OPTIMAL EXPERIMENTAL DESIGN IN CANONICAL EXPANSIONS WITH APPLICATIONS TO SIGNAL COMPRESSION

Roozbeh Dehghannasiri, Xiaoning Qian, and Edward R. Dougherty

Department of Electrical and Computer Engineering, Texas A&M University, College Station, USA
roozbehdn@tamu.edu, xqian@ece.tamu.edu, edward@ece.tamu.edu

ABSTRACT
In this paper, we introduce an experimental design framework for Karhunen-Loève compression. This method based on the concept of mean objective of uncertainty determines the best unknown parameter of the covariance matrix to be estimated first in order to improve the quality of the compressed signal. Moreover, we find the closed-form solution to the intrinsically Bayesian robust Karhunen-Loève compression that is required for experimental design and provides the optimal signal compression on average relative to the uncertainty class of covariance matrices. We verify the performance of the proposed experimental design method for the case in which the covariance matrix consists of disjoint blocks.

Index Terms— Canonical expansion, Karhunen-Loève compression, intrinsically Bayesian robust (IBR), optimal experimental design, mean objective cost of uncertainty (MOCU).

1. INTRODUCTION
In many real-world applications, model uncertainty is inherent due to systems complexity, lack of enough training data, perturbation, or noise. When there is no uncertainty, the aim is to design an optimal operator — be it filter, controller, or classifier — relative to the single perfectly known model. However, when the model knowledge is lacking, the aim becomes to design a robust operator. Roughly speaking, a robust operator shall achieve acceptable performance relative to all models compatible with the prior knowledge, these models comprising an uncertainty class. Two approaches have been used for designing robust operators: minimax and Bayesian. In the minimax approach, the goal is to design a robust operator with the best worst-case performance across the uncertainty class [1–3]. In the Bayesian approach, the robust operator achieves the best performance on average relative to a prior distribution that reflects our knowledge about uncertain parameters [4–7].

Considering the Bayesian approach, assume that there exists a class Ψ of operators such that the cost of applying an operator ψ ∈ Ψ to a model with parameter θ is characterized by a cost function ηθ(ψ). In this setting, the model-specific optimal operator is

\[ ψ(θ) = \arg \min_{ψ ∈ Ψ} η_θ(ψ), \]

and the robust operator, called an intrinsically Bayesian robust (IBR) operator, is

\[ ψ(Θ) = \arg \min_{ψ ∈ Ψ} \mathbb{E}_{θ \in Θ} [η_θ(ψ)], \]

where the expectation is taken relative to the prior distribution f(θ) when θ ∈ Θ, which denotes the uncertainty class.

Canonical expansions [8, 9] are effective tools for studying random processes particularly for applications such as data compression, signal estimation, and control. Canonical expansions have the following general form:

\[ X(t) = \mu_X(t) + \sum_{k=1}^{∞} Z_k x_k(t), \]

where \( \mu_X(t) \) is the mean function of random process \( X(t) \), \( x_k(t) \) are deterministic coordinate functions, and \( Z_k \) are uncorrelated zero-mean coefficient random variables. Canonical expansions enable us to work with a set of uncorrelated random variables instead of a sequence of correlated random variables indexed by a continuous variable \( t \).

The Karhunen-Loève (KL) expansion [10] is a widely used canonical expansion that plays a major role in data compression [11–17]. The KL-expansion, whose coefficients and coordinate functions are obtained based upon the covariance matrix, can be used for optimal signal compression. However, in many cases full knowledge of the covariance matrix might not be available. Hence, it is of interest to determine unknown parameters of the covariance matrix via additional experiments, acquiring more data by utilizing more sensors, antennas, etc. Since these operations can be expensive, it is prudent to come up with an experimental design strategy to find out which unknown parameter in the covariance matrix is better to be determined first to enhance the quality of the compressed signal.

The aim of this paper is to propose an experimental design method for uncertainty reduction in KL-compression.
The proposed method is based on the concept of mean objective cost of uncertainty (MOCU) for uncertainty quantification. MOCU [18] measures model uncertainty in terms of the expected deterioration in the operator performance, herein signal compression, caused by model uncertainty. MOCU has already been used to design objective-based experimental design methods for gene regulatory networks to enhance the performance of therapeutic intervention [19, 20]. One key step in the proposed experimental design method is to compute the intrinsically Bayesian robust (IBR) KL-compression, which provides minimum mean-square-error (MSE) on average. In this paper, we find a closed-form solution to the IBR KL-compression and show that it can be found in the same way as an ordinary KL-compression with the covariance matrix compression and demonstrate that it can be found in the same way in the context of KL-expansion along with proofs is provided in [21].

2. METHOD

Let $X = [X(1), ..., X(N)]^T$ be a discrete zero-mean random process defined over $N$ points whose covariance matrix is $K$ and $u_i = [u_i(1), u_i(2), ..., u_i(N)]^T$ and $\lambda_i$ be the $i$-th eigenvector and eigenvalue of $K$, respectively, i.e., $Ku_i = \lambda_i u_i$. Then $X(n)$ can be expressed in terms of the KL-expansion as follows [9]:

$$X(n) = \sum_{i=1}^{N} Z_i u_i(n),$$  

where $Z_i$ is the generalized Fourier coefficient of $X$ relative to $u_i$, i.e., $Z_i = X^T u_i$. Eigenvectors and eigenvalues are sorted such that $\lambda_1 \geq \lambda_2, ..., \geq \lambda_N$. Eigenvectors build an orthonormal system, meaning that for $i \neq j$, $< u_i, u_j > = 0$, $< ..., >$ being the inner product, and $\|u_i\| = 1$, $\| \|$ being the norm operator.

The main utility of the KL-expansion is in signal compression. Consider the case that we desire a compression to $m < N$ terms for $X$ as

$$X_m(n) = \sum_{i=1}^{m} Z_i u_i(n).$$

Then the MSE between the original process X and the $m$-term KL-expansion $X_m$ for a particular $n$ is obtained as

$$\text{MSE} < X(n), X_m(n) > = E \left[ \left| X(n) - X_m(n) \right|^2 \right] = \sum_{i=m+1}^{N} \lambda_i |u_i(n)|^2,$$

and the MSE over all $n$ is

$$\text{MSE} < X, X_m > = \sum_{n=1}^{N} E \left[ \left| X(n) - X_m(n) \right|^2 \right] = \sum_{i=m+1}^{N} \lambda_i. \quad (7)$$

Since the eigenvalues are decreasing, the best $m$ terms for compression are the first $m$ terms in the KL-expansion. As we select more terms, the MSE of the compression decreases. Therefore, if the full knowledge of the covariance matrix $K$ is available, then one can use it for optimal signal compression using the KL-expansion.

Now assume that the covariance matrix is not known a priori and $K^\theta$ is parameterized by $\theta = (\theta_1, \theta_2, ..., \theta_l) \in \Theta$, which corresponds to an uncertainty class of covariance matrices. We desire an experimental design method that guides us to find the unknown parameter to be determined first. As experimental design is used to reduce model uncertainty, the first step in the experimental design is to quantify uncertainty in an appropriate way.

The mean objective cost of uncertainty (MOCU) [18] quantifies model uncertainty in an objective-based manner by computing the expected difference between the best option in the absence of uncertainty (optimal operator) and the best option in the presence of uncertainty (IBR operator). Regarding compression to $m$ terms, let $X^\theta$, $X_m^\theta$, and $X_m^\theta$ be the random process corresponding to covariance matrix $K^\theta$, the $m$-term model-specific KL-expansion relative to $K^\theta$, and the $m$-term IBR KL-expansion that achieves the minimum average MSE relative to the prior distribution $f(\theta)$, respectively. Then, MOCU can be computed as

$$M(\Theta) = E_\theta \left[ \text{MSE} < X^\theta, X_m^\theta > - \text{MSE} < X^\theta, X_m^\theta > \right]. \quad (8)$$

As seen in (8), to compute MOCU, we need to compute the $m$-term IBR KL-expansion $X_m^\theta$. To do this, we find a covariance matrix $K^\theta$ that can represent the uncertainty class $\Theta$ and can be used to find $X_m^\theta$. To find such a $K^\theta$, first we should compute the MSE between a given process $X$ and a compressed process $X_m'$ obtained using an arbitrary covariance matrix $K'$. Note that $K'$ is not necessarily equal to the covariance matrix $K$ of process $X$. Therefore, $X_m'$ is obtained as

$$X_m'(n) = \sum_{i=1}^{m} Z_i'^* u_i(n),$$

where $u_i'$ is the eigenvector of $K'$ and $Z_i'$ is the generalized Fourier coefficient of $X$ relative to $u_i'$, i.e., $Z_i' = X^T u_i'$. The following theorem, which is proved in [21], paves the way for finding the IBR KL-compression.

**Theorem 1** If $X(n)$ and $X_m'(n)$ are defined as in $(4)$ and $(9)$, respectively, then

$$\text{MSE} < X, X_m' > = \sum_{i=1}^{N} \lambda_i - \sum_{i=1}^{m} (u_i')^T K u_i'. \quad (10)$$
Note that if $K = K'$, then $(u'_i)^T K u'_i = \lambda_i$ and (10) reduces to (7). Now we use (10) to calculate the average MSE of an $m$-term expansion obtained using an arbitrary covariance matrix $K'$ across the uncertainty class $\Theta$ as follows:

$$E_\Theta \left[ \text{MSE} \left< X^\theta, X'_m \right> \right] = \sum_{i=1}^{N} E_\Theta \left[ \lambda_i^\theta \right] - \sum_{i=1}^{m} (u'_i)^T E_\Theta \left[ K^\theta \right] u'_i. \quad (11)$$

To obtain the IBR KL-compression, we need to solve the following constrained optimization problem to find the $u'_i$ that minimizes (11):

$$\text{minimize} \quad E_\Theta \left[ \text{MSE} \left< X^\theta, X'_m \right> \right]
\text{subject to} \quad ||u'_i|| = 1, \quad i = 1, \ldots, m,$$

and denote the solution by $u^\theta_i$. We utilize the method of Lagrange multipliers to solve this optimization problem. The Lagrangian for this optimization is

$$L(u'_i, \zeta_i) = \sum_{i=1}^{N} E_\Theta \left[ \lambda_i^\theta \right] - \sum_{i=1}^{m} (u'_i)^T E_\Theta \left[ K^\theta \right] u'_i
- \sum_{i=1}^{m} \zeta_i \left( <u'_i, u'_i> - 1 \right). \quad (13)$$

Setting $\frac{\partial L(u'_i, \zeta_i)}{\partial u'_i(n)} = 0$ yields the following relation for $u^\theta_i$ that minimizes the constrained optimization in (12):

$$\sum_{n_1=1}^{N} E_\Theta \left[ K^\theta(n, n_1) \right] u^\theta_i(n_1) = \zeta_i u^\theta_i(n). \quad (14)$$

The relation in (14) suggests that the Lagrange multiplier $\zeta_i$ and $u^\theta_i$ are in fact the $i$-th eigenvalue and eigenvector of the expected covariance matrix $E_\Theta [K^\theta]$, denoted by $K^\theta$ hereafter. In other words, the IBR KL-expansion is found similarly to a standard KL-compression for a known covariance matrix, except that the covariance matrix $K$ is replaced by the expected covariance matrix $K^\theta$. It should be noted that this correspondence, in which the covariance matrix is replaced by the expected covariance matrix, is analogous to the case that the power spectra of the underlying processes are unknown and the IBR Wiener filter is obtained as a Wiener filter relative to the effective power spectra [4].

Returning to experimental design, we are interested in finding the best parameter among $\theta_1, \theta_2, \ldots, \theta_l$ to be estimated first. If the unknown parameter $\theta_j$ is assumed to be $\hat{\theta}_j$, then the remaining MOCU given $\theta_j = \hat{\theta}_j$, denoted by $M(\theta; \hat{\theta}_j)$, can be computed similarly to (8) for the remaining uncertainty class $\Theta|\hat{\theta}_j = \hat{\theta}_j$ obtained by assigning $\theta_j = \hat{\theta}_j$ in the models inside the uncertainty class and governed by the conditional distribution $f(\theta|\hat{\theta}_j = \hat{\theta}_j)$. Taking the expectation of $M(\Theta; \theta_j = \hat{\theta}_j)$ relative to the marginal distribution of parameter $\theta_j, f(\hat{\theta}_j)$, we find the expected remaining MOCU $M(\Theta; j)$ given parameter $\theta_j$ is determined by

$$M(\Theta; j) = E_{\hat{\theta}_j} \left[ M(\Theta; \theta_j = \hat{\theta}_j) \right]. \quad (15)$$

In fact, $M(\Theta; j)$ is the MOCU expected to remain after determining $\theta_j$. The parameter $\theta_j^*$ minimizing (15) is suggested to be determined first:

$$j^* = \arg \min_{j=1,2,\ldots,l} M(\Theta; j). \quad (16)$$

Using (7) that implies $\text{MSE} < X^\theta, X'_m > = \sum_{i=m+1}^{N} \lambda_i^\theta$, and (11) with $X'_m$ and $u^\theta_i$ in place of $X'_m$ and $u'_i$, respectively, it can be shown that the minimization problem in (16) reduces to

$$j^* = \arg \max_{j=1,2,\ldots,l} E_{\hat{\theta}_j} \left[ \sum_{i=1}^{m} \lambda_i^{\theta|\hat{\theta}_j = \hat{\theta}_j} \right]. \quad (17)$$

Thus, to evaluate which unknown parameter $\theta_j$ to determine first, we need to compute the conditional expected covariance matrix $K^{\theta|\hat{\theta}_j = \hat{\theta}_j} = E_{\theta|\hat{\theta}_j = \hat{\theta}_j} [K^\theta]$ for each possible realization $\hat{\theta}_j$ of the parameter, where the expectation is taken relative to $f(\theta|\theta_j = \hat{\theta}_j)$, and then obtain its eigenvalues $\lambda_i^{\theta|\hat{\theta}_j = \hat{\theta}_j}$.

### 3. Simulation Results

Assume that the true values of parameters $\theta = (\theta_1, \ldots, \theta_l)$ are $\mu = (\mu_1, \ldots, \mu_l)$. Determining the value of $\theta_j^*$ results in a smaller uncertainty class $\Theta|\theta_j^* = \mu_j^*$, which contains all covariance matrices in $\Theta$ whose parameter $\theta_j^*$ is $\mu_j^*$. To evaluate the effectiveness of the chosen parameter, we compute

$$\text{MSE} < X^\mu, X'_m|\theta_j^* = \mu_j^* > = \sum_{i=1}^{N} \lambda_i^\mu - \sum_{i=1}^{m} (u'_i|\theta_j^* = \mu_j^*)^T K^{\mu|\theta_j^* = \mu_j^*} u'_i. \quad (18)$$

which is the MSE between the $m$-term IBR KL-expansion obtained for the remaining uncertainty class $\Theta|\theta_j^* = \mu_j^*$ and the uncompressed signal with the true covariance matrix $K^\mu$. In this equation, $\lambda_i^\mu$ is the $i$-th eigenvalue of the true covariance matrix and $u'_i|\theta_j^* = \mu_j^*$ is the $i$-th eigenvector of $K^{\theta|\theta_j^* = \mu_j^*}$.

In our simulations, we consider a covariance matrix with disjoint blocks, in which there is no correlation between the random variables across different blocks. Among other applications, this model has been used to study feature selection in systems where blocks correspond to uncorrelated subsystems [22], for instance, in genomics, where each block represents the correlation between genes within a pathway and
genes from different pathways are assumed to be uncorrelated. We consider the following class of covariance matrices:

\[
\begin{bmatrix}
\sigma_1^2 & \rho_1 \sigma_1 \sigma_2 & 0 & 0 & 0 \\
\rho_1 \sigma_1 \sigma_2 & \sigma_2^2 & \rho_2 \sigma_2^2 & \rho_2 \sigma_2^2 & 0 \\
0 & \rho_2 \sigma_2^2 & \sigma_2^2 & \rho_2 \sigma_2^2 & 0 \\
0 & \rho_2 \sigma_2^2 & \rho_2 \sigma_2^2 & \sigma_2^2 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix},
\]

(19)
in which \(\sigma_1^2, \rho_1, \sigma_2^2, \rho_2\) are unknown and we are interested in finding which one is better to be determined first using the proposed experimental design method. Assume that the aim is to compress five random variables to one random variable, i.e., \(m = 1\). Fig. 1 presents the simulation results for the model in (19) when the unknown parameters are independent, uniformly distributed, and have the following nominal uncertainty intervals: \(\sigma_1^2 \in [0.1, 4], \rho_1 \in [-0.3, 0.3], \sigma_2^2 \in [0.1, 3]\), and \(\rho_2 \in [0.01, 0.4]\). In Fig. 1, the interval of one specific uncertain parameter changes while the other parameter intervals are fixed to the nominal intervals. The first row shows the value of the function in (17) for each parameter (the parameter with the highest value is chosen to be determined) and the second row shows the MSE as defined in (18) averaged over different assumed true covariance matrices. For example, if \(\sigma_1^2 \in [0.1, \sigma_1^2_{\text{max}}]\) and \(\sigma_1^2_{\text{max}}\) changes from 0.5 to 7, according to Fig. 1 (a), the parameter to be determined first is given by

\[
\begin{align*}
\sigma_1^2 & \quad \rho_1 \\
\sigma_1^2_{\text{max}} & \quad 1 \\
\sigma_1^2 & \quad 4.2 \leq \sigma_1^2_{\text{max}} \leq 5.3 \\
\rho_1 & \quad 5.3 \leq \sigma_1^2_{\text{max}}
\end{align*}
\]

(20)

We can see in this figure that the determination of the chosen parameter always achieves the lowest average MSE compared to the determination of other unknown parameters.

4. CONCLUSION

In this paper, we proposed an experimental design method for KL-compression based on the concept of mean objective cost of uncertainty. We emphasize the significance of finding the intrinsically Bayesian robust operator required for the experimental design framework. It should be noted that a detailed study of the proposed experimental design method for KL-compression can be found in [21] when Wishart priors are also considered for the unknown covariance matrix.
5. REFERENCES


