Scaled Canonical Coordinates for Compression and Transmission of Noisy Sensor Measurements

Yuan Wang, Haonan Wang, and Louis L. Scharf, Life Fellow, IEEE
Colorado State University, Fort Collins, CO 80523

Abstract—This paper is motivated by sensing and wireless communication, where data compression or dimension reduction may be used to reduce the required communication bandwidth. High-dimensional measurements are converted into low-dimensional representations through linear compression. Our aim is to compress a noisy sensor measurement, allowing for the fact that the compressed measurement will then be transmitted over a noisy channel. We give the closed-form expression for the optimal compression matrix that minimizes the trace or determinant of the error covariance matrix. We show that the solutions share a common architecture consisting of a canonical coordinate transformation, scaling by coefficients which account for canonical correlations and channel noise variance, followed by a coordinate transformation into the sub-dominant invariant subspace of the channel noise.

I. INTRODUCTION

In a distributed sensor network, one can precompress observations to lower-dimensional measurements before transmitting them around the network or to a fusion center. Such data compression or dimension reduction reduces the communication burden, but increases mean-squared error and reduces information rate. In this paper, we are interested in designing the linear compression matrix that minimizes the mean-square error or maximizes the information rate at the optimal compression ratio, under a power constraint.

Fig. 1. Linear compression of a noisy measurement \( x \) with channel noise \( v \).

The diagram of Fig. 1 frames the problem of interest in this paper. In this figure, \( \theta \in \mathbb{R}^p \) is a signal of interest. The signal \( \theta \) is carried through a sensor by a linear transformation \( H \in \mathbb{R}^{n \times p} \) and then observed as the noisy and transformed measurement \( x = H\theta + u \in \mathbb{R}^n \). This noisy measurement \( x \) is to be compressed with the linear transformation \( W \in \mathbb{R}^{m \times n} \) and then transmitted through a noisy channel. The channel transforms the measurement by a channel matrix \( D \in \mathbb{R}^{m \times m} \) and adds noise to produce a measurement \( z \in \mathbb{R}^m \). Our goal is to design the compressor \( W \) so that the noisy and compressed measurement \( z \) may be processed for an estimator of the signal \( \theta \) whose error covariance has minimum trace or minimum determinant. The minimum trace solution minimizes mean-squared error of the estimate and the minimum determinant solution minimizes volume of the error concentration ellipsoid. In the Gaussian case, it maximizes differential information rate.

We are not the first to consider this problem and its variants. In fact as we will show, this paper is an extension of the original work of Schizas, Giannakis, and Luo [1]. The innovation of this paper is this. First we replace the mean-squared error criterion of [1] with a maximum information rate criterion, and second we show that our designs and theirs may be cast as scaled and rotated canonical coordinate designs. This finding is important, for it generalizes the theory of canonical coordinates to a much more general class of problems than the class for which they were originally designed [12] and subsequently applied [3]-[6]. The maximum information rate designs of this paper require a different proof technique than the proof technique of Schizas, et al. [1].

Let us place our work in the context of prior art, by again making reference to Fig. 1. The problem addressed by Schizas, Giannakis, and Luo [1] is to design the compression matrix \( W \) so that the measurement \( z \) may be filtered to produce a minimum mean-squared error estimate of the signal \( \theta \). We generalize this problem to the maximization of information rate and show that canonical coordinates are central to both criteria. The solutions of [1] and this paper generalize the work on minimum mean-squared error and maximum information rate designs for reduced-rank filtering [2]-[6] and for precoders and equalizers (e.g., [7], [8] and many others). The result of [10] is a special case of pre-coding and equalizing.

So we may summarize by saying that the theory of canonical coordinates treats the problem of compression when there is noise at the input to the compressor and the theory of scaled and rotated canonical coordinates developed in this paper treats the problem of compression when there is noise at the input and the output of the compressor. Noise at the output brings an important element of design to the compression problem, for it forces a constraint on the power out of the compressor \( W \), a constraint that leads to rather complicated reasoning about Lagrangians and the KKT conditions for optimality, as for example in the prior work of [1], [7], and [8].

The rest of this paper is organized as follows. In Section II, we briefly introduce the problem of interest. In Section III, in the channel-noise-free case, the compression matrix returns half canonical coordinates for trace minimization, and full canonical coordinates for determinant minimization. In Sec-

This work was supported in part by NSF grant CFF-1018472, AFOSR contract FA-9550-10-1-0241, and DMS-1106975.
tion IV, when the compressed measurement is transmitted over a noisy channel, the compression matrix for trace or determinant minimization returns a scaled and rotated canonical coordinate design. Moreover, the scaling matrix, which accounts for canonical correlations and channel noise variance, has a mercury/waterfilling interpretation. Section V concludes the paper.

Notation. The set of length m real vectors is denoted by $\mathbb{R}^m$ and the set of $m \times n$ real matrices is denoted $\mathbb{R}^{m \times n}$. Bold upper case letters denote matrices, boldface lower case letters denote column vectors, and italics denote scalars. The scalar $x_i$ denotes the ith element of vector $x$, and $x_{ij}$ denotes the element of $X$ at row $i$ and column $j$. The diagonal matrix with diagonal elements $x$ is denoted $\text{diag}(x)$. The $n \times n$ identity matrix is denoted $I_n$, and the $m \times n$ matrix with elements all 0 is denoted by $0_{m \times n}$. The transpose, inverse, pseudo inverse, trace and determinant of a matrix are denoted by $(\cdot)^T$, $(\cdot)^{-1}$, $(\cdot)^+$, $\text{tr}(\cdot)$ and $\text{det}(\cdot)$, respectively.

A covariance matrix is denoted by bold upper case $Q$ with specified subscripts: $Q_{zz}$ denotes the covariance matrix of a random vector $z$; $Q_{zi}$ is the cross-covariance matrix between $z_i$ and $z_j$; and $Q_{zi}^+$ is the error covariance for the best linear unbiased estimator of $z_i$ from $z_j$. Moreover, $Q_{zi}^{1/2}$ is defined from the Cholesky decomposition $Q_{zi} = Q_{zi}^{1/2}Q_{zi}^{1/2}$. In this paper, we assume that the covariance matrices $Q_{\theta \theta}$, $Q_{xx}$, $Q_{xv}$ and the cross-covariance matrix $Q_{\theta x}$ are known.

In practice, the covariance matrices are determined from the physics of a problem or estimated from a two-channel experiment that generates realizations of $(\theta, x)$. Only the second moment are required, not the exact distribution of the random signals.

The performance of the compression matrix is determined by evaluating functions of the resulting error covariance $Q_{ee}$. See [13] for more detailed review and discussion of candidate functions. In Sections III and IV, we will focus on two classical criteria: $\text{tr}(Q_{ee})$ and $\text{det}(Q_{ee})$. The first measure $\text{tr}(Q_{ee})$ is the mean squared error of $\hat{\theta}_e$. This criterion has been studied by Schizas, et al. [1]. The second measure $\text{det}(Q_{ee})$ is the volume of the error concentration ellipsoid.

When $z$ and $\theta$ are jointly Gaussian distributed, minimizing $\text{det}(Q_{ee})$ is equivalent to maximizing the mutual information between $z$ and $\theta$, or the differential information rate at which measurement $z$ brings information about $\theta$ [11]. For simplicity, let us refer to the problems where we try to minimize $\text{tr}(Q_{ee})$ and $\text{det}(Q_{ee})$ as the min-trace and min-det problems, respectively.

II. PROBLEM STATEMENT

Suppose that $\theta \in \mathbb{R}^p$ is a random signal of interest. Consider the linear model, as depicted in Fig. 1,

$$
x = H\theta + u
$$

$$
z = DWx + v. \quad (1)
$$

Here $x \in \mathbb{R}^m$ is a noisy measurement of $\theta \in \mathbb{R}^p$, $W \in \mathbb{R}^{m \times n}$ ($m \leq n$) is the compression matrix, and $Wx$ is the signal to be transmitted over a noisy channel with a full-rank channel matrix $D \in \mathbb{R}^{m \times m}$ and random noise $v \in \mathbb{R}^m$. Note that the dimension of the signal $Wx$ is smaller than that of the original signal $x$. It is assumed that the channel noise $v$ has mean 0, and is independent of $\theta$, $u$ and $x$. Our objective is to design the compression matrix $W$ such that the compressed measurement $z$ is optimal according to a pre-specified performance metric.

In this paper, we use linear estimation which is optimal in the multivariate normal case. In particular, given a measurement $z$, the best linear unbiased estimator (BLUE) of $\theta$ is

$$
\hat{\theta}_z = \mu_\theta + Q_{\theta z}Q_{zz}^+(z - \mu_z),
$$

where $\mu_\theta, \mu_z$ are the means of $\theta$ and $z$ respectively, and $Q_{zz}^+$ is the pseudo inverse of $Q_{zz}$. The error covariance matrix of $\hat{\theta}_z$, denoted by $Q_{ee}$, is

$$
Q_{ee} = E[(\theta - \hat{\theta}_z)(\theta - \hat{\theta}_z)^T] = Q_{\theta \theta} - Q_{\theta z}Q_{zz}^+Q_{z \theta}.
$$

Under model (1), $Q_{ee}$ can also be written as a function of $W$; that is,

$$
Q_{ee} = Q_{\theta \theta} - Q_{\theta z}W^TD^T \times \left(DWQ_{xx}W^TD^T + Q_{xv}\right)^{-1}DWQ_{z \theta}. \quad (2)
$$

In this paper, we choose to re-use the variables $F, K$, and $G$ for both SVDs.

III. CHANNEL-NOISE-FREE COMPRESSION DESIGN

When there are no channel effects, then $z = Wx$, and the error covariance matrix is

$$
Q_{ee} = Q_{\theta \theta} - Q_{\theta z}W^TW^T - WQ_{z \theta}. \quad (3)
$$

The solutions of the min-trace and min-det problems can be obtained by directly applying the results on optimal reduced-rank filtering [2]-[6].

The basic idea is to transfer $(\theta, x)$ to canonical coordinates $(\hat{\theta}, \hat{x})$ which have a diagonal cross-covariance matrix. For the min-trace problem, we consider the Singular Value Decomposition (SVD) of the half coherence matrix ([2], [5], [6])

$$
Q_{\theta z}Q_{x \theta}^{-1/2} = FKG^T, \quad (4)
$$

where $K \in \mathbb{R}^{p \times n}$ is a diagonal matrix with diagonal elements $k_1 \geq \ldots \geq k_{\min(p,n)} \geq 0$, and $F \in \mathbb{R}^{p \times p}$ and $G \in \mathbb{R}^{n \times n}$ are orthogonal matrices. The vectors $\hat{\theta} = F^T\theta$ and $\hat{x} = G^TQ_{xx}^{-1/2}x$ are the half canonical coordinates for $\theta$ and $x$ respectively. Note that the cross-covariance matrix between $\hat{\theta}$ and $\hat{x}$ is the diagonal matrix $K$ given in (4).

For the min-det problem, the choice of the canonical coordinates is different. In this case, we consider an SVD of the coherence matrix ([3]-[6], [12])

$$
Q_{\theta z}^{-1/2}Q_{\theta z}Q_{x \theta}^{-1/2} = FKG^T, \quad (5)
$$

where $K \in \mathbb{R}^{p \times n}$ is a diagonal matrix with diagonal elements $k_1 \geq \ldots \geq k_{\min(p,n)} \geq 0$, and $F \in \mathbb{R}^{p \times p}$ and $G \in \mathbb{R}^{n \times n}$ are orthogonal matrices. Now, the vectors $\hat{\theta} = F^TQ_{\theta z}^{1/2}\theta$ and $\hat{x} = G^TQ_{xx}^{-1/2}x$ are the full canonical coordinates of $\theta$ and $x$ respectively. Note that for the simplicity of our notation, we choose to re-use the variables $F, K$, and $G$ for both SVDs.
The optimal compression matrix is given in Proposition 1, which is a re-statement of the results of [2] and [5].

**Proposition 1:** For the min-trace and min-det problems, the optimal compression matrix \( W_0^* \in \mathbb{R}^{m \times n} \) can be written as

\[
W_0^* = G_m^T Q_{xx}^{-1/2}
\]

where \( G_m \) consists of the first \( m \) columns of \( G \). The matrix \( G \) is defined in (4) for the min-trace problem and in (5) for the min-det problem. Moreover, for any \( m \times m \) nonsingular matrix \( T \), \( T W_0^* \) is also an optimal compression matrix.

Proposition 1 figures prominently in our derivation of scaled and rotated canonical coordinates for optimum compression with channel noise. It is also worth mentioning that \( W_0^* x \) returns the first \( m \) canonical coordinates in \( x \).

**IV. COMPRESSION DESIGN WITH SENSOR NOISE AND CHANNEL NOISE**

Now we extend the results in Section III by considering the linear compression of the noisy measurement to be transmitted over a noisy channel. We assume the channel noise \( v \) has mean zero and covariance matrix \( Q_{vv} \), and \( v \) is independent of \( \theta \) and \( x \). A significant feature of the design for noisy transmission is the need for a **power constraint** on the compression matrix, for otherwise the design problem is not well-defined.

In this paper, we restrict the compression matrix \( W \) subject to \( \text{tr}(W Q_{xx} W^T) \leq P \) for some pre-specified constant \( P \).

Define \( Q_{\omega \omega} = D^{-1}Q_{vv}(D^{-1})^T \) with the eigendecomposition \( Q_{\omega \omega} = U_{\omega} \Sigma_{\omega} U_{\omega}^T \), where \( U_{\omega} \) is an \( m \times m \) orthogonal matrix and \( \Sigma_{\omega} \in \mathbb{R}^{m \times m} \) is a diagonal matrix with diagonal elements \( 0 \leq \sigma_{\omega,1}^2 \leq \cdots \leq \sigma_{\omega,m}^2 \).

**A. Min-Trace Compression with Channel Noise**

Under the power constraint, Schizas, et al. [1] have derived the optimal compression matrix to minimize \( \text{tr}(Q_{\omega \omega}) \). In Theorem 1, we re-state their result as a scaled and rotated canonical coordinate design.

**Theorem 1:** An optimal compression matrix \( W_{tr}^* \) minimizing \( \text{tr}(Q_{\omega \omega}) \) is given by

\[
W_{tr}^* = U_{\omega} \Sigma_{tr}^* G^T Q_{xx}^{-1/2}
\]

Here the matrices \( G \) and \( K \) are given in (4), \( \Sigma_{tr}^* \) is an \( m \times n \) diagonal matrix with diagonal elements

\[
\sigma_{ii}^* = \begin{cases} \sqrt{k_i \sigma_{\omega,i} \sum_{\omega,i} - \sigma_{\omega,i}^2} & i = 1, \ldots, \kappa \\ 0 & i = \kappa + 1, \ldots, m, \end{cases}
\]

with \( \kappa \) the maximum integer between 1 and \( \text{rank}(K) \) such that \( \sigma_{ii}^2 > 0 \) for \( i = 1, \ldots, \kappa \), and

\[
\mu = \left( 1 + \sum_{i=1}^{\kappa} \sigma_{ii}^2 - \sum_{i=1}^{\kappa} \sigma_{ii} k_i \right)^2.
\]

It can be seen that \( W_{tr}^* \) factors into whitening \( Q_{xx}^{-1/2} \), canonical coordinate transformation \( G^T \), scaling \( \Sigma_{tr}^* \) and rotation \( U_{\omega} \) into the sub-dominant invariant subspace of \( Q_{\omega \omega} \).

**B. Min-Det Compression with Channel Noise**

The optimal compression matrix \( W \) to minimize \( \det(Q_{ee}) \) under the power constraint solves the optimization problem,

\[
W_{det}^* = \arg \min_{W \in \mathbb{R}^{m \times n}} \det(Q_{ee}) \text{ subject to } \text{tr}(W Q_{xx} W^T) \leq P.
\]

(9)

The matrix \( W \) has \( mn \) degrees of freedom. But let’s restrict \( W \) to a subset of \( \mathbb{R}^{m \times n} \), over which the local minimizer of \( \det(Q_{ee}) \) can be expressed explicitly. For a given \( n \times n \) orthogonal matrix \( V \), define \( \Omega_V \) = \{ \( U_{\omega} \Pi_m \Sigma_{\omega}^T V^T Q_{xx}^{-1/2} \), where \( \Pi_m \in \mathbb{R}^{m \times m}, \Sigma_{\omega} \in \mathbb{R}^{n \times n} \) are permutation matrices, and \( \Sigma \in \mathbb{R}^{m \times n} \) is diagonal with \( \sum_{i=1}^{m} \sigma_{ii}^2 \leq P \).

For fixed matrices \( U_{\omega} \) and \( Q_{xx}^{-1/2} \), the set \( \Omega_V \) is a subset of the constrained space of problem (9). Therefore, the local minimizer of \( \det(Q_{ee}) \) over \( W \in \Omega_V \) generally gives a suboptimal solution for problem (9). However, in Lemma 1, we show that, for a suitable choice of \( V \), the suboptimal solution on \( \Omega_V \) is a global optimal solution for problem (9).

**Lemma 1:** Suppose that \( G \) is the orthogonal matrix given in (5). Then,

\[
\min_{W \in \Omega_G} \det(Q_{ee}) = \det(Q_{ee}(W_{det}^*)).
\]

From Lemma 1, it can be seen that the local minimizer over \( \Omega_G \) is also a global minimizer of (9). For any \( W \in \Omega_G \), we have

\[
\det(Q_{ee}(W)) = \det(Q_{ee}(W_{det}^*) \times \det(I_n + \Pi_m^T \Gamma \Pi_n (I_n + \Sigma + \Sigma_{\omega}^T \Pi_m \Sigma^{-1} \Pi_m \Sigma)^{-1}.)
\]

(10)

Here \( \Gamma = K^T (I_p - K K^T)^{-1} K \), with \( K \) given in (5), is an \( n \times n \) positive semi-definite diagonal matrix with diagonal elements \( k_i^2 = k_i^2/(1 - k_i^2) \) for \( i = 1, \ldots, \min(n, p) \) and 0 otherwise. We require \( 0 \leq k_i < 1 \) for all \( i \). The permutation matrices \( \Pi_n \) and \( \Pi_m \) reorder the diagonal elements of \( \Gamma \) and \( \Sigma_{\omega}^{-1} \), respectively. In fact, for any \( W \in \Omega_G \), \( \Pi_n \) reorders the canonical coordinates \( G^T Q_{xx}^{-1/2} x \) and determines which \( m \) coordinates will be transmitted, and the permutation matrix \( \Pi_m \) reorders the selected coordinates and determines which subchannel the coordinate will be transmitted over. The optimal compression matrix can be obtained by minimizing \( \det(Q_{ee}(W)) \) with respect to the permutation matrices \( \Pi_m \), \( \Pi_n \) and the diagonal matrix \( \Sigma \). The computational complexity of this optimization has been greatly reduced since there are just \( 2m + n \) degrees of freedom in the permutation matrices \( \Pi_m \), \( \Pi_n \) and the diagonal matrix \( \Sigma \). We give in Theorem 2 the closed-form expression for the optimal compression matrix \( W_{det}^* \).

**Theorem 2:** The optimal compression matrix \( W_{det}^* \) solving problem (9) is

\[
W_{det}^* = U_{\omega} \Sigma_{det}^* G^T Q_{xx}^{-1/2}
\]

(11)

Here \( G \) is given in (5) where the matrix \( K \) contains singular values \( 0 \leq k_i < 1 \) for all \( i \); \( \Sigma_{det}^* \in \mathbb{R}^{m \times n} \) is a diagonal matrix.
with diagonal elements $\sigma_{11}^*, \ldots, \sigma_{mm}^*$ with
\[
\sigma_{ii}^2 = \begin{cases} 
\frac{1}{2} \sigma_{\omega,i}^2 \left( 2 - \gamma_i^2 + \sqrt{\gamma_i^4 + 4 \mu^2 \sigma_{\omega,i}^2} \right) & i \leq \kappa \\
0 & i > \kappa
\end{cases}
\] (12)
where $\kappa$ is the maximum integer between 1 and $\text{rank}(\Gamma)$ such that $\sigma_{ii}^2 > 0$ or equivalently $\sigma_{\omega,i}^2/k_i^2 < 1/\mu$ for $i = 1, \ldots, \kappa$. The value of $\mu$ is nonnegative and uniquely solves $\sum_{i=1}^\kappa \sigma_{ii}^2 = P$.

Simple calculation shows the minimum determinant of the error covariance to be
\[
\text{det}(Q_{ee}(W_{det}^*)) = \text{det} Q_{\theta \theta|x} \prod_{i=\kappa+1}^{\min\{n,p\}} \frac{1}{1-k_i^2} \times \prod_{i=1}^{\kappa} \left( 1 + \frac{2}{\sqrt{1+4(\gamma_i^2 \sigma_{\omega,i}^2/\mu)^{-1} - 1}} \right). \tag{13}
\]

The first term on the right hand side is the minimum volume of the error concentration ellipsoid with no dimension reduction; the second term scales this volume according to canonical correlations of discarded canonical coordinates; the third term scales the volume by a term that depends on the channel noise variance, the power $P$, and the full canonical correlations. The integer $\kappa$ is the number of subchannels assigned with positive power. In fact, $\kappa$ is the optimal compression rate for a given power $P$.

It is worth mentioning that, for a sufficiently large $P$, we have $1/\mu > \sigma_{\omega,i}^2/k_i^2$ (or equivalently $\sigma_{ii}^2 > 0$) for $i = 1, \ldots, m$. Consequently, the solution given in Theorem 2 is also an optimal compression for the channel-noise-free case. We simply let the nonsingular matrix $T$ in Proposition 1 be $T = U_\omega \text{diag}(\sigma_{11}, \ldots, \sigma_{mm})$. When $P$ goes to infinity, the third part in (13) goes to 1, and the minimum determinant of $Q_{ee}$ converges to the channel-noise-free case. On the other hand, the diagonal elements of $\Sigma_{det}^*$ go to infinity. Therefore, we can see that the optimization problem is ill-posed without a (finite) power constraint.

Finally, we comment on the canonical correlations. Under our current framework, all canonical correlations, $k_i$, are less than 1. In fact, the factorization in (11) still holds when $k_i = 1$ with a different scaling matrix $\Sigma_{det}^*$. In the sensor-noise-free case, suppose that $x = H\theta$ and the matrix $H \in \mathbb{R}^{n \times p}$ has rank $p$. The full canonical correlations between $\theta$ and $x$ are all 1. In this case, the compressor $W$ operates on $H\theta$ directly and the design of $W$ becomes a precoder design problem [7], [8], [10]. The optimal scaling matrix $\Sigma_{det}^*$ has diagonal elements $\sigma_{11}, \ldots, \sigma_{mm}$ such that
\[
\sigma_{ii}^2 = \begin{cases} 
\frac{1}{\mu} - \sigma_{\omega,i}^2 & \sigma_{\omega,i}^2 < 1/\mu \\
0 & \sigma_{\omega,i}^2 \geq 1/\mu
\end{cases}
\] (14)
where the value of $\mu$ is determined by the power constraint $\sum_{i=1}^{m} \sigma_{ii}^2 = P$.

C. A Mercury/Waterfilling Explanation for Min-Det Optimal Compressor

For the sensor-noise-free case, $x = H\theta$, the optimal compressor has been discussed in Section IV-B, with the factorization in (11) and $\Sigma_{det}^*$ given in (14). The scaling matrix $\Sigma_{det}^*$ distributes the power among all the $m$ subchannels according to a waterfilling policy [11]. In general, $x$ is a noisy measurement of $\theta$ and the full canonical correlations are strictly less than 1. Therefore, the optimal power allocation policy needs to be adjusted according to the canonical correlations. As a consequence of Theorem 2, the solution can be interpreted as a mercury/waterfilling policy, which is a three-step procedure that has been introduced in [15]:

1. For the $i$th vessel, fill in the solid base with height $\sigma_{\omega,i}^2/k_i^2$.

2. Compute $\mu$ from the power constraint. For the vessels with base height less than $1/\mu$, fill in mercury in the vessel until the height reaches
\[
\max \left\{ \frac{1}{\mu} - \frac{1}{2} \sigma_{\omega,i}^2 \left( -2 - \gamma_i^2 + \sqrt{\gamma_i^4 + 4 \sigma_{\omega,i}^2/\mu} \right), \frac{\sigma_{\omega,i}^2}{k_i^2} \right\}.
\]

3. Pour water into all vessels until the height of water in each vessel reaches $1/\mu$.

In this mercury/waterfilling policy, $1/\mu$ is the parameter in the formula for water volume $\sigma_{ii}^2$ that minimizes $\text{det}(Q_{ee})$ under the constraint that the total volume of water is $P$. Given the value of $\mu$, the determinant of the error covariance is minimized when the value of $\sigma_{ii}^2$ equals the height of water in the corresponding vessel.

The height of the solid base, $\sigma_{\omega,i}^2/k_i^2$, is the variance of the channel noise in the $i$th vessel divided by the $i$th squared canonical correlation. A higher solid base means a less informative channel with high channel noise and weak correlation with $\theta$. For any vessel with base height exceeding $1/\mu$, neither mercury nor water will be added, or equivalently, no power will be assigned to the corresponding subchannel.

While the base height determines whether water will be added, the mercury stage regulates the water level for each vessel. Without adding mercury, the optimal power allocation will have variable solid-plus-water levels among different vessels. The mercury is added to balance the sensor noise contained in $x$ and the channel noise added in transmission. Recall that no mercury is added in the special case when $x = \theta$.

D. Scaled and Rotated Canonical Coordinate Design

Theorems 1 and 2 suggest a common architecture for compression, which specializes to all previous designs for reduced-rank filtering and for reduced rank precoding and equalizing. The optimal compressor can be factored into four component matrices. As shown in Fig. 3, the first matrix $Q_{xx}^{-1/2}$ whitens the noisy measurement $x$. The second matrix $G^T$ transforms the whitened measurement into a canonical coordinate system. For the min-det problem, the full canonical coordinates, $G^T Q_{xx}^{-1/2} x$, are uncorrelated and have unit
variance. The third matrix $\Sigma^* \in \mathbb{R}^{m \times n}$ is diagonal. The role of $\Sigma^*$ is to extract the first $m$ full canonical coordinates and distribute power across the canonical channels. The $i$th canonical coordinate is scaled to have power $\sigma_i^2$. For the min-det problem, when $\gamma_{ii}^2 = 0$ (i.e., $k_i = 0$), the corresponding scaling is $\sigma_{ii} = 0$, which means those canonical coordinates uncorrelated or weakly correlated with $\theta$ will be eliminated. In general, the diagonal elements of $\Sigma^*$ have a mercury/waterfilling interpretation. The matrix $U_\omega$ rotates the compressed canonical coordinates into the sub-dominant invariant subspace of the matrix $Q_\omega$.

The difference between the trace and determinant designs is in the canonical coordinates and in the values of scaling constants in the diagonal scaling matrix.

V. Conclusion

In this paper we have considered the problem of compressing a noisy measurement for transmission over a noisy channel, introduced in [1]. This problem generalizes the problem of reduced rank filtering and the problem of reduced rank precoder and equalizer design [2]-[11], producing those designs as special cases. We have shown that designs for minimizing trace or determinant of an error covariance matrix share a common architecture. In this architecture, a noisy sensor measurement is first transformed into a system of canonical coordinates. These coordinates are then scaled and rotated into the sub-dominant subspace of the channel noise. The difference between the two designs resides in the definition of canonical coordinates and in the determination of the scaling constants.

REFERENCES


