Sparse Sensing with Low Precision for Multi-model System: An Information Theoretic Approach

Brijesh Kumar[1] and Mani Bhushan[2]

Abstract— **In this work, we propose symmetric Kullback Leibler divergence based sparse sensing with low precision for estimating the states of a multi-model system. Here, we consider the Kalman filter as an estimator for state estimation, with sensor precisions treated as decision variables. The resultant optimization formulation is Non-Linear Programming (NLP) problem. Solving this may result in a local optimal solution. Hence, we propose an equivalent reformulation of the problem as a Convex Programming (CP) problem using Semidefinite Programming. The global optimal solution can then be obtained by solving the resulting CP problem. The optimal sensor precision is obtained by minimizing the l_1 norm while ensuring the specified achievable performance bound is satisfied. The l_1 norm promotes sparsity in the solution. The utility of the proposed approach is demonstrated using the quadruple-tank process. The results show that sensor precision obtained from a single linearized model may lead to a suboptimal design compared with a multi-model approach.**

Keywords: **Sparse sensing, Sensor precision, Kalman filter, Convex programming problem, Quadruple tank process**

I. INTRODUCTION

The effectiveness of estimation and control system critically depends upon the information collection architecture [12]. In particular, information architecture includes (i) Instrument type: Selection of instruments like actuators and sensors for a system-specific task, (ii) Instrument precision: Optimal selection of signal-to-noise ratio, of measurement instrument ([5], [13]), and (iii) Instrument location: Optimal selection of variables to be manipulated or measured ([7], [9]). Here, we restrict ourselves to estimation with optimal design of instrument precision. In many processes and systems, it is not feasible to measure all variables or states because of resource and budget constraints [10]. Also, sensors are noisy and may not reflect true value of measured variables. As a result, the unmeasured states must be estimated and estimates of measured states can be improved using state observers or estimators such as the Luenberger observer [17], the Kalman filter [4], or the H_∞ filter [5]. Among these methods, the Kalman filter is the most widely used because it is relatively easy to implement and provides an unbiased optimal estimate under certain assumptions [15]. The quality of the state estimates obtained by the Kalman filter depends on the precision of the sensors used to measure the system states [4].

Higher sensor precision leads to more accurate state estimation. However, high-precision sensors require sophisticated manufacturing and advanced technology, making them more expensive and increasing the overall capital cost [5]. Furthermore, high-precision sensors can disturb the measurements of other sensors, such as RADAR systems, which may lead to degraded estimation performance [16]. Due to these limitations, it is desirable to minimize required sensor precision while ensuring that user specified estimation criteria are satisfied. In particular, the performance of Kalman filter is characterized by the covariance of the estimation error [15]. More commonly, in estimation problems, the literature employs various alphabetical optimality criteria such as A, D, E, and T optimality to characterize estimation accuracy ([4], [7]). These alphabetical optimality criteria have the following limitations [9]: (i) They rely only on the covariance of estimates rather than the full probability distribution. Hence, a given optimality criterion will yield the same performance of estimates even if density functions of estimates are different with identical covariance, (ii) Different optimality criteria may lead to different performance of estimates. Hence, the choice of the optimality criteria to be used for characterizing estimation performance is not obvious. In recent estimation literature, the Symmetric Kullback Leibler Divergence (SKLD) has been used, which helps to address the limitations mentioned above ([10], [14]).

Most real-world systems are described by nonlinear dynamical models. Although the literature on sensor placement design, optimization, estimation, and control is well established within linear system theory, it is comparatively less developed for nonlinear systems. A common practice is to linearize the nonlinear model around a specific operating point. Such linearization is typically performed for various online tasks, including monitoring, control, and real-time optimization. However, when the plant operates over a wide range of operating conditions, a single linearized model may not adequately capture the system behavior ([11], [18]). In such cases, multiple linear models obtained at different linearization points can be employed to better approximate the nonlinear dynamics. This scenario is considered in this work. To the best of our knowledge, no existing studies have considered sensor precision as a decision variable for multi-model systems within an information-theoretic framework.

In the current work, we propose symmetric Kullback Leibler divergence based sparse sensing with low precision of a multi-model system, with guarantees of a specified achievable performance bound. SKLD measures the distance between two Probability Density Functions (PDFs) [8]. The

[1] Brijesh Kumar is with Department of Chemical Engineering, Indian Institute of Technology Bombay, India brijesh19@iitb.ac.in
[2] Mani Bhushan is with Department of Chemical Engineering, Indian Institute of Technology Bombay, India mbhushan@iitb.ac.in

corresponding two PDFs in our case are of estimation error, namely: (i) Design PDF: This PDF is obtained based on selected sensor precision, and (ii) Target PDF: This is user defined PDF. In this work, we consider Kalman filter for state estimation, with sensor precision treated as decision variables. The corresponding decision variables are obtained by minimizing l_1 norm while satisfying performance constraint. Here, l_1 norm promotes sparsity of sensor precision. The proposed optimization problem for sparse sensing and the low precision problem is a Non-Linear Programming (NLP) problem, which is computationally intractable. Hence, we reformulated it into a computationally tractable Convex Programming (CP) problem formulation using Semidefinite programming. Solving the CP problem leads to global optimal sensor precision. The utility of the proposed approach is demonstrated using the quadruple tank process.

The rest of the paper is organized as follows: Section II contains preliminary material relevant to this paper. Section III presents the proposed sparse sensing and low precision problem formulation and its computationally tractable reformulation. Section IV presents the implementation of the proposed approach using a case study. Finally, Section V concludes the paper.

II. Preliminaries

A. Linear dynamical system

Consider the following discrete-time invariant linear system:

$$\mathbf{x}(k+1) = \boldsymbol{\phi}\mathbf{x}(k) + \mathbf{w}(k) \tag{1}$$

where, $\mathbf{x} \in \mathbb{R}^n$ is state vector. Further, $\boldsymbol{\phi}$ and $\mathbf{w} \in \mathbb{R}^n$ are matrix characterizing state dynamics and state noise respectively. The dynamical system (Eq. 1) is integrated with the measurement model as follows:

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) \tag{2}$$

where, $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{v} \in \mathbb{R}^n$ and $\mathbf{C} \in \mathbb{R}^{n\times n}$ are the measurement vector, measurement noise vector and measurement matrix. Further, we assume $\mathbf{w}$ and $\mathbf{v}$ are zero-mean and mutually independent white Gaussian stochastic state and measurement noises with covariance matrices $\mathbf{Q}$ and $\mathbf{R}$. Next, we summarize the Kalman filter for state estimation.

B. Kalman filter

The Kalman filter is a recursive filtering approach consisting of two steps [15]:

1) Prediction step: In this step, the conditional state density is propagated using dynamic model (Eq. 1) by propagating its mean and covariance as:

$$\widehat{\mathbf{x}}(k|k-1) = \boldsymbol{\phi}\widehat{\mathbf{x}}(k-1|k-1) \tag{3}$$

$$\mathbf{P}(k|k-1) = \boldsymbol{\phi}\mathbf{P}(k-1|k-1)\boldsymbol{\phi}^T + \mathbf{Q} \tag{4}$$

2) Update step: The mean and covariance of the conditional state density are updated with measurement available at k^{th} instant:

$$\widehat{\mathbf{x}}(k|k) = \widehat{\mathbf{x}}(k|k-1) + \mathbf{L}(k)\left(\mathbf{y}(k) - \mathbf{C}\widehat{\mathbf{x}}(k|k-1)\right) \tag{5}$$

$$\mathbf{P}(k|k) = \left((\boldsymbol{\phi}\mathbf{P}(k-1|k-1)\boldsymbol{\phi}^T + \mathbf{Q})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\right)^{-1} \tag{6}$$

where, $\mathbf{L}(k) = \mathbf{P}(k|k-1)\mathbf{C}^T(\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^T + \mathbf{R})^{-1}$ is Kalman gain matrix. The estimation error at k^{th} time instant is defined as:

$$\mathbf{e}(k|k) = \mathbf{x}(k) - \widehat{\mathbf{x}}(k|k) \tag{7}$$

where $\mathbf{x}(k)$ and $\widehat{\mathbf{x}}(k|k)$ are the true (unknown) state and estimated state at instant (k) respectively. $\mathbf{e}(k|k)$ is Gaussian random variable with zero mean and covariance $\mathbf{P}(k|k)$. Next, we discuss Theorem 1, which ensures that steady-state error covariance matrix does not depend on the initial condition and is given by the Riccati equation for discrete time-invariant system.

***Theorem** 1 ([1]):* When the dynamical system is time-invariant and state dynamic matrix $\boldsymbol{\phi}$ is stable, then the error covariance converges to unique limit $\mathbf{P}$ which is independent of the initial error covariance $\mathbf{P}(0|0)$ and satisfies the following steady-state Riccati equation :

$$\mathbf{P} = \left((\boldsymbol{\phi}\mathbf{P}\boldsymbol{\phi}^T + \mathbf{Q})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\right)^{-1} \tag{8}$$

C. Multi-Model system

Most real-world systems exhibit nonlinear behavior, and their dynamics can vary significantly across different operating conditions. Consequently, when a plant operates over a wide range, relying on a single linearized model may fail to capture these variations accurately, making it insufficient for reliable real-time operation, monitoring, and control. Therefore, multiple linear models at different operating points can be used to better approximate the nonlinear dynamics, which is the approach considered in this work. Accordingly, the linear time-invariant models corresponding to M linearization points are represented as:

$$\mathbf{x}(k+1) = \boldsymbol{\phi}_{(i)}\mathbf{x}(k) + \mathbf{w}_{(i)}(k) \ ,\forall i = 1,2,...,M \tag{9}$$

Furthermore, the steady-state covariance of the estimation error for the M models are given by:

$$\mathbf{P}_{(i)} = \left((\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}_{(i)}^T + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\right)^{-1}, \quad \forall i = 1,2,...,M \tag{10}$$

Next, we discuss the problem settings for the Sparse Sensor with Low Precision (SSLP) problem.

D. Problem setting for SSLP

The Kalman filter provides optimal estimates of the states of a linear dynamical system (Eq. 1). The quality of the estimate depends upon the choice of measurement as well as sensor precision. In this work, we assume that the set of measurements is available, while the sensor precision is unknown. Further, sensor precision is defined as the reciprocal of a squared measure such as power, energy, or standard deviation, of the measurement noise affecting the sensor output [5]. In particular, we consider the sensor precision as: $s_j = \frac{1}{\sigma_j^2},\ \forall j = 1, 2, ..., n$, where s_j and σ_j are the sensor precision and standard deviation of j^{th} sensor. In this work, we assume measurement noises are uncorrelated (i.e. $\mathbf{R}$ is a diagonal matrix). Now, $\mathbf{R}^{-1}$ in Eq. 8, can be re-written to reflect the error covariance of measurements corresponding to the selected choice of sensor precision as:

$$\mathbf{R}^{-1} = diag\bigg(\{s_j\}_{j=1,2,....n} \bigg) \tag{11}$$

where, $diag(.)$ represents a diagonal matrix. Further, we define sensor precision vectors as: $\mathbf{s} = [s_1, s_2, ..., s_n]^T$. The bound on the sensor precision vector can be incorporated into the optimization problem as discussed in Section III, i.e. $\mathbf{s} \le \mathbf{s}_{max}$, where $\mathbf{s}_{max}$ denotes the vector of maximum allowable precisions.

Now, the sparsity of the sensors is defined using the l_1 norm as [5]:

$$\|\mathbf{s}\|_{1,\mathbf{w}} = \sum_{j=1}^{n} w_j s_j \tag{12}$$

where, $\mathbf{w} = [w_1, w_2, ..., w_n]^T$, with $w_j > 0,\ \forall j = 1, 2, ..., n$. The SSLP problem aims to determine the low sensor precision by minimizing the l_1 norm while ensuring that the estimator performance meets a specified achievable performance level. Note that minimizing l_1 norm leads to a sparse solution for the sensor precision.

E. Symmetric Kullback Leibler Divergence

The Symmetric Kullback Leibler Divergence denoted as $SKLD(\mathcal{P}||\mathcal{Q})$ is a measure of how two Probability Density Functions (PDFs) are different from each other. It is defined as [14]:

$$SKLD(\mathcal{P}||\mathcal{Q}) = \frac{1}{2}\int_{-\infty}^{\infty} \Bigg\{ \mathcal{P}(\mathbf{z})\ ln\left(\frac{\mathcal{P}(\mathbf{z})}{\mathcal{Q}(\mathbf{z})}\right) + \mathcal{Q}(\mathbf{z})\ ln\left(\frac{\mathcal{Q}(\mathbf{z})}{\mathcal{P}(\mathbf{z})}\right) \Bigg\} d\mathbf{z} \tag{13}$$

where, $\mathcal{P}$ and $\mathcal{Q}$ are two PDFs of random variable $\mathbf{z} \in \mathbb{R}^\beta$. The symmetric KLD satisfies the following properties [14]:

1) $SKLD(\mathcal{P}||\mathcal{Q}) \ge 0$ for all $\mathcal{P}$ and $\mathcal{Q}$
2) $SKLD(\mathcal{P}||\mathcal{Q}) = 0$, if and only if $\mathcal{P} = \mathcal{Q}$
3) $SKLD(\mathcal{P}||\mathcal{Q}) = SKLD(\mathcal{Q}||\mathcal{P})$
4) $SKLD(\mathcal{P}||\mathcal{Q}) \le SKLD(\mathcal{P}||\mathcal{R}) + SKLD(\mathcal{R}||\mathcal{Q})$ for all $\mathcal{P}$, $\mathcal{Q}$ and $\mathcal{R}$, where, $\mathcal{R}$ is another PDF.

For the case when random variable $\mathbf{z}$ follow Gaussian Distribution with $\mathcal{P}(\mathbf{z}) = \mathcal{N}(\boldsymbol{\mu}_\mathcal{P}, \boldsymbol{\Sigma}_\mathcal{P})$ and $\mathcal{Q}(\mathbf{z}) = \mathcal{N}(\boldsymbol{\mu}_\mathcal{Q}, \boldsymbol{\Sigma}_\mathcal{P})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean and covariance and $\mathcal{N}(.)$ represents Gaussian distribution, then the integral in Eq. 13 has an analytical solution as:

$$SKLD(\mathcal{P}||\mathcal{Q}) = \frac{1}{4}\Big[tr\{\boldsymbol{\Sigma}_\mathcal{Q}\boldsymbol{\Sigma}_\mathcal{P}^{-1}\} + tr\{\boldsymbol{\Sigma}_\mathcal{P}\boldsymbol{\Sigma}_\mathcal{Q}^{-1}\} + (\boldsymbol{\mu}_\mathcal{Q} - \boldsymbol{\mu}_\mathcal{P})^T[\boldsymbol{\Sigma}_\mathcal{P}^{-1} + \boldsymbol{\Sigma}_\mathcal{Q}^{-1}](\boldsymbol{\mu}_\mathcal{Q} - \boldsymbol{\mu}_\mathcal{P}) - 2\beta \Big] \tag{14}$$

III. Problem formulation

A. Proposed formulation

This work proposes a Symmetric Kullback Leibler Divergence based sparse sensing with low precision approach for a multi-model system. The SKLD measures the distance between two PDFs of state estimation errors (Section II-E). In current work, these two PDF for i^{th} model are: (i) design PDF ($\mathcal{P}_{(i)}(\mathbf{e})$) which is steady state PDF corresponding to a specific choice of precision sensing $\mathbf{s}$, and (ii) a user specified target PDF ($\mathcal{Q}_{(i)}(\mathbf{e})$). Since estimation error ($\mathbf{e}$) is Gaussian, $\mathcal{P}_{(i)}(\mathbf{e}) \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{(i)})$. Further, $\mathcal{Q}_{(i)}(\mathbf{e})$ is also assumed to be Gaussian as $\mathcal{N}(\mathbf{0}, \mathbf{P}_{T,(i)})$. The design error covariance ($\mathbf{P}_{(i)}$) is computed using the steady state algebraic Riccati equation (Eq. 8) for a specific choice of precision sensing $\mathbf{s}$. The target error density covariance $\mathbf{P}_{T,(i)}$ is user-specified or computed using maximum permissible precision of sensor (best case scenario). Here, we used SKLD as a performance measure of estimation error for the multi-model scenario as:

$$\sum_{i=1}^{M} \frac{q_i}{4}\left[tr\{\mathbf{P}_{(i)}\mathbf{P}_{T,(i)}^{-1}\} + tr\{\mathbf{P}_{T,(i)}\mathbf{P}_{(i)}^{-1}\} - 2n \right] \le \alpha \tag{15}$$

where q_i denotes the weight assigned to SKLD for i^{th} model. Further, α is a user defined performance index satisfying $0 < \alpha < \infty$, ensuring broad applicability. Satisfaction of Eq. 15 ensures that the design PDF of the estimation error is close to the target PDF of the estimation error.

The Sparse Sensing with Low Precision (SSLP) problem is formulated as:

***Formulation** 1:*

$$\begin{aligned}
&\min_{\mathbf{s}} \ \ \|\mathbf{s}\|_{1,\mathbf{w}} \\
&s.t \sum_{i=1}^{M} \frac{q_i}{4}\left[tr\{\mathbf{P}_{(i)}\mathbf{P}_{T,(i)}^{-1}\} + tr\{\mathbf{P}_{T,(i)}\mathbf{P}_{(i)}^{-1}\} - 2n \right] \le \alpha \\
&\quad \mathbf{P}_{(i)} = \left[(\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}_{(i)}^T + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C} \right]^{-1}, \\
&\quad \forall\ i = 1, 2, .., M,\ \mathbf{R}^{-1} = diag\bigg(\{s_j\}_{j=1,2,....n} \bigg) \\
&\quad 0 < s_j \le s_{max}, \forall j = 1, 2, ..., n
\end{aligned}$$

Solving Formulation 1 leads to sparse sensing with low precision such that design PDF is close to the target PDF of estimation errors. Formulation 1 is a Non-Linear Programming (NLP) problem. Solving this NLP problem may lead to a local optimal solution. Under the assumption $\mathbf{P}_{(i)} \succeq \mathbf{P}_{T,(i)}, \forall i = 1, 2, ..., M$, for all feasible s_j values, we now discuss convex reformulation of Formulation 1 using Semidefinite programming. Solving this resultant formulation leads to a global optimal solution.

B. Computationally tractable reformulation of SSLP problem

Towards reformulating Formulation 1, we propose the following lemma and theorem:

***Lemma** 1:* The constraint 1 (Also represented in Eq. 15) in Formulation 1 is active at the optimal solution.

Proof: This lemma can be proved by contradiction. Assume that the constraint is not active at optimality, i.e.,

$$\sum_{i=1}^{M} \frac{q_i}{4}\left[\mathrm{tr}\{\mathbf{P}^*_{(i)}\mathbf{P}^{-1}_{T,(i)}\} + \mathrm{tr}\{\mathbf{P}_{T,(i)}(\mathbf{P}^*_{(i)})^{-1}\} - 2n\right] < \alpha, \tag{16}$$

where $\mathbf{P}^*_{(i)}, \forall i = 1, 2, \ldots, M$ is the design covariance of the estimation error at the optimal $\mathbf{s}^*$. The inequality in Eq. (16) can be converted into an equality by introducing a slack Δ such that

$$\sum_{i=1}^{M} \frac{q_i}{4}\left[\mathrm{tr}\{\mathbf{P}^*_{(i)}\mathbf{P}^{-1}_{T,(i)}\} + \mathrm{tr}\{\mathbf{P}_{T,(i)}(\mathbf{P}^*_{(i)})^{-1}\} - 2n\right] + \Delta = \alpha. \tag{17}$$

Now, distribute the slack Δ among the individual covariances as:

$$\mathbf{P}_{(i)} = \mathbf{P}^*_{(i)} + \Delta_{(i)}, \quad \Delta_{(i)} \in \Delta \tag{18}$$

where, $\Delta = \{\Delta_{(1)}, \Delta_{(2)}, ..., \Delta_{(M)}\}$. Further, each design covariance is increased by $\Delta_{(i)}$, which in turn reduces the sensor precision vector slightly by some ϵ, so that $\widetilde{\mathbf{s}} = \epsilon\, \mathbf{s}^*$, where $\widetilde{\mathbf{s}}$ is new optimal solution. Consequently, the optimizer would further reduce $\mathbf{s}$ to $\widetilde{\mathbf{s}}$, leading to a contradiction. Therefore, the constraint 1 in Formulation 1 must be active at optimality.

***Lemma** 2:* The symmetric Kullback Leibler Divergence (Constraints of Formulation 1) for i^{th} model is convex in matrix $\mathbf{P}_{(i)}$.

Proof: The proof is presented in Kumar et al. [10] and is omitted for brevity.

***Theorem** 2:* Constraint represented in Eq. 15 is a convex constraint with respect to matrix $\mathbf{P}_{(i)}$ for all $i = 1, 2, \ldots, M$.

Proof: To prove that the constraint in Eq. 15 is convex with respect to the matrix $\mathbf{P}_{(i)}$, it suffices to show that SKLD corresponding to the i^{th} model is convex in $\mathbf{P}_{(i)}$ for all $i = 1, 2, \ldots, M$. According to Lemma 2, the SKLD for the i^{th} model is convex in the matrix $\mathbf{P}_{(i)}$. Since a nonnegative weighted sum of convex functions preserves convexity, if $q_i \geq 0$ for all $i = 1, 2, \ldots, M$, then the function

$$\sum_{i=1}^{M} \frac{q_i}{4}\left[tr\{\mathbf{P}_{(i)}\mathbf{P}^{-1}_{T,(i)}\} + tr\{\mathbf{P}_{T,(i)}\mathbf{P}^{-1}_{(i)}\} - 2n\right] \tag{19}$$

is convex in $\mathbf{P}_{(i)}$. Therefore, the constraint in Eq. 15 is convex constraint.

***Theorem** 3:* If $\mathbf{Q}_{(i)}$ and $\boldsymbol{\phi}_{(i)}$ are invertible and $\mathbf{P}_{(i)} \succeq \mathbf{P}_{T,(i)}, \forall i = 1, 2, ..., M$, then Convex Programming (CP) problem formulation equivalent to NLP Formulation 1 is:

***Formulation** 2:*

$$\begin{aligned}
&\min_{\mathbf{s}} \quad \|\mathbf{s}\|_{1,\mathbf{w}} \\
&s.t \sum_{i=1}^{M} \frac{q_i}{4}\Big[tr\{\mathbf{P}_{(i)}\mathbf{P}^{-1}_{T,(i)}\} \\
&\qquad\qquad + tr\{\mathbf{P}_{T,(i)}(\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)})\} - 2n\Big] \leq \alpha \\
&\begin{bmatrix} \mathbf{X}_{(i)} & \mathbf{I} \\ \mathbf{I} & \boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}^T_{(i)} \end{bmatrix} \succeq 0, \forall i = 1, 2, ..., M \\
&\begin{bmatrix} \mathbf{X}_{(i)} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C} - \boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} & \mathbf{X}_{(i)} \\ \mathbf{X}_{(i)} & \mathbf{X}_{(i)} + \mathbf{Q}_{(i)} \end{bmatrix} \succeq 0 \\
&\forall\ i = 1, 2, ..., M,\ \mathbf{R}^{-1} = diag\Big(\ \{s_j\}_{j=1,2,...,n}\Big) \\
&0 < s_j \leq s_{max}, \forall j = 1, 2, ..., n
\end{aligned}$$

Proof: Introducing new auxiliary variables $\mathbf{X}_{(i)}$ to Formulation 1, leads to following equivalent formulation:

Formulation 2.1:

$$\begin{aligned}
&\min_{\mathbf{s}} \quad \|\mathbf{s}\|_{1,\mathbf{w}} \\
&s.t \sum_{i=1}^{M} \frac{q_i}{4}\Big[tr\{\mathbf{P}_{(i)}\mathbf{P}^{-1}_{T,(i)}\} \\
&\qquad\qquad + tr\{\mathbf{P}_{T,(i)}(\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)})\} - 2n\Big] \leq \alpha \\
&\mathbf{X}_{(i)} = (\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}^T_{(i)})^{-1} \quad \forall\ i = 1, 2, ..., M \\
&\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} = \Big[(\mathbf{X}^{-1}_{(i)} + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Big], \\
&\forall\ i = 1, 2, ..., M,\ \mathbf{R}^{-1} = diag\Big(\ \{s_j\}_{j=1,2,...,n}\Big) \\
&0 < s_j \leq s_{max}, \forall j = 1, 2, ..., n
\end{aligned}$$

The constraints $\mathbf{X}_{(i)} = (\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}^T_{(i)})^{-1}$, and $\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} = [(\mathbf{X}^{-1}_{(i)} + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}]$ are relaxed to $\mathbf{X}_{(i)} \succeq (\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}^T_{(i)})^{-1}$, and $\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} \preceq [(\mathbf{X}^{-1}_{(i)} + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}]$ respectively. The relaxation does not affect the optimal solution since the relaxed constraints will be active at optimality. This follows from the following observations: (i) Constraint 1 (Also represented in Eq. 15) in Formulation 1 is active at optimum (Lemma 1), (ii) Constraint 1 (Also represented in Eq. 15) in Formulation 1 is convex with $\mathbf{P}_{(i)}, \forall i = 1, 2, ..., M$ (Theorem 2), and (ii) $\mathbf{P}_{(i)} \geq \mathbf{P}_{T,(i)}, \forall i = 1, 2, ...M$ for all feasible solutions. Based on these observations, it is seen that the optimizer will try to decrease $\mathbf{P}_{(i)}$ so that it gets closer to $\mathbf{P}_{T,(i)}$ thereby ensuring that the relaxed constraints will be active. Now using Sherman-Woodbury-Morrison formula [3] and rearranging, one can write constraint $\boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} \preceq [(\mathbf{X}^{-1}_{(i)} + \mathbf{Q}_{(i)})^{-1} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}]$ as:

$$\begin{aligned}
&\mathbf{X}_{(i)} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C} - \boldsymbol{\phi}^T_{(i)}\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} - \\
&\qquad\qquad \mathbf{X}_{(i)}(\mathbf{X}_{(i)} + \mathbf{Q}^{-1}_{(i)})^{-1}\mathbf{X}_{(i)} \succeq 0
\end{aligned} \tag{20}$$

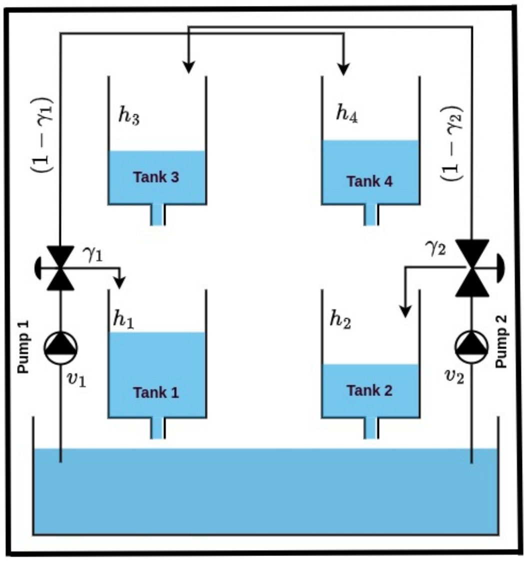

Fig. 1: Schematic diagram of QTP (adapted from [6])

Taking Schur's complement of Eq. 20 leads to:

$$\begin{bmatrix} \mathbf{X}_{(i)} + \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C} - \boldsymbol{\phi}_{(i)}^T\mathbf{X}_{(i)}\boldsymbol{\phi}_{(i)} & \mathbf{X}_{(i)} \\ \mathbf{X}_{(i)} & \mathbf{X}_{(i)} + \mathbf{Q}_{(i)}^{-1} \end{bmatrix} \succeq 0 \tag{21}$$

Further taking Schur's complement of constraint $\mathbf{X}_{(i)} \succeq (\boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}_{(i)}^T)^{-1}$ leads to

$$\begin{bmatrix} \mathbf{X}_{(i)} & \mathbf{I} \\ \mathbf{I} & \boldsymbol{\phi}_{(i)}\mathbf{P}_{(i)}\boldsymbol{\phi}_{(i)}^T \end{bmatrix} \succeq 0 \tag{22}$$

With these changes, Formulation 2.1 results in Formulation 2. Thus, CP problem Formulation 2 is equivalent to Formulation 1. This completes the proof.
The advantage of Formulation 2 is that the formulation can be solved to global optimality.

IV. Case study

In this section, we demonstrate the proposed approach for sparse sensing with low precision for multi-model system using Quadruple tank process. This work compares our proposed SKLD based SSLP problem formulation for multi-model scenario with single model. For ease of readability, we refer SSLP problem for Multi-Model as SSLP-MM and SSLP for Single-Model as SSLP-SM. The CP problem Formulation 2 is solved using MOSEK solver using YALMIP toolbox [2] in MATLAB. All computations were performed in MATLAB version R2023b running on a computer with 64 GB RAM and an Intel Core i9-13 Gen processor.

A. Quadruple tank process

The schematic diagram of Quadruple Tank Process (QTP) [6] is presented in Fig. 1. The process contains four states ($\mathbf{x} = [h_1\ h_2\ h_3\ h_4]^T$) and two process inputs ($\mathbf{u} = [v_1\ v_2]^T$). The QTP is a nonlinear dynamical system. The corresponding nonlinear model and details about the parameter are presented in Johansson [6]. In the current work, the proposed approach described in Section III is illustrated using the linearized model of the nonlinear QTP dynamical system. In particular, from an estimation and control perspective, the QTP is linearized at two operating points [6]: (a) minimum-phase, and (b) non-minimum-phase. The linearized models at the minimum-phase and non-minimum-phase operating points are referred to as Model 1 ($i = 1$) and Model 2 ($i = 2$), respectively. Hence, we consider $M = 2$. We consider the weights $q_1 = 0.7$ and $q_2 = 0.3$. This implies that Model 1 is used 70% of the time, while Model 2 is used 30% of the time. The state dynamics for Model 1 ($\boldsymbol{\phi}_1$) and Model 2 ($\boldsymbol{\phi}_2$) are:

$$\boldsymbol{\phi}_1 = \begin{bmatrix} 0.9233 & 0 & 0.1813 & 0 \\ 0 & 0.9462 & 0 & 0.1493 \\ 0 & 0 & 0.8112 & 0 \\ 0 & 0 & 0 & 0.8465 \end{bmatrix}, \text{and}$$
$$\boldsymbol{\phi}_2 = \begin{bmatrix} 0.9284 & 0 & 0.1051 & 0 \\ 0 & 0.9531 & 0 & 0.0686 \\ 0 & 0 & 0.8908 & 0 \\ 0 & 0 & 0 & 0.9298 \end{bmatrix} \tag{23}$$

The corresponding state noise covariance matrices for Model 1 and Model 2 are given by $\mathbf{Q}_1 = \mathbf{I}_4$ and $\mathbf{Q}_2 = 2 * \mathbf{I}_4$ respectively. The measurement matrix is chosen as $\mathbf{C} = \mathbf{I}_4$ and weight ($\mathbf{w}$) for each sensor is choosen as $w = 1$. Further, the target estimation error covariance matrices corresponding to Model 1 and Model 2 are given as follows:

$$\mathbf{P}_{T,(1)} = \begin{bmatrix} 1.2383 & 0 & 0.0830 & 0 \\ 0 & 1.2566 & 0 & 0.0727 \\ 0.0830 & 0 & 1.0876 & 0 \\ 0 & 0.0727 & 0 & 1.1241 \end{bmatrix}, \text{and}$$
$$\mathbf{P}_{T,(2)} = \begin{bmatrix} 1.5983 & 0 & 0.0624 & 0 \\ 0 & 1.6117 & 0 & 0.0536 \\ 0.0624 & 0 & 1.4928 & 0 \\ 0 & 0.0536 & 0 & 1.5199 \end{bmatrix} \tag{24}$$

The maximum allowable permissible precision for each sensor is chosen as $s_{max} = 3$. We consider five different values of the achievable performance parameter α : $0.0001,\ 0.01,\ 1,\ 3,\ 5$. In this work, we consider three scenarios:

1) **SSLP-MM**: In this scenario, we consider our proposed multi-model formulation.
2) **SSLP-SM1**: In this scenario, we address the SSLP using a single-model approach, specifically Model 1.
3) **SSLP-SM2**: In this scenario, we address the SSLP using a single-model approach, specifically Model 2.

B. Results and discussions

In this work, we compare the results obtained using the SSLP-MM scenario with those from SSLP-SM1 and SSLP-SM2. Fig. 2 presents the low sensor precision $\|\mathbf{s}\|_{1,\mathbf{w}}$ versus the achievable performance (α) for three different scenarios. It can be observed from the figure that there is a tradeoff between the achievable performance and the low sensor precision. Furthermore, if the multi-model approach is not

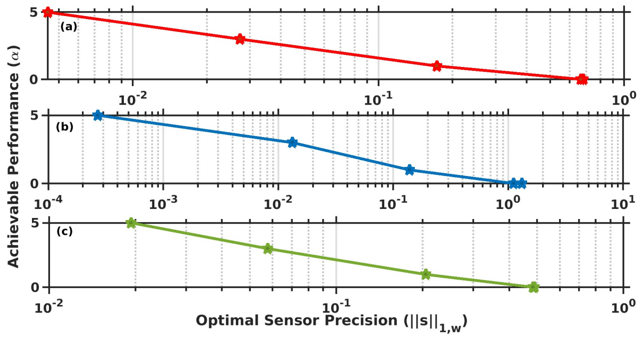

Fig. 2: Low sensor precision ($\|\mathbf{s}\|_{1,\mathbf{w}}$) vs achievable performance (α): (a) SSLP-MM, (b) SSLP-SM1, and (c) SSLP-SM2

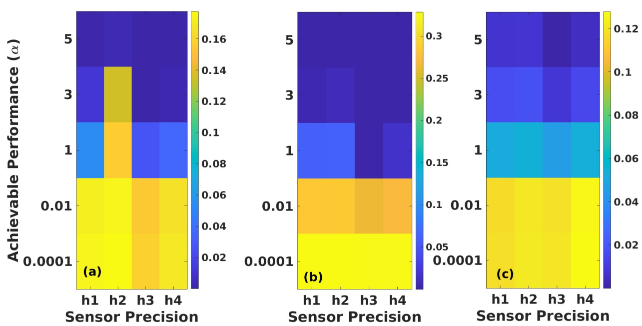

Fig. 3: Sparse low sensor precision vs achievable performance (α): (a) SSLP-MM, (b) SSLP-SM1, and (c) SSLP-SM2

considered during the design procedure, different sensor precisions are obtained. For the same achievable performance, the low sensor precisions vary across the scenarios, as shown in Fig. 3, where the low precision values for the four sensors are presented. The sensor precisions obtained under the three scenarios differ significantly, indicating that neglecting the multi-model configuration can lead to a suboptimal design. Additionally, it can be observed from the figure that as the achievable performance increases, the precision required from individual sensors decreases. This implies that the distance between the design and the target density increases. Consequently, lower-precision sensors can be used to achieve the desired performance, resulting in a reduction of the overall capital cost. A consistent trend is observed across all scenarios.

V. Conclusions

In this work, we presented a symmetric Kullback Leibler divergence based approach for sparse sensing with low precision for multi-model system. The resultant optimization problem is Non-Linear Programming (NLP) problem. Solving this NLP problem may yield local optimal solution. Hence, we proposed an equivalent Convex Programming (CP) problem formulation. Solving this CP problem ensures that optimal solution will be globally optimal. The proposed approach is demonstrated using the quadruple tank process. The results for the multi-model approach are compared with those from a single linearized model for both minimum and nonminimum phase operating points. The results shows that the sensor precision obtained from a single linearized model may lead to a suboptimal design compared with a multi-model approach. In the future, this approach will be extended to multi-model subject to model uncertainty.

References

[1] Brian D O Anderson and John B Moore. *Optimal control*. Courier Corporation, 2007.

[2] Mosek ApS. MOSEK– a MATLAB software package for semidefinite programming. https://www.mosek.com, 2025.

[3] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[4] Niladri Das and Raktim Bhattacharya. Optimal sensor precision for multirate sensing for bounded estimation error. *IEEE Transactions on Aerospace and Electronic Systems*, 58(2):844–854, 2021.

[5] Vedang M Deshpande and Raktim Bhattacharya. Guaranteed robust performance of H_∞ filters with sparse and low precision sensing. *IEEE Transactions on Automatic Control*, 69(2):1029–1036, 2023.

[6] Karl Henrik Johansson. The quadruple-tank process: A multivariable laboratory process with an adjustable zero. *IEEE Transactions on Control Systems Technology*, 8(3):456–465, 2002.

[7] Siddharth Joshi and Stephen P Boyd. Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2):451–462, 2008.

[8] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1997.

[9] Brijesh Kumar and Mani Bhushan. Information theory-based sensor placement for data reconciliation with uncertain models. *Industrial & Engineering Chemistry Research*, 64(50):24093–24105, 2025.

[10] Brijesh Kumar, Garima Patel, and Mani Bhushan. Symmetric Kullback Leibler divergence based robust sensor placement design for linear dynamical system subject to bounded uncertainties. In *2025 IEEE 64th Conference on Decision and Control (CDC)*, pages 3714–3719. IEEE, 2025.

[11] K Kishore Kumar and Sachin C Patwardhan. Nonlinear predictive control of systems exhibiting input multiplicities using the multi-model approach. *Industrial & Engineering Chemistry Research*, 41(13):3186–3198, 2002.

[12] Faming Li, Maurício C. De Oliveira, and Robert E Skelton. Integrating information architecture and control or estimation design. *SICE Journal of Control, Measurement, and System Integration*, 1(2):120–128, 2008.

[13] Jianbo Lu and Robert E Skelton. Economic design problem: Integrating instrumentation and control. In *Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251)*, volume 6, pages 4321–4325. IEEE, 1999.

[14] Arjun M and Nabil Magbool Jan. Convex optimization approach to design sensor networks using information theoretic measures. *AIChE Journal*, 70(2):e18267, 2024.

[15] Peter S Maybeck. *Stochastic models, estimation, and control*. Academic Press. INC, 1982.

[16] Philipp Reitz, Christian Künzle, Norman Franchi, and Maximilian Lübke. Evaluation of the interference performance of FMCW radar sensors in dense indoor environments. *IEEE Access*, 12:46834–46850, 2024.

[17] Jun-Wei Wang, Ya-Qiang Liu, and Chang-Yin Sun. Luenberger observer design for state estimation of a linear parabolic distributed parameter system with discrete measurement sensors. In *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pages 1123–1128. IEEE, 2016.

[18] Zhong Zhao, Xiaohua Xia, Jingchun Wang, Jian Gu, and Yihui Jin. Nonlinear dynamic matrix control based on multiple operating models. *Journal of Process Control*, 13(1):41–56, 2003.