

Sensor Importance towards Observability Degree via Shapley Values

Vishal Chalapadi Ravindra*

Abstract—Sensor selection is an often under-appreciated aspect of state estimator or Kalman filter design. The basic minimum requirement for the choice of a sensor set while designing Kalman filters is that all states are observable. In addition, the sensors should be chosen with a view towards estimating the states with a desired accuracy. Often observability is treated as true/false check during filter design. Beyond observability — the observability degree — which measures *how observable* the states are, has been used as the metric of choice to for sensor selection or placement applications. The higher the degree of observability, the better the possibility of designing Kalman filters that achieve the desired state estimation accuracy and consistency requirements. When a wide variety of sensors are available, sometimes with cost and physical constraints involved, sensor selection plays a crucial role in filter design. In such situations it is important to know the *expected* contribution of each sensor towards observability degree. Shapley values, developed in cooperative game theory for fair allocation of the payout of a multi-player game to individual players, are widely used in machine learning to assess feature importance. This paper shows that Shapley values can indeed be leveraged to quantify the *expected marginal* contribution of each sensor in any given sensor set towards the observability degree. This quantification of the fair contribution of each sensor towards the observability degree can be leveraged by filter designers for sensor selection, placement and filter (state estimator) design.

I. INTRODUCTION

A deterministic state space system is said to be *fully observable* if its initial state can be fully and uniquely recovered from a finite number of observations of the output and knowledge of the input. As a deterministic system is not affected by noise, knowledge of the initial state is equivalent to the knowledge of the state at any instance of time [1]. Generalizing this further, suppose the state dynamics and the observations of the output are modeled as driven by zero mean additive white Gaussian noise, it is well known that the Kalman filter ([4]) provides the optimal minimum mean square estimate of the unknown state [1]. When the state is modeled as stochastic, observability is not just a sufficient but also a necessary condition to design a filter that can converge over time.

The two most common ways to check if a system is observable is via the observability matrix and the observability Gramian. The system is observable if the observability matrix

is full column rank [1] or if the observability Gramian is not singular [3]. These two approaches however tell us only if or not if the system is observable, and do not quantify the *degree of observability*. Degree of observability¹ is defined as a measure of *how observable an observable system is*. For fully observable systems, the degree of observability is an important criteria for applications such as sensor selection/placement [6], [13], path planning [9] and Kalman filter design [10]. A combination of sensors is more desirable from the point of view of filter design if it significantly improves the degree of observability over alternate combinations. Several metrics based on the observability Gramian matrix, such as trace, rank, log determinant, condition number, etc. have been proposed and evaluated in [5]–[12]. The essence being the evaluation of *how close* the observability Gramian matrix is to being singular.

Problems like sensor selection, sensor placement and virtual sensing ([8], [13], [14]) focus on the selection of the “optimal” set of sensors that are able to achieve a desired degree of observability. Implications of the modularity properties of observability Gramian based metrics are discussed in [13],² where it is shown that simple greedy approaches can be used to solve NP-hard optimal sensor placement problems if the metrics are submodular in nature. A Kalman filter measurement update equation based approach was proposed in [10], for attributing the impact of each sensor on the estimation of individual states. However, this approach is predicated upon the running of a Kalman filter using a given sensor set.

The aim of the present paper is not the solution of the “optimal” sensor selection or placement problem, but to introduce a way for fair quantification of the contribution of each sensor in an overall coalition or set of sensors being considered for filter design, towards selected observability degree measures. Sometimes it is important to *exactly* know what is the individual contribution of each sensor, especially in cases such as virtual sensing where the goal is to replace sensors of higher cost with virtual sensors or state estimates. Furthermore, the method proposed is not conditional on filter design like in [10], but rather can be performed before it.

The present paper proposes to leverage a concept first proposed in cooperative game theory for *fair attribution* of the payoff of a multi-player game to each individual

* Vishal Chalapadi Ravindra is with Intuit Credit Karma, Bengaluru, India vishal.ravindra@creditcharma.com
© 2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. This paper has been accepted for publication in the Proceedings of the Eleventh Indian Control Conference (ICC 2025), held in Bengaluru, India on December 18-20, 2025.

¹Observability degree and degree of observability are used interchangeably in this paper.

²The author acknowledges the anonymous reviewer who pointed them towards this reference.

player [11]. The Shapley value of a player³ is the marginal contribution of a player to the overall payout which is characterized by a value function. Shapley values are widely used in machine learning [7] to understand feature importance in order to explain individual predictions from a classification or a regression model. The contribution of this paper is to show that Shapley values can *also* be applied to observability degree measures (based on the observability Gramian) in order to arrive at fair and exact contributions of individual sensors. Shapley values can be derived based on any observability degree metric, irrespective of their modularity properties ([13]) as long as certain smoothness conditions are met [7]. The appeal of this approach is two fold: (a) avoiding the need to run a Kalman filter *a priori* and (b) *interpretability* – as it is known that Shapley values are a fair attribution metric [7], [11]. As the present paper is intended to serve as a concept introduction paper, the scope is limited to linear time invariant (LTI) systems. Using two simple simulation scenarios, the ability of Shapley values to learn interaction effects between individual sensors that enhance the observability degree is highlighted.

Section II provides a concise mathematical background for LTI state space systems, observability and measures of characterizing observability. Section III introduces the concept of degree of observability and describes two commonly used measures for it. Section IV introduces Shapley values from cooperative game theory and describes the four axioms that are satisfied by them in order for them to be considered fair allocations of payouts for each player in a multi-player game. It has to be noted that the complexity of exact Shapley value computation grows exponentially with the number of sensors in the full coalition being considered. However, leveraging of fast approximation techniques are left out of scope of this paper and can instead be found in [7]. Section V describes the key contribution of this paper, which is to apply Shapley values to quantify the individual contribution of each sensor towards observability degree. Section VI describes two simulation scenarios and results that demonstrate the idea described in the present paper, followed by Conclusions.

II. STATE SPACE SYSTEM AND OBSERVABILITY

Consider the following (deterministic) discrete time linear time invariant (LTI) system

$$\begin{aligned} \mathbf{x}_{k+1} &= A\mathbf{x}_k \\ \mathbf{y}_k &= C\mathbf{x}_k \end{aligned} \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ denotes the state vector, $\mathbf{y} \in \mathbb{R}^p$ denotes the measurement vector and $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{p \times n}$ denote the state transition and measurement matrices, respectively. The time index is denoted by k and it can take the values $k = 0, 1, \dots, K$, where $k = 0$ represents the initial condition and K denotes the horizon.

A discrete time LTI system such as the one described in (1) is said to be observable if the initial state $\mathbf{x}_0$ can be

uniquely recovered from the sequence of measurements (or sensors)⁴ $\{\mathbf{y}_k\}_{k=0}^K$, where $K \geq n - 1$ [1].

A. Observability matrix

From (1), the measurement at any time index k can be expressed in terms of the initial state as follows

$$\mathbf{y}_k = CA^k \mathbf{x}_0 \quad (2)$$

For a finite horizon of $K \geq n - 1$ measurements, (2) can be expressed in vector and matrix forms as

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_K \end{bmatrix} &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^K \end{bmatrix} \mathbf{x}_0 \\ \mathcal{Y}_K &= \mathcal{O}_K \mathbf{x}_0 \end{aligned} \quad (3)$$

forming a linear system where $\mathbf{x}_0$ is the unknown (initial) state to be estimated. From linear system theory [3], it is well known that $\mathbf{x}_0$ has a unique solution, if and only if the matrix $\mathcal{O}_K$ – known as the *observability matrix* – has full column rank (i.e., $\rho\{\mathcal{O}_K\} = n$).

B. Observability Gramian

An alternative approach to deduce observability is via the observability Gramian matrix. Substituting for $\mathbf{y}_k$ from the measurement equation given in (2), the energy of the output (measurement) over a horizon K is given by

$$\begin{aligned} \mathcal{E}_y &= \sum_{k=0}^K \|\mathbf{y}_k\|^2 = \sum_{k=0}^K \mathbf{x}_0^T (A^T)^k C^T C A^k \mathbf{x}_0 \\ &= \mathbf{x}_0^T \left(\sum_{k=0}^K (A^T)^k C^T C A^k \right) \mathbf{x}_0 \\ &= \mathbf{x}_0^T \mathcal{W}_O \mathbf{x}_0 \end{aligned} \quad (4)$$

Thus,

$$\mathcal{W}_O = \sum_{k=0}^K (A^T)^k C^T C A^k \quad (5)$$

where $\mathcal{W}_O \in \mathbb{R}^{n \times n}$ is known as the *observability Gramian* matrix. As the right hand side of (4) forms a quadratic form, it can be inferred that, for all directions of the initial state $\mathbf{x}_0$ to produce an *observable* output, the energy of the output should be positive, i.e., $\mathcal{E}_y > 0$. This implies that for the system to be observable, the observability Gramian matrix must be positive definite. Furthermore, using the alternate matrix form defined in (3) and applying it to the energy equation in (4), it is straight forward to derive the relationship between the observability Gramian and observability matrices as

$$\mathcal{W}_O = \mathcal{O}_K^T \mathcal{O}_K \quad (6)$$

Hence, for the LTI system in (1) to be observable, the observability Gramian $\mathcal{W}_O$ should have full row and column rank n , while the observability matrix $\mathcal{O}_K$ should have full column rank n .

³A player is analogous to a sensor, while a multi-player game is analogous to the sensor set being considered.

⁴Measurements and sensors are used interchangeably in this paper.

C. Additive property of the observability Gramian

A property that is useful in the ensuing sections is the additive property of the observability Gramian. The measurement matrix C in (1) can be expressed in terms of individual measurements (or sensors) as follows

$$C = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_p \end{bmatrix} \quad (7)$$

where $C_i \in \mathbb{R}^{1 \times n}$ represents the i^{th} row of C and $i = \{1, \dots, p\}$. Substituting from (7) and using the expression $C^T C = \sum_{i=1}^p C_i^T C_i$, in the expression of the observability Gramian from (5), it is straightforward to derive the following relationship

$$\mathcal{W}_O = \sum_{k=0}^K (A^T)^k C^T C A^k = \sum_{i=1}^p \mathcal{W}_O^i \quad (8)$$

where

$$\mathcal{W}_O^i = \sum_{k=0}^K (A^T)^k C_i^T C_i A^k \quad (9)$$

represents the observability Gramian of the LTI sub-system (A, C_i) , where C_i is the row of the measurement matrix C corresponding to measurement (or sensor) i . Hence, from (9) it can be seen that the observability Gramian of the full LTI system is additive with respect to the observability Gramian of the sub-systems derived from individual measurements.

III. DEGREE OF OBSERVABILITY

Several candidates for observability degree based on functions of the observability Gramian are presented in [5], [13]. Observability degree measures can be broadly divided into three categories [13]: (a) modular functions, i.e., linear functions that map subsets or coalitions of sensors to the observability Gramian matrix, (b) submodular functions, i.e., where the contribution that each sensor makes to an existing set of sensors diminishes with increasing cardinality of the sensor set and (c) non-submodular function, where no such diminishing contribution property exists.⁵ Table I lists one commonly used observability degree metric from each of the three function types mentioned.

Metric	Expression	Type
Trace	$\text{tr} \{\mathcal{W}_O\}$	Modular
Log determinant	$\log \det \{\mathcal{W}_O\}$	Submodular
Min. eigenvalue	$\lambda_{\min} \{\mathcal{W}_O\}$	Non-submodular

TABLE I

COMMONLY USED MEASURES ON THE OBSERVABILITY GRAMIAN MATRIX, ONE EACH FROM THE THREE MODULARITY CATEGORIES DISCUSSED IN [13].

⁵It is shown in [13] that efficient approximation strategies exist for submodular functions that significantly reduce computational complexity for sensor selection or placement posed as maximization problems.

The trace of the observability Gramian measures the total energy captured by the finite sequence of outputs, as shown in (4). As the output energy is shown to take a quadratic form, geometrically, the observability Gramian forms an ellipse. From (4), it can be seen that larger the trace, larger the energy of the output, implying a higher degree of observability. However, there can be cases where a measurement (sensor) or a coalition of sensors lead to “strong” observability of certain directions while leaving other directions as unobservable. Such cases (also known as partially observable systems) can result in a large trace even while having one or more zero eigenvalues. The minimum eigenvalue of the observability Gramian, however, is a direct indicator of the weakest observable direction [2], with zero indicating unobservability. It is also a much better indicator of observability degree as farther the minimum eigenvalue is away from zero, the “more observable” the system is. Hence, for most applications as well as from a numerical robustness aspect, it is more desirable to maximize the minimum eigenvalue of the observability Gramian than its trace during sensor selection for filter design.

Remark 1: For the purpose of the present paper $\log \det \{\mathcal{W}_O\}$ is not considered, as it is undefined for sensor coalitions that result in the observability Gramian matrix becoming singular (loss of observability).

IV. SHAPLEY VALUES – AN OVERVIEW

In cooperative game theory, when a game is played by a group (or coalition) of players and there is a well defined payout from the game, Shapley values [11] are the fair way to allocate payout contributions to individual players. Shapley value of a player i is the expected contribution the player makes to the payout received by the full coalition of players, marginalized over all the coalitions that the player can join in. In other words, Shapley values are a way to quantify the impact of a player in a multi-player game based on the value they bring to each coalition of players they can join in. The contribution of this paper is to leverage Shapley values to fairly and exactly quantify individual sensors’ contributions towards overall observability degree of an LTI system. Due to the axioms that underpin Shapley values, the contributions thus computed can be thought of as the “fair” allocation of importance towards observability degree for each sensor. In the rest of the section we provide an overview of the building blocks of Shapley value theory.

Suppose there is a set of players $N = \{1, 2, \dots, p\}$, the Shapley value for a player $i \in N$ is defined via a value function $v(\cdot)$ that characterizes the payout of the game. It is defined as

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (p - |S| - 1)!}{p!} [v(S \cup \{i\}) - v(S)] \quad (10)$$

where ϕ_i denotes the Shapley value of player i , S denotes a coalition of players that is a subset of the superset N and $S \setminus \{i\}$ represents a coalition S that does not include player i . The value function $v(S)$ represents the payout attributable to

coalition S .⁶ Marginal contribution of a player i to a coalition S is given by $v(S \cup \{i\}) - v(S)$, i.e., the value added by the player i to the coalition S . Hence, the Shapley value of player i is nothing but the average change in payout that i causes by joining all possible coalitions, as compared to the payouts when it is not part of them. The weight term in (10) is the fraction of all possible ways in which the players in S can precede i .

A Shapley value is the only attribution method (or the *fair* attribution method) that satisfies the following four axioms [11]:

1. *Efficiency*: The total value of the game is fully distributed, i.e., the Shapley values of the individual players should add up to the overall payout of the game:

$$\sum_{i \in N} \phi_i(v) = v(N) \quad (11)$$

where $N = \{1, 2, \dots, p\}$ is the set of all players. This makes Shapley values very interpretable – each Shapley value is simply the decomposition of the total payout to individual players.

2. *Symmetry*: Players who contribute equally, get equal value. If,

$$v(S \cup \{j\}) = v(S \cup \{k\}) \quad (12)$$

$\forall S \subseteq N \setminus \{j, k\}$, then $\phi_j(v) = \phi_k(v)$.

3. *Dummy*: A player who adds no value to any coalition gets zero. If,

$$v(S \cup \{j\}) = v(S) \quad (13)$$

$\forall S \subseteq N \setminus \{j\}$, then $\phi_j(v) = 0$. The interpretation is that a player that does not add value to *any* coalition contributes nothing to the total payout of the game. Such players are called dummy players.

4. *Additivity*: If a game with payout or value v^* can be expressed as a sum of two individual games with payouts v_1 and v_2 , i.e., $v^* = v_1 + v_2$, then the Shapley value of the player j can also be expressed as

$$\begin{aligned} \phi_j(v_1 + v_2) &= \phi_j(v_1) + \phi_j(v_2) \\ \phi_j(v^*) &= \phi_j(v_1) + \phi_j(v_2) \end{aligned} \quad (14)$$

as the sum of the Shapley values of the player in those individual games. If the value function v^* can be shown to be additive over the value functions of individual models (for e.g., ensemble approaches such as bagging), then it can be verified that the additivity axiom is met [7].

The present paper proposes the use of Shapley values to *fairly* allocate contributions to individual sensors (players) towards the degree of observability (overall payout) based on certain value functions (discussed in Section III) that characterize the degree of observability of LTI systems.

⁶The payoff from the game if only players in the subset or coalition of players S were taking part.

V. INDIVIDUAL SENSOR CONTRIBUTION TOWARDS DEGREE OF OBSERVABILITY

In this section, we show how the two of the observability degree metrics presented in Table I, i.e., the trace and the minimum eigenvalue of the observability Gramian, can both be used as value functions to derive Shapley values. We omit the log determinant of the observability Gramian from the scope of the present paper as, even though it is a submodular function with known approximation properties, it is undefined for sensor coalitions that result in loss of rank.

We now show how the chosen observability degree measures can be “broken down” into allocations that provide a fair quantification of the contribution of individual sensors – characterised by rows C_i of the measurement matrix as shown in (7) – via Shapley values.

A. Trace of the observability Gramian

As the trace of a matrix is a linear operator and using the additive property of the observability Gramian in (8), it can be seen that the trace of the observability Gramian of the overall system (A, C) is the sum of the traces of the observability Gramian of sub-systems formed by individual sensors (A, C_i) :

$$\begin{aligned} \text{tr}\{\mathcal{W}_O\} &= \text{tr}\left\{\sum_{i=1}^p \mathcal{W}_O^i\right\} \\ &= \sum_{i=1}^p \text{tr}\{\mathcal{W}_O^i\} \end{aligned} \quad (15)$$

Due to the additivity property of the trace, as shown in (15), and since $\text{tr}\{A\} \in \mathbb{R}^1$ when $A \in \mathbb{R}^{n \times n}$ is a well defined and smooth value function, Shapley values derived by substituting the following value function in (10)

$$v(S) = \text{tr}\{\mathcal{W}_O^S\} \quad (16)$$

follow all the axioms listed in Section IV, as applicable. The term $\mathcal{W}_O^S$ denotes the observability Gramian formed by the system (A, C_S) where C_S is the measurement matrix formed by the coalition of sensors $S \subseteq \{1, 2, \dots, p\}$.

Incidentally, due to its modular nature of this observability degree metric [13], the following Proposition results:

Proposition 1: When the value function over the full sensor set N is additive over value functions over individual sensors j , i.e.,

$$v(N) = \sum_{j=1}^p v(\{j\}) \quad (17)$$

then the Shapley value of the sensor j is the same as the value function over sensor j , i.e.,

$$\phi_j(v) = v(\{j\}) \quad (18)$$

Proof: Applying the definition of the Shapley value of sensor j from (10), we get

$$\phi_j(v) = \sum_{S \subseteq N \setminus \{j\}} w_S [v(S \cup \{j\}) - v(S)] \quad (19)$$

where $w_S = \frac{|S|!(p-|S|-1)!}{p!}$ is the weight attached to coalition S . From (17), it follows that

$$v(S \cup \{j\}) = v(S) + v(\{j\}) \quad (20)$$

Substituting from (20) in (19), we get

$$\phi_j(v) = \sum_{S \subseteq N \setminus \{j\}} w_S v(\{j\}) \quad (21)$$

Further simplifying

$$\begin{aligned} \phi_j(v) &= v(\{j\}) \sum_{S \subseteq N \setminus \{j\}} w_S \\ &= v(\{j\}) \end{aligned} \quad (22)$$

as $\sum_{S \subseteq N \setminus \{j\}} w_S = 1$. This is because sensor j can join all possible coalitions S that it is not a part of, without any limitations. Furthermore, as a corollary from (22), the *Efficiency* axiom can also be verified as follows:

$$\begin{aligned} \sum_{j=1}^p \phi_j(v) &= \sum_{j=1}^p v(\{j\}) \\ &= v(N) \end{aligned} \quad (23)$$

due to the additivity property in (17). Proposition 1 thus implies that the Shapley values derived from the trace of the observability Gramian, *do not take into account interactions between individual sensors*. We now investigate if value functions based on minimum eigenvalue of the observability Gramian, yields Shapley values that take into account interaction effects where, even though a single sensor by itself might not make the system observable when part of coalitions with other sensors it can both make the system observable and improve the degree of observability.

B. Minimum eigenvalue of the observability Gramian

As discussed in Section III, for the observability Gramian to be positive definite, its minimum eigenvalue should be positive, i.e., $\lambda_{\min} \{\mathcal{W}_O\} > 0$. From (8), we get

$$\lambda_{\min} \{\mathcal{W}_O\} = \lambda_{\min} \left\{ \sum_{i=1}^p \mathcal{W}_O^i \right\} \quad (24)$$

However, in general, due to the non-additive nature of eigenvalues

$$\lambda_{\min} \left\{ \sum_{i=1}^p \mathcal{W}_O^i \right\} \neq \sum_{i=1}^p \lambda_{\min} \{\mathcal{W}_O^i\} \quad (25)$$

As a result, contrary to Proposition 1, we cannot prove that the Shapley values of the individual sensors are identical to the value functions over individual sensors. Hence, they *do* consider interaction effects with other sensors and *reward situations where a sensor increases the minimum eigenvalue of the resulting observability Gramian by joining one or more coalitions*. Furthermore, as is shown in [13], since the minimum eigenvalue of the observability Gramian is a non-submodular function, it does not follow the property of diminishing returns. This implies that there could be scenarios where marginal contribution of a given sensor

could be higher for a larger coalition of sensors than a smaller one. As a result, this metric is able to identify directions in the state space whose unobservable mode(s) become observable due to the addition of a given sensor, better. Also, as $\lambda_{\min} \{\mathcal{W}_O\}$ is a valid value function that maps a matrix in $\mathbb{R}^{n \times n}$ to a scalar in $\mathbb{R}^1$ due to its symmetricity and because it is real, finite and well defined for all sensor coalitions considered, Shapley values can be derived and the axioms listed in Section IV hold where applicable.

VI. SIMULATIONS AND RESULTS

Scenario I considers the case where no individual sensor (or measurement), by itself, results in full observability, but the overall system is still observable. Scenario II considers an example where some sensors, by themselves, yield full observability while some do not.

A. Scenario I: Collaborative sensors

Consider the following two-state, two-measurement system

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; \quad C = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (26)$$

where both the “sensors” C_1 and C_2 (the two rows of the C matrix) observe a linear combination of the two states. It is straightforward to verify that this system is fully observable even though, if the two sensors (C_1 and C_2) are used individually, the observability Gramian suffers a rank loss. This is a classic case of both sensors being equally important, as individually even though they may not have a high degree of observability, but together they make the system observable.⁷

Sensor	Value Function	Standalone Value	Shapley Value
C_1	Trace	20.0	20.0
C_2	Trace	20.0	20.0
C_1	Min Eigenvalue	0	10.0
C_2	Min Eigenvalue	0	10.0

TABLE II

STANDALONE AND SHAPLEY VALUES FOR TWO SENSORS USING TWO DIFFERENT OBSERVABILITY METRICS WITH HORIZON $N = 10$.

From Table II, we can see that; (a) as is expected from Proposition 1, the standalone sensor traces and the Shapley values are identical and (b) the minimum eigenvalues if the two sensors were used individually result in a loss of observability, but their Shapley values are identical⁸ and show their fair importances towards the eigenvalue based observability degree metric. This shows that the minimum eigenvalue metric captures the interaction effect between the two sensors whereas the trace of the observability Gramian does not. It can also be verified that the Shapley values of the two metrics sum up to the overall observability degrees, respectively, illustrating the effect of the *Efficiency* axiom.

⁷This class of sensors are also known as complementary sensors.

⁸*Symmetry* axiom of Shapley values – due to equal importance.

B. Scenario II: Standalone, collaborative and redundant sensors

Consider a slightly more complex example in the form of the following three state, four sensor system

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} ; \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (27)$$

From visual inspection of (27), it can be inferred that sensor 1 alone, i.e., $C_1 = [1 \ 0 \ 0]$, is enough to get full observability. Furthermore, individually used, sensors $C_2 = [0 \ 1 \ 0]$ and $C_4 = [0 \ 0 \ 1]$ would lead to a rank loss. Sensor 3, i.e., $C_3 = [1 \ 1 \ 0]$ can be thought of as a redundant sensor to the C_1, C_2 pair as it observes a linear combination of the first two states. Intuitively, it can be said that as C_1 leads to the observability of the system by itself, it should have a high Shapley value. Also, C_3 is a redundant sensor that can be used in combination with C_1 to eliminate the need for both C_2 and C_3 . Sensor C_4 , though, adds the least value to any well constructed observability index as the first three sensors are enough to observe the third state.

Sensor	Value Function	Standalone Value	Shapley Value
C_1	Trace	3187.0	3187.0
C_2	Trace	295.0	295.0
C_3	Trace	5312.0	5312.0
C_4	Trace	10.0	10.0
C_1	Min Eigenvalue	1.3920	1.5129
C_2	Min Eigenvalue	0.0000	0.2306
C_3	Min Eigenvalue	0.6405	0.7089
C_4	Min Eigenvalue	0.0000	0.0243

TABLE III

STANDALONE AND SHAPLEY VALUES FOR FOUR SENSORS USING TRACE AND MINIMUM EIGENVALUE OF OBSERVABILITY GRAMIAN WITH HORIZON $N = 10$.

From Table III, it can be verified that, as expected from Proposition 1, the Shapley values based on trace as the value function do not take into account interaction effects. If trace is used as the value function, the Shapley value of C_3 turns out to be the highest which is not intuitive as discussed earlier. Shapley values based on the minimum eigenvalue value function, however, achieve a very intuitive allocation of sensor importances. Firstly, the Shapley value of each sensor shows a slight increase over the standalone value functions. This is due to the interaction effects that they capture which the trace based value function does not. Secondly, aligning with intuition, C_1 has the highest Shapley value followed by C_3 . Note that the orders are reversed when trace is used as the value function. The Shapley values of C_2 and C_4 also show a small increase, as even if (by themselves) they do not make the system observable but when joined in coalitions with other sensors, they contribute towards improving the degree of observability. Due to the *Efficiency* axiom, the sum of the Shapley values based on the minimum eigenvalue can be verified to be the minimum eigenvalue of the overall system, i.e., 2.477, with Shapley values of sensors C_1 and C_3 alone

responsible for 90% of the observability degree even though they make up just 50% of the full sensor set.

VII. CONCLUSIONS

This paper presents an approach for fair allocation of importance of individual sensors towards overall degree of observability. This has applications in sensor selection/placement problems and filter design for state estimation. It is shown that fair allocation of contributions from each sensor in an LTI system towards degree of observability can be obtained by deriving Shapley values. This is illustrated using two commonly used observability degree metrics, (a) trace of the observability Gramian and (b) minimum eigenvalue of the observability Gramian. The relative merits of using either of the two metrics are highlighted via connection to the Shapley value axioms. It is shown that minimum eigenvalue of the Gramian is more reliable for both overall observability degree measurement as well as sensor contribution allocation via Shapley values as it measures interaction effects between sensors.

REFERENCES

- [1] Y. Bar-Shalom, X.R. Li and T. Kirubarajan, *Estimation with applications to tracking and navigation*, Wiley Inter-Science, 2001.
- [2] F.M. Ham and R.G. Brown, "Observability, eigenvalues and Kalman filtering, *IEEE Transactions on Aerospace and Electronic Systems*, Issue 2, pp 269–273, Vol. 19, 1983.
- [3] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
- [4] R.E. Kalman, "A new approach to linear filtering and prediction problems", *Transactions of the ASME – Journal of Basic Engineering*, Vol. 82, Series D, pp 35–45, 1960.
- [5] A.J. Krener and K. Ide, "Measures of unobservability", in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC)*, Shanghai, China, 15–18 Dec. 2019.
- [6] V. Lystianingrum, V.G. Agelidis and B. Hredzak, "Observability degree criteria evaluation for temperature observability in a battery string towards optimal thermal sensors placement", in *IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Singapore, 21–24 Apr. 2014.
- [7] C. Molnar, *Interpretable machine learning: A guide for making black box models explainable*, 3rd edition, christophm.github.io/interpretable-ml-book/
- [8] C.D. Peterson, R. Fraanje, B.S. Cazzalato, A.C. Zander and C.H. Hansen, "A Kalman filter approach to virtual sensing for active noise control", *Mechanical Systems and Signal Processing*, Vol. 22, Issue 2, pp. 490–508, 2008.
- [9] M. Rafieisakhaei, S. Chakravorty and P.R. Kumar, "On the use of the observability gramian for partially observed robotic path planning problems", in *IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, Victoria, Australia, 12–15 Dec. 2017.
- [10] J. Ruggaber and J. Brembeck, "A novel Kalman filter design and analysis method considering observability and dominance properties of measurands applied to vehicle state estimation", *Sensors*, 21(14), 4750.
- [11] L.S. Shapley, "A value for n -person games", *Contributions to the theory of games*, vol. 2, pp. 307–317, 1953.
- [12] J. Staiger, L. Mazzanti and F. Naetz, "State-oriented evaluation of observability and sensor placement for mechanical estimation applications", in *2023 IEEE International Conference on Mechatronics*, Loughborough, UK, Apr. 2023.
- [13] T.H. Summers, F.L. Cortesi and J. Lygeros, "On submodularity and controllability in complex dynamical networks", *IEEE Transactions on Control of Network Systems*, Vol. 3, Issue 1, pp/ 91–101, 2015.
- [14] T.A. Wenzel, K. Burnham, M.V. Blundell and R.A. Williams, "Kalman filter as a virtual sensor: Applied to automotive stability systems", *Transactions of the Institute of Measurement and Control*, Vol. 29, Issue 2, pp. 95–11, 2007.