Fifty-first Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 2 - 3, 2013

Optimal Jamming Over Additive Noise: Vector Source-Channel Case

Emrah Akyol and Kenneth Rose

***Abstract*— This paper extends the recent analysis of optimal zero-delay jamming problem, specialized to scalar sources over scalar additive noise channels, to vector spaces. Particularly, we find the saddle point solution to the jamming problem in "vector spaces", within its most general, non-Gaussian setting; the problem has been open even for the important special case of Gaussian source and channel. Similar to the scalar setting, linearity conditions for encoding and decoding mappings play a pivotal role in jamming, in the sense that the optimal jamming strategy is to effectively force both transmitter and receiver to default to linear mappings, i.e., the jammer ensures, whenever possible, that the transmitter and receiver cannot benefit from non-linear strategies. The optimal strategy is then "randomized" linear encoding for transmitter and generating independent noise for the jamming. Moreover, the optimal jammer allocates power according to the well-known "water-filling" over the eigenvalues of the channel noise, and the density of the jamming noise that render the optimal encoding and decoding mappings linear, is determined by the power constraints and the source density.**

I. Introduction

We consider the problem of optimal jamming, by a power constrained agent, over additive noise channel, in vector spaces. This problem was solved in [1], [2] for scalar Gaussian source and scalar channel. The saddle point solution to this zero-sum game, derived for Gaussian source-channel pair, involves randomized linear mapping for the transmitter and generating independent, Gaussian noise as the jammer output and a linear decoder. We recently extended this work to non-Gaussian scalar sources and scalar channels [3], and showed that the linearity of encoding and decoding mappings is essential while Gaussianity of the jammer is merely to satisfy the linearity conditions for the Gaussian source and channel. In this paper, by leveraging the recent results on conditions for linearity of optimal estimation and communication mappings, [4], [5], we extend the "scalar" analysis to vector sources and channels. The contributions of this paper are:

- We derive the necessary and sufficient condition (called the "vector matching condition") on the jamming noise density to ensure linearity of the optimal transmitter and the receiver, within the vector setting. The condition is much more involved than the scalar one, due to dependencies among the source and channel components. The jammer aims to render both encoder and decoder linear mappings, while allocates jamming power over the channel eigenvalues by the well known waterfilling solution.
- We discover an analogy between zero-delay vector jamming and vector rate-distortion and vector channel capacity, in terms of power/rate/jamming power allocation as waterfilling.

This work is supported in part by the NSF under grants CCF-1016861 and CCF-1118075.

Emrah Akyol and Kenneth Rose are with Department of Electrical and Computer Engineering, University of California at Santa Barbara, CA 93106, USA {eakyol, rose}@ece.ucsb.edu

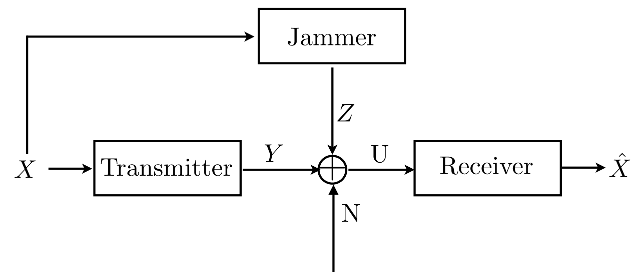

Fig. 1. The jamming problem.

This paper is organized as follows. In Section II, we present the problem definition and preliminaries. In Section III, we review the prior results related to jamming, estimation and communication problems. In Section IV, we derive the linearity result and then the main result on optimal jamming. Finally, we discuss the future directions in Section V.

II. Problem Definition

Let $\mathbb{R}$ and $\mathbb{R}_+$ denote the respective sets of real numbers and positive real numbers. Let $\mathbb{E}(\cdot)$, $\mathbb{P}(\cdot)$ and $*$ denote the expectation, probability and convolution operators, respectively. Let $Bern(p)$ denote the Bernoulli random variable, taking values in $\{-1, 1\}$ with probability $\{p, 1-p\}$. The Gaussian density with mean μ and variance σ^2 is denoted as $\mathcal{N}(\mu, \sigma^2)$. Let $f^{'}(x) = \frac{df(x)}{dx}$ denote the first order derivative of the function $f(\cdot)$. All logarithms in the paper are natural logarithms and may in general be complex, and the integrals are, in general, Lebesgue integrals. Let us define the set $\mathcal{S}$ as the set of Borel measurable $\mathbb{R}^m \to \mathbb{R}^k$ mappings.

In general, lowercase letters (e.g., c) denote scalars, boldface lowercase (e.g., $\boldsymbol{x}$) vectors, uppercase (e.g., C, X) matrices and random variables. I denotes the identity matrix. R_X, and R_{XZ} denote the auto-covariance of $\boldsymbol{X}$ and cross covariance of $\boldsymbol{X}$ and $\boldsymbol{Z}$ respectively. A^T denotes the transpose of matrix A. $(x)^+$ denotes the function $\max(0, x)$. ∇ denotes the gradient and ∇_x denotes the partial gradient with respect to $\boldsymbol{x}$.

Throughout this paper, we assume that the source is an m-dimensional vector with zero mean and covariance R_X. The channel noise is additive k-dimensional Gaussian, of zero mean and covariance R_N. Covariance matrices R_X and R_N allow the diagonalization

$$R_X = Q_X \Lambda_X Q_X^T, \text{ and } R_N = Q_N \Lambda_N Q_N^T \quad (1)$$

where $Q_X Q_X^T = Q_N Q_N^T = I$ and Λ_X and Λ_N are diagonal matrices, having ordered eigenvalues as entries, i.e., $\Lambda_X = \mathrm{diag}\{\boldsymbol{\lambda}_X\}$ and $\Lambda_N = \mathrm{diag}\{\boldsymbol{\lambda}_N\}$ where $\boldsymbol{\lambda}_X$ and $\boldsymbol{\lambda}_X$ are ordered (descending) eigenvalues and Q_X^T and Q_N^T are the KLT matrices of the source and the channel, respectively. We will make use of the following auxiliary lemma, see eg. [6] for a proof.

Lemma 1. *Let $\boldsymbol{\lambda}_X$ and $\boldsymbol{\lambda}_N$ be two ordered vectors in $\mathbb{R}_+^m$ with descending entries ($\lambda_X(1) \geq \lambda_X(2), \dots, \lambda_X(m)$) and Π denote any permutation of the indices $\{1, 2, \dots, m\}$, then*

$$\min_{\Pi} \sum_{i=1}^{m} \boldsymbol{\lambda}_X(\Pi(i)) \boldsymbol{\lambda}_N(i) = \sum_{i=1}^{m} \boldsymbol{\lambda}_X(i) \boldsymbol{\lambda}_N(m-i) \quad (2)$$

and

$$\max_{\Pi} \sum_{i=1}^{m} \boldsymbol{\lambda}_X((i)) \boldsymbol{\lambda}_N(\Pi(i)) = \sum_{i=1}^{m} \boldsymbol{\lambda}_X(i) \boldsymbol{\lambda}_N(i) \quad (3)$$

We consider the general communication system whose block diagram is shown in Figure 1. Source $\boldsymbol{X} \in \mathbb{R}^m$ is mapped into $\boldsymbol{Y} \in \mathbb{R}^k$ by function $\boldsymbol{g_T(\cdot)} \in \mathcal{S}$ and transmitted over an additive noise channel. The adversary receives the same signal $\boldsymbol{X}$ and generates the jamming signal $\boldsymbol{Z}$ through function $\boldsymbol{g_A(\cdot)} \in \mathcal{S}$ which is added to the channel output, and aims to compromise the transmission of the source. The received signal $\boldsymbol{U} = \boldsymbol{Y} + \boldsymbol{Z} + \boldsymbol{N}$ is mapped by the decoder to an estimate $\hat{\boldsymbol{X}}$ via function $\boldsymbol{h}(\cdot) \in \mathcal{S}$. The zero mean noise $\boldsymbol{N}$ is assumed to be independent of the source $\boldsymbol{X}$. The source density is denoted $f_X(\cdot)$ and the noise density is $f_N(\cdot)$ with characteristic functions $F_X(\boldsymbol{\omega})$ and $F_N(\boldsymbol{\omega})$, respectively.

The overall cost, measured as the mean squared error (MSE) between source X and its estimate at the decoder $\hat{X}$, is a function of the transmitter, jammer and the receiver mappings:

$$J(\boldsymbol{g}_T(\cdot), \boldsymbol{g}_A(\cdot), \boldsymbol{h}(\cdot)) = \mathbb{E}\{||\boldsymbol{X} - \hat{\boldsymbol{X}}||^2\}. \quad (4)$$

Transmitter $\boldsymbol{g}_T(\cdot) : \mathbb{R}^m \to \mathbb{R}^k$ and receiver $\boldsymbol{h}(\cdot) : \mathbb{R}^m \to \mathbb{R}^k$ seek to minimize this cost while the adversary (jammer) seeks to maximize it by appropriate choice of $\boldsymbol{g}_A(\cdot) : \mathbb{R}^m \to \mathbb{R}^k$. Power constraints must be satisfied by the transmitter

$$\mathbb{E}\{||\boldsymbol{g}_T(\boldsymbol{X})||^2\} \leq P_T, \quad (5)$$

and jammer

$$\mathbb{E}\{||\boldsymbol{g}_A(\boldsymbol{X})||^2\} \leq P_A. \quad (6)$$

The conflict of interest underlying this problem implies that the optimal transmitter-receiver-adversarial policy is the saddle point solution $(g_T^*(\cdot), g_A^*(\cdot), h^*(\cdot))$ satisfying the set of inequalities

$$J(\boldsymbol{g}_T^*, \boldsymbol{g}_A, \boldsymbol{h}^*) \leq J(\boldsymbol{g}_T^*, \boldsymbol{g}_A^*, \boldsymbol{h}^*) \leq J(\boldsymbol{g}_T, \boldsymbol{g}_A^*, \boldsymbol{h}). \quad (7)$$

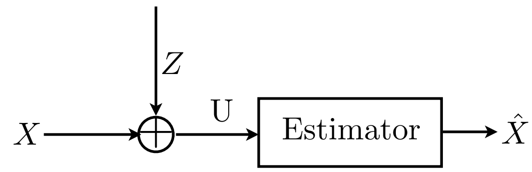

Fig. 2. The estimation problem.

III. Prior Work

The jamming problem, in the form defined here, was studied in [1], [2], for "scalar" Gaussian sources and channels. The problem of interest is intrinsically connected to the fundamental problems of estimation theory and the theory of zero-delay source-channel coding. In particular, conditions for linearity of optimal estimation [4] and optimal mappings in communications [5] are relevant to our problem here. We start with the estimation problem.

A. Estimation Problem

Consider the setting in Figure 2. The estimator receives U, the noisy version of the source X and generates the estimate $\hat{X}$ by the function $h : \mathbb{R} \to \mathbb{R}$ such that MSE, $\mathbb{E}\{(X - \hat{X})^2\}$ is minimized. It is well known that, when a Gaussian source is contaminated with Gaussian noise, a linear estimator minimizes MSE. Recent work [4] analyzed, more generally, the conditions for linearity of optimal estimators. Here, we present the basic result pertaining to the jamming problem considered here. Specifically, we present the necessary and sufficient condition for source and channel distributions such that the linear estimator $h(U) = \frac{\kappa}{\kappa+1} U$ is optimal where $\kappa = \frac{\sigma_X^2}{\sigma_Z^2}$ is the SNR.

Theorem 1 ([4]). *Given SNR level κ, and noise Z with characteristic function $F_Z(\omega)$, there exists a source X for which the optimal estimator is linear* if and only if

$$F_X(\omega) = F_Z^\kappa(\omega). \quad (8)$$

Given a valid characteristic function $F_Z(\omega)$, and for some $\kappa \in \mathbb{R}^+$, the function $F_Z^\kappa(\omega)$ may or may not be a valid characteristic function, which determines the existence of a matching source. For example, matching is guaranteed for integer κ and it is also guaranteed for infinitely divisible Z. More comprehensive discussion of the conditions on κ and $F_Z(\omega)$ for $F_Z^\kappa(\omega)$ to be a valid characteristic function can be found in [4].

Extension of the conditions to the vector case is nontrivial due to the dependencies across components of the source and noise. Formally, we consider the problem of estimating the vector source $\boldsymbol{X} \in \mathbb{R}^m$ given the observation $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{Z}$, where $\boldsymbol{X}$ and $\boldsymbol{Z} \in \mathbb{R}^m$ are independent, as shown in Figure 2. Let Q be the eigenmatrix of $R_X R_Z^{-1}$, and $U = Q^{-1}$ and let eigenvalues $\lambda_1, ..., \lambda_m$ be the elements of the diagonal matrix Λ, i.e., the following holds:

$$R_X R_Z^{-1} = U^{-1} \Lambda U \quad (9)$$

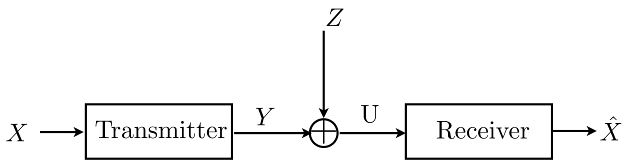

Fig. 3. The communication setting.

We are looking for the conditions on $F_X(\boldsymbol{\omega})$ and $F_Z(\boldsymbol{\omega})$ such that $\boldsymbol{h}(\boldsymbol{Y}) = K\boldsymbol{Y}$ with $K = R_X(R_X + R_Z)^{-1}$ minimizes the estimation error $\mathbb{E}\{||\boldsymbol{X} - \boldsymbol{h}(\boldsymbol{Y})||^2\}$.

By following a similar approach to the scalar case, the necessary and sufficient condition for linearity was derived in [4], reproduced below:

Theorem 2 ([4])**.** *Let the characteristic functions of the transformed source and noise ($U\boldsymbol{X}$ and $U\boldsymbol{Z}$) be $F_{UX}(\boldsymbol{\omega})$ and $F_{UZ}(\boldsymbol{\omega})$. The necessary and sufficient condition for linearity of optimal estimation is:*

$$\frac{\partial \log F_{UX}(\boldsymbol{\omega})}{\partial \omega_i} = \lambda_i \frac{\partial \log F_{UZ}(\boldsymbol{\omega})}{\partial \omega_i}, 1 \leq i \leq m \tag{10}$$

B. Communication Problem

In [5], a communication scenario whose block diagram is shown in Figure 3 was studied. In this setting, a scalar source $X \in \mathbb{R}$ is mapped into $Y \in \mathbb{R}$ by function $g \in \mathcal{S}$, and transmitted over an additive noise channel. The channel output $U = Y + Z$ is mapped by the decoder to the estimate $\hat{X}$ via function $h \in \mathcal{S}$. The zero mean noise Z is assumed to be independent of the source X. The source density is denoted $f_X(\cdot)$ and the noise density is $f_Z(\cdot)$ with characteristic functions $F_X(\omega)$ and $F_Z(\omega)$, respectively.

The objective is to minimize, over the choice of encoder g and decoder h, the distortion

$$D = \mathbb{E}\{(X - \hat{X})^2\}, \tag{11}$$

subject to the average transmission power constraint,

$$\mathbb{E}\{g^2(X)\} \leq P_T. \tag{12}$$

The necessary and sufficient condition for linearity for both mappings is given by the following theorem.

Theorem 3 ([5])**.** *For a given power constraint P_T, noise Z with variance σ_Z^2 and characteristic function $F_Z(\omega)$, source X with variance σ_X^2 and characteristic function $F_X(\omega)$, the optimal encoder and decoder mappings are linear if and only if*

$$F_X(\alpha\omega) = F_Z^{\kappa}(\omega) \tag{13}$$

where $\kappa = \frac{P_T}{\sigma_Z^2}$ and $\alpha = \sqrt{\frac{P_T}{\sigma_X^2}}$.

Here, we first extend this result to vector spaces, in Section III-b. Towards, deriving this extension, we need the optimal encoding and decoding transforms for a general communication problem with source and channel noise covariances R_X and R_N and total encoding power limit P_T, here we reproduce the classical result due to [7] (see also [8], [9] for alternative derivations of this result.).

Theorem 4 ([7]–[9])**.** *The encoding transform that minimizes the MSE distortion subject to the power constraint P_T is*

$$C = Q_N \Sigma Q_X^T \tag{14}$$

where Σ is diagonal power allocation matrix. Moreover the total distortion is

$$D = \frac{\left(\sum\limits_{i=1}^{w} (\sqrt{\lambda_X(i)\lambda_N(m-i)}\right)^2}{P_T + \sum\limits_{i=1}^{w} \lambda_N(m-i)} + \sum_{w+1}^{m} \lambda_X(i) \tag{15}$$

where w is the number of active channels determined by the power P_T.

Remark 1. *Distortion expression (15) has an interesting interpretation of power allocation as "reverse water-filling" over the source eigenvalues. As we will show in the next section, the optimal jammer also performs power allocation as water-filling over the channel eigenvalues.*

Remark 2. *Note that the ordering of the eigenvalues are so that the largest source eigenvalue is multiplied with the smallest of the and so on, which physically means that the encoder uses the best channel for the smallest variance source component, and so on. This is simply due to Lemma 1.*

Assumption 1: In this paper, we assume the source and the channel have matched dimensions, i.e., $m = k$ (which means bandwidth compression or expansion). This assumption is essential in the sense that when $m \neq k$ the jammer cannot ensure linearity of $g(X)$ and $h(Y)$. A well known example is when $m = 2$ and $k = 1$, the optimal mappings are highly nonlinear, even in the case of Gaussian source and channel (see eg. [5], [10]and there references there in for details).

Assumption 2: Throughout this paper, we assume that P_T is high enough, so that all the channels are active, (each channel is allocated strictly positive power), i.e., $w = m$, hence (15) can be rewritten as

$$J(\boldsymbol{\lambda}_X, \boldsymbol{\lambda}_N) = \frac{\left(\sum\limits_{i=1}^{m} (\sqrt{\lambda_X(i)\lambda_N(m-i)}\right)^2}{P_T + \sum\limits_{i=1}^{m} \lambda_N(i)} \tag{16}$$

This assumption is not necessary but significantly simplifies the results.

C. Gaussian Scalar Jamming Problem

The problem of transmitting independent and identically distributed Gaussian random variables over a Gaussian channel in the presence of an additive jammer was considered in [1], [2]. In [2] a game theoretic approach was developed and it was shown that the problem admits a mixed saddle point solution where the optimal transmitter and receiver employ a "randomized" strategy. The randomization information can be sent over a side channel between transmitter and receiver or it could be viewed as the information generated by the third party and observed by both transmitter and receiver [1]. Surprisingly, the optimal jamming strategy ignores the input to the jammer and merely generates "Gaussian" noise, independent of the source.

Theorem 5 ([1], [2]). *The optimal encoding function for the transmitter is randomized linear mapping:*

$$Y(i) = \gamma(i)\alpha_T X(i), \tag{17}$$

where $\gamma(i)$ is i.i.d. Bernoulli ($\frac{1}{2}$) over the alphabet $\{-1, 1\}$

$$\gamma(i) \sim Bern(\frac{1}{2}) \tag{18}$$

and $\alpha_T = \sqrt{\frac{P_T}{\sigma_X^2}}$. The optimal jammer generates i.i.d. Gaussian output $Z(i)$

$$Z(i) \sim \mathcal{N}(0, P_A) \tag{19}$$

where $Z(i)$ is independent of the source $X(i)$.

The proof of Theorem 5 relies on the well known fact that for a Gaussian source over a Gaussian channel, zero-delay linear mappings achieve the performance of the asymptotically high delay optimal source-channel communication system [11]. This fact is unique to the Gaussian source-channel pair, hence it might be tempting to conclude that the saddle point solution in Theorem 5 can only be obtained in the "all Gaussian" setting. Perhaps surprisingly, in [3], we showed that there are infinitely many source-noise pairs that yield a saddle point solution similar to Theorem 5. We also proved that the linearity property of the optimal transmitter and receiver at the saddle point solution still holds, while the Gaussianity of the jammer output in the early special case was merely a means to satisfy this linearity condition, and does not hold in general. Here, we reproduce the main result in [3] for the scalar, non-Gaussian jamming problem.

Theorem 6 ([3]). *For the jamming problem, the optimal encoding function for the transmitter is randomized linear mapping:*

$$Y(i) = \gamma(i)\alpha_T X(i), \tag{20}$$

where $\gamma(i)$ is i.i.d. Bernoulli ($\frac{1}{2}$) over the alphabet $\{-1, 1\}$

$$\gamma(i) \sim Bern(\frac{1}{2}) \tag{21}$$

and $\alpha_T = \sqrt{\frac{P_T}{\sigma_X^2}}$. The optimal jamming function is to generate i.i.d. output $Z(i)$ with characteristic function

$$F_Z(\omega) = \frac{F_X^{\beta}(\alpha_T \omega)}{F_N(\omega)} \tag{22}$$

where $Z(i)$ is independent of the adversarial input $X(i)$ and $\beta = \frac{P_A + \sigma_N^2}{P_T}$.

[1] In practice, randomization can be achieved by (pseudo)random number generators at the transmitter and receiver using the same seed.

IV. Main Results

A. An Upper Bound on Distortion Based on Linear Mappings

In this section, we present a new lemma that is used to upper bound the distortion of any zero-delay communication system by that of the fixed, best linear encoder and decoder. A scalar version of this lemma appeared in [3], here we extend this result to vector spaces.

Lemma 2. *Consider the problem setting in Figure 1. For any given jammer $\boldsymbol{Z}$, the distortion achievable by the transmitter-receiver, D, is upper bounded by the distortion achieved by linear encoder and decoder*

$$J_u = \frac{\left(\sum_{i=1}^{m} \sqrt{\lambda_X(i) \max(\theta, \lambda_N(m-i))} \right)^2}{P_T + \sum_{i=1}^{m} \max(\theta, \lambda_N(i))} \tag{23}$$

where θ satisfies the water-filling condition:

$$\sum_{i=1}^{m} (\theta - \lambda_N(i))^+ = P_A. \tag{24}$$

Note that this upper bound is determined by only second order statistics, regardless of the normalized densities.

Proof: The proof is based on the fact that the transmitter and receiver can always use the linear mappings that satisfy the power constraints. The jammer will try to make all effective noise eigenvalues identical since

$$\min_{\Pi} J(\boldsymbol{\lambda}_X(\Pi), \boldsymbol{\lambda}_N) \tag{25}$$

is maximized when $\boldsymbol{\lambda}_N$ is uniform (see eg. [12] for a proof based on majorization). The effective channel eigenvalues are upper bounded by the vector $\boldsymbol{\sigma} = \boldsymbol{\lambda}_N + \boldsymbol{\lambda}_Z$ where $\boldsymbol{\lambda}_Z$ is the ascending ordered eigenvalues of the jammer noise Z. To achieve this upper bound, the jammer sets $R_Z = Q_N \Lambda_Z Q_N^T$, i.e., the eigenvectors of the noise and the jammer must be identical. If P_A is not large enough, then the solution of is known to be the "waterfilling solution"; intuitively the jammer aims to make entries of $\boldsymbol{\sigma}$ as close to uniform as possible.

Remark 3. *Note that the optimal jammer performs water-filling over the channel eigenvalues while the encoder allocates power according to reverse water-filling over the source eigenvalues. This observation parallels the information theoretical (asymptotically high delay) water-filling duality, where the rate distortion optimal vector encoding scheme allocates total rate by reverse water-filling over source eigenvalues and vector channel capacity achieving scheme allocates power over channel eigenvalues by water-filling.*

Remark 4. *Lemma 2 is the key result that connects the recent results on "linearity" of optimal estimation and communication mappings to the jamming problem. Lemma 2 implies that the optimal strategy for a jammer which can* *only control the density of the additive noise channel, is to force the transmitter and receiver to use linear mappings.*

B. Conditions for Linearity of Communication Mappings in Vector Spaces

For source $\boldsymbol{X} \in \mathbb{R}^m$ and channel $\boldsymbol{Z} \in \mathbb{R}^m$, we derive the necessary and sufficient condition for simultaneous linearity of optimal encoder and decoder. In [5], it was shown that the communication problem is convex in the density of the channel input for the scalar setting. It is easy to extend this convexity result to matched dimensions, see Assumption 1. Moreover, by simply extending the related scalar result in [3], it is straightforward to show that the optimal decoder is linear if and only if the optimal decoder is linear. Hence, we will investigate the conditions for linearity of optimal decoder given that the encoder is linear. Towards deriving our results, we will make use of the following auxiliary lemma from matrix analysis.

Lemma 3. *Given a function $f : \mathbb{R}^n \to \mathbb{R}$, matrix $A \in \mathbb{R}^{n\times m}$ and vector $\boldsymbol{x} \in \mathbb{R}^m$*

$$\nabla_x f(A\boldsymbol{x}) = A^T \nabla f(A\boldsymbol{x}) \tag{26}$$

See Appendix II. Next, we present the necessary and sufficient condition for linearity of optimal decoder for a given, optimal linear decoder.

Theorem 7. *Let the characteristic functions of the transformed source and noise ($\Sigma Q_X^T \boldsymbol{X}$ and $Q_Z^T \boldsymbol{Z}$) be $F_{\Sigma Q_X^T X}(\boldsymbol{\omega})$ and $F_{Q_Z^T Z}(\boldsymbol{\omega})$. The necessary and sufficient condition for linearity of optimal estimation is:*

$$\frac{\partial \log F_{\Sigma Q_X^T X}(\boldsymbol{\omega})}{\partial \omega_i} = S_i \frac{\partial \log F_{Q_Z^T Z}(\boldsymbol{\omega})}{\partial \omega_i}, 1 \le i \le m \tag{27}$$

where $S = \Sigma\Lambda_X \Sigma\Lambda_Z^{-1}$.

C. Main Result

Our main result concerns the optimal strategy for the transmitter, the adversary and the receiver in Figure 1 for the transmission index i.

Theorem 8. *For the jamming problem, the optimal encoding function for the transmitter is:*

$$\boldsymbol{Y}(i) = \gamma(i) C \boldsymbol{X}(i), \tag{28}$$

where $\gamma(i)$ is i.i.d. Bernoulli ($\frac{1}{2}$) over the alphabet $\{-1, 1\}$

$$\gamma(i) \sim Bern(\frac{1}{2}) \tag{29}$$

and $C = Q_N \Sigma Q_X^T$. The optimal jamming function is to generate i.i.d. output $\boldsymbol{Z}(i)$, independent of $\boldsymbol{X}(i)$, that satisfies:

$$\frac{\partial \log F_{\Sigma Q_X^T X}(\boldsymbol{\omega})}{\partial \omega_i} = S_i \frac{\partial \log F_{Q_N^T (N+Z)}(\boldsymbol{\omega})}{\partial \omega_i}, 1 \le i \le m \tag{30}$$

for $S = \Sigma\Lambda_X \Sigma\Lambda_Z^{-1}$ and

$$R_Z = Q_N \lambda_Z Q_N^T \tag{31}$$

where

$$\lambda_Z(i) = (\theta - \lambda_N(i))^+ \tag{32}$$

and θ satisfies the water-filling condition:

$$\sum_{i=1}^{m} (\theta - \lambda_N(i))^+ = P_A \tag{33}$$

The optimal receiver is

$$\boldsymbol{h}(\boldsymbol{U}(i)) = R_X C^T (C R_X C^T + R_N + R_Z)^{-1} \boldsymbol{U}(i), \tag{34}$$

and total cost is

$$J = \frac{\left(\sum\limits_{i=1}^{m} \sqrt{\lambda_X(i) \max(\theta, \lambda_N(m-i))}\right)^2}{P_T + \sum\limits_{i=1}^{m} \max(\theta, \lambda_N(i))} \tag{35}$$

Moreover, this saddle point solution is (almost surely) unique.

Proof: The proof follows from verification of the saddle point inequalities given in (7), following the approach in [13], and is merely a simple extension of the scalar result in [3]. The essential idea of the proof is that the jammer renders the channel noise to match the source in a way that the optimal encoding and decoding mappings are linear. The condition for such matching is given in Theorem 7. Moreover, R_Z shares the same eigenvectors as R_N as explained in the proof of Lemma 2.

V. Discussion

In this paper, we studied the vector jamming problem. Similar to scalar setting, linearity conditions for encoding and decoding mappings play the key role in jamming, in the sense that the optimal jamming strategy is to effectively force both transmitter and receiver to default to linear mappings. Hence, the optimal strategy is "randomized" linear encoding for transmitter and generating independent noise for the jamming. The eigenvalues of the optimal jamming noise are allocated according to water-filling over the eigenvalues of the channel noise, and the density of the jamming noise is matched to source and the channel to render the mappings linear. We derived the matching condition to be satisfied by the jammer and also the second order statistics of the jamming noise. The power allocation solutions in the zero-delay problems (water-filling for jammer and reverse water-filling for the transmitter) nicely parallels the resource allocation strategies in asymptotically high delay (Shannon type) problems, such as rate allocation in rate-distortion (reverse water-filling) and power allocation in channel capacity (water-filling).

Throughout this paper, we assumed that a matching jamming noise, that forces the transmitter and receiver to use linear mappings, can always be generated by the jammer. The case for which this source-channel-jammer matching is not possible was analyzed for scalar sources and channels in our recent work [3]. Intuitively, the jammer approximates the matching solution in some polynomials which are orthogonal under the measure of the channel output. The analysis for vector settings can be carried out similarly, and is omitted since it is orthogonal to the focus of the current paper.

APPENDIX I
DERIVATION OF THE MATCHING CONDITION

Let us rewrite the MSE optimal decoder, $h(\boldsymbol{y}) = \mathbb{E}\{\boldsymbol{X}|\boldsymbol{y}\}$ using Bayes' rule and independence of $\boldsymbol{X}$ and $\boldsymbol{Z}$:

$$\boldsymbol{h}(\boldsymbol{y}) = \frac{\int \boldsymbol{x} f_X(\boldsymbol{x}) f_Z(\boldsymbol{y} - C\boldsymbol{x})\, \boldsymbol{d}x}{\int f_X(\boldsymbol{x}) f_Z(\boldsymbol{y} - C\boldsymbol{x})\, \boldsymbol{d}x} \tag{36}$$

Plugging $\boldsymbol{h}(\boldsymbol{y}) = K\boldsymbol{y}$ in (36) we obtain,

$$K\boldsymbol{y} \int f_X(\boldsymbol{x}) f_Z(\boldsymbol{y} - C\boldsymbol{x})\, \boldsymbol{d}x = \int \boldsymbol{x} f_X(\boldsymbol{x}) f_Z(\boldsymbol{y} - C\boldsymbol{x})\, \boldsymbol{d}x \tag{37}$$

Taking the Fourier transform of both sides

$$jK\nabla_\omega \left[F_X(C\boldsymbol{\omega})F_Z(\boldsymbol{\omega})\right] = jC^{-1}[\nabla_\omega F_X(C\boldsymbol{\omega})]F_Z(\boldsymbol{\omega}) \tag{38}$$

and rearranging terms, we get

$$\left(C^{-1} - K\right) \frac{1}{F_X(C\boldsymbol{\omega})} \nabla F_X(C\boldsymbol{\omega}) = K\frac{1}{F_Z(\boldsymbol{\omega})} \nabla F_Z(\boldsymbol{\omega}) \tag{39}$$

Using $\nabla \log F_X(C\boldsymbol{\omega}) = \frac{1}{F_X(C\boldsymbol{\omega})} \nabla F_X(C\boldsymbol{\omega})$,

$$\nabla \log F_X(C\boldsymbol{\omega}) = (C^{-1} - K)^{-1} K \nabla \log F_Z(\boldsymbol{\omega}) \tag{40}$$

Note that (see eg. [14])

$$K = R_X C^T (C R_X C^T + R_Z)^{-1} \tag{41}$$

Let use define $\beta \triangleq (C^{-1} - K)^{-1} K$ then,

$$\beta = (C^{-1} - R_X C^T (C R_X C^T + R_Z)^{-1})^{-1} \times R_X C^T (C R_X C^T + R_Z)^{-1} \tag{42}$$

It is easy to see that

$$\beta^{-1} = R_Z (C R_X C^T)^{-1} \tag{43}$$

and hence

$$\beta = (C R_X C^T) R_Z^{-1} \tag{44}$$

Plugging (14) in (45), we have

$$\beta = Q_Z (\Sigma \Lambda_X \Sigma \Lambda_Z^{-1}) Q_Z^T \tag{45}$$

and hence

$$\nabla \log F_X(C\boldsymbol{\omega}) = Q_Z (\Sigma \Lambda_X \Sigma \Lambda_Z^{-1}) Q_Z^T \nabla \log F_Z(\boldsymbol{\omega}) \tag{46}$$

defining $S \triangleq (\Sigma \Lambda_X \Sigma \Lambda_Z^{-1})$, we have

$$Q_Z^T \nabla_\omega \log F_X(C\boldsymbol{\omega}) = S Q_Z^T \nabla \log F_Z(\boldsymbol{\omega}) \tag{47}$$

Let us define $\tilde{\boldsymbol{\omega}} \triangleq Q_Z^T \boldsymbol{\omega}$, hence $\boldsymbol{\omega} = Q_Z \tilde{\boldsymbol{\omega}}$. Plugging this in (47), we have

$$Q_Z^T \nabla_{Q_Z\tilde{\omega}} \log F_X(C Q_Z \tilde{\boldsymbol{\omega}}) = S Q_Z^T \nabla_{Q_Z \tilde{\omega}} \log F_Z(Q_Z \tilde{\boldsymbol{\omega}}) \tag{48}$$

Using Lemma 3, we can rewrite (48) as

$$\nabla_{\tilde{\omega}} \log F_X(C^T Q_Z \tilde{\boldsymbol{\omega}}) = S \nabla_{\tilde{\omega}} \log F_Z(Q_Z \tilde{\boldsymbol{\omega}}) \tag{49}$$

noting that $C^T Q_Z = Q_X \Sigma$ and the characteristic functions of the source and noise after transformation can be written in terms of the known characteristic functions $F_X(\boldsymbol{\omega})$ and $F_Z(\boldsymbol{\omega})$, specifically $F_{\Sigma Q_X^T X}(\boldsymbol{\omega}) = F_X(\Sigma Q_X^T \boldsymbol{\omega})$ and $F_{Q_Z^T Z}(\boldsymbol{\omega}) = F_Z(Q_Z \boldsymbol{\omega})$, we have

$$\nabla_{\tilde{\omega}} \log F_{\Sigma Q_X^T X}(\tilde{\boldsymbol{\omega}}) = S \nabla_{\tilde{\omega}} \log F_{Q_Z^T Z}(\tilde{\boldsymbol{\omega}}) \tag{50}$$

Using the fact that S is diagonal, we convert (50) to the set of m scalar differential equations of (30).

Converse can be shown by retracing the steps in the derivation of the necessity. Note that none of these steps, (36)-(47), introduce any loss of generality, hence retracing back from (47) to (36), we show that if (47) is satisfied, the optimal decoder is linear given the encoder is linear.

APPENDIX II
PROOF OF LEMMA 3

By the chain rule we have,

$$\begin{aligned}
\frac{\partial f(A\boldsymbol{x})}{\partial x_i} &= \sum_{k=1}^{n} \frac{\partial f(A\boldsymbol{x})}{\partial [A\boldsymbol{x}]_k} \frac{\partial [A\boldsymbol{x}]_k}{\partial [\boldsymbol{x}]_i} && (51)\\
&= \sum_{k=1}^{n} \frac{\partial f(A\boldsymbol{x})}{\partial [A\boldsymbol{x}]_k} \frac{\partial ([A]_k{}^T \boldsymbol{x})}{\partial [\boldsymbol{x}]_i} && (52)\\
&= \sum_{k=1}^{n} \frac{\partial f(A\boldsymbol{x})}{\partial [A\boldsymbol{x}]_k} [A]_{ki} && (53)\\
&= \sum_{k=1}^{n} \partial_k f(A\boldsymbol{x}) [A]_{ki} && (54)\\
&= [A]_i{}^T \nabla f(A\boldsymbol{x}) && (55)
\end{aligned}$$

It follows from (55) that $\nabla_x f(A\boldsymbol{x}) = A^T \nabla f(A\boldsymbol{x})$.

REFERENCES

[1] T. Başar, "The Gaussian test channel with an intelligent jammer," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 152–157, 1983.

[2] T. Başar and Y.W. Wu, "A complete characterization of minimax and maximin encoder-decoder policies for communication channels with incomplete statistical description," *IEEE Transactions on Information Theory,*, vol. 31, no. 4, pp. 482–489, 1985.

[3] E. Akyol, K. Rose, and T. Başar, "On optimal jamming over an additive noise channel," *draft, available online at: http://arxiv.org/abs/1303.3049*.

[4] E. Akyol, K. Viswanatha, and K. Rose, "On conditions for linearity of optimal estimation," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3497–3508, 2012.

[5] E. Akyol, K. Viswanatha, K. Rose, and Ramstad T., "On zero-delay source-channel coding," *IEEE Transactions on Information Theory, in review, available online at: http://arxiv.org/abs/1302.3660*.

[6] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[7] K.H. Lee and D. Petersen, "Optimal linear coding for vector channels," *IEEE Transactions on Communications*, vol. 24, no. 12, pp. 1283–1290, 1976.

[8] T. Başar, B. Sankur, and H. Abut, "Performance bounds and optimal linear coding for discrete-time multichannel communication systems (corresp.)," *IEEE Transactions on Information Theory*, vol. 26, no. 2, pp. 212–217, 1980.

[9] E. Akyol and K. Rose, "On linear transforms in zero-delay gaussian source channel coding," in *IEEE International Symposium on Information Theory Proceedings (ISIT),*. IEEE, 2012, pp. 1543–1547.

[10] F. Hekland, P.A. Floor, and T.A. Ramstad, "Shannon-Kotelnikov mappings in joint source-channel coding," *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 94–105, 2009.

[11] T. M Cover and J. A. Thomas, *Elements of information theory*, Wiley-interscience, 2006.
[12] AW Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* , Academic Press, New York, 1979.
[13] T. Başar and Y.W. Wu, "Solutions to a class of minimax decision problems arising in communication systems," *Journal of optimization theory and applications*, vol. 51, no. 3, pp. 375–404, 1986.
[14] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall Upper Saddle River, NJ, 2000.