse response of the ideal brickwall filter having the half-Nyquist zone $\mathcal{F}^+$ as the support, i.e., the Fourier transform $G(f) \triangleq \mathscr{F}\{g(t)\}$ is given by

$$G(f) = \begin{cases} 1, & \forall f \in \mathcal{F}^+ \\ 0, & \text{elsewhere.} \end{cases} \tag{17}$$

Fig. 2-(a) shows how a p-FRESH vectorizer works in the time domain, where a scalar-valued signal $Y(t)$ is converted to a $2L$-by-1 vector-valued signal $\boldsymbol{Y}(t)$. The p-FRESH vectorizer modulates $Y(t)$ and its complex conjugate $Y(t)^*$ each with L different carriers, of which frequencies are integer multiples of the reference rate $1/T$, and filters to form $\boldsymbol{Y}(t)$, of which entries are all strictly band-limited to the half-Nyquist zone $\mathcal{F}^+$.

Fig. 2-(b) shows how a p-FRESH vectorizer works in the frequency domain when a deterministic signal $s(t)$ with $L = 2$ is converted to a vector-valued signal $\boldsymbol{s}(t)$. Let $\boldsymbol{s}(f)$ be the element-wise Fourier transform of $\boldsymbol{s}(t)$. Then, to generate the lth entry $[\boldsymbol{s}(f)]_l$ of $\boldsymbol{s}(f)$, the p-FRESH vectorizer left-shifts the Fourier transform $S(f)$ of $s(t)$ by f_l for $l \le L$, while $S(-f)^*$ by f_{l-L} for $l > L$, and leaves only the signal component in $\mathcal{F}^+$.

Note that, given a signal, different reference pairs result in different p-FRESH vectorizations in general. It can be seen that, given an arbitrary signal with bandwidth B_0, no information is lost during the vectorization as long as the reference bandwidth B satisfies

$$B_0 \le B, \tag{18}$$

regardless of the reference rate $1/T$. Note also that the p-FRESH vectorizer processes both $Y(t)$ and $Y(t)^*$, like linear-conjugate linear filters [33] or, equivalently, widely linear filters [34] that are devised to process improper-complex random signals. The major difference is that, as it can be easily seen from Fig. 2-(b), the p-FRESH vectorizer filters out half the signal components in $Y(t)$ and $Y(t)^*$ to completely remove the redundancy during the vectorization.

[3]The abbreviation FRESH appears first in [30], where the cyclic Wiener filtering that is a generalization of the Wiener filtering for cyclostationary random processes is proposed.

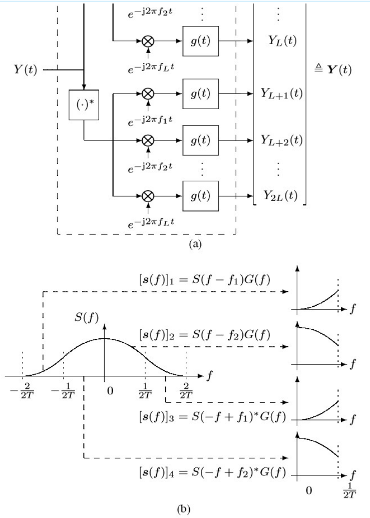


Fig. 2. (a) p-FRESH vectorization viewed in the time domain. (b) p-FRESH vectorization viewed in the frequency domain for a deterministic signal $s(t)$ with $L = 2$.

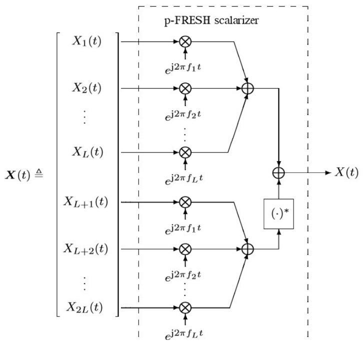
Second, we define an LCL-TV operator that converts a vector-valued signal to a scalar-valued one.

Definition 6: The p-FRESH scalarization of a $2L$-by-1 vector-valued signal $\boldsymbol{X}(t)$ with reference pair $(B, 1/T)$ is defined as

$$X(t) \triangleq \sum_{l=1}^{L} \left\{ X_l(t)e^{j2\pi f_l t} + \left(X_{L+l}(t)e^{j2\pi f_l t}\right)^* \right\}. \tag{19}$$

Fig. 3 shows how a p-FRESH scalarizer works in the time domain, where a $2L$-by-1 vector-valued random process $\boldsymbol{X}(t)$ is converted to a scalar-valued random process $X(t)$. In the frequency domain, the p-FRESH scalarizer right-shifts the lth input element $X_l(t)$ of $\boldsymbol{X}(t)$ by f_l, while the complex conjugate of the $(L + l)$th element by $-f_l$, both for $l = 1, 2, ..., L$, to construct the output $X(t)$.

If the input to the p-FRESH scalarizer is the p-FRESH vectorized version of an SOCS random process with a reference bandwidth $B \ge B_0$ and an arbitrary reference rate



the case with the HSR unless the input is proper-complex.[4]

Let $\boldsymbol{R_Y}(f) \triangleq \mathscr{F}\{\boldsymbol{r_Y}(\tau)\}$ be the $2L$-by-$2L$ element-wise Fourier transform of the auto-correlation matrix, i.e., $\boldsymbol{R_Y}(f)$ is the PSD matrix of the p-FRESH vectorizer output $\boldsymbol{Y}(t)$. From *Lemma 1*, it can be seen that a p-FRESH vectorizer with $T = KT_0$ not only generates a sufficient statistic that allows a perfect reconstruction of the SOCS input $Y(t)$ but also tactically aligns its frequency components in such a way that the vector-valued output $\boldsymbol{Y}(t)$ has no correlation among different frequency components and no complementary correlation at all on $f \in \mathcal{F}^+$. The next lemma points out one important property of the PSD matrix $\boldsymbol{R_Y}(f)$ of $\boldsymbol{Y}(t)$.

Lemma 2: Given a reference pair $(B, 1/T)$, the PSD matrix $\boldsymbol{R_Y}(f)$ of the p-FRESH vectorizer output $\boldsymbol{Y}(t)$ of an SOCS random process $Y(t)$ with cycle period T is a Hermitian-symmetric positive semi-definite matrix for all $f \in \mathcal{F}^+$. ■

Proof: See Appendix B.

This property is used in the frequency-domain reformulation of the problem, as shown in the next section. The

$\mathrm{Cov}\{\underline{z}\} = E\{(\underline{z}-\mu_z)(\underline{z}-\mu_z)^H\} = \leftarrow \mathrm{Cov}(\underline{X}), \mathrm{Cov}(\underline{X},\underline{Y}), \mathrm{Cov}(\underline{Y})$

$\widetilde{\mathrm{Cov}}\{\underline{z}\} = E\{(\underline{z}-\mu_z)(\underline{z}-\mu_z)^T\} =$

□ Def. proper $\Leftrightarrow$ $\widetilde{\mathrm{Cov}}\{\underline{z}\} = 0$ $\Leftrightarrow$ $\mathrm{Cov}(\underline{X}) = \mathrm{Cov}(\underline{Y})$, $\mathrm{Cov}(\underline{X},\underline{Y}) = 0$

□ Def. (i) circular $\Leftrightarrow$ $\underline{z}e^{j\theta}$ has the same pdf as $\underline{z}$ for all $\theta \in [0, 2\pi)$.

(ii) circular $\Leftrightarrow$ $(\underline{z}-\mu_z)e^{j\theta}$

□ Lemma. If $\underline{z}$ is circular, then $\underline{z}$ is proper.

□ Def $\underline{z} = \underline{X} + j\underline{Y}$ is complex Gaussian if $\begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix}$ has Gaussian distribution.

□ Def If $\underline{z} \in \mathbb{C}^N$ a complex-Gaussian is proper then

$\xrightarrow{N=1}$ $f_z(z) = \frac{1}{\pi\sigma^2}\exp\left(-\frac{|z-\mu_z|^2}{\sigma^2}\right)$

$f_{\underline{z}}(\underline{z}) = \frac{1}{\pi^N \sigma_1^2\sigma_2^2\cdots\sigma_N^2}$

$\exp\left(-(\underline{z}-\mu_z)^H C^{-1}(\underline{z}-\mu_z)\right)$

For complex proper Gaussian

$f_{\underline{z}}(\underline{z}) = \frac{1}{\pi^N \det(C_{\underline{zz}})}\exp\left(-(\underline{z}-\mu_z)^H C_{\underline{zz}}^{-1}(\underline{z}-\mu_z)\right)$

when C is invertible.

For real Gaussian,

$f_{\underline{X}}(\underline{x}) = \frac{1}{\sqrt{(2\pi)^N \det(C)}}\exp\left(-\frac{(\underline{x}-\mu_x)^T C_{\underline{xx}}^{-1}(\underline{x}-\mu_x)}{2}\right)$

$\underline{X} \sim N(\mu_x, C_{\underline{xx}})$

□ Q Pdf of improper Gaussian Random Vector?

proper.

$\underline{z} \sim CN(\mu_z, C_{\underline{zz}})$

$\underline{z} \sim CPN(\mu_z, C_{\underline{zz}})$

$\boxed{\underline{z} \sim CN(\mu_z, C_{\underline{zz}}, \tilde{C}_{\underline{zz}})}$

$(\underline{z}-\mu_z)e^{j\theta}$ $\quad$ $\underline{z} \sim CN(\mu_z, C_{\underline{zz}}, 0)$

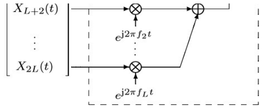


Fig. 3. p-FRESH scalarization viewed in the time domain.

$1/T$, then the p-FRESH scalarizer with the same reference pair as the vectorizer just performs the inverse operation of the p-FRESH vectorization. In this sense, a p-FRESH vectorization of the SOCS process may not be more than one of many possible representations that convert a scalar-valued signal or random process to an equivalent vector-valued one. However, the following lemma shows that an appropriate choice of the reference rate $1/T$ is crucial to capture the correlation structure in an SOCS process. In what follows, we define superscripts ^{T} and $^\mathcal{H}$ as the transposition and the Hermitian transposition, respectively, and $\mathbf{0}_{2L}$ and $\mathbf{0}_{2L\times 2L}$ as the $2L$-by-1 all-zero vector and the $2L$-by-$2L$ all-zero matrix, respectively.

Lemma 1: If the reference rate $1/T$ of a p-FRESH vectorizer is chosen such that a zero-mean SOCS random process input $Y(t)$ with cycle period T_0 is also SOCS with cycle period T, then the output $\boldsymbol{Y}(t)$ always becomes a proper-complex zero-mean vector-WSS process, i.e., if

$$T = KT_0 \tag{20}$$

for some positive integer K, then

$$\boldsymbol{\mu_Y}(t) \triangleq \mathbb{E}\{\boldsymbol{Y}(t)\} = \mathbf{0}_{2L}, \forall t, \tag{21a}$$
$$\boldsymbol{r_Y}(t,s) \triangleq \mathbb{E}\{\boldsymbol{Y}(t)\boldsymbol{Y}(s)^\mathcal{H}\} = \boldsymbol{r_Y}(t-s), \forall t, \forall s, \text{and} \tag{21b}$$
$$\tilde{\boldsymbol{r}}_{\boldsymbol{Y}}(t,s) \triangleq \mathbb{E}\{\boldsymbol{Y}(t)\boldsymbol{Y}(s)^T\} = \mathbf{0}_{2L\times 2L}, \forall t, \forall s, \tag{21c}$$

where $\boldsymbol{r_Y}(\tau) \triangleq \boldsymbol{r_Y}(\tau, 0)$ is a $2L$-by-$2L$ matrix-valued function.

Proof: See Appendix A. ■

As far as the auto-correlation function $\boldsymbol{r_Y}(t,s)$ is concerned, this lemma applying the p-FRESH vectorizer to an SOCS input with an arbitrary K is a mere extension of the result in [9], where the HSR is applied to a WSCS input with $K = 1$. However, it additionally shows that, regardless of the propriety of the SOCS input, the complementary auto-correlation function $\tilde{\boldsymbol{r}}_{\boldsymbol{Y}}(t,s)$ always vanishes, which is not

$\boldsymbol{R_Y}(f)$ of the p-FRESH vectorizer output $\boldsymbol{Y}(t)$ of an SOCS random process $Y(t)$ with cycle period T is a Hermitian-symmetric positive semi-definite matrix for all $f \in \mathcal{F}^+$.

Proof: See Appendix B. ■

This property is used in the frequency-domain re-formulation of the problem, as shown in the next section. The following proposition shows how the p-FRESH vectorization of the channel output can be rewritten in terms of the p-FRESH vectorization of the channel input and that of the channel impulse response.

Proposition 1: The p-FRESH vectorization $\boldsymbol{Y}(t)$ of the channel output $Y(t)$ can be rewritten as

$$\boldsymbol{Y}(t) = \mathrm{diag}\{\boldsymbol{h}(t)\} * \boldsymbol{X}(t) + \boldsymbol{N}(t), \tag{22}$$

where $\boldsymbol{Y}(t)$, $\boldsymbol{h}(t)$, $\boldsymbol{X}(t)$, and $\boldsymbol{N}(t)$ are the p-FRESH vectorizations of $Y(t)$, $h(t)$, $X(t)$, and $N(t)$, respectively, and $\mathrm{diag}\{\boldsymbol{h}(t)\}$ denotes the diagonal matrix having the lth diagonal entry equal to the lth entry of $\boldsymbol{h}(t)$.

Proof: By the definition (16b), the lth entry $[\boldsymbol{Y}(t)]_l = Y_l(t)$ of $\boldsymbol{Y}(t)$ can be rewritten, for $l \le L$, as

$$Y_l(t) \triangleq \left(Y(t)e^{-j2\pi f_l t}\right) * g(t) \tag{23a}$$
$$= \left((h(t) * X(t) + N(t))e^{-j2\pi f_l t}\right) * g(t) \tag{23b}$$
$$= \left(\int_{-\infty}^{\infty} h(\tau)X(t-\tau)d\tau \cdot e^{-j2\pi f_l t}\right) * g(t) + N_l(t) \tag{23c}$$
$$= \int_{-\infty}^{\infty} h(\tau)e^{-j2\pi f_l \tau} X_l(t-\tau)d\tau + N_l(t) \tag{23d}$$
$$= [\boldsymbol{h}(t)]_l * [\boldsymbol{X}(t)]_l + [\boldsymbol{N}(t)]_l, \tag{23e}$$

where (23d) follows since the convolution with $g(t)$ acts only on signals parameterized by t, and (23e) follows since $g(t)$ is the impulse response of the ideal brickwall filter. Similarly, it can be shown that (23e) also holds for $l > L$. Therefore, we have (22). ■

This implies that, by placing a p-FRESH scalarizer and a p-FRESH vectorizer, respectively, at the input and the output of the SISO channel as shown in Fig. 4-(a), the SISO channel depicted in Fig. 1 can be converted to an equivalent MIMO channel shown in Fig. 4-(b). This equivalent MIMO channel is simpler than general MIMO channels, in that the channel matrix is always square and diagonal. However, the additive proper-complex Gaussian noise $\boldsymbol{N}(t)$ is in general colored.

[4] As it will be discussed in the second paragraph of Section V, the selection of the center frequency is very important in down-converting a real bandpass random process to make it SOCS.

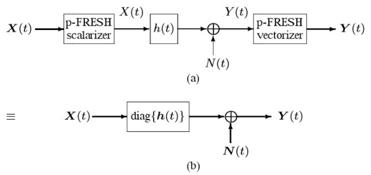


Fig. 4. (a) Conversion of a SISO channel to an equivalent MIMO channel by using a p-FRESH scalarizer and a p-FRESH vectorizer. (b) Equivalent MIMO channel with vector-valued colored noise.

Since the capacity of MIMO additive proper-complex white Gaussian noise channels is well known [25], we consider whitening the vector-WSS noise $\boldsymbol{N}(t)$ of the equivalent MIMO channel as shown in the next section.

IV. Cyclic Water Filling and Capacity of SOCS Gaussian Noise Channel

In this section, the problem is re-formulated in the frequency domain and then the channel capacity is derived. Throughout the section, it is assumed that

$$B = B_0 \text{ and } T = KT_0 \tag{24}$$

(c) The degree of freedom $\mathcal{N}_p(f)$ at f is defined as the length of an effective Fourier transform of the p-FRESH vectorizer output.

When we design a band-limited signal $s(t)$ or a band-limited SOCS random process $X(t)$, such a removal of the entries in $\boldsymbol{s}(f)$ or $\boldsymbol{R_X}(f)$ is necessary because every entry is not always a free variable unless $L = 2BT$. From (25a) and (25b), it can be seen that $\mathcal{N}_p(f)$ is given by

$$\mathcal{N}_p(f) \triangleq \begin{cases} 2L, & \text{for } f \in [0, B-(L-1)/(2T)), \\ 2L-2, & \text{otherwise} \end{cases} \tag{26a}$$

when L is odd, while

$$\mathcal{N}_p(f) \triangleq \begin{cases} 2L-2, & \text{for } f \in [0, L/(2T)-B), \\ 2L, & \text{otherwise} \end{cases} \tag{26b}$$

when L is even. Thus, an effective Fourier transform of the p-FRESH vectorizer output becomes an $\mathcal{N}_p(f)$-by-1 vector and an effective PSD matrix becomes an $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix at each $f \in \mathcal{F}^+$. For simplicity, we do not introduce new notations for the effective Fourier transforms and PSD matrices.

Definition 8: Given a reference pair $(B, 1/T)$, the element-wise inverse Fourier transform of an $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix-valued function of $f \in \mathcal{F}^+$ is defined as a $2L$-by-$2L$ matrix-valued function of t obtained as follows: First, construct a $2L$-by-$2L$ matrix-valued function $\boldsymbol{W}(f)$ of f as if the $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix-valued function is the effective PSD matrix of

$\underline{z} \sim \mathcal{CN}(\mu_{\underline{z}}, C_{\underline{z}\underline{z}}, \mathbb{O})$

$(z-\mu_z)e^{j\theta}$

$ze^{j\theta}$

1-D circularity → N-D $\underline{z}e^{j\theta}$, $(\underline{z}-\mu_{\underline{z}})e^{j\theta}$

→ N-D $U\underline{z}$, $U(\underline{z}-\mu_{\underline{z}})$ ... w/ unitary matrix U.

☐ Theorem.

$\underline{y} = H\underline{x} + \underline{n}$

where $\underline{x}$ input vector $\in \mathbb{C}^M$

$\underline{y}$ output " $\in \mathbb{C}^N$

H ch. matrix $\in \mathbb{C}^{M\times N}$

$\underline{n} \sim \mathcal{CN}(\underline{0}, C_{\underline{n}\underline{n}})$

☐ SVD (Singular-Value Decomposition) ...

of $H = U\Sigma V^H$

$$\begin{cases} U^H U = UU^H = I \\ V^H V = VV^H = I \\ [\Sigma]_{i,j} = \sigma_i^2 \delta_{ij}, \ 1 \le i \le M, \ 1 \le j \le N \end{cases}$$

(Recall $\Sigma \in \mathbb{C}^{M\times N}$, $H \in \mathbb{C}^{M\times N}$)

☐ Corollary $H = \mathrm{diag}(\underline{h}) = \begin{bmatrix} |h_1|e^{j\phi_1} & & & \\ & |h_2|e^{j\phi_2} & & \\ & & \ddots & \\ & & & |h_M|e^{j\phi_M} \end{bmatrix}$

arbitrary choice

$= \begin{bmatrix} e^{j\phi_1} & & & \\ & e^{j\phi_2} & & \\ & & e^{j\phi_3} & \\ & & & e^{j\phi_4+\theta_4} \end{bmatrix} \begin{bmatrix} |h_1| & & & \\ & |h_2| & & \\ & & \ddots & \\ & & & |h_M| \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & e^{-j3\phi_3} & \\ & & & e^{j\theta_4} \end{bmatrix}^H$

GAUSSIAN NOISE CHANNEL

In this section, the problem is re-formulated in the frequency domain and then the channel capacity is derived. Throughout the section, it is assumed that

$$B = B_0 \text{ and } T = KT_0 \tag{24}$$

for some positive integer K, unless otherwise specified. Thus, the p-FRESH vectorization of $Y(t)$ becomes a sufficient statistic and the SOCS noise becomes a vector-WSS process.

A. Noise Whitening and Problem Re-Formulation

To find a whitening filter, we need the following notions of effective Fourier transform of the p-FRESH vectorizer output, effective PSD matrix, degree of freedom, and effective element-wise inverse Fourier transform.

Definition 7: Suppose that a reference pair $(B, 1/T)$ is given for a p-FRESH vectorizer and that $L \triangleq \lceil 2BT \rceil$ is defined as before.

(a) The effective Fourier transform of the p-FRESH vectorizer output of a deterministic signal $s(t)$ is defined as a variable-length vector-valued function of f that is obtained by removing the Lth and the $2L$th entries of $\boldsymbol{s}(f) \triangleq \mathscr{F}\{\boldsymbol{s}(t)\}$ for

$$B - \frac{L-1}{2T} \le f < \frac{1}{2T} \tag{25a}$$

when L is odd, and the 1st and the $(L+1)$th entries for

$$0 \le f < \frac{L}{2T} - B \tag{25b}$$

when L is even.

(b) The effective PSD matrix of the p-FRESH vectorizer output of an SOCS random process $X(t)$ with cycle period T is defined as a variable-size matrix-valued function of f that is obtained by removing the Lth and the $2L$th rows and columns of $\boldsymbol{R_X}(f) \triangleq \mathscr{F}\{\boldsymbol{r_X}(\tau)\}$ for f satisfying the condition (25a) when L is odd, while the 1st and the $(L+1)$th rows and columns for f satisfying the condition (25b) when L is even.

Definition 8: Given a reference pair $(B, 1/T)$, the element-wise inverse Fourier transform of an $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix-valued function of $f \in \mathcal{F}^+$ is defined as a $2L$-by-$2L$ matrix-valued function of t obtained as follows: First, construct a $2L$-by-$2L$ matrix-valued function $\boldsymbol{W}(f)$ of f as if the $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix-valued function is the effective PSD matrix of $\boldsymbol{W}(f)$. Second, construct a $2L$-by-$2L$ matrix-valued function $\boldsymbol{W}(t)$ of t by applying the inverse Fourier transform to every entry of $\boldsymbol{W}(f)$.

Note that, by definition, the Lth and the $2L$th row and column of $\boldsymbol{W}(f)$ have all zero entries for f satisfying the condition (25a), while the 1st and the $(L+1)$th row and column have all zero entries for f satisfying the condition (25b).

Proposition 2: Let $\boldsymbol{W}(t)$ be the element-wise inverse Fourier transform of the inverse square root of the effective PSD matrix of $N(t)$, i.e.,

$$\boldsymbol{W}(t) = \mathscr{F}^{-1}\{\boldsymbol{R_N}(f)^{-\frac{1}{2}}\}. \tag{27}$$

Then, the $2L$-by-$2L$ MIMO LTI system with impulse response $\boldsymbol{W}(t)$ whitens the noise $\boldsymbol{N}(t)$.

Proof: Define the output of the MIMO LTI system as

$$\tilde{\boldsymbol{N}}(t) \triangleq \boldsymbol{W}(t) * \boldsymbol{N}(t). \tag{28}$$

Then, its auto-correlation matrix $\boldsymbol{r}_{\tilde{\boldsymbol{N}}}(\tau)$ can be written as

$$\boldsymbol{r}_{\tilde{\boldsymbol{N}}}(\tau) \triangleq \mathbb{E}\{\tilde{\boldsymbol{N}}(t)\tilde{\boldsymbol{N}}(t-\tau)^{\mathcal{H}}\} \tag{29a}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{W}(t)\boldsymbol{r_N}(\tau - t + s)\boldsymbol{W}(s)^{\mathcal{H}} dtds. \tag{29b}$$

Thus, the PSD matrix $\boldsymbol{R}_{\tilde{\boldsymbol{N}}}(f)$ is given by

$$\boldsymbol{R}_{\tilde{\boldsymbol{N}}}(f) \triangleq \mathscr{F}\{\boldsymbol{r}_{\tilde{\boldsymbol{N}}}(\tau)\} \tag{30a}$$

$$= \boldsymbol{R_N}(f)^{-\frac{1}{2}}\boldsymbol{R_N}(f)\boldsymbol{R_N}(f)^{-\frac{1}{2}} = \boldsymbol{I}_{\mathcal{N}_p(f)}, \tag{30b}$$

where we used the fact that $\boldsymbol{R_N}(f)^{-\frac{1}{2}}$ exists for all $f \in \mathcal{F}^+$ because the PSD matrix $\boldsymbol{R}_{\boldsymbol{N_I}}(f)$ of the interference $N_I(t)$ is positive semi-definite $\forall f \in \mathcal{F}^+$ by *Lemma 2* and the PSD matrix $\boldsymbol{R}_{\boldsymbol{N_A}}(f)$ of the ambient noise $N_A(t)$ is a scaled

identity matrix $\forall f \in \mathcal{F}^+$. Therefore, the conclusion follows. ■

This MIMO whitening filter plays a crucial role in formulating and solving the optimization problem to find the channel capacity in the frequency domain. In what follows, all the functions of $f \in \mathcal{F}^+$ are effective ones, unless otherwise specified.

Theorem 1: The optimization problem to find the capacity of an SOCS Gaussian noise channel is given by

$$\underset{(\boldsymbol{R_X}(f))_f}{\text{maximize}} \int_{\mathcal{F}^+} \log_2 \det\left\{\boldsymbol{I}_{\mathcal{N}_p(f)} + \tilde{\boldsymbol{H}}(f)\boldsymbol{R_X}(f)\tilde{\boldsymbol{H}}(f)^{\mathcal{H}}\right\}df \tag{31a}$$

$$\text{subject to} \int_{\mathcal{F}^+} \text{tr}\{\boldsymbol{R_X}(f)\}df = P, \tag{31b}$$

where tr$\{\cdot\}$ denote the trace, the PSD matrix $\boldsymbol{R_X}(f)$ of $\boldsymbol{X}(t)$ is a positive semi-definite $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ matrix for all $f \in \mathcal{F}^+$, and the noise equivalent channel $\tilde{\boldsymbol{H}}(f)$ at $f \in \mathcal{F}^+$ is defined as

$$\tilde{\boldsymbol{H}}(f) \triangleq \boldsymbol{R_N}(f)^{-\frac{1}{2}}\boldsymbol{H}(f), \tag{32}$$

where

$$\boldsymbol{H}(f) \triangleq \text{diag}\{\boldsymbol{h}(f)\}. \tag{33}$$

Proof: If we place the whitening filter $\boldsymbol{W}(t)$ defined by (27) at the output of the p-FRESH vectorizer in Fig. 4-(a), then the observation model (22) can be converted equivalently to

$$\tilde{\boldsymbol{Y}}(t) = \tilde{\boldsymbol{H}}(t) * \boldsymbol{X}(t) + \tilde{\boldsymbol{N}}(t), \tag{34}$$

where $\tilde{\boldsymbol{N}}(t)$ is defined in (28), and $\tilde{\boldsymbol{Y}}(t)$ and $\tilde{\boldsymbol{H}}(t)$ are defined as $\tilde{\boldsymbol{Y}}(t) \triangleq \boldsymbol{W}(t) * \boldsymbol{Y}(t)$ and $\tilde{\boldsymbol{H}}(t) \triangleq \boldsymbol{W}(t) * \text{diag}\{\boldsymbol{h}(t)\}$, respectively. Thus, the MIMO channel shown in Fig. 4-(a) is equivalent to that in Fig. 5-(a). Since the additive Gaussian noise $\tilde{\boldsymbol{N}}(t)$ is white as shown in *Proposition 2*, the optimization problem to find the capacity of the SOCS

have independence among $\boldsymbol{N}^{(k)}(t)$, for $k = 1, 2, ..., K$. It is well known that the capacity of parallel Gaussian channels with independent noise components can be achieved by using independent channel codes [21]. Thus, the optimal input PSD matrix $\boldsymbol{R_X}(f)$ becomes a block diagonal matrix. If we multiply complex exponential functions to $\boldsymbol{Y}^{(k)}$, for $k = 1, 2, ..., K$, such that $\boldsymbol{Y}^{(k)}$ is shifted in the frequency domain to the interval $f \in [(k-1)/(2KT_0), k/(2KT_0))$, the mutual information to be maximized becomes

$$\sum_{k=1}^{K}\int_{\frac{k-1}{2KT_0}}^{\frac{k}{2KT_0}} \log_2 \det\left\{\boldsymbol{I}_{\mathcal{N}_p(f)} + \tilde{\boldsymbol{H}}(f)\boldsymbol{R_X}(f)\tilde{\boldsymbol{H}}(f)^{\mathcal{H}}\right\}df, \tag{36}$$

which is identical to (31a) with $K = 1$, where $\mathcal{N}_p(f)$, $\tilde{\boldsymbol{H}}(f)$, and $\boldsymbol{R_X}(f)$ are all defined for $K = 1$. Since the optimization problems have the same objective functions and constraints, the conclusion follows. ■

B. Optimal Solution and Its Property

In this subsection, the optimal solution to (31) is derived and its property is investigated. For this, consider the singular value decomposition

$$\tilde{\boldsymbol{H}}(f) = \boldsymbol{U}(f)\boldsymbol{\Gamma}(f)\boldsymbol{V}(f)^{\mathcal{H}} \tag{37}$$

of the $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ noise equivalent channel matrix $\tilde{\boldsymbol{H}}(f)$ for each $f \in \mathcal{F}^+$, where $\boldsymbol{U}(f)$ and $\boldsymbol{V}(f)$ are unitary matrices, and $\boldsymbol{\Gamma}(f)$ is a diagonal matrix having the singular values $\gamma_n(f)$, for $n = 1, 2, ..., \mathcal{N}_p(f)$, as the diagonal entries.

Theorem 2: Let $\boldsymbol{R}_{\tilde{\boldsymbol{X}},\text{opt}}(f)$ be an $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ diagonal matrix whose nth diagonal entry is given by

$$[\boldsymbol{R}_{\tilde{\boldsymbol{X}},\text{opt}}(f)]_{n,n} \triangleq \left[\nu_{\text{opt}} - \frac{1}{\gamma_n(f)^2}\right]^+, \tag{38a}$$

where ν_{opt} is the unique solution to

$e^{j\phi_4+\theta_4}$ $|h_M|$ $e^{j\phi_3}$ $e^{j\theta_4}$

□ Thm. If $\underline{n} \sim CN(\underline{0}, \sigma^2 I)$

then

$$C \text{ (bits/use)} = \max_{p_i \ge 0} \log\left(\prod \frac{[p_i - \sigma_i^2]^+}{\sigma^2}\right)$$

s.t. $\sum p_i = P = E\{\|\underline{x}\|^2\}$.

□ Cor. If $\underline{n} \sim CN(\underline{0}, C)$

then ... $C \succ 0$

$C = WDW^H$ $\quad D = \text{diag}(d_1, d_2, \dots, d_M)$

$W = [w_1\ w_2 \dots w_M]$, $WW^H = W^HW = I$

normalized eigenvector of C.

$$\underline{x} \to \boxed{H} \to \oplus \to \underline{y} \to \boxed{C^{-\frac{1}{2}}} \to$$

$\underline{n}$

$\underline{z} = C^{-\frac{1}{2}}H\underline{x} + C^{-\frac{1}{2}}\underline{n}$ $\to CN(0, I)$

where $C^{-\frac{1}{2}} = WD^{-\frac{1}{2}}W^H = W\begin{bmatrix} \frac{1}{\sqrt{d_1}} & & & \\ & \frac{1}{\sqrt{d_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{d_M}} \end{bmatrix}W^H$

$\underline{z} = \boxed{\tilde{H}}\underline{x} + \tilde{\underline{n}}$

□ $C^{-\frac{1}{2}}\,\text{diag}(\underline{h}) \ne \text{diag}$

$$[c_1\ c_2]\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = [c_1\ c_2]\begin{bmatrix} a e_1^T \\ b e_2^T \end{bmatrix} = a c_1 e_1^T + b c_2 e_2^T$$

□

where $\boldsymbol{N}(t)$ is defined in (28), and $\boldsymbol{Y}(t)$ and $\boldsymbol{H}(t)$ are defined as $\tilde{\boldsymbol{Y}}(t) \triangleq \boldsymbol{W}(t) * \boldsymbol{Y}(t)$ and $\tilde{\boldsymbol{H}}(t) \triangleq \boldsymbol{W}(t) * \mathrm{diag}\{\boldsymbol{h}(t)\}$, respectively. Thus, the MIMO channel shown in Fig. 4-(a) is equivalent to that in Fig. 5-(a). Since the additive Gaussian noise $\tilde{\boldsymbol{N}}(t)$ is white as shown in *Proposition 2*, the optimization problem to find the capacity of the SOCS Gaussian noise channel can be formulated in exactly the same way as that of a MIMO channel with a proper-complex white Gaussian noise [26], where the proper-complex Gaussianity of the channel input is necessary to achieve the capacity. Therefore, the conclusion follows. ■

Before we move to the next subsection where the solution to (31) is derived, we examine whether a different choice of K may result in a different value of the channel capacity or not.

Proposition 3: For any positive integer K, the problem described in (31) is equivalent to that with $K = 1$.

Proof: Suppose that $\boldsymbol{Y}(t)$ is the output of the p-FRESH vectorizer when the input is $Y(t)$ and $T = KT_0$ for some $K > 1$. By pre-multiplying a permutation matrix, the elements of $Y(t)$ can be rearranged as

$$\bar{\boldsymbol{Y}}(t) \triangleq \begin{bmatrix} \boldsymbol{Y}^{(1)}(t) \\ \boldsymbol{Y}^{(2)}(t) \\ \vdots \\ \boldsymbol{Y}^{(K)}(t) \end{bmatrix}, \tag{35}$$

where the lth entry of $\boldsymbol{Y}^{(k)}(t)$, for $k = 1, 2, ..., K$, is defined as $[\boldsymbol{Y}^{(k)}(t)]_l \triangleq Y_{(l-1)K+k}(t)$. If the permuted vector-valued noise $\bar{\boldsymbol{N}}(t)$ and its elements $\boldsymbol{N}^{(k)}(t)$, for $k = 1, 2, ..., K$, are similarly defined, then, by *Lemma 1* with $K = 1$, we

Theorem 2: Let $\boldsymbol{R}_{\hat{\boldsymbol{X}},\mathrm{opt}}(f)$ be an $\mathcal{N}_p(f)$-by-$\mathcal{N}_p(f)$ diagonal matrix whose nth diagonal entry is given by

$$[\boldsymbol{R}_{\hat{\boldsymbol{X}},\mathrm{opt}}(f)]_{n,n} \triangleq \left[\nu_{\mathrm{opt}} - \frac{1}{\gamma_n(f)^2}\right]^+, \tag{38a}$$

where ν_{opt} is the unique solution to

$$\int_{\mathcal{F}^+} \sum_{n=1}^{\mathcal{N}_p(f)} \left[\nu_{\mathrm{opt}} - \frac{1}{\gamma_n(f)^2}\right]^+ df = P \tag{38b}$$

with $[x]^+ \triangleq (x+|x|)/2$ denoting the positive part of x.[5] Then, the optimal solution to (31) is given by

$$\boldsymbol{R}_{\boldsymbol{X},\mathrm{opt}}(f) = \boldsymbol{V}(f)\boldsymbol{R}_{\hat{\boldsymbol{X}},\mathrm{opt}}(f)\boldsymbol{V}(f)^{\mathcal{H}}, \tag{39}$$

and, consequently, the capacity of the SOCS Gaussian noise channel is given by

$$C_{\mathrm{SOCS}} = \int_{\mathcal{F}^+} \sum_{n=1}^{\mathcal{N}_p(f)} \log_2\left(1 + [\gamma_n(f)^2 \nu_{\mathrm{opt}} - 1]^+\right) df. \tag{40}$$

Proof: By linearly pre-coding the signal $\boldsymbol{X}(t)$ as $\hat{\boldsymbol{X}}(t) \triangleq \mathscr{F}^{-1}\{\boldsymbol{V}(f)^{\mathcal{H}}\} * \boldsymbol{X}(t)$ and transforming the whitened observation $\tilde{\boldsymbol{Y}}(t)$ as $\hat{\boldsymbol{Y}}(t) \triangleq \mathscr{F}^{-1}\{\boldsymbol{U}(f)^{\mathcal{H}}\} * \tilde{\boldsymbol{Y}}(t)$, we obtain $\hat{\boldsymbol{Y}}(t) = \mathscr{F}^{-1}\{\boldsymbol{\Gamma}(f)\} * \hat{\boldsymbol{X}}(t) + \hat{\boldsymbol{N}}(t)$, where $\hat{\boldsymbol{N}}(t) \triangleq \mathscr{F}^{-1}\{\boldsymbol{U}(f)^{\mathcal{H}}\} * \tilde{\boldsymbol{N}}(t)$. Thus, the MIMO channel shown in Fig. 5-(a) is equivalent to that in Fig. 5-(b). Because the proper-complex noise process $\tilde{\boldsymbol{N}}(t)$ is white and $\boldsymbol{U}(f)$ is unitary, the noise process $\hat{\boldsymbol{N}}(t)$ is also proper-complex and white. Thus, the integrand of the objective function in (31a)

[5] If $\gamma_{n_0}(f) = 0$ for some n_0, then $[\boldsymbol{R}_{\hat{\boldsymbol{X}},\mathrm{opt}}(f)]_{n_0,n_0} = 0$ and n_0 is excluded from the summation in (38b).