Kullback-Leibler divergence region in MIMO radar detection problems

Emanuele Grossi and Marco Lops
DIEI, Università degli Studi di Cassino e del Lazio Meridionale, Italy 03043
Email: e.grossi@unicas.it, lops@unicas.it

Abstract—One of the most relevant performance measures in detection problems is the pair of Kullback-Leibler divergences between the densities of the observations under the two hypotheses. In this work, the problem of space-time code design for a multiple-input, multiple-output detection systems is addressed, and the set of achievable divergence pairs is characterized when different constraints are imposed on the code matrix. Knowledge of this region permits to achieve the desired performance level in the detection system (e.g., the time for taking a decision in a sequential probability ratio test or the probability of miss in a likelihood ratio test) and, if a cost function is assigned, to carry out optimal waveform design.

I. INTRODUCTION

In multiple-input, multiple-output (MIMO) remote sensing systems [1], [2] space-time (ST) coding is a key ingredient to achieve diversity and, in general, for performance optimization purposes. Different cost/merit functions have been introduced and studied, the most used being the received signal-to-noise ratio (SNR), the minimum mean-square error in estimating the unknown target response, the mutual information between the received signal and the target response, and the Chernoff’s bound on the detection probability [2]–[7]. However, from a detection point of view, the more appropriate and versatile appears to be the Kullback-Leibler divergence between the densities of the observations under the two hypotheses [8]–[12]. When a physical constraint is imposed on the transmitted code matrix, the knowledge of the region of achievable divergence pairs permits to obtain the desired performance level in the detection system – e.g., in the likelihood ratio test (LRT), the divergences are linked to the asymptotic probability of miss, while, in the sequential probability ratio test (SPRT), they determine the average sample number (ASN) under the two hypotheses – and, in general, to carry out code optimization. In the next section, the MIMO detection architecture under investigation is described and the code design problem is presented, while, in Sec. III, the set of achievable Kullback-Leibler divergences is presented under different system constraints. Finally, Sec. IV is devoted to concluding remarks, while mathematical proofs are deferred to the Appendices.

II. SYSTEM MODEL AND CODE MATRIX DESIGN

Consider a MIMO detection architecture composed of \( n_t \) transmit and \( n_r \) receive antennas. At each epoch \( \ell \), a linear combination of \( N \) orthogonal, narrowband waveforms (a frame from now on) is emitted from each transmit antenna, and the echo from a prospective target is received. After standard A/D conversion (e.g., see [11], and references therein), the \( N \)-dimensional observations on the \( n_r \) receive antennas corresponding to a given resolution element (e.g., pointing angle, range, and Doppler shift) and to the \( \ell \)-th transmitted frame can be cast in the following \( N \times n_r \) complex matrix

\[
R(\ell) = \begin{cases} CH(\ell) + N(\ell), & \text{under } H_1 \\ N(\ell), & \text{under } H_0 \end{cases}
\]

where: \( H_1 \) and \( H_0 \) denote the target presence and absence hypothesis, respectively; \( C \in \mathbb{C}^{N \times n_t} \) is the ST code matrix; \( H(\ell) \in \mathbb{C}^{n_r \times n_r} \) is the scattering matrix, i.e., \( H_{ij}(\ell) \) is the scattering from antenna \( i \) towards antenna \( j \); and \( N(\ell) \in \mathbb{C}^{N \times n_r} \) represents the overall disturbance (thermal noise and reverberation from the environment) on the receive antennas. The entries of \( H(\ell) \) and \( N(\ell) \) are modeled as circularly symmetric, Gaussian random variables, with

\[
E[H_{hk}(\ell)H_{jk}^*(m)] = L_{ij}\delta_{hk}\delta_{\ell m} \\
E[H_{hk}(\ell)N_{jk}^*(m)] = M_{ij}\delta_{hk}\delta_{\ell m}.
\]

The total transmitted energy is \( \text{tr}(CC^H) \), and the received SNR per receive antenna is \( \text{tr}(CLC^HM^{-1}) \), \( \text{tr}(\cdot) \) denoting the trace operator, where \( L = [L_{ij}]_{i,j \in \{1,\ldots,n_r\}} \), and \( M = [M_{ij}]_{i,j \in \{1,\ldots,n_r\}} \) is full rank. For future reference, the following eigenvalue decompositions are introduced

\[
M = U_M \text{diag} \{\lambda_N(M), \ldots, \lambda_1(M)\} U_M^H \\
L = V_L \text{diag} \{\lambda_L(L), \ldots, \lambda_1(L)\} V_L^H
\]

where \( U_M \in \mathbb{C}^{N \times N} \) and \( V_L \in \mathbb{C}^{n_r \times n_r} \) are unitary, \((\cdot)^H \) denote conjugate transpose, and diag\{\(a_1, \ldots, a_n\)\} is the diagonal, \( n \times n \) matrix with \( a_1, \ldots, a_n \) on the main diagonal.

Based on a certain number of observations sampled from the process \( \{R(\ell)\}_{\ell \in \mathbb{N}} \) – deterministic or random, according to whether a fixed sample-size (FSS) or a sequential test is being considered –, the receiver has to decide if a target is present or if measurements are generated by noise alone. The code matrix \( C \) represents a degree of freedom, which can be exploited at the transmitter side to improve the overall system performance. Key parameters in the detection problem are the Kullback-Leibler divergences between the densities of the observations.

1Notice that no space correlation is taken into account at the receiver side, and that the frame-to-frame independence amount to a Swerling I target fluctuation, which can be achieved through frequency agile sensors.
$\mathbf{R}(\ell)$ in (1) under the two hypotheses. These densities are

\begin{align}
 f_0(\mathbf{R}(\ell)) &= \frac{e^{-\text{tr}(\mathbf{R}(\ell)\mathbf{M}^{-1}\mathbf{R}(\ell))}}{(\pi^N \det(\mathbf{M}))^{\nu_r}} \\
 f_1(\mathbf{R}(\ell)) &= \frac{e^{-\text{tr}(\mathbf{R}(\ell)(\mathbf{M} + \mathbf{CLC}^H)^{-1}\mathbf{R}(\ell))}}{(\pi^N \det(\mathbf{M} + \mathbf{CLC}^H))^{\nu_r}}
\end{align}

under $H_0$ and $H_1$, respectively, and, as shown in [11], [12], the divergences are

\begin{align}
 D(f_1\|f_0) &= n_r \sum_{i=1}^{\Delta} [\rho_i - \ln(1 + \rho_i)] \\
 D(f_0\|f_1) &= n_r \sum_{i=1}^{\Delta} [\ln(1 + \rho_i) - \frac{\rho_i}{1 + \rho_i}]
\end{align}

where $\Delta = \min\{n_r, N\}$, and

$$
\rho_i = \lambda_i(\mathbf{M}^{-1/2}\mathbf{CLC}^H\mathbf{M}^{-1/2}), \quad i = 1, \ldots, \Delta.
$$

Throughout the paper, the notation $\mathbf{A}^{1/2}$ is used to denote the unique, positive semidefinite square root of the Hermitian, positive semidefinite matrix $\mathbf{A}$, and $\lambda_i(\mathbf{B})$ to denote the $i$-th eigenvalue of a square matrix $\mathbf{B}$; if $\mathbf{B}$ is Hermitian, the eigenvalues are sorted in descending order. Therefore, $\rho_i$, can be interpreted as the SNR under $H_1$ along the $i$-th eigenmode of the MIMO channel.

To fully understand the role in the detection problem of the divergences in (3), consider, e.g., the LRT (which, by the Neyman-Pearson Lemma [13], is the most powerful test with bounded probability of false alarm based on a fixed sample-size) or the SPRT (which, by the Wald-Wolfowitz optimum property [13], minimizes the ASN under both hypotheses in the class of tests, FSS or sequential, with bounded strength). Letting

$$
P_{fa} = \mathbb{P}($accept $H_1|H_0)$
$$

$$
P_{miss} = \mathbb{P}($accept $H_0|H_1)$
$$

Stein’s Lemma ensures that the probability of miss in the LRT with fixed probability of false alarm decays at an exponential rate dictated by $D(f_0\|f_1)$, i.e.,

$$
\lim_{n \to \infty} \frac{1}{n} \ln P_{miss} = -D(f_0\|f_1).
$$

In the SPRT, instead, the average sample number is proportional to the inverse of $D(f_0\|f_1)$ and $D(f_1\|f_0)$ when $(P_{fa}, P_{miss})$ is fixed, i.e.,

$$
\text{ASN}_{H_0} \simeq \frac{P_{fa} \ln \frac{P_{fa}}{P_{miss}} + (1 - P_{fa}) \ln \frac{1 - P_{fa}}{P_{miss}}}{D(f_0\|f_1)} \\
\text{ASN}_{H_1} \simeq \frac{P_{miss} \ln \frac{P_{miss}}{P_{fa}} + (1 - P_{miss}) \ln \frac{1 - P_{miss}}{P_{fa}}}{D(f_1\|f_0)}
$$

where the approximations hold neglecting overshoots [13].

Therefore, knowledge of the region of possible values that the divergences in (3) can assume when the the code matrix $\mathbf{C}$ is a design parameter permits to achieve the desired performance level in the detection system and, if a cost function is assigned, to carry out optimal waveform design. To give an example, let $Q$ be a decreasing cost function of the pair $(D(f_1\|f_0), D(f_0\|f_1))$, then the optimum code matrix $\mathbf{C}$ is the one corresponding to the point on the border of the achievable region that touches the level set $\{(x, y) : Q(x, y) = c\}$ corresponding to the smallest $c \in \mathbb{R}$ (see Fig. 1). Reasonable choices for the cost function can be $Q(x, y) = -\alpha x - (1 - \alpha) y$, $\alpha \in [0, 1]$, i.e., proportional to the convex combination of the divergences, in which case the level set are parallel lines, or $Q(x, y) = \alpha x + \beta y$, $\alpha, \beta \in [0, 1]$, i.e., proportional to the convex combination of the ASN’s in (4), in which case the level set are hyperbolas.

### III. Achievable Kullback-Leibler Divergences

The set of achievable divergences under a physical constraint on the code matrix $\mathbf{C}$ is studied here. Two following two scenarios will be discussed.

- In remote sensing applications, and in particular when the sensing system is ground-based and has surveillance tasks, it is often required that a standard target (i.e., possessing a given average radar cross-section and fluctuating according to a Swerling model) located at a given distance produce the required SNR level [14]–[16]. In this case, a constraint is imposed on the SNR per receive antenna, $\text{tr}(\mathbf{CLC}^H \mathbf{M}^{-1})$.

- In monitoring applications with small, battery-powered transmit nodes the main limitation to system performance is the transmitted energy [17]–[20]. In this case, a constraint is imposed on the transmitted energy, $\text{tr}(\mathbf{C}^H \mathbf{C})$.

#### A. Constrained receive SNR

The set of all of the divergence pairs achievable when the SNR is fixed at the level $\nu$ is described here, i.e.,

$$
\left\{ (D(f_0\|f_1), D(f_1\|f_0)) : \text{tr}(\mathbf{CLC}^H \mathbf{M}^{-1}) = \nu \right\}.
$$

This is needed as targets with a certain threat level should not pass undetected inside a specified minimum range. Moreover, the measurement accuracy is tied to the received SNR, and excessively low values of this parameter lead to track degradation and, ultimately, to unexpected track termination [15].
As shown in Appendix A, the achievable region is the set whose border is described by the curves
\[ \{ \varphi_m(t) \}_{m=2}^{\Delta-1}, \quad t \in [0, 1/m] \]
and
\[ \varphi_{\Delta}(t), \quad t \in [0, 1] \]
where \( \varphi_m : [0, 1] \rightarrow \mathbb{R}^2 \) is defined, for \( m \in \{2, \ldots, \Delta\} \), as
\[
\varphi_m(t) = (m-1) \left( \frac{(1-t)m}{m-1} - \ln \left( \frac{1 + (1-t)m}{m-1} \right) - \frac{(1-t)m}{m-1} \right) + \left( t\nu - \ln(1+t\nu) \right) \frac{1}{\ln(1+t\nu) - \frac{t\nu}{1+t\nu}}. \tag{5}
\]
The divergence pairs on the border originate from the points
\[ \rho(t) = \nu \left( \frac{1 - t}{m-1} \cdots 1 - t}{m-1} t \ 0 \cdots 0 \right)^T \]
and the cusps correspond to
\[ \rho = \left( \frac{\nu/m}{m} \cdots \frac{\nu/m}{m} 0 \cdots 0 \right)^T, \quad m = 1, \ldots, \Delta. \]

Fig. 2 gives an example of such a region when \( \Delta = 4 \) and \( \nu = 10 \) dB.

**B. Constrained transmit energy**

Here the goal is to describe the set of the achievable divergence pairs when the transmit energy is fixed to the level \( \mathcal{E} \), i.e., the set
\[ \left\{ (D(f_0\|f_1), D(f_1\|f_0)) : \text{tr}(CC^H) = \mathcal{E} \right\}. \]
This problem is more challenging than the one tackled in Sec. III-A, and only a bound is given. This bound is partly based on the solution (given in [12]) of the following

**Problem 1:** Find the code matrix \( C \) which solves
\[
\max_{C \in \mathbb{C}^{N \times n} } \left\{ \alpha D(f_1\|f_0) + (1-\alpha)D(f_0\|f_1) \right\}
\text{ s. t. } \text{tr}(CC^H) \leq \mathcal{E}
\]
with \( \alpha \in [0,1] \) and \( \mathcal{E} > 0 \).

In particular, the region of achievable divergence pairs is included in the set reported in Figure 3, whose extreme points are listed below: the shaded area is the achievable region, obtained through numerical procedures, and the solid lines inside, themselves obtained through numerical procedures, correspond to the case where the singular vectors of the code matrix are matched to the eigenvectors of \( M \) and \( L \).

- The lower and upper curves are
\[
\varphi_{\text{low}}(t) = \begin{pmatrix} t - \ln(1+t) \\ \ln(1+t) - t/(1+t) \end{pmatrix},
\varphi_{\text{up}}(t) = \Delta \begin{pmatrix} t - \ln(1+t) \\ \ln(1+t) - t/(1+t) \end{pmatrix},
\]
t \geq 0, and correspond to rank-1 coding with SNR maximization and rank-\( \Delta \) coding with SNR’s equalization, respectively.

- Points A and D are
\[
A = \varphi_{\text{up}} \left( \sum_{i=1}^{\Delta} \frac{\lambda_{N-\Delta+i}(M)}{\lambda_i(L)} \right),
D = \varphi_{\text{low}} \left( \frac{\sum_{i=1}^{\Delta} \lambda_{\pi(i)}(M)}{\lambda_N(M)} \right).
\]
All of the points between the origin and A on \( \varphi_{\text{up}}(t) \) and between the origin and D on \( \varphi_{\text{low}}(t) \) are achievable. The code matrix corresponding to A is
\[
C = U_M^H \Omega V_L^H \tag{7}
\]
where \( U_M^\pi \) is the matrix obtained through the permutation \( \pi(i) = (N-\Delta+i) \mod N \), \( i = 1, \ldots, N \), of the columns of \( U_M \), and \( \Omega \in \mathbb{C}^{N \times n} \) is diagonal, with
\[
\Omega_{ii} = \begin{cases} \sqrt{\frac{\sum_{i=1}^{\Delta} \lambda_{\pi(i)}(M)/\lambda_i(L)}} & \text{if } i \leq \Delta \\ 0 & \text{otherwise}. \end{cases} \tag{8}
\]

898
while the code matrix corresponding to $D$ is
\[ C = U_M \Sigma V_L^H \]  
where $\Sigma \in \mathbb{R}^{N \times n_1}$ is a diagonal matrix with \{\sqrt{E}, 0, \ldots, 0\} on the main diagonal.

- The curve $\varphi_{\text{null}}$ is the convex hull obtained solving Problem 1 for all $\alpha \in [0, 1]$: $\alpha = 1$ corresponds to point $D$, while $\alpha = 0$ to point $C$, as $D(f_0 || f_1)$ is maximized under a transmit energy constraint. Observe that $C$ collapses on $D$ when $E$ is small, as shown in [12], and that not all of the divergence pairs on $\varphi_{\text{null}}$ are achievable, as the region is not convex in this area (e.g., see the zoom around point $D$ in Fig. 3).

- Finally, $B$ is the point on curve $\varphi_{\text{up}}$ corresponding to the same value of $D(f_0 || f_1)$ as point $C$, and the segment $B - C$ represents an upper bound on $D(f_0 || f_1)$.

See Appendix B for the proof.

IV. CONCLUSION
In the framework of ST code design for MIMO detection, the set of achievable Kullback-Leibler divergences between the densities of the observations under the two hypotheses has been characterized under different constraints on the transmitted code matrix: in particular, an analytical form of the region has been found for constrained receive SNR, while for the case of constrained transmit energy a bound on the achievable region has been determined. This region allows to control the performance of the detection architecture and, ultimately, to carry out optimal waveform design.

APPENDIX A
CONSTRAINED RECEIVE SNR
First, the extreme values of $D(f_1 || f_0)$ are to be derived under the constraint $\sum_{i=1}^{\Delta} \rho_i = \nu$. From (3a), $D(f_1 || f_0)$ is Schur-convex in $\rho$, and then
\[ \nu - \Delta \ln(1 + \nu/\Delta) \leq D(f_1 || f_0) \leq \nu - \ln(1 + \nu). \]
The minimum is achieved at $\rho = (\nu/\Delta \cdots \nu/\Delta)^T$ while the maximum at $\rho = (\nu \ 0 \cdots 0)^T$. In order to find the achievable region of divergence pairs, the minimum and the maximum value of $D(f_0 || f_1)$ will be sought under the constraint that $D(f_1 || f_0) = c$, for some $c \in [\nu - \Delta \ln(1 + \nu/\Delta), \nu - \ln(1 + \nu)]$. Since $D(f_0 || f_1)$ is a continuous function, every point between the maximum and the minimum is also achievable. The minimum will be first derived.

From the expression of the divergences in (3a) and (3b), the problem is
\[
\min_{\rho \in \mathbb{R}^\Delta} \sum_{i=1}^{\Delta} \left[ \ln(1 + \rho_i) - \frac{\rho_i}{1 + \rho_i} \right]
\]
s.t. \[\sum_{i=1}^{\Delta} \rho_i = \nu\]
\[\rho_i \geq 0, \ i = 1, \ldots, \Delta\]
\[\sum_{i=1}^{\Delta} \rho_i - \ln(1 + \rho_i) = c\]  
for some $c \in [\nu - \Delta \ln(1 + \nu/\Delta), \nu - \ln(1 + \nu)]$. After eliminating the SNR constraint, the problem becomes
\[
\min_{\rho \in \mathbb{R}^\Delta} \left\{ \frac{1}{1 + \nu - \sum_{i=1}^{\Delta-1} \rho_i} + \sum_{i=1}^{\Delta-1} \frac{1}{1 + \rho_i} \right\}
\]
s.t. \[\ln\left(1 + \nu - \sum_{i=1}^{\Delta-1} \rho_i\right) + \sum_{i=1}^{\Delta-1} \ln(1 + \rho_i) = \nu - c\]
\[\rho_i \geq 0, \ i = 1, \ldots, \Delta\]
\[\nu - \sum_{i=1}^{\Delta-1} \rho_i \geq 0.\]

Neglect for a moment the inequality constraints: their validity will be checked on the candidate solution. Introducing the Lagrangian
\[
\frac{1}{1 + \nu - \sum_{i=1}^{\Delta-1} \rho_i} + \sum_{i=1}^{\Delta-1} \frac{1}{1 + \rho_i} + \mu \left[ \ln\left(1 + \nu - \sum_{i=1}^{\Delta-1} \rho_i\right) + \sum_{i=1}^{\Delta-1} \ln(1 + \rho_i) - (\nu - c) \right]
\]
the first-order conditions gives
\[
\mu = \frac{1}{1 + \nu - \sum_{i=1}^{\Delta-1} \rho_i} + \frac{1}{1 + \rho_i}, \ \forall i
\]
i.e., $\rho_i = \rho_j$, for any $i, j \in \{1, \ldots, \Delta\}$. Take
\[
\rho_i = \frac{1 - \gamma}{\Delta - 1} \nu
\]
so that
\[
\mu = \frac{1}{1 + \gamma \nu} + \frac{1}{1 + \frac{\gamma \nu}{\Delta - 1}}
\]
where $\gamma \in \mathbb{R}$ satisfies the constraint, i.e.,
\[
\ln(1 + \gamma \nu) + (\Delta - 1) \ln\left(1 + \frac{1 - \gamma}{\Delta - 1} \nu\right) = \nu - c.
\]
Letting $\phi : (-1/\nu, (\Delta - 1)/\nu + 1) \to \mathbb{R}$ be defined as
\[
\phi(\gamma) = \ln(1 + \gamma \nu) + (\Delta - 1) \ln\left(1 + \frac{1 - \gamma}{\Delta - 1} \nu\right)
\]
one notice that
i. $\phi$ is continuous, and it goes to $-\infty$ on the border of its domain;
ii. $\phi$ is increasing in $(-1/\nu, 1/\Delta)$ and decreasing in $(1/\Delta, (\Delta - 1)/\nu + 1)$, and its maximum is $\phi(1/\Delta) = \Delta \ln(1 + \nu/\Delta) > 0$;
iii. $\phi(0) = (\Delta - 1) \ln(1 + \nu/\Delta) > \ln(1 + \nu) = \phi(1)$; and
iv. $\phi(1) \leq \nu - c \leq \phi(1/\Delta)$.
Thus, $\phi(\gamma) = \nu - c$ has two solutions: $\gamma_1 \in [1/\Delta, 1]$ and $\gamma_2 \in [-1/\nu, 1/\Delta]$ (see Figures 4), which gives the minimum and the maximum of the objective function. It will be shown that the minimum corresponds to $\gamma_1$. Indeed, the Hessian matrix evaluated in the points $\rho$ and $\mu$ of (12) and (13), respectively,
the maximum. Finally, since \( \gamma \) inequality constraints are satisfied. Therefore, the solution is \( a > \) which is positive whenever \( \{ \}
\]

In order to have a minimum, the Hessian must be positive definite in \( \{ x \in \mathbb{R}^{\Delta - 1} : \sum_{i=1}^{\Delta - 1} x_i = 0 \} \). On this set, it results

\[
x^T (aI_\Delta + b1_\Delta) x = a \| x \|^2 + b \left( \sum_{i=1}^{\Delta - 1} x_i \right)^2 = a \| x \|^2
\]

which is positive whenever \( a > 0 \). From (15a), \( a > 0 \) for \( \gamma = \gamma_1 \) and \( a < 0 \) for \( \gamma = \gamma_2 \), so that \( \gamma_1 \) corresponds to the maximum. Finally, since \( \gamma_1 \in [1/\Delta, 1] \), from (12), the inequality constraints are satisfied. Therefore, the solution is

\[
\rho = \nu \left( \frac{1 - \gamma_1}{\Delta - 1} \cdots \frac{1 - \gamma_1}{\Delta - 1} \right)^T, \quad \gamma_1 \in [1/\Delta, 1]
\]

which give rise to the curve \( \varphi_\Delta(t), t \in [1/\Delta, 1] \), in (5).

At this point, the maximum value of \( \text{D}(\varphi_0||\varphi_f) \) remains to be found. To this end, the set of achievable \( \text{D}(\varphi_0||\varphi_f) \) is partitioned in \( \Delta - 1 \) non-overlapping intervals

\[
\mathcal{D}_m = \left[ \nu - m \left( 1 + \frac{\nu}{m} \right), \nu - (m - 1) \ln \left( 1 + \frac{\nu}{m - 1} \right) \right]
\]

\( m \in \{2, \cdots, \Delta\} \). Then, for all \( m \in \{2, \cdots, \Delta\} \), \( \text{D}(\varphi_0||\varphi_f) \) is to be maximized under the constraint \( \text{D}(\varphi_0||\varphi_f) = c_m \), for some \( c_m \in \mathcal{D}_m \). Taking (10), with max instead of min, the \( m \)-th problem is

\[
\begin{align*}
\max_{\rho \in \mathbb{R}^{\Delta}} & \quad \sum_{i=1}^{\Delta} \frac{1}{1 + \rho_i} \\
\text{s.t.} & \quad \sum_{i=1}^{\Delta} \rho_i = \nu \\
& \quad \rho_i \geq 0, \quad i = 1, \ldots, \Delta \\
& \quad \sum_{i=1}^{\Delta} \ln(1 + \rho_i) = \nu - c_m
\end{align*}
\]

Since the objective function is decreasing, \( \rho_i \) can be set equal to zero for \( i = m + 1, \ldots, \Delta \), when \( m < \Delta \). After eliminating the SNR constraint, the problem becomes

\[
\begin{align*}
\min_{x \in \mathbb{R}^m} & \quad \left\{ \frac{1}{1 + \nu - \sum_{i=1}^{m-1} x_i} + \sum_{i=1}^{m-1} \frac{1}{1 + x_i} \right\} \\
\text{s.t.} & \quad \ln \left( 1 + \nu - \sum_{i=1}^{m-1} x_i \right) + \sum_{i=1}^{m-1} \ln(1 + x_i) = \nu - c_m \\
& \quad x_i \geq 0, \quad i = 1, \ldots, m \\
& \quad \nu - \sum_{i=1}^{m-1} x_i \geq 0
\end{align*}
\]

which has been already solved: set \( \Delta = m \) in Problem in (11) and take the critical point that guarantees negative definiteness, i.e., \( \gamma = \gamma_2 \). Since \( c_m \in \mathcal{D}_m \), it results that

\[
\nu - c > (m - 1) \ln \left( 1 + \frac{\nu}{m - 1} \right) = \varphi(0)
\]

(see Figure 4 with \( \Delta = m \)) and then \( \gamma_2 \in (0, 1/m] \); this value satisfies the inequality constraints. Therefore the solution is

\[
\rho = \nu \left( \frac{1 - \gamma_2}{m - 1} \cdots \frac{1 - \gamma_2}{m - 1} \gamma_2 \right)^T
\]

for \( \gamma_2 \in (0, 1/m] \), which give rise to the curves \( \varphi_m(t), t \in [0, 1/m] \), in (5).

**APPENDIX B**

**CONSTRAINED TRANSMIT ENERGY**

From (3a), \( \text{D}(\varphi_1||\varphi_0) \) is Schur-convex in \( \rho \). Therefore, letting \( t = \sum_{i=1}^{\Delta} \rho_i \), it results that

\[
t - \Delta \ln(1 + t/\Delta) \leq \text{D}(\varphi_1||\varphi_0) \leq t - \ln(1 + t).
\]

The lower bound is achieved at \( \rho = (t/\Delta \cdots t/\Delta)^T \) while the upper bound at \( \rho = (0 \cdots 0)^T \). Evaluating also \( \text{D}(\varphi_0||\varphi_1) \) on this two points, for \( t \geq 0 \), the curves \( \varphi_{\text{up}} \) and \( \varphi_{\text{low}} \) in (6) are obtained.

In order to find points \( A \) on \( \varphi_{\text{up}} \) and \( D \) on \( \varphi_{\text{low}} \), the maximum value for the constant \( t \) must be found under the transmit energy constraint. In particular, for point \( D \), the problem to be solved is

\[
\begin{align*}
\max & \quad \lambda_1(M^{-1/2}CCLC^H M^{-1/2}) \\
\text{s.t.} & \quad \text{tr}(CC^H) \leq \mathcal{E}
\end{align*}
\]
The maximum is \( \mathcal{E}\lambda_i(L)/\lambda_N(M) \), and it is obtained with the code matrix in (9). As to point \( A \), the problem is

\[
\max \ t \\
\text{s.t.} \quad \lambda_i(M^{-1/2}CLC^HM^{-1/2}) = \frac{t}{\Delta},
\]

for \( i \in \{1, \ldots, \Delta\} \)

\[
\text{tr}(CC^H) \leq \mathcal{E}.
\]

Let \( n = \text{rank}(L) \), \( V_L \in \mathbb{C}^{n \times n} \) be the matrix containing the first \( n \) columns of \( V_L \), and \( T = V_L \Xi^V_L^H \), where \( \Xi \in \mathbb{R}^{n_t \times n_t} \) is diagonal, and

\[
\Xi_{ii} = \begin{cases} 
1/\lambda_i(L), & \text{if } i \leq n \\
1/\lambda_n(L), & \text{if } n < n_t \text{ and } i > n.
\end{cases}
\]

From the constraint, \( \text{rank}(C) \) must be equal to \( m^* \), and, for \( i = 1, \ldots, m^* \), it results

\[
\lambda_i(C^HC) \geq \lambda_i(V_L^H C^H C V_L)
\]

\[= \lambda_i(V_L^H V_L C^H C V_L V_L^H)
\]

\[= \lambda_i(T^{1/2}/L^{1/2}C^H C L^{1/2}/T^{1/2})
\]

where the inequality follows from [21, Corollary 4.3.16]. This implies

\[
\text{tr}(CC^H) \geq \text{tr}(L^{1/2}C^H C L^{1/2} T)
\]

\[\geq \sum_{i=1}^{n_t} \lambda_i(C^H C L^{1/2} T) \lambda_{n_t-i+1}(T)
\]

\[= \sum_{i=1}^{m^*} \lambda_i(C^H C L^{1/2} T) \lambda_{i+1-n_t}(T)
\]

\[\geq \sum_{i=1}^{m^*} \frac{\lambda_i(C^H C L^{1/2} T)}{\lambda_i(L)}
\]

(16)

where the second inequality follows from [22, Theorem 9.H.1.h]. Since the constraint imposes that

\[
M^{-1/2}CLC^HM^{-1/2} = (t/\Delta)WW^H
\]

for some \( W \in \mathbb{C}^{n \times \Delta} \) such that \( W^H W = I_{\Delta} \), it results that

\[
\lambda_i(C^H C L^{1/2} T) \geq \lambda_i(M^{1/2}W^{1/2} M^{-1/2} W^{1/2})
\]

\[= \lambda_i(W^{1/2} M W)
\]

\[\geq \frac{\lambda_N - \Delta + 1}{\Delta} \lambda_{N - \Delta + 1}(M)
\]

(17)

for \( i = 1, \ldots, \Delta \), where the inequality follows from [21, Corollary 4.3.16]. Finally, (16) and (17) give the lower bound

\[
\text{tr}(CC^H) \geq \sum_{i=1}^{\Delta} \frac{\lambda_N - \Delta + 1(M)}{\lambda_i(L)}.
\]

Since \( \text{tr}(CC^H) \leq \mathcal{E} \), the maximum value for \( t \) is

\[
t^* = \sum_{i=1}^{\Delta} \frac{\Delta \mathcal{E}}{\lambda_N - \Delta + 1(M)}/\lambda_i(L)
\]

and the code matrix that achieves \( t^* \) is that given in (7)-(8). Finally, (16) and (17) give the lower bound

\[
\text{tr}(CC^H) \leq \mathcal{E}.
\]

Finally point \( C \) and \( \varphi_{\text{hull}} \) are obtained solving Problem 1 for all \( \alpha \in [0, 1] \). Indeed, letting \( \alpha D^*(f_0, f_0) \) be the solution of Problem 1 for fixed \( \alpha \) and \( c^* \) be the corresponding optimal value of the objective function, then \( \alpha D^*(f_0, f_0) + (1-\alpha) D^*(f_0, f_1) = c^* \) is the tangent line of the achievable region in the point \( \alpha D^*(f_0, f_0) \).

REFERENCES


