On the relationship between maximin information and common knowledge

Aditya Mahajan

Abstract

In a recent paper, Nair ("A nonstochastic information theory for communication and state estimation," IEEE Trans. on Automatic Control, 2013) introduced the concept of maximin information between uncertain variables. Maximin information characterizes the zero-error capacity of a communication channel, and provides fundamental limits on estimating the state of a linear time-invariant plant over a noisy channel with exponential uniform accuracy. In this note, we show that maximin information is related to common knowledge as defined by Aumann ("Agreeing to disagree", Annals of Statistics, 1976 and "Interactive epistemology I: Knowledge", International Journal of Game Theory, 1999).

I. Introduction

In a recent paper, Nair [1] introduced the concept of *maximin information* between two uncertain variables. Maximin information is defined in terms of taxicab connectedness, and is related to *common random variable* as defined by Wolf and Wullschleger [2], captures the most refined common value that can be agreed upon based on individually observing two variables. Maximin information was shown to be fundamental in characterizing the performance of state estimation of a deterministic linear time invariant systems over noisy channels.

The purpose of this note is to show that maximin information defined in [1] is related to *common knowledge* defined by Aumann [3] in game theory. We refer the reader to [4] for a overview of the literature on common knowledge.

The rest of this note is organized as follows. We present the notion of maximin information and common knowledge in Section II and show that they are related in Section III.

II. Preliminaries

A. Mathematical preliminaries

Definition 1 (σ-field) Given a set $\mathcal{S}$, a collection $\mathfrak{S}$ of subsets of $\mathcal{S}$ is called a σ-field if

1) $\mathcal{S} \in \mathfrak{S}$.
2) If $\mathcal{S}_n \in \mathfrak{S}$, $n = 1, 2, \dots$, then

$$\bigcup \mathcal{S}_n \in \mathfrak{S} \quad \text{and} \quad \bigcap \mathcal{S}_n \in \mathfrak{S}$$

3) For any $\mathcal{S}' \in \mathfrak{S}$, we have that $\mathcal{S} \setminus \mathcal{S}' \in \mathfrak{S}$. □

Definition 2 (Measurable space) A set $\mathcal{S}$ together with a σ-field $\mathfrak{S}$ of its subsets is a measurable space, and is denoted by $(\mathcal{S}, \mathfrak{S})$. □

Definition 3 (σ-field generated by a collection) Given a measurable space $(\mathcal{S}, \mathfrak{S})$ and a subset $\mathfrak{E}$ of $\mathfrak{S}$, $\sigma(\mathfrak{E})$ denotes the smallest sub-σ-field of $\mathfrak{S}$ containing $\mathfrak{E}$. Equivalently, $\sigma(\mathfrak{E})$ is the intersection of all σ-fields containing $\mathfrak{E}$. □

Throughout this note, we use $(\Omega, \mathfrak{F})$ to denote the uncertainty space and call each $\omega \in \Omega$ a *state of nature* or simply as a *state*.

Definition 4 (Uncertain variables [1]) An *uncertain variable* X is a measurable mapping from $(\Omega, \mathfrak{F})$ to a measurable space $(\mathcal{X}, \mathfrak{X})$ where $\mathfrak{X}$ contains all singletons. □

Remark 1 Our definition of uncertain variables is slightly different from that of [1]. The development in [1] does not assume a measurable space. However, as we will see later, measurable spaces make it easier to connect to existing results on common knowledge. As in [1], our treatment is non-stochastic and we do not assume that a probability measure is associated with the measurable space. □

Definition 5 (Reachable set of uncertain variables) For two uncertain variables X and Y defined on the same uncertainty space $(\Omega, \mathfrak{F})$, let

$$[\![X]\!] = \{X(\omega) : \omega \in \Omega\},$$

$$[\![X, Y]\!] = \{(X(\omega), Y(\omega)) : \omega \in \Omega\}$$

denote the reachable sets of X and (X, Y). □

Definition 6 Let $\sigma(X)$ denote the σ-field generated by X, i.e.,

$$\sigma(X) = \big\{\{\omega : X(\omega) \in B\} : B \in \mathfrak{X}\big\}.$$

Furthermore, let $\mathcal{D}_X$ denote the *information partition* of Ω generated by X, i.e.,

$$\mathcal{D}_X = \{X^{-1}(x) \in \Omega : x \in \mathcal{X}\}$$

□

It is easy to verify that $\sigma(\mathcal{D}_X) \subseteq \sigma(X) \subseteq \mathfrak{F}$. Moreover, when $\mathcal{X}$ is finite or countable, $\sigma(X) = \sigma(\mathcal{D}_X)$.

B. Taxicab connectedness and maximin information

Definition 7 (Taxicab connectedness/isolation [1])

1) A pair of points (x, y) and $(x', y') \in [\![X, Y]\!]$ is called *taxicab connected* if there is a *taxicab sequence* connecting them, i.e., a finite sequence $\{(x_i, y_i)\}_{i=1}^{n}$ of points in $[\![X, Y]\!]$ such that $(x, y) = (x_1, y_1)$, $(x_n, y_n) = (x', y')$ and each point differs in at most one coordinate from its predecessor, i.e., $y_i = y_{i-1}$ and/or $x_i = x_{i-1}$, for each $i \in [2, \dots, n]$.

2) A set $\mathcal{A} \in [\![X, Y]\!]$ is called *taxicab connected* if every pair of points $\mathcal{A}$ is taxicab connected in $[\![X, Y]\!]$.
3) A pair of sets $\mathcal{A}$, $\mathcal{B}$ is called *taxicab isolated* if no point in $\mathcal{A}$ is taxicab connected in $[\![X, Y]\!]$ to any point in $\mathcal{B}$.
4) A *taxicab-isolated partition* (of $[\![X, Y]\!]$) is a cover of $[\![X, Y]\!]$ such that every pair of distinct sets in the cover is taxicab isolated.
5) A *taxicab partition* (of $[\![X, Y]\!]$) is a taxicab-isolated partition of $[\![X, Y]\!]$ each member of which is taxicab connected. □

Proposition 1 (From [1]) *The set $[\![X, Y]\!]$ has a unique taxicab partition $\mathcal{T}[X, Y]$.* □

The maximin information I_* is defined in [1] using a different concept, but then is shown to be equal to

$$I_*[X; Y] = \log|\mathcal{T}[X; Y]|$$

C. Knowledge and common knowledge

Assume that a rational agent observes the uncertain variable X. The value of X gives some information to the agent about the uncertainty space $(\Omega, \mathfrak{F})$. This information is captured by the information partition and knowledge function. For simplicity, we use "agent X" to denote an agent who observes X.

Definition 8 (Information partition and knowledge field [5]) Any uncertain variable X induces an equivalence relation $\sim_X$ on Ω defined as

$$\omega \sim_X \xi \iff X(\omega) = X(\xi)$$

The quotient space of this equivalent relation is the *information partition* of agent X. The σ-field generated by events in the information partition is the *knowledge field* of agent X. Note that the information partition of agent X is given by $\mathcal{D}_X$ and the knowledge field of agent X is given by $\sigma(X)$. □

Remark 2 The knowledge field is defined as $\sigma(X)$ rather than $\sigma(\mathcal{D}_X)$ because in probabilistic frameworks, it is more natural to work with $\sigma(X)$ rather than $\sigma(\mathcal{D}_X)$. When $\mathcal{X}$ is finite or countable, both these representation are equal. When $\mathcal{X}$ is uncoutable, $\sigma(X)$ is not the same as $\sigma(\mathcal{D}_X)$.

With either representation, two almost-surely identical random variables may generate different information fields. This technicality was resolved by Nielsen [6] by defining equivalence class of *generalized events* (i.e., set of events that differ on a null set) and considering the *Boolean σ-field* (see [7]) of generalized events. □

Definition 9 (Knowledge operator [5]) Define the knowledge operator of agent X, $K_X : \mathfrak{F} \to \mathfrak{F}$, as follows: for any $E \in \mathfrak{F}$, $K_X(E)$ is the largest element of $\sigma(X)$ included in E. Equivalently, $K_X(E)$ is the union of all elements in $\mathcal{D}_X$ included in E. □

Informally, $K_X(E)$ is the set of states ω at which agent X knows that the event E has occurred.

Definition 10 (Mutual knowledge [5]) Given two agents, one observing X and the other observing Y, define the *mutual knowledge* operator $K^1 : \mathfrak{F} \to \mathfrak{F}$ by

$$K^1(E) = K_X(E) \cap K_Y(E).$$

□

$K^1(E)$ is the set of states in which the event E is mutually known to agents X and Y, i.e., X knows that E has occurred and Y knows that E has occurred.

Definition 11 (Higher level mutual knowledge [5]) Define *m-th level mutual knowledge* K^m by

$$K^{m+1}(E) = K^1(K^m(E)), \quad m = 1, 2, \ldots.$$

□

$K^2(E)$ denotes the set of states in which all agents know E, and all know that all know E. $K^3(E)$ is the set of states in which all agents know E, all know that all know E, and all know that all know that all know E. Higher order mutual knowledge is similarly defined.

Definition 12 (Common knowldege operator [5]) The common knowledge operator between two agents X and Y is given by

$$K^\infty(E) = \bigcap_{m=1}^{\infty} K^m(E).$$

□

Thus, $K^\infty(E)$ denotes the set of states in which all agents know E, all know that all know E, all know that all know that all know E, ad infinitum.

Proposition 2 (Common knowledge field [5]) *$K^\infty(E)$ is the largest event in $\sigma(X) \cap \sigma(Y)$ that is included in E. For this reason, $\sigma(X) \cap \sigma(Y)$ is called the* common knowledge field *between X and Y.* □

III. Main result

Our main result is to show that the taxicab partition $\mathcal{T}[X;Y]$ of two uncertain variables is related to the common knowledge field. More precisely:

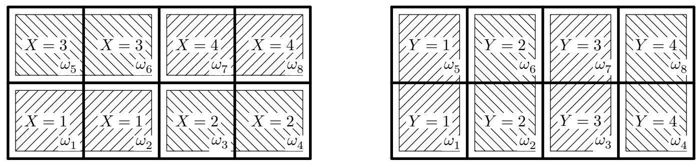
Theorem 1 *Given two uncertain variables X and Y, define an equivalence relation $\sim$ on Ω as follows: for ω, ξ in Ω, $\omega \sim \xi$ iff $(X(\omega), Y(\omega))$ and $(X(\xi), Y(\xi))$ are taxicab connected (i.e., belong to the same element of the taxicab partition $\mathcal{T}[X;Y]$). Let $\mathcal{P}$ be the quotient space of this equivalence relation. If $I_*[X;Y] < \infty$, then $\sigma(\mathcal{P})$ is the common knowledge field between X and Y.* □

Before presenting a proof, let us consider an example. Let $\Omega = \{\omega_1, \ldots, \omega_8\}$ and $\mathfrak{F}$ be the power-set of Ω. Define the uncertain variables X and Y as

$$X(\omega_i) = 1 + \lfloor (i-1)/2 \rfloor, \quad Y(\omega_i) = 1 + (i - 1 \bmod 4), \quad i = 1, \ldots, 8.$$

The information partition corresponding to these uncertain variables are shown in Fig 1. The taxicab partition is given by

$$\mathcal{T}[X;Y] = \big\{\{p(\omega_1), p(\omega_2), p(\omega_5), p(\omega_6)\}, \{p(\omega_3), p(\omega_4), p(\omega_7), p(\omega_8)\}\big\}$$

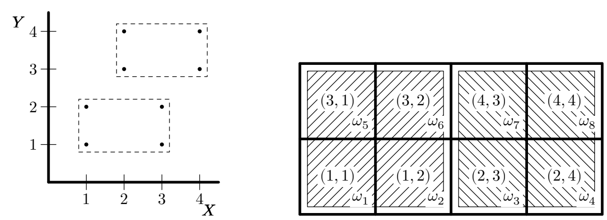


(a) $X(\omega_i) = 1 + \lfloor (i-1)/2 \rfloor$ (b) $Y(\omega_i) = 1 + (i - 1 \bmod 4)$

Fig. 1. The information partitions corresponding to two uncertain variables X and Y.

where $p(\omega)$ is a short hand notation for $(X(\omega), Y(\omega))$. This taxicab partition corresponds to the atoms of $\sigma(X) \cap \sigma(Y)$ (see Fig 2):

$$\sigma(X) \cap \sigma(Y) = \sigma\big(\{\omega_1, \omega_2, \omega_5, \omega_6\}, \{\omega_3, \omega_4, \omega_7, \omega_8\}\big).$$



(a) (b)

Fig. 2. (a) The taxicab partition and (b) the atoms of the common knowledge field corresponding to the uncertain variables X and Y.

PROOF (PROOF OF THEOREM 1) For any $\omega \in \Omega$, let $D_X(\omega)$ denote the element of the information partition $\mathcal{D}_X$ containing ω. Following [5], we say that two states ω and ξ in Ω are X-*adjacent* if $\xi \in D_X(\omega)$. Two states ω and ξ are *adjacent* if they are either X-adjacent or Y-adjacent. Adjacency is symmetric and reflexive. Define the *distance* $d(\omega, \xi)$ as the minimal length chain from ω to ξ in which successive elements are adjacent; if no such chain exists then the distance is infinite. Let $B(\omega, n)$ denote an *open* ball with center ω and radius n, set $\{\xi \in \Omega : d(\omega, \xi) < n\}$.

The notion of taxicab connectedness defined in [1] is related to adjacency. Specifically, we have the following.

Lemma 1 *Each element of* $\mathcal{P}$ *is of form* $B(\omega, \infty)$. □

PROOF If $d(\omega, \xi) = 1$, then either $X(\omega) = X(\xi)$ or $Y(\omega) = Y(\xi)$. Therefore, $(X(\omega), Y(\omega))$ and $(X(\xi), Y(\xi))$ are taxicab connected. More generally, if $d(\omega, \xi) = n$, then $(X(\omega), Y(\omega))$ and $(X(\xi), Y(\xi))$ are connected by a taxicab sequence of length n. Thus, two points $(X(\omega), Y(\omega))$ and $(X(\xi), Y(\xi))$ in $[\![X, Y]\!]$ are taxicab connected iff the distance $d(\omega, \xi)$ between them is less than infinity, i.e.,

$$\omega \sim \xi \iff d(\omega, \xi) < \infty;$$

or, equivalently,

$$\omega \sim \xi \iff \xi \in B(\omega, \infty).$$

Therefore, each element of $\mathcal{P} = \Omega/\sim$ is of the form $B(\omega, \infty)$. ■

As shown in [5], the definition of adjacency defined above is also related to common knowledge. In particular,

Lemma 2 (Lemma A.33 in [5]) *For any $E \in \Omega$ and $n \in \mathbb{N}$, $\omega \in K^n(E)$ if and only if $E \supseteq B(\omega, n+1)$.* □

By taking union over all n, we get that for any E, $\omega \in K^\infty(E)$ if and only if $E \supseteq B(\omega, \infty)$. Therefore, the sets $B(\omega, \infty)$ (and, equivalently, the elements of $\mathcal{P}$) are the smallest events of the common knowledge field. Since $I_*[X;Y]$ is finite, $\mathcal{P}$ has a finite number of elements. Therefore, $\sigma(\mathcal{P})$ is equivalent to the common knowledge field. ■

IV. Discussion

In this technical note, we have shown that the notion of maximin information defined in [1] is related to common knowledge. Such a relation was alluded to in [1]. In particular, it was stated in [1] that the taxicab partition $T[X;Y]$ represents the finest posterior knowledge that can be agreed on from individually observing X and Y. Theorem 1 formally proves the above claim when knowledge is defined in terms of knowledge operators of [3] and agreement is in the sense of common knowledge.

Both taxicab partitions and common knowledge are related to the notion of *common random variable* between two *finite* valued *random variables* due to Wolf and Wullschleger [2]. Their definition can be easily generalized to finite valued uncertain variables as follows. Let G be the bipartite graph with vertices $\mathcal{X} \cup \mathcal{Y}$; there is an edge between $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ if $(x, y) \in [\![X, Y]\!]$. The *common uncertain variable* between X and Y is the connected component of G containing both X and Y. As pointed out in [1], by construction the taxicab partitions correspond to the connected components of the bipartite graph G. This bipartite graph is also related to common knowledge. The partition of Ω obtained by the connected components of G corresponds to the atoms of $\sigma(X) \cap \sigma(Y)$.

Nair [1] proves that maximin information satisfies two properties also exhibited by Shannon information, namely

1) *More data cannot hurt*. For any uncertain variables $I_*[X;Y] \le I_*[X;Y,Z]$.
2) *Data processing inequality*. If $X \leftrightarrow Y \leftrightarrow Z$ forms a Markov uncertainty chain ([1, Definition 2.2]), then $I_*[X;Z] \le I_*[X;Y]$.

The first property is an immediate consequence of the fact that $\sigma(Y) \subseteq \sigma(Y, Z)$. Therefore, $\sigma(X) \cap \sigma(Y) \subseteq \sigma(X) \cap \sigma(Y, Z)$. The second property, is an immediate consequence of the first property and the definition of uncertainty chains (see [1]).

Based on these properties Nair [1] characterized fundamental limits for networked control systems (zero-error capacity of a channel and exponential stability of estimating the state of a linear time-invariant system over a noisy channel) using maximin information.

There is a substantial literature on analysing the information flow in decentralized systems using common information [8]–[13]. In recent years, the notion of common knowledge has also been used to obtain dynamnic programming decomposition for decentralized stochastic control problems [14]. See [15, Chapter 12] for an overview of the role of common information in decentralized control. We hope that the relationship between maximin information and common knowledge established in this paper will be useful to get more insights for the design of networked control systems.

REFERENCES

[1] G. N. Nair, "A nonstochastic information theory for communication and state estimation," *IEEE Trans. Autom. Control*, vol. 58, no. 6, pp. 1497–1510, Jun. 2013, early Access.

[2] S. Wolf and J. Wullschleger, "Zero-error information and applications in cryptography," in *Information Theory Workshop, 2004. IEEE*. IEEE, 2004, pp. 1–6.

[3] R. J. Aumann, "Agreeing to disagree," *Annals of Statistics*, no. 4, pp. 1236–1239, 1976.

[4] P. Vanderschraaf and G. Sillari, "Common knowledge," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Spring 2009.

[5] R. J. Aumann, "Interactive epistemology I: Knowledge," *International Journal of Game Theory*, vol. 28, pp. 263–300, 1999.

[6] L. T. Nielsen, "Common knowledge, communication, and convergence of beliefs," *Mathematical Social Sciences*, vol. 8, no. 1, pp. 1–14, 1984.

[7] R. Sikorski, *Boolean algebras*. Springer New York, 1969, vol. 2.

[8] V. Borkar and P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Autom. Control*, vol. 27, no. 3, pp. 650–655, Jun. 1982.

[9] J. D. Geanakoplos and H. M. Polemarchakis, "We can't disagree forever," *Journal of Economic Theory*, vol. 28, no. 1, pp. 192–200, 1982.

[10] R. B. Washburn and D. Teneketzis, "Asymptotic agreement among communicating decisionmakers," *Stochastics*, vol. 13, pp. 103–129, 1984.

[11] D. Teneketzis and P. Varaiya, "Consensus in distributed estimation with inconsistent beliefs," *Systems and Control Letters*, vol. 4, pp. 217–221, Jun. 1984.

[12] J. Tsitsiklis and M. Athans, "Convergence and asymptotic agreement in distributed decision problems," *IEEE Trans. Autom. Control*, vol. 29, no. 1, pp. 42–50, 1984.

[13] D. A. Castanon and D. Teneketzis, "Further results on the asymptotic agreement problem," *IEEE Trans. Autom. Control*, vol. 33, no. 6, pp. 515–523, 1988.

[14] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1644–1658, Jul. 2013.

[15] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.