This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LCOMM.2014.2338316

Active Interference Cancellation for OFDM Spectrum Sculpting: Linear Processing is Optimal

Jorge F. Schmidt, *Member, IEEE,* Daniel Romero, *Student Member, IEEE*, and Roberto López-Valcarce, *Member, IEEE*

***Abstract*— Active interference cancellation (AIC) is a multi-carrier spectrum sculpting technique which reduces the power of undesired out-of-band emissions by adequately modulating a subset of reserved cancellation subcarriers. In most schemes online complexity is a concern, and thus cancellation subcarriers have traditionally been constrained to linear combinations of the data subcarriers. Recent AIC designs truly minimizing out-of-band emission shift complexity to the offline design stage, motivating the consideration of more general mappings to improve performance. We show that there is no loss in optimality incurred by constraining these mappings to the set of linear functions.**

***Index Terms*— Active Interference Cancellation, Out-of-band radiation, Spectrum Sculpting, Cognitive OFDM.**

I. Introduction

Orthogonal Frequency Division Multiplexing (OFDM) has become the modulation format of choice in modern high-speed wireless and wireline systems, due to its many well-known qualities. Nevertheless, one shortcoming of OFDM resides in the large sidelobes of the Inverse Discrete Fourier Transform (IDFT), which result in substantial leakage across subcarriers with the ensuing adjacent channel interference. This issue is often dealt with by deactivating a number of guard subcarriers at the edges of the signal spectrum, with the consequent penalty in data rate. In order for OFDM to be adopted by future high-performance systems, e.g., 5G, a number of enhancements will become necessary to overcome this and other drawbacks [1], [2]. The leakage problem is also of concern in wideband OFDM-based cognitive systems in which deep notches must be sculpted in the spectrum in order to avoid interfering to narrowband licensed users [3].

An appealing approach to IDFT leakage reduction is active interference cancellation (AIC), first proposed in [4]: undesired emission is reduced by judiciously modulating a number of *cancellation subcarriers* (CSs), while using the remaining *data subcarriers* (DSs) for transmission as usual. Thus, operation is transparent to the receiver, which just discards the CSs after demodulation. The advantage of AIC resides in that the number of CSs required to achieve a given level of undesired emission is typically much smaller than the number of guard subcarriers to be turned off in the traditional approach.

Several AIC designs have been subsequently proposed [5]-[9]. These works minimize w.r.t. the CS values a cost function given by the magnitude of the instantaneous signal spectrum for a given OFDM symbol at a number of frequencies within the protected band, subject to different constraints to control the power allocated to CSs. These designs suffer from one main drawback. Finding the optimal CS values requires solving a constrained optimization problem for each OFDM symbol, as the solution is dependent on the specific DS values; this results in significant online complexity.

In this context, AIC has been recast in terms of minimization of the average undesired emission power under a total power constraint, assuming a linear map from DS to CS values [10]. This formulation drastically reduces the online computational cost of AIC, which is a main concern in practice. Specifically, the resulting matrix defining the optimal mapping is independent of the instantaneous data, so no optimization problem has to be solved on the fly: online complexity remains low, without sacrificing performance, and with tight control on the transmit power. In view of this, as a next step it is reasonable to ask whether performance could be further improved by allowing more general (nonlinear) relations between DS and CS values under the framework of [10]. Were the answer affirmative, then it would make sense to approach the design of such nonlinear mappings in an optimal way.

Our contribution is to answer this question: we prove that performance cannot improve by incorporating nonlinear dependencies. Hence, there is no loss of optimality by restricting the mapping from DS to CS values to the class of linear functions, which have the advantage of simple implementation. We note that the fact that certain earlier AIC designs such as [6] and [9] directly result in a linear relation for the optimum CS values does not readily imply that this should also be the case for the setting in [10] because, as stated above, the use of instantaneous values (as opposed to statistical averages) in the cost function and constraints results in structurally different problems. As it turns out, the corresponding proof is not trivial due to a number of constraints inherent to the problem.

The letter is organized as follows. Sec. II describes the psd-based AIC design from [10]. Optimality of linear processing is established in Sec. III, and conclusions are drawn in Sec. IV.

II. PSD-based Active Interference Cancellation

Consider the transmission of an OFDM signal with N subcarriers. The power radiated in some band $\mathcal{B}$, covered by N_P contiguous subcarriers within the transmission bandwidth, is to be minimized[1]. The AIC scheme reserves $N_A =$

The authors are with the Department of Signal Theory and Communications, University of Vigo, 36310 Vigo, Spain {jschmidt, dromero, valcarce}@gts.uvigo.es.
Work supported by the Spanish Government, ERDF funds (TEC2013-47020-C2-1-R COMPASS, CONSOLIDER-INGENIO 2010 CSD2008-00010 COMONSENS), FPU Grant AP2010-0149, and the Galician Regional Government (CN 2012/260 AtlantTIC).

[1] If $\mathcal{B}$ is outside the transmission bandwidth, as would be the case for out-of-band radiation minimization, then $N_P = 0$, as no system subcarriers overlap with the target band.

$N_P + N_C$ subcarriers for cancellation, whereas the remaining $N_D = N - N_A$ subcarriers are unaffected and used for data transmission (in practice, $N_A \ll N_D$). The vector modulating the system subcarriers for the m-th OFDM symbol is then

$$\boldsymbol{x}_m = [\ x_0^{(m)} \quad x_1^{(m)} \quad \cdots \quad x_{N-1}^{(m)}\ ]^T = \boldsymbol{S}\boldsymbol{d}_m + \boldsymbol{T}\boldsymbol{c}_m, \quad (1)$$

with $\boldsymbol{d}_m \in \mathbb{C}^{N_D}$ the data vector, and $\boldsymbol{c}_m \in \mathbb{C}^{N_A}$ comprising the cancellation coefficients. Matrices $\boldsymbol{S} \in \mathbb{C}^{N \times N_D}$, $\boldsymbol{T} \in \mathbb{C}^{N \times N_A}$ contain different columns of $\boldsymbol{I}_N$, and respectively map the elements of $\boldsymbol{d}_m$ and $\boldsymbol{c}_m$ to their corresponding subcarriers. It is assumed that $\{\boldsymbol{d}_m\}$ is a zero-mean i.i.d. process with $E\{\boldsymbol{d}_m \boldsymbol{d}_{m-l}^H\} = \boldsymbol{I}_{N_D}$ if $l = 0$, and $\boldsymbol{0}$ otherwise[2].

As described in [11], the discrete-time baseband signal $s[n]$ is generated by means of IDFT modulation, pulse shaping and guard interval insertion as follows (note that n, k and m denote respectively the discrete time index, the discrete frequency index, and the symbol index):

$$s[n] = \sum_{m=-\infty}^{\infty} \sum_{k=0}^{N-1} x_k^{(m)} h[n - mL] e^{j\frac{2\pi}{N}k(n-mL)}, \quad (2)$$

with L the symbol length in samples and $h[n]$ the shaping pulse. The analog signal $s(t)$ is obtained by D/A conversion with an interpolation filter $g(t)$ and sampling period T_s:

$$s(t) = \sum_{n=-\infty}^{\infty} s[n] g(t - nT_s). \quad (3)$$

It is assumed that the cancellation coefficients $\boldsymbol{c}_m$ are obtained by means of some (possibly nonlinear) mapping applied to $\boldsymbol{d}_m$, i.e., $\boldsymbol{c}_m = \mathcal{G}\{\boldsymbol{d}_m\}$. As a result, $\boldsymbol{x}_m$ and $\boldsymbol{x}_n$ are statistically independent for $n \neq m$. Let $\boldsymbol{\mu} \triangleq E[\boldsymbol{c}_m]$, so that we can write $\boldsymbol{c}_m = \tilde{\boldsymbol{c}}_m + \boldsymbol{\mu}$, where $\tilde{\boldsymbol{c}}_m$ has zero mean; and let $\tilde{\boldsymbol{x}}_m = \boldsymbol{S}\boldsymbol{d}_m + \boldsymbol{T}\tilde{\boldsymbol{c}}_m$, so that $\boldsymbol{x}_m = \tilde{\boldsymbol{x}}_m + \boldsymbol{T}\boldsymbol{\mu}$. Then, following analogous steps to those in [11] and [12], one finds that the psd of $s(t)$ is given by

$$\begin{aligned}\mathsf{S}_s(f) &= \boldsymbol{\phi}^H(f) E[\tilde{\boldsymbol{x}}_m \tilde{\boldsymbol{x}}_m^H] \boldsymbol{\phi}(f) \\ &\quad + \frac{|\boldsymbol{\phi}^H(f)\boldsymbol{T}\boldsymbol{\mu}|^2}{LT_s} \sum_{k=-\infty}^{\infty} \delta\left(f - \tfrac{k}{LT_s}\right), \end{aligned} \quad (4)$$

where we have introduced

$$\boldsymbol{\phi}(f) \triangleq [\ \phi_0(f) \quad \phi_1(f) \quad \cdots \quad \phi_{N-1}(f)\ ]^T, \quad (5)$$

$$\phi_k(f) \triangleq \frac{G(f)}{T_s} H(e^{j2\pi(f-k\Delta_f)T_s}), \quad (6)$$

with $\Delta_f = \frac{1}{NT_s}$ the subcarrier spacing, and where $H(e^{j\omega})$ and $G(f)$ denote the Fourier transforms of $h[n]$ and $g(t)$ respectively. Note that the type of time guard interval employed (cyclic-prefix or zero-padding) affects the particular pulse $h[n]$ [11], but nevertheless (4)-(6) remain valid in both cases.

The problem can be stated as finding the mapping $\mathcal{G}$ such that the power radiated in the protected band $\mathcal{B}$, given by

$$P_{\mathcal{B}} \triangleq \int_{\mathcal{B}} \mathsf{S}_s(f) \mathrm{d}f, \quad (7)$$

[2]By adopting a prewhitening step, the results in this letter readily apply to the case $E\{\boldsymbol{d}_m \boldsymbol{d}_m^H\} \neq \boldsymbol{I}_{N_D}$.

is minimized, subject to adequate design constraints. A reasonable approach is to constrain the total transmit power:

$$\min_{\mathcal{G}} P_{\mathcal{B}} \quad \text{s.t.} \int_{-\infty}^{\infty} \mathsf{S}_s(f)\mathrm{d}f \le P_{\max}. \quad (8)$$

It is readily seen from (4) that the optimum value of the mean $\boldsymbol{\mu}$ for Problem (8) is $\boldsymbol{\mu} = \boldsymbol{0}$, since this choice simultaneously minimizes the objective $P_{\mathcal{B}}$ and the constraint. Thus, hereafter we will assume that $\boldsymbol{c}_m = \tilde{\boldsymbol{c}}_m$ is a zero-mean process.

From (1) and (4), and defining $\boldsymbol{\Phi}(f) \triangleq \boldsymbol{\phi}(f)\boldsymbol{\phi}^H(f)$, the psd of the OFDM signal (with $\boldsymbol{\mu} = \boldsymbol{0}$) can be written as

$$\begin{aligned}\mathsf{S}_s(f) &= \mathsf{S}_0(f) + \mathrm{Tr}\left\{\boldsymbol{T}^T\boldsymbol{\Phi}(f)\boldsymbol{T}E\left[\boldsymbol{c}_m\boldsymbol{c}_m^H\right]\right\} \\ &\quad + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}(f)\boldsymbol{T}E\left[\boldsymbol{c}_m\boldsymbol{d}_m^H\right]\right\}, \end{aligned} \quad (9)$$

where $\mathsf{S}_0(f) \triangleq \mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}(f)\boldsymbol{S}\right\}$ is the psd obtained by making $\boldsymbol{c}_m = \boldsymbol{0}\ \forall m$, i.e., by turning off the reserved subcarriers.

If $\mathcal{G}$ is constrained to the class of linear mappings, then $\boldsymbol{c}_m = \boldsymbol{\Theta}\boldsymbol{d}_m$ for some fixed $\boldsymbol{\Theta} \in \mathbb{C}^{N_A \times N_D}$, which becomes the optimization parameter. In that case, (1) becomes $\boldsymbol{x}_m = (\boldsymbol{S} + \boldsymbol{T}\boldsymbol{\Theta})\boldsymbol{d}_m$, and the psd (9) simplifies to

$$\begin{aligned}\mathsf{S}_s(f) &= \mathsf{S}_0(f) + \mathrm{Tr}\left\{\boldsymbol{\Theta}^H\boldsymbol{T}^T\boldsymbol{\Phi}(f)\boldsymbol{T}\boldsymbol{\Theta}\right\} \\ &\quad + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}(f)\boldsymbol{T}\boldsymbol{\Theta}\right\}. \end{aligned} \quad (10)$$

With this choice, (8) becomes a convex problem in $\boldsymbol{\Theta}$, which can be easily solved by means of the generalized singular value decomposition [10], [13]. Note that the problem (and hence its solution) depends only on system parameters, and not on the specific values taken by the data vector $\boldsymbol{d}_m$.

III. Optimality of Linear PSD-Based AIC

Consider now a general (possibly nonlinear) mapping $\mathcal{G}$. We drop the symbol index m since it does not play any role in the development. Let $r_W \triangleq \mathrm{rank}\, E[\boldsymbol{c}\boldsymbol{c}^H] \le N_A$, and consider a factorization $E[\boldsymbol{c}\boldsymbol{c}^H] = \boldsymbol{W}\boldsymbol{W}^H$ where $\boldsymbol{W} \in \mathbb{C}^{N_A \times r_W}$ has full column rank r_W. Let us introduce the random vector $\boldsymbol{z} = \boldsymbol{W}^\dagger \boldsymbol{c}$, with $(\cdot)^\dagger$ denoting the pseudoinverse. Then $E[\boldsymbol{z}] = \boldsymbol{0}$ and $E[\boldsymbol{z}\boldsymbol{z}^H] = \boldsymbol{I}_{r_W}$; moreover, one has the following.

Lemma 1: It holds that $\boldsymbol{c} = \boldsymbol{W}\boldsymbol{z}$.

Proof: Consider a full column rank matrix $\boldsymbol{P} \in \mathbb{C}^{N_A \times (N_A - r_W)}$ with columns orthogonal to those of $\boldsymbol{W}$, i.e., $\boldsymbol{P}^H\boldsymbol{W} = \boldsymbol{0}$. Then $\boldsymbol{A} \triangleq [\,\boldsymbol{W}\ \boldsymbol{P}\,]$ is invertible with

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{W}^\dagger \\ \boldsymbol{P}^\dagger \end{bmatrix}. \quad (11)$$

Let now

$$\boldsymbol{y} = \boldsymbol{A}^{-1}\boldsymbol{c} = \begin{bmatrix} \boldsymbol{W}^\dagger\boldsymbol{c} \\ \boldsymbol{P}^\dagger\boldsymbol{c} \end{bmatrix} = \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{P}^\dagger\boldsymbol{c} \end{bmatrix}, \quad (12)$$

and note that $E[(\boldsymbol{P}^\dagger\boldsymbol{c})(\boldsymbol{P}^\dagger\boldsymbol{c})^H] = (\boldsymbol{P}^\dagger\boldsymbol{W})(\boldsymbol{P}^\dagger\boldsymbol{W})^H = \boldsymbol{0}$, so that we must have $\boldsymbol{P}^\dagger\boldsymbol{c} = \boldsymbol{0}$. Therefore,

$$\boldsymbol{c} = \boldsymbol{A}\boldsymbol{y} = [\,\boldsymbol{W}\ \boldsymbol{P}\,] \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{0} \end{bmatrix} = \boldsymbol{W}\boldsymbol{z}, \quad (13)$$

as was to be shown. ∎

Therefore, the cross-covariance matrix of cancellation and data coefficients is given by

$$E[\boldsymbol{c}\boldsymbol{d}^H] = \boldsymbol{W}E[\boldsymbol{z}\boldsymbol{d}^H] = \boldsymbol{W}\boldsymbol{Q} \quad \text{with} \quad \boldsymbol{Q} \triangleq E[\boldsymbol{z}\boldsymbol{d}^H], \quad (14)$$

so that the psd in (9) can be expressed as

$$\mathsf{S}_s(f) = \mathsf{S}_0(f) + \mathrm{Tr}\left\{\boldsymbol{W}^H\boldsymbol{T}^T\boldsymbol{\Phi}(f)\boldsymbol{T}\boldsymbol{W}\right\} + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}(f)\boldsymbol{T}\boldsymbol{W}\boldsymbol{Q}\right\}. \quad (15)$$

Thus, our free parameters are $\boldsymbol{W}$ and $\boldsymbol{Q}$. Note that whereas $\boldsymbol{W}$ is arbitrary, this is not the case for $\boldsymbol{Q}$, which is the cross-covariance of two spatially white random vectors. The following lemma shows a key property of these matrices:

Lemma 2: Let $\boldsymbol{z} \in \mathbb{C}^p$ and $\boldsymbol{d} \in \mathbb{C}^q$ be zero-mean with $E\left[\boldsymbol{z}\boldsymbol{z}^H\right] = \boldsymbol{I}_p$ and $E\left[\boldsymbol{d}\boldsymbol{d}^H\right] = \boldsymbol{I}_q$, and let $\boldsymbol{Q} = E\left[\boldsymbol{z}\boldsymbol{d}^H\right]$. Then $\boldsymbol{I}_p - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}$ and $\boldsymbol{I}_q - \boldsymbol{Q}^H\boldsymbol{Q} \geq \boldsymbol{0}$, or equivalently, all singular values of $\boldsymbol{Q}$ belong in $[0, 1]$.

Proof: The linear MMSE estimate of $\boldsymbol{z}$ based on $\boldsymbol{d}$ is $\hat{\boldsymbol{z}} = \boldsymbol{Q}\boldsymbol{d}$, and the estimation error $\boldsymbol{e} = \boldsymbol{z} - \hat{\boldsymbol{z}}$ has covariance matrix $E\left[\boldsymbol{e}\boldsymbol{e}^H\right] = \boldsymbol{I}_p - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}$. Reversing the roles of $\boldsymbol{z}$ and $\boldsymbol{d}$ yields $\boldsymbol{I}_q - \boldsymbol{Q}^H\boldsymbol{Q} \geq \boldsymbol{0}$. Noting that the positive singular values of $\boldsymbol{Q}$ are the square roots of the nonzero eigenvalues of $\boldsymbol{Q}\boldsymbol{Q}^H$ (or of $\boldsymbol{Q}^H\boldsymbol{Q}$) concludes the proof. ■

Our approach is to replace Problem (8) by

$$\min_{\boldsymbol{W},\boldsymbol{Q}} P_{\mathcal{B}} \quad \text{s.t.} \quad \begin{cases} \boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}, \\ \int_{-\infty}^{\infty} \mathsf{S}_s(f)\mathrm{d}f \leq P_{\max}. \end{cases} \quad (16)$$

Note that, in principle, the feasible set of Problem (16) may be larger than that of Problem (8), since the converse of Lemma 2 has not been proved; in other words, it has not been shown whether given $\boldsymbol{Q}$ with $\boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}$, a zero mean random vector $\boldsymbol{z}$ with covariance $\boldsymbol{I}_{r_W}$ can always be found in order to meet $E[\boldsymbol{z}\boldsymbol{d}^H] = \boldsymbol{Q}$. Nevertheless, we will show that for the matrix $\boldsymbol{Q}$ solving Problem (16), such random vector can be obtained. Specifically, such optimal $\boldsymbol{Q}$ happens to have orthonormal rows, i.e., $\boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H = \boldsymbol{0}$, so that the random vector $\boldsymbol{z}$ can be constructed as $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{d}$, since in that case it holds that $E[\boldsymbol{z}] = \boldsymbol{0}$, $E[\boldsymbol{z}\boldsymbol{z}^H] = \boldsymbol{Q}\boldsymbol{Q}^H = \boldsymbol{I}_{r_W}$, and $E[\boldsymbol{z}\boldsymbol{d}^H] = \boldsymbol{Q}$, as desired. Therefore, it follows that the corresponding solution is also the solution to the original Problem (8).

As a first step, in Sec. III-A we consider a simpler version of Problem (16) in which the total power constraint is removed, and show that for this "unconstrained" problem, linear processing is optimal. The proof will be extended in Sec. III-B in order to take the total power constraint into account.

A. Psd-based AIC solution without total power constraint

Consider then the minimization of $P_{\mathcal{B}}$ subject only to $\boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}$. From (15), letting $\boldsymbol{\Phi}_{\mathcal{B}} \triangleq \int_{\mathcal{B}} \boldsymbol{\Phi}(f)\mathrm{d}f$ and noting that $\mathsf{S}_0(f)$ does not depend on $\boldsymbol{W}$, $\boldsymbol{Q}$, such problem is equivalent to

$$\min_{\boldsymbol{W},\boldsymbol{Q}} \quad \mathrm{Tr}\left\{\boldsymbol{W}^H\boldsymbol{T}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}\right\} + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}\boldsymbol{Q}\right\} \quad \text{s.t.} \quad \boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}. \quad (17)$$

Let us focus on the minimization w.r.t. $\boldsymbol{Q}$ for a given $\boldsymbol{W}$. Defining $\boldsymbol{F}^H \triangleq \boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}$, this amounts to solving

$$\min_{\boldsymbol{Q}} 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{F}^H\boldsymbol{Q}\right\} \quad \text{s.t.} \quad \boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}. \quad (18)$$

The solution to this problem is given by the following result.

Lemma 3: Let $\boldsymbol{F} \in \mathbb{C}^{r_W \times N_D}$, $r_F = \mathrm{rank}\,\boldsymbol{F} \leq r_W$, and consider the SVD $\boldsymbol{F} = \boldsymbol{U}_F\boldsymbol{\Sigma}_F\boldsymbol{V}_F^H$ where $\boldsymbol{\Sigma}_F$ is $r_F \times r_F$ diagonal with the nonzero singular values. Then the minimum value of the objective in Problem (18) is $-2\|\boldsymbol{F}\|_*$, where $\|\boldsymbol{F}\|_* = \mathrm{Tr}\,\boldsymbol{\Sigma}_F$ denotes the nuclear norm of $\boldsymbol{F}$. This minimum is attained by any $\boldsymbol{Q}$ of the form

$$\boldsymbol{Q} = -\boldsymbol{U}_F\boldsymbol{V}_F^H + \tilde{\boldsymbol{U}}_F\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{V}}_F^H, \quad (19)$$

where $\tilde{\boldsymbol{U}}_F$, $\tilde{\boldsymbol{V}}_F$ are such that $[\,\boldsymbol{U}_F\ \tilde{\boldsymbol{U}}_F\,] \in \mathbb{C}^{r_W \times r_W}$ is unitary, $[\,\boldsymbol{V}_F\ \tilde{\boldsymbol{V}}_F\,] \in \mathbb{C}^{N_D \times r_W}$ has orthonormal columns, and $\tilde{\boldsymbol{Q}} \in \mathbb{C}^{r_\Delta \times r_\Delta}$ is such that $\boldsymbol{I}_{r_\Delta} - \tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H \geq \boldsymbol{0}$, where $r_\Delta = r_W - r_F$.

Proof: Von Neumann's trace inequality [14] states that given two matrices $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{C}^{m\times n}$ with singular values $\sigma_1(\boldsymbol{A}) \geq \sigma_2(\boldsymbol{A}) \geq \cdots \geq 0$ and $\sigma_1(\boldsymbol{B}) \geq \sigma_2(\boldsymbol{B}) \geq \cdots \geq 0$, then

$$\mathrm{Re}\,\mathrm{Tr}\{\boldsymbol{A}^H\boldsymbol{B}\} \leq \sum_{k=1}^{\min\{m,n\}} \sigma_k(\boldsymbol{A})\sigma_k(\boldsymbol{B}), \quad (20)$$

with equality holding iff there exist SVDs of $\boldsymbol{A}$ and $\boldsymbol{B}$ with common left- and right-singular vectors, i.e., iff there exist unitary matrices $\boldsymbol{U} \in \mathbb{C}^{m\times m}$, $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ such that

$$\boldsymbol{A} = \boldsymbol{U}\,\mathrm{diag}\{\,\sigma_1(\boldsymbol{A})\ \sigma_2(\boldsymbol{A}) \cdots\}\boldsymbol{V}^H, \quad (21)$$

$$\boldsymbol{B} = \boldsymbol{U}\,\mathrm{diag}\{\,\sigma_1(\boldsymbol{B})\ \sigma_2(\boldsymbol{B}) \cdots\}\boldsymbol{V}^H. \quad (22)$$

Applying (20) to $\boldsymbol{A} = -\boldsymbol{F}$, $\boldsymbol{B} = \boldsymbol{Q}$ yields

$$2\,\mathrm{Re}\,\mathrm{Tr}\{\boldsymbol{F}^H\boldsymbol{Q}\} \geq -2\sum_{k=1}^{r_W} \sigma_k(\boldsymbol{F})\sigma_k(\boldsymbol{Q}), \quad (23)$$

since the singular values of $-\boldsymbol{F}$ are those of $\boldsymbol{F}$. The right-hand side of (23) is minimized w.r.t $\sigma_k(\boldsymbol{Q}) \in [0, 1]$ if $\sigma_k(\boldsymbol{Q}) = 1$ for $k = 1, \ldots, r_W$, yielding $-2\|\boldsymbol{F}\|_*$. In addition, in view of (21), the bound is attained iff $\boldsymbol{Q}$ is of the form (19). ■

The key fact from Lemma 3 is that we can take $\tilde{\boldsymbol{Q}} = \boldsymbol{I}_{r_\Delta}$, so that for every $\boldsymbol{W}$ there exists an optimal $\boldsymbol{Q}_\star(\boldsymbol{W}) = -\boldsymbol{U}_F\boldsymbol{V}_F^H + \tilde{\boldsymbol{U}}_F\tilde{\boldsymbol{V}}_F^H$ with orthonormal rows: $\boldsymbol{Q}_\star(\boldsymbol{W})\boldsymbol{Q}_\star^H(\boldsymbol{W}) = \boldsymbol{I}_{r_W}$. This allows us to rewrite (17) as

$$\min_{\boldsymbol{W},\boldsymbol{Q}} \mathrm{Tr}\left\{\boldsymbol{Q}^H\boldsymbol{W}^H\boldsymbol{T}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}\boldsymbol{Q}\right\} + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}\boldsymbol{Q}\right\} \quad \text{s.t.} \quad \boldsymbol{Q}\boldsymbol{Q}^H = \boldsymbol{I}_{r_W}. \quad (24)$$

Note that the objective function in (24) depends only on the product $\boldsymbol{W}\boldsymbol{Q}$. As any matrix $\boldsymbol{\Theta} \in \mathbb{C}^{N_A \times N_D}$ with $\mathrm{rank}\,\boldsymbol{\Theta} = r_W \leq N_A$ can be written as[3] $\boldsymbol{\Theta} = \boldsymbol{W}\boldsymbol{Q}$ with $\boldsymbol{W} \in \mathbb{C}^{N_A \times r_W}$, $\boldsymbol{Q} \in \mathbb{C}^{r_W \times N_D}$ with $\boldsymbol{Q}\boldsymbol{Q}^H = \boldsymbol{I}_{r_W}$, then (24) is equivalent to

$$\min_{\mathrm{rank}\,\boldsymbol{\Theta} = r_W} \quad \mathrm{Tr}\left\{\boldsymbol{\Theta}^H\boldsymbol{T}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{\Theta}\right\} + 2\,\mathrm{Re}\,\mathrm{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{\Theta}\right\}. \quad (25)$$

The optimal solution corresponds to the integer $r_W \in \{1, 2, \ldots, N_A\}$ yielding the smallest value of the minimized cost in (25). Clearly, this is the problem obtained when the cancellation coefficients are linearly related to the data (see (10)), i.e., $\boldsymbol{c} = \boldsymbol{\Theta}\boldsymbol{d}$, with no rank constraint on $\boldsymbol{\Theta}$. Thus, it is concluded that *linear processing is optimal for this problem.*

[3]To see this, consider the SVD $\boldsymbol{\Theta} = \boldsymbol{U}_\Theta\boldsymbol{\Sigma}_\Theta\boldsymbol{V}_\Theta^H$, and take $\boldsymbol{W} = \boldsymbol{U}_\Theta\boldsymbol{\Sigma}_\Theta\boldsymbol{Z}^H$ and $\boldsymbol{Q} = \boldsymbol{Z}\boldsymbol{V}_\Theta^H$, where $\boldsymbol{Z}$ is an arbitrary $r_W \times r_W$ unitary matrix.

B. Psd-based AIC with total power constraint

Consider now Problem (16), focusing again on minimization w.r.t. $\boldsymbol{Q}$ for $\boldsymbol{W}$ fixed. With $\boldsymbol{F}^H \triangleq \boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{B}}\boldsymbol{T}\boldsymbol{W}$ as before, let

$$\boldsymbol{\Phi}_{\mathcal{T}} \triangleq \int_{-\infty}^{\infty} \boldsymbol{\Phi}(f)\mathrm{d}f, \qquad \boldsymbol{G}^H \triangleq \boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{T}}\boldsymbol{T}\boldsymbol{W}, \tag{26}$$

$$\bar{P} \triangleq P_{\text{max}} - \operatorname{Tr}\left\{\boldsymbol{S}^T\boldsymbol{\Phi}_{\mathcal{T}}\boldsymbol{S}\right\} - \operatorname{Tr}\left\{\boldsymbol{W}^H\boldsymbol{T}^T\boldsymbol{\Phi}_{\mathcal{T}}\boldsymbol{T}\boldsymbol{W}\right\}. \tag{27}$$

Then, it is seen that the optimal $\boldsymbol{Q}$ is the solution to

$$\min_{\boldsymbol{Q}} 2\operatorname{Re}\operatorname{Tr}\left\{\boldsymbol{F}^H\boldsymbol{Q}\right\} \text{ s.t. } \begin{cases} \boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}, \\ 2\operatorname{Re}\operatorname{Tr}\left\{\boldsymbol{G}^H\boldsymbol{Q}\right\} \leq \bar{P}. \end{cases} \tag{28}$$

Problem (28) has a linear objective, a linear inequality constraint, and a convex quadratic inequality constraint, so it is convex. From the discussion in Sec. III-A for the problem without total power constraint, if we can show that the optimal $\boldsymbol{Q}$ solving (28) has orthonormal rows, then we can conclude that linear processing is also optimal for this constrained case.

To this end, and since $\boldsymbol{F}$, $\boldsymbol{G}$ and $\bar{P}$ all depend on $\boldsymbol{W}$, we shall discuss the following three cases. First, let $r_G = \operatorname{rank}\boldsymbol{G}$, and consider the SVD

$$\boldsymbol{G} = \boldsymbol{U}_G\boldsymbol{\Sigma}_G\boldsymbol{V}_G^H, \tag{29}$$

together with $\tilde{\boldsymbol{U}}_G$, $\tilde{\boldsymbol{V}}_G$ such that $[\boldsymbol{U}_G\ \tilde{\boldsymbol{U}}_G] \in \mathbb{C}^{r_W\times r_W}$ is unitary and $[\boldsymbol{V}_G\ \tilde{\boldsymbol{V}}_G] \in \mathbb{C}^{N_D\times r_W}$ has orthonormal columns.

Case 1: $\bar{P} < -2\|\boldsymbol{G}\|_*$. In that case, it follows from Lemma 3 that no $\boldsymbol{Q}$ simultaneously satisfies both constraints in (28), so that the $\boldsymbol{W}$ defining Problem (28) is infeasible.

Case 2: $\bar{P} > -2\|\boldsymbol{G}\|_*$. From Lemma 3, it is seen that the matrix $\tilde{\boldsymbol{Q}} = -\boldsymbol{U}_G\boldsymbol{V}_G^H + \tilde{\boldsymbol{U}}_G\tilde{\boldsymbol{V}}_G^H$ satisfies the first constraint in (28) with equality, and the second constraint with strict inequality. Then there exists $\alpha < 1$ such that $\boldsymbol{Q} = \alpha\tilde{\boldsymbol{Q}}$ satisfies *both* constraints with strict inequality. Therefore, Slater's condition is satisfied, so that strong duality holds [15]. In that case (28) is equivalent to

$$\begin{aligned} \min_{\boldsymbol{Q}} \mathcal{L}(\boldsymbol{Q},\lambda_\star) &= 2\operatorname{Re}\left[\operatorname{Tr}\left\{(\boldsymbol{F}^H - \lambda_\star\boldsymbol{G}^H)\boldsymbol{Q}\right\}\right] - \lambda_\star\bar{P} \\ \text{s.t.} \qquad & \boldsymbol{I}_{r_W} - \boldsymbol{Q}\boldsymbol{Q}^H \geq \boldsymbol{0}, \end{aligned} \tag{30}$$

where $\mathcal{L}(\boldsymbol{Q},\lambda)$ is the Lagrangian and $\lambda_\star$ is the Lagrange multiplier maximizing the dual function [15]. As the term $-\lambda_\star\bar{P}$ does not depend on $\boldsymbol{Q}$, we can apply Lemma 3 to conclude that there exists a matrix $\boldsymbol{Q}_\star$ with orthonormal rows which solves (30), and therefore (28) as well.

Case 3: $\bar{P} = -2\|\boldsymbol{G}\|_*$. By virtue of Lemma 3, the only feasible values of $\boldsymbol{Q}$ are given by

$$\boldsymbol{Q} = -\boldsymbol{U}_G\boldsymbol{V}_G^H + \tilde{\boldsymbol{U}}_G\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{V}}_G^H, \tag{31}$$

with $\tilde{\boldsymbol{Q}} \in \mathbb{C}^{r'_\Delta\times r'_\Delta}$ such that $\boldsymbol{I}_{r'_\Delta} - \tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H \geq \boldsymbol{0}$, where now $r'_\Delta = r_W - r_G$. Thus, if $r'_\Delta = 0$, then the only feasible $\boldsymbol{Q}$ is $\boldsymbol{Q} = -\boldsymbol{U}_G\boldsymbol{V}_G^H$, which has orthonormal rows. On the other hand, if $r'_\Delta > 0$, then the only free parameter is $\tilde{\boldsymbol{Q}}$. Substituting (31) in the objective of Problem (28) and making $\tilde{\boldsymbol{F}}^H \triangleq \tilde{\boldsymbol{V}}_G^H\boldsymbol{F}^H\tilde{\boldsymbol{U}}_G$ results in

$$\min_{\tilde{\boldsymbol{Q}}} 2\operatorname{Re}\operatorname{Tr}\left\{\tilde{\boldsymbol{F}}^H\tilde{\boldsymbol{Q}}\right\} \text{ s.t. } \boldsymbol{I}_{r'_\Delta} - \tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^H \geq \boldsymbol{0}. \tag{32}$$

Hence, applying again Lemma 3, it is concluded that it is possible to find an optimal $\tilde{\boldsymbol{Q}}_\star$ with $\tilde{\boldsymbol{Q}}_\star\tilde{\boldsymbol{Q}}_\star^H = \boldsymbol{I}_{r'_\Delta}$. Substituting its value back in (31) we obtain $\boldsymbol{Q}_\star = -\boldsymbol{U}_G\boldsymbol{V}_G^H + \tilde{\boldsymbol{U}}_G\tilde{\boldsymbol{Q}}_\star\tilde{\boldsymbol{V}}_G^H$, which again has orthonormal rows and solves (28).

The above considerations imply that for every feasible $\boldsymbol{W}$ it is possible to pick an optimum $\boldsymbol{Q}_\star(\boldsymbol{W})$ with orthonormal rows. Therefore, by the same argument as in Sec. III-A, under a total power constraint, linear processing remains optimal.

IV. Conclusions

It has been shown that, regardless of the data distribution, there is no loss of optimality by constraining the relation between data and cancellation subcarriers to the class of linear mappings with a constant coefficient matrix. This is true with and without a total transmit power constraint. Thus, the low complexity of online implementation associated to this type of mappings comes with no penalty in terms of performance.

In practice, the solution to (8) may allocate too much power to cancellation subcarriers, resulting in spectrum overshoot. As shown in [16], this phenomenon can be tackled by introducing an additional parameter α controlling the fraction of total power allocated to data and cancellation subcarriers: the result of this letter is also applicable to such α-AIC design from [16].

References

[1] G. Berardinelli, K. Pajukoski, *et al.*, "On the potential of OFDM enhancements as 5G waveforms," in *IEEE Vehicular Technol. Conf. (VTC Spring)*, May 2014.
[2] J. G. Andrews, S. Buzzi, *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, 2014, in press.
[3] H. Arslan, H. A. Mahmoud, and T. Yucek, "OFDM for cognitive radio: merits and challenges," in *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems*, H. Arslan, Ed. Springer, 2007.
[4] H. Yamaguchi, "Active interference cancellation technique for MB-OFDM cognitive radio," in *34th European Microw. Conf.*, vol. 2, Oct. 2004, pp. 1105–08.
[5] S. Brandes, I. Cosovic, and M. Schnell, "Reduction of out-of-band radiation in OFDM systems by insertion of cancellation carriers," *IEEE Commun. Letters*, vol. 10, no. 6, pp. 420–422, Jun. 2006.
[6] J. Balakrishnan, "Dummy tone insertion for spectral sculpting of the multi-band OFDM UWB system," in *Int. Conf. Commun. Syst. Networks (COMSNETS)*, Jan. 2009.
[7] E. H. M. Alian and P. Mitran, "Improved active interference cancellation for sidelobe suppression in cognitive OFDM systems," in *IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2012, pp. 1460–65.
[8] E. H. M. Alian, H. E. Saffar, and P. Mitran, "Cross-band interference reduction trade-offs in SISO and MISO OFDM-based cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2436–45, Jul. 2012.
[9] P. Kryszkiewicz and H. Bogucka, "Out-of-band power reduction in NC-OFDM with optimized cancellation carriers selection," *IEEE Commun. Letters*, vol. 17, no. 10, pp. 1901–1904, 2013.
[10] J. F. Schmidt, S. Costas-Sanz, and R. López-Valcarce, "Choose your subcarriers wisely: Active interference cancellation for cognitive OFDM," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 3, no. 4, pp. 615–625, Dec. 2013.
[11] T. van Waterschoot, V. Le Nir, J. Duplicy, and M. Moonen, "Analytical expressions for the power spectral density of CP-OFDM and ZP-OFDM signals," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 371–374, Apr. 2010.
[12] J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2nd ed. Prentice-Hall, 2001.
[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, Oct. 1996.
[14] J. von Neumann, "Some matrix inequalities and metrization of matrix-space," *Tomsk Univ. Rev.*, vol. 1, pp. 286–300, 1937, reprinted in Collected Works (Pergamon Press, 1962), IV, pp. 205-219.
[15] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.
[16] J. F. Schmidt and R. López-Valcarce, "OFDM spectrum sculpting with active interference cancellation: keeping spectral spurs at bay," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May. 2014.