

# Geometry of information structures, strategic measures and associated stochastic control topologies\*

Naci Saldi

*Department of Mathematics, Bilkent University, Çankaya, Ankara, Turkey*  
e-mail: [naci.saldi@bilkent.edu.tr](mailto:naci.saldi@bilkent.edu.tr)

and

Serdar Yüksel

*Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada*  
e-mail: [yuksel@queensu.ca](mailto:yuksel@queensu.ca)

**Abstract:** In many areas of applied mathematics, decentralization of information is a ubiquitous attribute affecting how to approach a stochastic optimization, decision and estimation, or control problem. In this review article, we present a general formulation of information structures under a probability theoretic and geometric formulation. We define information structures, place various topologies on them, and study closedness, compactness and convexity properties on the strategic measures induced by information structures and decentralized control/decision policies under varying degrees of relaxations with regard to access to private or common randomness. Ultimately, we present existence and tight approximation results for optimal decision/control policies. We discuss various lower bounding techniques, through relaxations and convex programs ranging from classically realizable and classically non-realizable (such as quantum theoretic and non-signaling) relaxations. For each of these, we establish closedness and convexity properties and also a hierarchy of correlation structures. As a further theme, we review and introduce various topologies on decision/control strategies defined independent of information structures, but for which information structures determine whether the topologies entail utility in arriving at existence, compactness, convexification or approximation results. These approaches, which we term as the strategic measures approach and the control topology approach, lead to complementary results on existence, approximations and upper and lower bounds in optimal decentralized stochastic decision, estimation, and control problems.

**MSC2020 subject classifications:** Primary 93E03, 60B10; secondary 90B99.

Received October 2020.

## Contents

1 Introduction . . . . .	452
--------------------------	-----

\*This research was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the BAGEP Award of the Science Academy.

1.1	Summary of results and outline of the article . . . . .	453
2	Information structures . . . . .	455
2.1	Witsenhausen's characterization of information structures . . . . .	455
2.2	Independent-measurements reduction of sequential teams . . . . .	457
3	Decentralized strategic measures . . . . .	458
3.1	Measurable policies as a subset of randomized policies and strategic measures . . . . .	459
3.2	Sets of strategic measures for static teams . . . . .	460
3.3	Sets of strategic measures for dynamic teams in the absence of static reduction . . . . .	462
3.4	Measurability properties of sets of strategic measures . . . . .	464
4	Relative compactness and closedness of strategic measures, and existence of optimal policies . . . . .	464
5	Convex relaxations . . . . .	472
5.1	Non-signaling relaxation . . . . .	473
5.2	Quantum-correlated relaxation . . . . .	475
5.3	Relations between convex relaxations . . . . .	483
5.4	Linear programming formulation and its dual . . . . .	487
6	Finite approximations of information structures via quantization . . . . .	489
6.1	Finite measurement approximate models . . . . .	493
7	Product topology induced by individual policies and relations with information structures . . . . .	496
7.1	Some remarks on classical deterministic and stochastic control and Young measures . . . . .	496
7.2	Trading-off continuity with compactness: a product space approach for individual control policies . . . . .	498
7.3	Policies defined by conditional independence given measurements . . . . .	500
7.4	Weak-* topology on randomized policies . . . . .	501
7.5	Exchangeability, infinite products of individual policies and optimality of symmetric randomized policies for mean-field teams . . . . .	505
7.6	Extended weak convergence, topology of information, and adapted topologies . . . . .	510
8	Revisiting (and avoiding a subtle potential error in) relaxed policies for partially observed stochastic control . . . . .	511
8.1	Discrete-time case . . . . .	513
8.2	A counterexample . . . . .	515
9	Conclusion, topics left out and some open problems . . . . .	517
9.1	Concluding remarks . . . . .	517
9.2	Discussion on existence and some related results . . . . .	518
9.3	Some open problems . . . . .	519
	Acknowledgements . . . . .	523
	References . . . . .	524

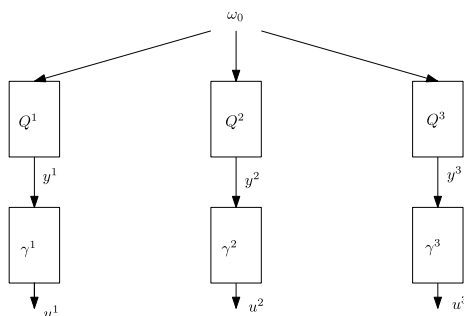


FIG 1. An example of a (decentralized) information structure.

## 1. Introduction

In statistical decision theory, stochastic control theory, information theory, game theory, economics, quantum physics and computer science; information structures determine which decision maker/agent/controller knows what information. To help the reader gain an appreciation of what the paper entails, consider the simple setup depicted in Figure 1.

Here, three decision makers have access to some local information (e.g.  $y^2$ ) and each one selects an action (e.g.,  $u^2$ ). This selection has to be based only on the available information (and perhaps with some additional independent randomization device).

One question of interest in many of the aforementioned disciplines is to study the geometric and topological properties of sets of all possible correlation structures with regard to the random variables  $(\omega_0, y^1, y^2, y^3, u^1, u^2, u^3)$  (which will be termed as a *strategic measure*). This setup can also be generalized to *dynamic* interactions among the decision makers where, e.g., the action  $u^1$  can affect the measurement of  $y^2$  and to setups where the decision makers have access to common randomness. Further correlation structures include those that are described by quantum mechanics (and, thus, with quantum entanglement). Furthermore, one may wish to provide convex relaxations leading to what we will study as non-signaling relaxations.

Suppose further that these decision makers wish to minimize a cost criterion of the form  $E[c(\omega_0, u^1, u^2, u^3)]$ . What can be said about the optimal solutions, their existence, approximations, convexity properties and convexifications, and numerical methods for exact solutions or upper or lower bounds?

The operational goal in our article is, for setups including that depicted in Figure 1 but also under much more general *information structures*, to develop a probability theoretic and topological approach to information structures and decision/control policies towards arriving at: (i) convexity properties, (ii) continuity and compactness properties, (iii) existence results, (iv) convex relaxations to facilitate numerical and analytical methods, and (v) approximation results for optimal policies under decentralized information structures.

Different scientific communities have their own notation, terminology, and machinery to study such problems. While in our analysis we will often have a bias towards a stochastic control theoretic angle and language, we will be broadly touching on ideas from various disciplines and often explain parallel contributions in the literature. Accordingly, the scope of our article entails results from a very broad literature across a variety of disciplines (ranging from game theory, quantum physics, information theory, probability, and control). Some of the results presented here have not been published elsewhere, some are based on the authors' prior work, but much of the material is a re-interpretation and unification of related results in the literature (both across various disciplines and across time, meaning that classical results are re-interpreted with more modern findings) whose connections have not previously been made explicit or reported, to our knowledge.

### ***1.1. Summary of results and outline of the article***

The article has two main themes: In the first theme, we study topological and geometric properties of correlation structures induced on strategic measures under decision/control policies and information structures. These strategic measures are probability measures defined on the product spaces of local measurements and actions, as well as external random variables; these will be made precise in the following sections.

In Section 2 we introduce information structures in decentralized decision making. In Section 3, we introduce probability measures (i.e., strategic measures) induced on the product spaces of state, measurements and actions under admissible, and some relaxed, control policies. In Section 4, we study compactness properties of the set of strategic measures and state existence results on optimal policies. In Section 5, we introduce convex relaxations on strategic measures, including those that are generated under non-signaling and quantum information theoretic relaxations. In Section 6, we study finite measurement information structures obtained via quantization of continuous-space measurement spaces, and establish their near optimality.

With regard to these various strategic measures introduced in these sections, we establish the hierarchies depicted, and also with a summary of the results reported, in Figure 2. Here, the sets are generated subject to progressively more relaxed informational or action-set constraints, as one progresses from the inner-most correlation set to the outer-most set along the depictions presented in Figure 2.

As our second main theme, we present a *control topologies* approach, where we will review or introduce various topologies on decision/control strategies (defined independently from information structures), but for which information structures determine whether the topologies entail utility in arriving at existence, compactness, convexification or approximation results. These topologies include those generated by the product metric of individual controllers viewed as subsets of appropriate probability measures, relaxed control policies and wide-sense admissible control policies, weak\* topology on randomized policies, policies

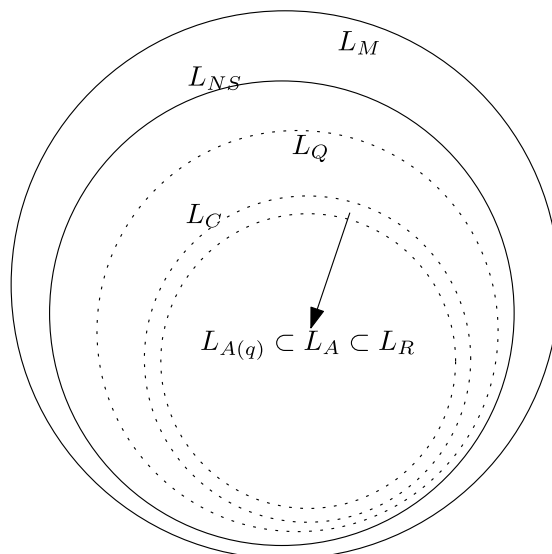


FIG 2. Hierarchies and a summary of the results reported in the paper on decentralized strategic measures and correlations: (i)  $L_{A(q)}(\mu)$ , quantized admissible strategic measures, is not closed, not convex, dense in  $L_A(\mu)$ . (ii)  $L_A(\mu)$ , set of strategic measures induced by admissible measurable functions, is not closed, not convex (under independent static reduction  $L_A(\mu)$  is dense in  $L_R(\mu)$  under mild conditions). (iii)  $L_R(\mu)$ , those generated with local private (relaxed policies) randomness, is typically not closed (under independent static reduction:  $L_R(\mu)$  is closed; which then leads to a general existence result on optimization problems with lower semi-continuous cost functions), not convex. (iv)  $L_C(\mu)$ , those with arbitrary common randomness, convex, typically not closed but compact when  $L_R(\mu)$  is. (v)  $L_Q(\mu)$ , quantum strategic measures, convex, not closed (but closed under dimension constraints). (vi)  $L_{NS}(\mu)$ , non-signaling strategic measures, is closed and convex. (vii)  $L_M(\mu)$ , those with local-Markov conditional independence, is closed and convex.

satisfying conditional independence properties leading to a universal dynamic program, and topologies that satisfy a closedness property under weak convergence of finite dimensional marginals.

In Section 7, we study this alternative approach of placing topologies directly on individual policies. Several examples and topologies are provided with various complementary results on existence reported.

These two approaches, which we name as *the strategic measures approach* vs. *the control topology approach*, lead to complementary results on existence, approximations and upper and lower bounds and solution methods in optimization problems.

In Section 8, we review and revisit partially observed Markov decision processes, and highlight that when a relaxation of control policies is allowed, this may lead to a strict improvement of optimal performance, thus invalidating the main purpose of relaxations.

We finally present, together with concluding remarks, a number of future

directions and a collection of some open problems, in Section 9.

## 2. Information structures

### 2.1. Witsenhausen's characterization of information structures

Hans Witsenhausen's contributions [141, 139, 137, 142, 140] to stochastic control theory, and his characterization of information structures in decentralized stochastic control have been crucial in our modern understanding of decentralized stochastic control and decision theory. In this section, we introduce the characterizations as laid out by Witsenhausen, termed *the Intrinsic Model* [141]; see [149] and [100] for a more comprehensive overview and further characterizations and classifications of information structures. In this model (described in discrete time), any action applied at any given time is regarded as applied by an individual decision maker/agent, who acts only once. One particular advantage of this model, in addition to its generality, is that the characterizations regarding information structures can be concisely described.

Suppose that in the decentralized system considered below, there is a predefined order in which the decision makers act. Such systems are called *sequential teams* (for non-sequential teams, we refer the reader to Andersland and Teneketzis [5], [6] and Teneketzis [130], in addition to Witsenhausen [139] and [149, p. 113]). Suppose that in the following, the action and measurement spaces are standard Borel spaces; that is, Borel subsets of Polish (complete, separable, and metric) spaces. In the context of a sequential system, the *Intrinsic Model* has the following components:

- A collection of *measurable spaces*

$$\{(\Omega, \mathcal{F}), (\mathbb{U}^i, \mathcal{U}^i), (\mathbb{Y}^i, \mathcal{Y}^i), i \in \mathcal{N}\},$$

with  $\mathcal{N} := \{1, 2, \dots, N\}$ , specifying the system's distinguishable events, and the control and measurement spaces of decision makers (DMs). Here  $N = |\mathcal{N}|$  is the number of control actions taken, and each of these actions is taken by an individual (different) DM (hence, even a DM with perfect recall can be regarded as a separate decision maker every time it acts). The pair  $(\Omega, \mathcal{F})$  is a measurable space (on which an underlying probability may be defined). The pair  $(\mathbb{U}^i, \mathcal{U}^i)$  denotes the Borel space with its Borel  $\sigma$ -algebra from which the action  $u^i$  of DM  $i$  is selected. The pair  $(\mathbb{Y}^i, \mathcal{Y}^i)$  denotes the observation/measurement space with its Borel  $\sigma$ -algebra for DM  $i$ .

- A *measurement constraint* which establishes the connection between the observation variables and the system's distinguishable events. The  $\mathbb{Y}^i$ -valued observation variables are given by

$$y^i = \eta^i(\omega, \mathbf{u}^{[1, i-1]}),$$

where  $\mathbf{u}^{[1, i-1]} = \{u^k, k \leq i-1\}$  and  $\eta^i$  is a measurable function. Hence, the information variable  $y^i$  induces a  $\sigma$ -field, denoted by  $\sigma(y^i)$ , over  $\Omega \times \prod_{k=1}^{i-1} \mathbb{U}^k$

- A *design constraint* which restricts the set of admissible  $N$ -tuple control laws

$$\underline{\gamma} = \{\gamma^1, \gamma^2, \dots, \gamma^N\},$$

also called *designs* or *policies*, to the set of all measurable control functions, so that  $u^i = \gamma^i(y^i)$ , with  $y^i = \eta^i(\omega, \mathbf{u}^{[1, i-1]})$  and  $\gamma^i$  is a measurable function. Let  $\Gamma^i$  denote the set of all admissible policies for DM  $i$  and let  $\mathbf{\Gamma} = \prod_{i=1}^N \Gamma^i$ .

We note that, the intrinsic model of Witsenhausen gives a set-theoretic characterization of information fields, however, for standard Borel spaces, the model above is equivalent to that of Witsenhausen's. Additionally, we introduce a fourth component:

- A *probability measure*  $P$  defined on  $(\Omega, \mathcal{F})$  which describes the uncertainty on the random events in the model.

Under this intrinsic model, an Information structure (IS) is *dynamic* if the information available to at least one DM is affected by the action of at least one other DM. An IS is *static*, if the information available to every decision maker is only affected by exogenous disturbances (i.e., state of the nature)  $\omega \in \Omega$ ; that is no other decision maker can affect the information at any given decision maker. Figure 1 is a depiction for a static team problem, where  $Q^i(y^i \in \cdot | \omega_0) := P(\eta^i(\omega) \in \cdot | \omega_0)$ ,  $i = 1, 2, 3$ .

ISs can also be classified as *classical*, *quasi-classical* or *nonclassical*. An IS is *classical* if  $y^i$  contains all of the information available to DM  $k$  for  $k < i$ ; that is, information is expanding (also known as the *perfect-recall* property). An IS is *quasi-classical* or *partially nested*, if whenever  $u^k$ , for some  $k < i$ , affects  $y^i$  through the measurement function  $\eta^i$ ,  $y^i$  contains  $y^k$  (that is  $\sigma(y^k) \subset \sigma(y^i)$ ). An IS which is not partially nested is *nonclassical*.

For any  $N$ -tuple of policies  $\underline{\gamma} = \{\gamma^1, \dots, \gamma^N\}$  let a cost function be defined as:

$$J(\underline{\gamma}) = E\mathcal{L}[c(\omega_0, \mathbf{u})] = E[c(\omega_0, \gamma^1(y^1), \dots, \gamma^N(y^N))], \quad (1)$$

for some non-negative measurable loss (or cost) function  $c : \Omega_0 \times \prod_{i=1}^N \mathcal{U}^i \rightarrow \mathbb{R}_+$ . Here, we have the notation  $\mathbf{u} = \{u^i, i \in \mathcal{N}\}$ , and  $\omega_0$  may be viewed as the cost function relevant exogenous variable contained in  $\omega$ .

**Definition 2.1.** For a given stochastic team problem with a given information structure, an  $N$ -tuple of policies  $\underline{\gamma}^* := (\gamma^{1*}, \dots, \gamma^{N*}) \in \mathbf{\Gamma}$  is an optimal team decision rule (team-optimal decision rule or simply team-optimal solution) if

$$J(\underline{\gamma}^*) = \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}) =: J^*. \quad (2)$$

The expected cost achieved by this strategy  $J^*$  is the optimal team cost.

In the following, we will denote by bold letters the ensemble of random variables across the DMs; that is  $\mathbf{y} = \{y^i, i = 1, \dots, N\}$  and  $\mathbf{u} = \{u^i, i = 1, \dots, N\}$ .

## 2.2. Independent-measurements reduction of sequential teams

Following Witsenhausen [142, Eqn (4.2)], as reviewed in [149, Section 3.7], we say that two information structures are equivalent if: (i) The policy spaces are equivalent/isomorphic in the sense that policies under one information structure are realizable under the other information structure, (ii) the costs achieved under equivalent policies are identical, and (iii) if there are constraints in the admissible policies, the isomorphism among the policy spaces preserves the constraint conditions.

A large class of sequential dynamic team problems admit an equivalent information structure which is static. This is called the *static reduction* of a dynamic team problem.

For some, but not all, results to be presented in our paper, we need to go beyond a static reduction, and we will need to make the measurements independent of each other as well as  $\omega_0$ . This is not possible for every team which admits a static reduction, for example quasi-classical team problems with LQG models [73] do not generally admit such a further reduction, since the measurements are partially nested: for partially nested (or quasi-classical) information structures, static reduction has been studied by Ho and Chu in the specific context of LQG systems [73] and for a class of non-linear systems satisfying restrictive invertibility properties [74].

Witsenhausen refers to such an information structure as *independent static* in [142, Section 4.2(e)]. One can also reduce a static team problem into an independent static form.

Note that the Intrinsic Model is equivalent to the following model [142]. The probability space  $(\Omega, \mathcal{F}, P)$  is the product of  $N+1$  probability spaces  $(\Omega_i, \mathcal{F}_i, P_i)$ ,  $i = 0, \dots, N$ ; that is,  $\Omega = \prod_{i=0}^N \Omega_i$  and  $P(d\omega) = \prod_{i=0}^N P_i(d\omega^i)$ . Each DM  $i$  measures  $y^i = \eta^i(\omega_0, \omega_i, u^1, \dots, u^{i-1})$  and the decisions are generated by  $u^i = \gamma^i(y^i)$ , with  $1 \leq i \leq N$ . Here  $\omega = (\omega_0, \omega_1, \dots, \omega_N)$  are primitive (exogenous) variables. Under this equivalent model, we can view  $\eta^i(\omega_0, \omega_i, u^1, \dots, u^{i-1})$  as a measurement channel with input  $(\omega_0, u^1, \dots, u^{i-1})$  and output  $y^i$ , where  $\omega_i$  is the noise; that is

$$g^i(dy^i|\omega_0, u^1, u^2, \dots, u^{i-1})$$

is a (controlled) stochastic kernel (to be defined later). Equivalently, through standard stochastic realization results (see [65, Lemma 1.2] or [35, Lemma 3.1]), we can represent any (controlled) stochastic kernel  $g^i(dy^i|\omega_0, u^1, u^2, \dots, u^{i-1})$  in a functional form  $y^i = \eta^i(\omega_0, \omega_i, u^1, u^2, \dots, u^{i-1})$  for some independent  $\omega_i$  and measurable  $\eta^i$ .

This team model admits an *independent static* reduction provided that the following absolute continuity condition holds:

(AC): For every  $i \in \mathcal{N}$ , there exists a reference probability measure  $Q^i$  on  $\mathbb{Y}^i$  and a measurable function  $f_i$  such that for all Borel  $S \subset \mathbb{Y}^i$ :

$$\begin{aligned} g^i(y^i \in S|\omega_0, u^1, u^2, \dots, u^{i-1}) \\ = \int_S f_i(y^i, \omega_0, u^1, u^2, \dots, u^{i-1}) Q^i(dy^i). \end{aligned}$$



Under this absolute continuity condition, since the action of each DM is determined by the measurement variables under a policy, we can write

$$P(d\omega_0, d\mathbf{y}, d\mathbf{u}) = P(d\omega_0) \prod_{i=1}^N \left( f_i(y^i, \omega_0, u^1, u^2, \dots, u^{i-1}) Q^i(dy^i) 1_{\{\gamma^i(y^i) \in du^i\}} \right).$$

The cost function  $J(\underline{\gamma})$  can then be written as

$$J(\underline{\gamma}) = \int P(d\omega_0) \prod_{i=1}^N \left( f_i(y^i, \omega_0, u^1, u^2, \dots, u^{i-1}) Q^i(dy^i) \right) c(\omega_0, \mathbf{u}),$$

where  $u^i = \gamma^i(y^i)$  for  $1 \leq i \leq N$ . Now, the measurement variables can be regarded as independent from each other and also from  $\omega_0$ , and by incorporating the density functions  $\{f_i\}$  into  $c$ , we can obtain an equivalent *independent static team* problem. Hence, the essential step is to appropriately adjust the probability space and the cost function. The new cost function may now explicitly depend on the measurement values; that is,

$$c_s(\omega_0, \mathbf{y}, \mathbf{u}) = c(\omega_0, \mathbf{u}) \prod_{i=1}^N f_i(y^i, \omega_0, u^1, u^2, \dots, u^{i-1}). \quad (3)$$

Here we can reformulate even a static team to one which is, clearly still static, but now with independent measurements which are also independent from the cost relevant exogenous variable  $\omega_0$ . Such a condition is in general not restrictive. Indeed, as Witsenhausen notes, a static reduction always holds when the measurement variables take values from countable set since a reference measure as in  $Q^i$  above can be always constructed on the measurement space  $\mathbb{Y}^i$  (e.g.,  $Q^i(z) = \sum_{j \geq 1} 2^{-j} 1_{\{z=m_j\}}$  where  $\mathbb{Y}^i = \{m_j, j \in \mathbb{N}\}$ ) so that the absolute continuity condition always holds. We refer the reader to [44] for relations with classical continuous-time stochastic control where the relation with Girsanov's classical measure transformation [66][26] is recognized, and [149, p. 114] for further discussions. For discrete-time partially observed stochastic control, similar arguments had been presented, e.g. by Borkar in [36], [38].

### 3. Decentralized strategic measures

For classical stochastic control problems, strategic measures were defined (see [125], [115], [55] and [60]) as the set of probability measures induced on the product (sequence) spaces of the states, measurements, and actions; that is, given an initial state distribution and a policy, one can uniquely define a probability measure on the product space of the states, measurements, and actions. Certain measurability, compactness, and convexity properties of strategic measures for classical stochastic control problems were studied in [55, 115, 60, 31].

In [151], strategic measures for decentralized stochastic control problems were introduced and many of their properties were established. For decentralized

stochastic control problems, considering the set of strategic measures along with compactification and/or convexification of these sets of measures through introducing private and/or common randomness allow one to place operationally flexible topologies (such as those leading to a standard Borel space, e.g., weak convergence topology, among others) on the set of strategic measures, as we will study in the following.

### 3.1. Measurable policies as a subset of randomized policies and strategic measures

A common method in control theory is to view a measurable policy as a special case of *relaxed* policies where relaxation is often employed by randomization. Such an approach has been ubiquitously adopted in various fields often with different terminology (e.g., relaxed controls (Young topology) in optimal deterministic control [103] [146], distributional strategies in economics [105] [104], local hidden variables in quantum information theory [21, 48, 135], optimal quantization [150] etc.).

We recall here the following representation result [35] (see also Section 7.1). Let  $\mathbb{X}, \mathbb{M}$  be Borel spaces. Let the notation  $\mathcal{P}(\mathbb{X})$  denote the set of probability measures on  $\mathbb{X}$ . Consider the set of probability measures

$$\Theta := \left\{ \zeta \in \mathcal{P}(\mathbb{X} \times \mathbb{M}) : \zeta(dx, dm) = P(dx) Q^f(dm|x), \right. \\ \left. Q^f(\cdot|x) = 1_{\{f(x) \in \cdot\}}, f : \mathbb{X} \rightarrow \mathbb{M} \right\}, \quad (4)$$

on  $\mathbb{X} \times \mathbb{M}$  having fixed input marginal  $P$  on  $\mathbb{X}$  and the stochastic kernel from  $\mathbb{X}$  to  $\mathbb{M}$  is realized by some measurable function  $f : \mathbb{X} \rightarrow \mathbb{M}$ . We equip this set with the weak convergence topology. This set is the (Borel measurable) set of the extreme points of the set of probability measures on  $\mathbb{X} \times \mathbb{M}$  with a fixed marginal  $P$  on  $\mathbb{X}$ . For compact  $\mathbb{M}$ , the Borel measurability of  $\Theta$  follows from [114] since the set of probability measures on  $\mathbb{X} \times \mathbb{M}$  with a fixed marginal  $P$  on  $\mathbb{X}$  is a convex and compact set in a complete separable metric space, and therefore, the set of its extreme points is Borel measurable. Moreover, the non-compact case holds by [35, Lemma 2.3]. Furthermore, given a fixed marginal  $P$  on  $\mathbb{X}$ , any stochastic kernel  $Q$  from  $\mathbb{X}$  to  $\mathbb{M}$  can be identified by a probability measure  $\xi \in \mathcal{P}(\Theta)$  such that

$$Q(\cdot|x) = \int_{\Theta} \xi(dQ^f) Q^f(\cdot|x). \quad (5)$$

In particular, a stochastic kernel can thus be viewed as an integral representation over probability measures induced by deterministic policies.

For a team setup, for any DM  $k$ , let

$$\Theta^k := \left\{ \zeta \in \mathcal{P}(\mathbb{Y}^k \times \mathbb{U}^k) : \zeta = P_k Q^{\gamma^k}, \right. \\ \left. Q^{\gamma^k}(\cdot|y^k) = 1_{\{\gamma^k(y^k) \in \cdot\}}, \gamma^k \in \Gamma^k, P_k(\cdot) = P(y^k \in \cdot) \right\}.$$

For a static team,  $P_k$  would be fixed; that is, independent of the policies of the preceding DMs. Therefore, in the static case, in view of (5), any element  $\zeta \in \mathcal{P}(\mathbb{Y}^k \times \mathbb{U}^k)$  with fixed marginal  $P_k$  on  $\mathbb{Y}^k$  can be expressed as a mixture of  $\Theta^k$

$$\zeta(A) = \int_{\Theta^k} \xi^k(dQ) Q(A), \quad A \in \mathcal{B}(\mathbb{Y}^k \times \mathbb{U}^k), \quad (6)$$

for some  $\xi \in \mathcal{P}(\Theta^k)$ . In the sequel, we generalize this idea to the set of strategic measures induced by measurable policies and define various relaxed policies that are obtained as a mixture of measurable policies. Indeed, instead of viewing  $N$ -tuple of policies as the joint strategy of DMs, we regard the induced probability distribution on the product space of state, measurements, and actions as the joint strategy and name it the *strategic measure*. However, we will recognize that the  $N$ -tuple view above is also of both operational and mathematical interest as will be detailed in Section 7.

### 3.2. Sets of strategic measures for static teams

Consider a static team problem defined under Witsenhausen's intrinsic model. In the following,  $B = B^0 \times \prod_{k=1}^N (A^k \times B^k)$  are used to denote the cylindrical Borel sets in  $\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$ .

Let  $L_A(\mu)$  be the set of strategic measures induced by all admissible measurable policies with  $(\omega_0, \mathbf{y}) \sim \mu$ ; that is,  $P \in L_A(\mu) \subset \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right)$  if and only if

$$P(B) = \int_{B^0 \times \prod_{k=1}^N A^k} \mu(d\omega_0, d\mathbf{y}) \prod_{k=1}^N 1_{\{u^k = \gamma^k(y^k) \in B^k\}}, \quad (7)$$

for all cylindrical  $B \in \mathcal{B}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right)$ , where  $\gamma^k \in \Gamma^k$  for  $k = 1, \dots, N$ . Let  $L_A(\mu, \underline{\gamma})$  be the strategic measure under a particular strategy  $\underline{\gamma} \in \mathbf{\Gamma}$ .

The first relaxation is obtained via individual randomization of policies. Namely, let  $L_R(\mu)$  be the set of strategic measures induced by all individually randomized team policies where  $\omega_0, \mathbf{y} \sim \mu$ ; that is,

$$L_R(\mu) := \left\{ P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right) : P(B) = \int_B \mu(d\omega_0, d\mathbf{y}) \prod_{k=1}^N \Pi^k(du^k | y^k) \right\},$$

where  $\Pi^k$  is a stochastic kernel from  $\mathbb{Y}^k$  to  $\mathbb{U}^k$  for each  $k = 1, \dots, N$ .

Another relaxation, which is stronger than the former one, is obtained by taking the mixture of the elements of  $L_A(\mu)$ . To this end, define  $\Upsilon = [0, 1]^N$ . We then let

$$L_C(\mu) := \left\{ P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right) :$$

$$P(B) = \int \eta(dz) L_A(\mu, \underline{\gamma}(z))(B), \eta \in \mathcal{P}(\Upsilon) \Big\},$$

where  $\underline{\gamma}(z)$  denotes a collection of team policies measurably parametrized by  $z \in \Upsilon$  so that the map  $L_A(\mu, \underline{\gamma}(\cdot)) : \Upsilon \rightarrow L_A(\mu)$  is Borel measurable as  $L_A(\mu)$  is a Borel subset of  $\mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right)$  under weak convergence topology (as we will see in Theorem 3.3).

Let  $L_{CR,\eta}$  denote the set of strategic measures that are induced by some fixed but common independent randomness (with probability measure  $\eta$ ) and arbitrary private independent randomness; that is,

$$L_{CR,\eta}(\mu) := \left\{ P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right) : \right. \\ \left. P(B) = \int_{B \times \Upsilon} \eta(dz) \mu(d\omega_0, d\mathbf{y}) \prod_k \Pi^k(du^k | y^k, z) \right\},$$

where  $\Pi^k$  takes place from the set of stochastic kernels from  $\mathbb{Y}^k \times \Upsilon$  to  $\mathbb{U}^k$  for each  $k = 1, \dots, N$ . Here, the common randomness  $\eta$  is fixed.

Let  $L_{CCR}$  denote the set of strategic measures that are induced by some arbitrary but common independent randomness and arbitrary private independent randomness, as in  $L_C(\mu)$ ; that is,

$$L_{CCR}(\mu) := \left\{ P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right) : \right. \\ \left. P(B) = \int_{B \times \Upsilon} \eta(dz) \mu(d\omega_0, d\mathbf{y}) \prod_k \Pi^k(du^k | y^k, z), \eta \in \mathcal{P}(\Upsilon) \right\},$$

where  $\Pi^k$  takes place from the set of stochastic kernels from  $\mathbb{Y}^k \times \Upsilon$  to  $\mathbb{U}^k$  for each  $k = 1, \dots, N$ . Here, the common randomness  $\eta$  is arbitrary, unlike  $L_{CR,\eta}(\mu)$ . The following result, essentially from [151], states some structural results about above-defined sets of strategic measures. In particular, it establishes convexity related properties of these sets.

**Theorem 3.1.** *Consider a static team problem. Then, we have the following characterizations.*

(i)  $L_R(\mu)$  has the following representation:

$$L_R(\mu) = \left\{ P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right) : P(B) = \int U(dz) L_A(\mu, \underline{\gamma}(z))(B), \right. \\ \left. U \in \mathcal{P}(\Upsilon), U(dv_1, \dots, dv_N) = \prod_s \eta_k(dv_k), \eta_k \in \mathcal{P}([0, 1]) \right\};$$

that is,  $U \in \mathcal{P}(\Upsilon)$  is constructed by the product of  $N$  independent random variables on  $[0, 1]$ .

(ii)  $L_C(\mu) = L_{CCR}(\mu)$  and this is a convex set. The set of extreme points of  $L_C(\mu)$  is  $L_A(\mu)$ . Furthermore,  $L_R(\mu) \subset L_C(\mu)$ .

(iii) We have the following equalities:

$$\inf_{\gamma \in \Gamma} J(\gamma) = \inf_{P \in L_A(\mu)} \int P(ds)c(s) = \inf_{P \in L_R(\mu)} \int P(ds)c(s) = \inf_{P \in L_C(\mu)} \int P(ds)c(s).$$

In particular, deterministic policies are optimal among the randomized class. In other words, individual and common randomness does not improve the optimal team cost.

(iv) The sets  $L_R(\mu)$  and  $L_{CR,\eta}(\mu)$  are not convex. In particular, the presence of independent or (fixed) common randomness does not convexify the set of strategic measures.

(v)  $L_R(\mu)$  and  $L_C(\mu)$  are not necessarily weakly closed.

With  $\mathbb{X}$  standard Borel space, let us recall the following: we say that a sequence of probability measures (on  $\mathbb{X}$ )  $P_n \rightarrow P$  weakly if

$$\int P_n(dx)f(x) \rightarrow \int P(dx)f(x)$$

for all  $f \in C_b(\mathbb{X})$ . If this holds for every  $f$  measurable and bounded, we say convergence holds *setwise*.

To establish (v), we will present two counterexamples below in Theorem 4.2 and Theorem 4.3.

**Remark 3.1.** (i) In the economics and game theory literature, information structures have been studied extensively. Stochastic team problems are called identical interest games. In this literature,  $L_C(\mu)$  appears in the analysis of Aumann's correlated equilibrium [13]. Common and independent randomness discussions appear in the analysis of comparison of information structures [97]. For further discussions, including a multi-stage generalization known as communication equilibria, see [64]. For a detailed treatment, we refer the reader to [104, p. 131].

(ii) In quantum mechanics, the set  $L_C(\mu)$  corresponds to the set of local hidden variable correlations: in a well-known paper [56], Einstein, Podolsky, and Rosen argued that due to the probabilistic nature of quantum mechanics, the description of randomness should be classically explained by a sufficiently large local hidden variable, where the local hidden variable describes the uncertainties in an experimental setup. This will be discussed later in the article when quantum correlations will be presented.

### 3.3. Sets of strategic measures for dynamic teams in the absence of static reduction

Note that if the dynamic team setup admits a static reduction (in particular independent static reduction), then one can define strategic measures by considering the equivalent static problem and characterize the convexity properties

of the set of strategic measures, as done in the previous section. In this section, we suppose that dynamic team does not admit a static reduction. Let  $\mu$  be the distribution of  $\omega_0$ . Recall that in the dynamic setup, the distribution of measurements  $\mathbf{y}$  is not fixed as opposed to the static case. In this case, we present the following characterization for strategic measures in dynamic sequential teams. Let, for all  $n \in \mathcal{N}$ ,

$$h_n = \{\omega_0, y^1, u^1, \dots, y^{n-1}, u^{n-1}, y^n, u^n\},$$

and  $p_n(dy^n|h_{n-1}) := P(dy^n|h_{n-1})$  be the transition kernel characterizing the measurements of DM  $n$  according to the intrinsic model. We note that this may be obtained by the relation:

$$\begin{aligned} & p_n(y^n \in \cdot | \omega_0, y^1, u^1, \dots, y^{n-1}, u^{n-1}) \\ & := P\left(\eta^i(\omega, \mathbf{u}^{[1, i-1]}) \in \cdot \middle| \omega_0, y^1, u^1, \dots, y^{n-1}, u^{n-1}\right) \\ & = P\left(g^n(\omega_0, \omega_n, u^1, \dots, u^{n-1}) \in \cdot \middle| \omega_0, y^1, u^1, \dots, y^{n-1}, u^{n-1}\right). \end{aligned} \quad (8)$$

Note that once a policy is fixed,  $p_n(dy^n|h_{n-1})$  represents the conditional distribution of  $y^n$  given the past history  $h_{n-1}$ . Let  $L_A(\mu)$  be the set of strategic measures induced by measurable policies and let  $L_R(\mu)$  be the set of strategic measures induced by individually randomized policies for the dynamic team. We have the following characterizations of  $L_A(\mu)$  and  $L_R(\mu)$  that are quite useful when establishing the closedness of these sets.

**Theorem 3.2** ([151, Theorem 2.2]). *Consider a dynamic team problem that does not admit a static reduction. Then, we have the following characterizations.*

(i) *A probability measure  $P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right)$  is a strategic measure induced by a measurable policy (that is in  $L_A(\mu)$ ) if and only if, for every  $n = 1, \dots, N$ , we have*

$$\int P(dh_{n-1}, dy^n) g(h_{n-1}, y^n) = \int P(dh_{n-1}) \left( \int_{\mathbb{Y}^n} g(h_{n-1}, z) p_n(dz|h_{n-1}) \right)$$

and

$$\int P(dh_n) g(h_{n-1}, y^n, u^n) = \int P(dh_{n-1}, dy^n) \left( \int_{\mathbb{U}^n} g(h_{n-1}, y^n, a) 1_{\{\gamma^n(y^n) \in da\}} \right),$$

for all continuous and bounded function  $g$  with appropriate arguments, where  $P(d\omega_0) = \mu(dw_0)$  and  $\gamma^n \in \Gamma^n$ .

(ii) *A probability measure  $P \in \mathcal{P}\left(\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)\right)$  is a strategic measure induced by a individually randomized policy (that is in  $L_R(\mu)$ ) if and only if, for every  $n = 1, \dots, N$ , we have*

$$\int P(dh_{n-1}, dy^n) g(h_{n-1}, y^n) = \int P(dh_{n-1}) \left( \int_{\mathbb{Y}^n} g(h_{n-1}, z) p_n(dz|h_{n-1}) \right) \quad (9)$$

and

$$\begin{aligned} & \int P(dh_n) g(h_{n-1}, y^n, u^n) \\ &= \int P(dh_{n-1}, dy^n) \left( \int_{\mathbb{U}^n} g(h_{n-1}, y^n, a^n) \Pi^n(da^n | y^n) \right) \end{aligned} \quad (10)$$

for all continuous and bounded function  $g$  with appropriate arguments, where  $P(d\omega_0) = \mu(dw_0)$  and  $\Pi^n$  is a stochastic kernel from  $\mathbb{Y}^n$  to  $\mathbb{U}^n$ .

**Remark 3.2.** A result similar to Theorem 3.1 can also be stated for the dynamic case, in particular with regard to  $L_A(\mu)$  being the set of extreme points of the convex hull of  $L_R(\mu)$ . The reader is referred to [151, Theorem 2.3] which essentially establishes this; see also [59, Theorem 1.c] for related discussions.

A celebrated result in economics theory, known as Kuhn's theorem [86], notes that the convex hull of admissible (i.e. those in  $L_A(\mu)$ ) strategic measures (hence  $L_C(\mu)$ ) is equivalent to  $L_R(\mu)$  when the information structure is classical. We can thus state that this does not apply in the absence of classical-ness, as  $L_R(\mu)$  would not be convex (if the information structure is not classical, then convexity fails [151, p.12]), but the convex hull of admissible policies is, by definition, convex; but the convex hull of  $L_R(\mu)$  is  $L_C(\mu)$ .

### 3.4. Measurability properties of sets of strategic measures

As noted earlier, the set  $L_A(\mu)$  is a Borel subset of  $\mathcal{P}(\Omega_0 \times \prod_k (\mathbb{Y}^k \times \mathbb{U}^k))$  under weak convergence topology. The same is true for  $L_R(\mu)$ , which is stated in the following theorem. This result will be crucial in the analysis to follow.

**Theorem 3.3** ([151, Theorem 2.10]). *Consider a sequential (static or dynamic) team.*

- (i) *The set of strategic measures  $L_R(\mu)$  is Borel when viewed as a subset of the space of probability measures on  $\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  under the topology of weak convergence.*
- (ii) *The set of strategic measures  $L_A(\mu)$  is Borel when viewed as a subset of the space of probability measures on  $\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  under the topology of weak convergence.*

For further properties of the sets of strategic measures, see [151].

## 4. Relative compactness and closedness of strategic measures, and existence of optimal policies

Existence of optimal policies for static and a class of sequential dynamic teams have been studied recently in [69, 151, 119]. More specific setups have been studied in [144], [137], [150] and [149]. Existence of optimal team policies has been established in [45] for a class of continuous-time decentralized stochastic

control problems. For a class of teams which are convex, one can reduce the search space to a smaller parametric class of policies, such as linear policies for quasi-classical linear quadratic Gaussian problems [117, 85, 73].

The following theorem states a general existence result for static teams and for dynamic teams admitting static reduction. Its proof depends on Weierstrass Extreme Value Theorem.

**Theorem 4.1.** *Consider a static team or the static reduction of a dynamic team with  $c$  denoting the cost function. Let  $c$  be lower semi-continuous in  $\mathbf{u}$  for every fixed  $\omega_0, \mathbf{y}$  and  $L_R(\mu)$  or  $L_C(\mu)$  be a compact set under weak convergence topology. Then, there exists an optimal team policy. This policy can be chosen deterministic and hence induces a strategic measure in  $L_A(\mu)$ .*

**Remark 4.1.** *Since the cost function  $c_s$  in an independent static reduction of a dynamic team also depends on the measurements  $\mathbf{y}$ , we include  $\mathbf{y}$  as an argument to the cost function  $c$  in the previous theorem.*

By Theorem 4.1, to prove the existence of an optimal team policy, it is sufficient to establish the lower semi-continuity of the cost function  $c$  and compactness of  $L_R(\mu)$  or  $L_C(\mu)$  or any other subset of  $L_C(\mu)$  that is sufficient for optimality. However, we recall that unless certain conditions are imposed, the conditional independence property dictated by IS is not preserved under weak convergence (indeed this is also true even for setwise convergence which is a stronger convergence notion). Hence,  $L_R(\mu)$  and  $L_C(\mu)$  are in general not closed and so compact. The following theorem establishes that  $L_R(\mu)$  is not closed under the weak convergence topology.

**Theorem 4.2** ([151, Theorem 2.7]). *Consider a sequence of probability measures  $P_n \in \mathcal{P}(\mathbb{U}^1 \times \mathbb{Y} \times \mathbb{U}^2)$  so that for all  $n$ :*

$$P_n(du^1|y, u^2) = P_n(du^1|y).$$

*If  $P_n \rightarrow P$  setwise (and thus also weakly), it is not necessarily the case that*

$$P(du^1|y, u^2) = P(du^1|y).$$

*That is, conditional independence of  $u^1$  and  $u^2$  given  $y$  is not preserved under setwise convergence. In particular,  $L_R(\mu)$  is not (weakly or setwise) closed.*

Before we present the proof of Theorem 4.2, let us note that, for each  $n$ , we can view  $P_n$  as a strategic measure of the following static team problem. In this team problem,  $\Omega$  is a degenerate space and we have two DMs that are sharing the same measurement  $y$ . Therefore, Theorem 4.2 states that the set of strategic measures  $L_R(\mu)$  is not closed under setwise (and so weak convergence) topology.

*Proof.* It suffices to provide a counterexample. We build on an example from [150] (used in a different context) in the following. Let  $\mathbb{Y} = [0, 1]$ ,  $\mathbb{U}^1 = \mathbb{U}^2 = \{0, 1\}$ , and  $y \sim m$  where  $m$  the Lebesgue measure (uniform distribution) on  $[0, 1]$ . Let

$$L_{nk} = \left[ \frac{2k-2}{2n}, \frac{2k-1}{2n} \right), \quad R_{nk} = \left[ \frac{2k-1}{2n}, \frac{2k}{2n} \right) \tag{11}$$



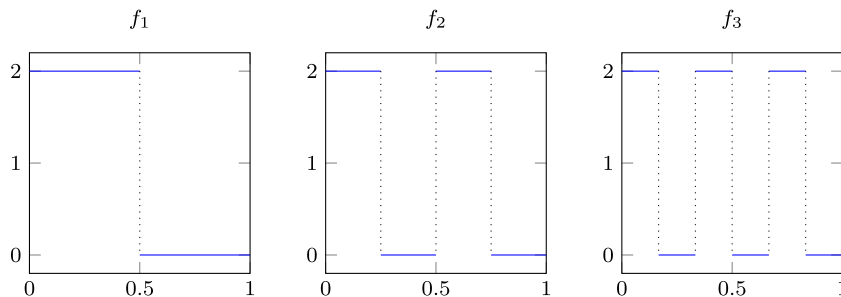


FIG 3. Plots of probability density functions  $f_1$ ,  $f_2$ , and  $f_3$ .

and define the *square wave function*

$$h_n(t) = \sum_{k=1}^n (1_{\{t \in L_{nk}\}} - 1_{\{t \in R_{nk}\}}).$$

Define further  $f_n(t) = h_n(t) + 1$ .

Let  $B_{n,1} = \bigcup_{k=1}^n L_{nk}$  and  $B_{n,2} = [0, 1] \setminus B_{n,1}$ . Define  $\{Q_n\}$  as the sequence of 2-cell quantizers given by

$$Q_n(1|y) = 1_{\{y \in B_{n,1}\}}, \quad Q_n(0|y) = 1_{\{y \in B_{n,2}\}}.$$

Let

$$P_n(u^1 = a|y) = P_n(u^2 = a|y) = Q_n(a|y),$$

where  $a \in \{0, 1\}$ . Define  $P \in \mathcal{P}(\mathbb{U}^1 \times \mathbb{Y} \times \mathbb{U}^2)$  as  $P(a, A, b) = 1_{\{a=b\}} \frac{1}{2} m(A)$ , where  $a, b \in \{0, 1\}$  and  $A \in \mathcal{B}([0, 1])$ .

By the proof of the Riemann-Lebesgue lemma ([136], Thm. 12.21), observe that for all  $A \in \mathcal{B}([0, 1])$ ,

$$\lim_{n \rightarrow \infty} \int_A Q_n(1|y) m(dy) = \lim_{n \rightarrow \infty} \int_A \frac{1}{2} f_n(t) m(dt) = \frac{1}{2} m(A),$$

and thus for all  $A \in \mathcal{B}([0, 1])$

$$\begin{aligned} & \lim_{n \rightarrow \infty} P_n(u^1 = 1, y \in A, u^2 = 1) \\ &= \lim_{n \rightarrow \infty} \int_A P_n(u^1 = 1|y) P_n(u^2 = 1|y) m(dy) \\ &= \lim_{n \rightarrow \infty} \int_A P_n(u^1 = 1|y) m(dy) \\ &= \frac{1}{2} m(A) \\ &= P(1, A, 1) \end{aligned} \tag{12}$$

A similar property applies for  $(u^1, u^2) = (0, 0), (0, 1)$  and  $(1, 0)$  so that

$$\lim_{n \rightarrow \infty} P_n(u^1 = a, y \in A, u^2 = b) \rightarrow P(a, A, b) = 1_{\{a=b\}} \frac{1}{2} m(A)$$

Thus,  $P_n \rightarrow P$  setwise. But even though  $P_n$  satisfies the conditional independence property that  $P_n(u^1 = 1|y, u^2) = Q_n(1|y)$ ,  $P$  does not satisfy the conditional independence property of  $u^1$  and  $u^2$  given  $y$ : Under  $P$ ,  $u^1$  and  $y$  are independent but  $u^1 = u^2$  and thus  $P(u^1 = a|y, u^2 = b) = 1_{\{a=b\}} \neq \frac{1}{2} = P(u^1 = a|y)$ . Thus, setwise (and hence weak) convergence does not preserve the conditional independence property.  $\square$

We can also show via counterexample that  $L_C(\mu)$  is not closed either with respect to the weak convergence topology.

**Theorem 4.3.**  *$L_C(\mu)$  is not weakly or setwise closed.*

*Proof.* Consider the same example as in the proof of Theorem 4.2, modified as follows. Let  $\mathbb{Y} = [0, 1] \cup (1, 2]$ ,  $\mathbb{U}^1 = \mathbb{U}^2 = \{0, 1, 2, 3\}$ . Let  $x$  be a  $\{0, 1\}$ -valued uniformly distributed Bernoulli random variable and let  $y \sim 1_{\{x=0\}}v + 1_{\{x=1\}}(v+1)$  where  $v \sim m$  with  $m$  being the Lebesgue measure (uniform distribution) on  $[0, 1]$ . Let  $L_{nk}$  and  $R_{nk}$  be as given in (11) and define the *square wave* function, as before,

$$h_n(t) = \sum_{k=1}^n (1_{\{t \in L_{nk}\}} - 1_{\{t \in R_{nk}\}}).$$

Define also  $f_n(t) = h_n(t) + 1$ . Let  $B_{n,1} = \bigcup_{k=1}^n L_{nk}$  and  $B_{n,2} = [0, 1] \setminus B_{n,1}$ . Define  $\{Q_n\}$  as the sequence of 4-cell quantizers given by

$$Q_n(1|y) = 1_{\{y \in B_{n,1}\}}, \quad Q_n(0|y) = 1_{\{y \in B_{n,2}\}}, \quad y \in [0, 1]$$

and

$$Q_n(3|y) = 1_{\{y \in 1+B_{n,1}\}}, \quad Q_n(2|y) = 1_{\{y \in 1+B_{n,2}\}}, \quad y \in (1, 2].$$

Let

$$P_n(u^1 = a|y) = P_n(u^2 = a|y) = Q_n(a|y),$$

where  $a \in \{0, 1, 2, 3\}$ . Define  $P \in \mathcal{P}(\mathbb{U}^1 \times \mathbb{Y} \times \mathbb{U}^2)$  as

$$P(a, A, b) = \frac{1}{4} 1_{\{a=b, a \in \{0,1\}\}} m(A \cap [0, 1]) + \frac{1}{4} 1_{\{a=b, a \in \{2,3\}\}} m(A \cap (1, 2])$$

where  $a, b \in \{0, 1, 2, 3\}$  and  $A \in \mathcal{B}([0, 2])$ .

We can show that  $P_n \rightarrow P$  setwise (and so weakly). In the limit, it is the case that  $u^1 = u^2$  almost surely. However, notice that  $P \notin L_C(\mu)$ :  $P$  is not a mixture of conditionally independent random variables given  $y$ , with the mixture being independent of  $y$ . Here, the mixture representation of  $P$  in terms of the two independent  $\{0, 1\}$ -valued and  $\{2, 3\}$ -valued random variables is so that the mixing random variable is correlated with  $y$  (and gives information on  $y$ ).  $\square$

We refer the reader to [2] [70] [14, Theorem 1.1] [150], [20] for further related results on such intricacies on conditional independence properties. A sufficient condition for compactness of  $L_R(\mu)$  under the weak convergence topology was reported in [69]. We re-state this result in a brief and different form below for reader's convenience.

**Theorem 4.4** ([69]). *Consider a static team where the action sets  $\mathbb{U}^i$  are compact. Furthermore, the measurements*

$$P(d\mathbf{y}|\omega_0) = \prod_{i=1}^N g^i(dy^i|\omega_0)$$

*satisfy  $g^i(dy^i|\omega_0) = \xi^i(y^i, \omega_0) \nu^i(dy^i)$  for some positive measure  $\nu^i$  and continuous  $\xi^i$  such that, for every  $\epsilon > 0$ , there exists  $\delta > 0$  so that for  $\rho_i(a, b) < \delta$  (where  $\rho_i$  is a metric on  $\mathbb{Y}^i$ )*

$$|\xi^i(b, \omega_0) - \xi^i(a, \omega_0)| \leq \epsilon h^i(a, \omega_0),$$

*with  $\sup_{\omega_0} \int h^i(a, \omega_0) \nu^i(dy^i) < \infty$ . Hence, static team admits independent static reduction. Then, the set  $L_R(\mu)$  is compact under the weak convergence topology. Therefore, if  $c(\omega_0, \mathbf{u})$  is lower semi-continuous in  $\mathbf{u}$  for any  $\omega_0$ , then there exists an optimal team policy (which is deterministic and hence in  $L_A(\mu)$ ).*

The results in [69] also apply to static reduction of sequential dynamic teams, and a class of teams with non-compact action spaces that however has moment-type cost function leading to a tightness condition on the set of strategic measures with a finite cost. In particular, the existence result applies to the counterexample of Witsenhausen [137].

However, the result above can be significantly relaxed where the relaxation is not only in the compactness condition, this can be modified by the usual tightness conditions which will also be presented below. The generalization is with respect to the topology considered: in the following, we present the most general conditions to our knowledge on existence of optimal policies. In the following, we slightly strengthen [148, Theorem 5.2] to allow for lower semi-continuity of the cost function only in the actions.

Before we present a general result on existence, we recall the  $w$ - $s$  topology introduced by Schäl [125]: The  $w$ - $s$  topology on the set of probability measures  $\mathcal{P}(\mathbb{X} \times \mathbb{U})$  is the coarsest topology under which  $\int f(x, u) \nu(dx, du) : \mathcal{P}(\mathbb{X} \times \mathbb{U}) \rightarrow \mathbb{R}$  is continuous for every measurable and bounded  $f$  which is continuous in  $u$  for every  $x$  (but unlike weak topology,  $f$  does not need to be continuous in  $x$ ).

**Theorem 4.5** ([148, Theorem 5.2]). *(i) Consider a static or a dynamic team that admits an independent static reduction. Let the action sets  $\mathbb{U}^i$  be compact. Then, the set  $L_R(\mu)$  is compact under weak convergence topology, and is also in the  $w$ - $s$  topology (weak in actions, strong in measurements). Then, the set  $L_R(\mu)$  is compact under weak convergence topology, and is also in the  $w$ - $s$  topology (weak in actions, strong in measurements).*

*(ii) Consider a static or a dynamic team that admits an independent static reduction with the loss denoted with  $c$  (for either case). Let  $c$  be lower semi-continuous in  $\mathbf{u}$  for any  $\omega_0, \mathbf{y}$ . Suppose further that  $\mathbb{U}^i$  is  $\sigma$ -compact (that is,  $\mathbb{U}^i = \cup_n K_n$  for a countable collection of increasing compact sets  $K_n$ ) and, without any loss of generality, the control laws can be restricted to those with  $E[\phi^i(u^i)] \leq M$  for some lower semi-continuous  $\phi^i : \mathbb{U}^i \rightarrow \mathbb{R}_+$*

which satisfies  $\lim_{n \rightarrow \infty} \inf_{u^i \notin K_n} \phi^i(u^i) = \infty$ . Then, an optimal team policy exists.

*Proof.* (i) For each  $m$ , consider the strategic measure:

$$P(d\omega_0) \prod_{k=1}^N (Q^k \gamma^k)_m(dy^k, du^k).$$

Suppose that  $(Q^k \gamma^k)_m(dy^k, du^k)$  converges to  $(Q^k \gamma^k)(dy^k, du^k)$  weakly. By [113, p. 57], the joint product measure  $P(d\omega_0) \prod_{k=1}^N (Q^k \gamma^k)_m(dy^k, du^k)$  will converge weakly to  $P(d\omega_0) \prod_{k=1}^N (Q^k \gamma^k)(dy^k, du^k)$ ; see also [126, Section 5].

Since the marginals on  $\prod_k \mathbb{Y}^k$  are fixed, [125, Theorem 3.10] (or [16, Theorem 2.5]) establishes that the set of all probability measures with a fixed marginal on  $\prod_k \mathbb{Y}^k$  is (sequentially) compact under the  $w$ - $s$  topology.

(ii) If the function  $c_s(\omega_0, \mathbf{y}, \cdot)$  were continuous, this in turn would ensure that the function (by a truncation and then a limiting argument)

$$\int P(d\omega_0) \prod_{k=1}^N (Q^k \gamma^k)(dy^k, du^k) c_s(\omega_0, \mathbf{y}, \mathbf{u}),$$

is lower semi-continuous under the  $w$ - $s$  topology. Now, by approximating the lower semi-continuous function  $c_s(\omega_0, \mathbf{y}, \cdot)$  from below by continuous functions, applying the argument above, and taking the limit, we conclude that the lower semi-continuity also applies when  $c_s(\omega_0, \mathbf{y}, \cdot)$  is only lower semi-continuous.

Finally, since the set of admissible strategic measures is sequentially compact under the  $w$ - $s$  topology, existence of an optimal team policy follows. The proof for the dynamic case follows analogously by considering its static reduction.  $\square$

**Remark 4.2.** Building on [151, Theorems 2.3 and 2.5] and [69, p. 1691] (due to Blackwell's theorem on irrelevant information [30, 32], [149, p. 457]), an optimal policy, when exists, can be assumed to be deterministic.

Note that Theorem 4.5 provide weaker conditions when compared with Theorem 4.4. So far, we present existence results for static or dynamic teams that admit independent static reduction. In the following, we present existence results for teams that do not admit independent static reduction.

**Theorem 4.6** ([151, Theorem 2.9]). Consider a sequential team with a classical information structure with the further property that  $\sigma(\omega_0) \subset \sigma(y^1)$  (under every policy,  $y^1$  contains  $\omega_0$ ). Suppose further that  $\prod_{k=1}^N \mathbb{U}^k$  is compact. If  $c$  is lower semi-continuous and each of the kernels  $p_n$  (defined in (8)) is weakly continuous so that

$$\int f(y^n) p_n(dy^n | \omega_0, y^1, \dots, y^{n-1}, u^1, \dots, u^{n-1}) \tag{13}$$

is continuous in  $\omega_0, y^1, \dots, y^{n-1}, u^1, \dots, u^{n-1}$  for every continuous and bounded  $f$ , then there exists an optimal team policy which is deterministic.

A further existence result along similar lines, for a class of static teams, is presented next.

**Theorem 4.7** ([148, Theorem 5.6]). *Consider a static team with a classical information structure (that is, with an expanding information structure so that  $\sigma(y^n) \subset \sigma(y^{n+1})$ ,  $n \geq 1$ ). Suppose further that  $\prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  is compact. If*

$$\tilde{c}(y^1, \dots, y^N, u^1, \dots, u^N) := E[c(\omega_0, \mathbf{u}) | \mathbf{y}, \mathbf{u}]$$

*is lower semi-continuous in  $\mathbf{u}$  for every  $\mathbf{y}$ , then there exists an optimal team policy which is deterministic.*

*Proof.* Different from Theorem 4.6, we eliminate the use of  $\omega_0$ , and study the properties of the set of strategic measures. Different from [151, Theorem 3.5], we relax weak continuity. Once again, instead of the weak topology, we will use the  $w$ - $s$  topology [124].

Note that when  $\prod_{k=1}^N \mathbb{U}^k$  is compact, the set of all probability measures on  $\prod_{k=1}^N \mathbb{Y}^k \times \mathbb{U}^k$  forms a weakly compact set. Since the marginals on  $\prod_{k=1}^N \mathbb{Y}^k$  is fixed, [124, Theorem 3.10] (or [16, Theorem 2.5]) establishes that the set of all probability measures with a fixed marginal on  $\prod_{k=1}^N \mathbb{Y}^k$  is relatively compact under the  $w$ - $s$  topology. Therefore, it suffices to ensure the closedness of the set of strategic measures, which leads to the sequential compactness of the set under this topology. To facilitate such a compactness condition, we first *expand* the information structure so that DM  $k$  has access to all the previous actions  $u^1, \dots, u^{k-1}$  as well. Later on, we will see that this expansion is redundant. With this expansion, any  $w$ - $s$  converging sequence of strategic measures will continue satisfying (10) in the limit due to the fact that there is no conditional independence property in the sequence since all the information is available at DM  $k$ . That is,  $P_n(du^n | y^n, y_{[0,n-1]}, u_{[0,n-1]})$  satisfies the conditional independence property trivially as all the information is available. On the other hand, for each element in the sequence of conditional probability measures, the conditional probability for the measurements writes as  $P(dy^n | y_{[0,n-1]}, u_{[0,n-1]}) = P(dy^n | y^{n-1})$ . We wish to show that this also holds for the  $w$ - $s$  limit measure. Now, we have that for every  $n$ ,  $y^n \leftrightarrow y^{n-1} \leftrightarrow h_{n-1}$  forms a Markov chain. By considering the convergence properties only on continuous functions and bounded  $f$ , as in (9), with  $P_m \rightarrow P$  weakly, we have that

$$\begin{aligned} & \int P(dy^n | y^{n-1}) P_m(dh_{n-1}) f(y^n, h_{n-1}) \\ &= \int \left( P(dy^n | y^{n-1}) f(y^n, h_{n-1}) \right) P_m(dh_{n-1}) \\ &\rightarrow \int \left( P(dy^n | y^{n-1}) f(y^n, h_{n-1}) \right) P(dh_{n-1}) \\ &= \int P(dy^n | y^{n-1}) P(dh_{n-1}) f(y^n, h_{n-1}) \end{aligned}$$

Here,  $\left( P(dy^n|y^{n-1})f(y^n, h_{n-1}) \right)$  is not continuous in  $y^{n-1}$ , but it is in  $h_{n-2}$  and  $u^{n-1}$  by an application of the dominated convergence theorem, and since

$$P_m(dy^1, \dots, dy^{n-1}, du^1, \dots, du^{n-1}) \rightarrow P(dy^1, \dots, dy^{n-1}, du^1, \dots, du^{n-1}),$$

in the  $w$ -s sense (setwise in the measurement variable coordinates), convergence holds. Thus, (9) is also preserved. Hence, for any  $w$ -s converging sequence of strategic measures satisfying (9)-(10) so does the limit since the team is static and with perfect-recall. By [124, Theorem 3.7], and the generalization of Portmanteau theorem for the  $w$ -s topology, the lower semi-continuity of  $\int \mu(dy, du)\tilde{c}(y, u)$  over the set of strategic measures leads to the existence of an optimal strategic measure. As a result, the existence follows from similar steps to that of Theorem 4.1. Now, we know that an optimal policy will be deterministic (see Remark 4.2). Thus, a deterministic policy may not make use of randomization, therefore DM  $k$  having access to  $\{y^k, y^{k-1}, y^{k-2}, \dots\}$  is informationally equivalent to him having access to  $\{y^k, (y^{k-1}, u^{k-1}), (y^{k-2}, u^{k-2})\}$  for an optimal policy. Thus, an optimal team policy exists.  $\square$

We can report that Theorem 4.5 (for static teams or dynamic teams with an independent static reduction) and Theorems 4.6 and 4.7 (for sequential teams that do not allow an independent static reduction) are the most general existence results, to our knowledge, for sequential team problems considered in this paper. These results complement each other and cover a very large class of decentralized stochastic control problems.

We finally state the following result.

**Theorem 4.8.** [75]  $L_C(\mu)$  is compact whenever  $L_R(\mu)$  is.

*Proof.* Let  $\{\phi_k\}$  be an arbitrary sequence of information structures in  $L_C(\mu)$  defined with  $B \in \mathcal{B}(\Omega_0 \times \prod_{k=1}^n (\mathbb{Y}^k \times \mathbb{U}^k))$

$$\phi_k(B) = \int_B \left( \int \nu_k(d\theta)\theta(dz) \right),$$

Following the relationship between  $L_{CCR}(\mu)$  and  $L_R(\mu)$ , as well as Theorem 3.2, we can view  $L_C(\mu)$  as the mixture of elements of  $L_R$ . The convergence properties of the sequence  $\{\phi_k\}$  can be studied by the convergence of the family of measures:

$$\phi_k(B) = \int_B f_m(z) \left( \int \nu_k(d\theta)\theta(dz) \right),$$

for a fixed countable collection of *weak-convergence determining* functions  $f_m : \Omega \times (\prod_{i=1}^n \mathbb{X}^i \times \mathbb{Y}^i) \rightarrow \mathbb{R}$  [57, Theorem 3.4.5].

Now, using Fubini's Theorem, we can rewrite the above as:

$$\int \nu_k(d\theta) \int f_m(z)\theta(dz),$$

and can note that the inner integral is continuous and bounded in  $\theta$ .

Thus we can write this as:

$$\int \nu_k(d\theta)F_m(\theta),$$

for the continuous and bounded  $F_m$  on  $L_R(\mu)$  given with  $F_m(\theta) := \int f_m(z)\theta(dz)$ .

Since  $L_R(\mu)$  is weakly compact, the set of probability measures on  $\mathcal{B}(L_R(\mu))$  is weakly compact [27]; and accordingly there exists a subsequence of  $\{\nu_k\}$ ,  $\{\nu_{k_l}\}$ , which converges weakly to some  $\nu$  in  $\mathcal{P}(L_R(\mu))$ . Since  $F_m$  is continuous and bounded on  $L_R(\mu)$ , it follows then that  $\int \nu_{k_l}(d\theta)F_m(\theta)$  will converge to  $\int \nu(d\theta)F_m(\theta)$ . This applies for each  $m$ , and therefore, we conclude that  $\{\phi_k\}$  has a convergent subsequence in  $L_C(\mu)$ . Thus,  $L_C(\mu)$  is compact.  $\square$

In the following, we will discuss lower and upper bounds on the optimal costs, numerical programs, and convex relaxations.

## 5. Convex relaxations

In this section, we introduce convex relaxations for static teams or dynamic teams that admit independent static reduction through relaxing conditional independence among actions. Therefore, in this section, the joint distribution of  $\omega_0, \mathbf{y}$  is in product form  $\mu(d\omega_0, d\mathbf{y}) = P(d\omega_0) \prod_{k=1}^N Q^k(dy^k)$ . These relaxations can be classified in increasing order as *quantum-correlated relaxation* and *non-signaling relaxation*. These classes are new to team decision theory and we believe that these new classes provide novel perspectives and results to team decision theory and decentralized stochastic control in the future.

We saw earlier in the article that individual or common (independent) randomness does not improve optimal team cost, whereas we will see that quantum-correlated and non-signaling relaxations in general improve the optimal cost. Moreover, the optimization problem associated with the non-signaling case can be written as a linear program.

A related hierarchy of policies was introduced in [12] to study games with local information. In [12], the advantages of quantum-correlated and non-signaling equilibria over classical ones were discussed and it was established that quantum-correlated and non-signaling equilibria are socially more beneficial. However, our aim is not to show benefits of quantum-correlated and non-signaling policies over classical ones, instead our main motivation is to obtain an approximation to the classical team problem by using these new classes of policies.

The non-signaling type relaxation for team problems was studied in [151] with a different name, building on [4] where it was proved that the set of strategic measures corresponding to the extreme points of non-signaling policies is a strict superset of the set of strategic measures corresponding to measurable policies. This set arises in applications in information theory in the context of converse theorems in multi-terminal source coding (see e.g. Berger-Tung inner-outer bounds [132]). In the quantum information theory literature [41], this correlation structure has evidently also been studied extensively, as we will present later in this section.

**5.1. Non-signaling relaxation**

A joint conditional distribution of actions given observations is *non-signaling* if, for any  $i \in \mathcal{N}$ , the marginal distribution of the action of DM  $i$  given its observation does not give any information about the observations of other agents [102, 21]. Non-signaling has been investigated in quantum mechanics due to its close connection to the foundations of quantum mechanics and relativity [116].

Let  $L_{NS}(\mu)$  denote the set of non-signaling strategic measures. Formally, non-signaling strategic measures are defined as follows. An element  $P \in L_{NS}(\mu)$  if  $P(d\omega_0, d\mathbf{y}) = \mu(d\omega_0, d\mathbf{y})$ , and for any subset  $\{k_1, \dots, k_M\}$  of  $\{1, \dots, N\}$ , the actions of the agents in  $\{k_1, \dots, k_M\}$  given their measurements are independent of measurements of agents in  $\{1, \dots, N\} \setminus \{k_1, \dots, k_M\}$  and  $\omega_0$ ; that is, for any  $\{k_1, \dots, k_M\} \subset \{1, \dots, N\}$ , we have

$$P(du^{k_1}, \dots, du^{k_M} | \omega_0, \mathbf{y}) = P(du^{k_1}, \dots, du^{k_M} | y^{k_1}, \dots, y^{k_M}). \tag{14}$$

We also define a more relaxed version of  $L_{NS}(\mu)$ , denoted by  $L_M(\mu)$  (which we will call *local-Markov correlations*), as follows: An element  $P \in L_M(\mu)$  if  $P(d\omega_0, d\mathbf{y}) = \mu(d\omega_0, d\mathbf{y})$ , and for any  $k \in \{1, \dots, N\}$ , the action  $u^k$  of agent  $k$  given the measurement  $y^k$  is independent of measurements of agents in  $\{1, \dots, N\} \setminus \{k\}$  and  $\omega_0$ ; that is, we have

$$P(du^k | \omega_0, \mathbf{y}) = P(du^k | y^k). \tag{15}$$

At first sight, it is tempting to claim that  $L_C(\mu)$  is the same as the set of non-signaling strategic measures  $L_{NS}(\mu)$  or  $L_M(\mu)$ . Indeed, in [4] this question was raised: a counterexample to the claim that  $L_C(\mu)$  is equivalent to the set  $L_{NS}(\mu)$  was given to establish that the set of extreme points of non-signaling policies is not  $L_A(\mu)$ , which would imply that non-signaling policies are more general than randomized ones as  $L_A(\mu)$  is the set of extreme points of  $L_C(\mu)$  (see Theorem 3.1). In the quantum information literature, this result was known long ago, as we will review later.

It turns out that the non-signaling condition (14) can be derived from fewer linear constraints, which will be described below. These constraints indeed enable us to write the optimization problem associated with non-signaling policies as a linear program.

**Lemma 5.1** ([102, Section II-A]). *An element  $P$  with  $P(d\omega_0, d\mathbf{y}) = \mu(d\omega_0, d\mathbf{y})$  is a non-signaling strategic measure if and only if it satisfies the following condition:*

(N) *For each  $k \in \{1, \dots, N\}$ , the marginal distribution of actions excluding  $u^k$  is independent of the observation  $y^k$  and  $\omega_0$ :*

$$P(d\mathbf{u}^{-k} | \omega_0, \mathbf{y}) = P(d\mathbf{u}^{-k} | \mathbf{y}^{-k}). \tag{16}$$

*Proof.* Note that each constraint (16) is linear in  $P$ . The equivalence of (14) and (N) can be established as follows. First, it is immediate that (14) implies (N). Conversely, let  $P$  satisfies (N). Fix any subset  $\{k_1, \dots, k_M\}$  of  $\{1, \dots, N\}$ .



Let  $\{l_1, \dots, l_T\} := \{1, \dots, N\} \setminus \{k_1, \dots, k_M\}$ . Then, using the condition **(N)**, we can prove that

$$P(du^{k_1}, \dots, du^{k_M} \mid \omega_0, \mathbf{y}) = P(du^{k_1}, \dots, du^{k_M} \mid \hat{\omega}_0, \hat{y}^{l_1}, \dots, \hat{y}^{l_T}, y^{k_1}, \dots, y^{k_M}),$$

for all  $\omega_0, y^{l_1}, \dots, y^{l_T}$  and  $\hat{\omega}_0, \hat{y}^{l_1}, \dots, \hat{y}^{l_T}$ , which implies that

$$P(du^{k_1}, \dots, du^{k_M} \mid \omega_0, \mathbf{y}) = P(du^{k_1}, \dots, du^{k_M} \mid y^{k_1}, \dots, y^{k_M}).$$

Indeed, we have

$$\begin{aligned} P(du^{k_1}, \dots, du^{k_M} \mid \omega_0, \mathbf{y}) &= \int_{\mathbb{U}^{l_2} \times \dots \times \mathbb{U}^{l_T}} P(d\mathbf{u}^{-l_1} \mid \omega_0, \mathbf{y}) \\ &= \int_{\mathbb{U}^{l_2} \times \dots \times \mathbb{U}^{l_T}} P(d\mathbf{u}^{-l_1} \mid \hat{\omega}_0, \hat{y}^{l_1}, \mathbf{y}^{-l_1}) \text{ (by **(N)**)} \\ &= P(du^{k_1}, \dots, du^{k_M} \mid \hat{\omega}_0, \hat{y}^{l_1}, \mathbf{y}^{-l_1}) \\ &= \int_{\mathbb{U}^{l_1} \times \mathbb{U}^{l_3} \times \dots \times \mathbb{U}^{l_T}} P(d\mathbf{u}^{-l_2} \mid \hat{\omega}_0, \hat{y}^{l_1}, \mathbf{y}^{-l_1}) \\ &= \int_{\mathbb{U}^{l_1} \times \mathbb{U}^{l_3} \times \dots \times \mathbb{U}^{l_T}} P(d\mathbf{u}^{-l_2} \mid \hat{\omega}_0, \hat{y}^{l_1}, \hat{y}^{l_2}, \mathbf{y}^{-l_1, -l_2}) \text{ (by **(N)**)} \\ &= P(du^{k_1}, \dots, du^{k_M} \mid \hat{\omega}_0, \hat{y}^{l_1}, \hat{y}^{l_2}, \mathbf{y}^{-l_1, -l_2}) \\ &\quad \vdots \\ &= P(du^{k_1}, \dots, du^{k_M} \mid \hat{\omega}_0, \hat{y}^{l_1}, \dots, \hat{y}^{l_T}, y^{k_1}, \dots, y^{k_M}). \end{aligned}$$

This completes the proof.  $\square$

The following result establishes convexity and topological properties of the sets  $L_{NS}(\mu)$  and  $L_M(\mu)$ .

**Theorem 5.1.** (i)  $L_M(\mu)$  is a convex set and it is closed under the weak convergence topology.

(ii)  $L_{NS}(\mu)$  is a convex set and it is closed under the weak convergence topology.

*Proof.* Convexity of both sets follow since both  $L_M(\mu)$  and  $L_{NS}(\mu)$  are intersections of convex sets.

Let  $P_n \in L_M(\mu)$  converge to  $P$  weakly. Then, for any continuous and bounded  $g$ :

$$\int P_n(dy^i, d\mathbf{y}^{-i}, d\omega_0, u^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \rightarrow \int P(dy^i, d\mathbf{y}^{-i}, d\omega_0, u^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i)$$

Since  $P_n \in L_M(\mu)$ , we have that

$$\begin{aligned} &\int P_n(dy^i, d\mathbf{y}^{-i}, d\omega_0, u^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \\ &= \int \int P_n(d\mathbf{y}^{-i}, d\omega_0 \mid y^i)P_n(dy^i, du^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i). \end{aligned}$$

But since the information structure is static, we write

$$\begin{aligned} & \int P_n(dy^i, d\mathbf{y}^{-i}, d\omega_0, u^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \\ &= \int \int \mu(d\mathbf{y}^{-i}, d\omega_0|y^i)P_n(dy^i, du^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i). \end{aligned}$$

Now,

$$h(y^i, u^i) := \left( \int \mu(d\mathbf{y}^{-i}, d\omega_0|y^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \right)$$

is continuous in  $u^i$  as  $g$  is continuous by dominated convergence theorem. Since the marginal on  $y^i$  is fixed, as noted earlier, [124, Theorem 3.10] (or [16, Theorem 2.5]) establishes that the set of all tight probability measures with a fixed marginal on  $\prod_k \mathbb{Y}^k$  is relatively compact under the  $w$ -s topology. This implies that  $P_n$  converges to  $P$  in  $w$ -s topology as the sequence  $\{P_n\}$  is tight. Hence,

$$\begin{aligned} & \int P_n(dy^i, d\mathbf{y}^{-i}, d\omega_0, u^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \rightarrow \int h(y^i, u^i)P(dy^i, du^i) \\ &= \int \int \mu(d\mathbf{y}^{-i}, d\omega_0|y^i)P(dy^i, du^i)g(y^i, \mathbf{y}^{-i}, \omega_0, u^i) \end{aligned}$$

and hence  $P \in L_M(\mu)$  also.

By viewing each subset  $\{k_1, \dots, k_m\}$  of  $\{1, \dots, N\}$  as a single decision maker with a collective action  $\mathbf{u}_m := \{u^{k_1}, \dots, u^{k_m}\}$  and  $\mathbf{y}_m := \{y^{k_1}, \dots, y^{k_m}\}$ , applying the same analysis above shows that each set is closed, and since intersection of closed sets is closed, the closedness of  $L_{NS}(\mu)$  follows. As noted above using the condition (N), without any loss, we can only consider the sets of cardinality  $N - 1$ .  $\square$

Since the constraints  $P(d\mathbf{u}^{-i}|y^i, \mathbf{y}^{-i}, \omega_0) = P(d\mathbf{u}^{-i}|\mathbf{y}^{-i})$ ,  $i \in \mathcal{N}$ , in  $L_{NS}(\mu)$  for  $P$  are linear, the optimization problem associated with non-signaling strategic measures can be written as a linear program over appropriate vector spaces. One such formulation will be given in Section 5.4.

## 5.2. Quantum-correlated relaxation

To introduce quantum-correlated strategic measures, we briefly introduce the mathematical formalism necessary to discuss quantum operations. We refer the reader to books [111, 135, 76] for basics of quantum information and computation. In this section, we suppose that all Hilbert spaces are complex and separable.

Quantum physical systems are described by complex Hilbert spaces. For a Hilbert space  $\mathcal{H}$  with an inner product  $\langle \cdot, \cdot \rangle$ , let  $\mathcal{L}(\mathcal{H})$  and  $\mathcal{D}(\mathcal{H})$  denote the set of Hermitian and positive Hermitian operators with unit trace. A state  $\rho$  of a quantum system, living in  $\mathcal{H}$ , is an element of  $\mathcal{D}(\mathcal{H})$ ; that is, it is a positive Hermitian operator with unit trace. A quantum state  $\rho$  is said to be *pure* if it

has rank equal to 1. Equivalently,  $\rho$  is pure state if there exists a unit vector  $u \in \mathcal{H}$  such that

$$\rho = u u^*,$$

where  $u^*$  is the complex conjugate of the vector  $u$ . By the spectral decomposition, every quantum state can be written as a mixture of pure states. Hence, the extreme points of  $\mathcal{D}(\mathcal{H})$  coincides with the set of pure states.

A measurement on a quantum system is given by Positive Operator Valued Measure (POVM) [76]; that is, given any measurable outcome space  $(\mathbb{X}, \mathcal{X})$ , a POVM is a mapping  $M : \mathcal{X} \rightarrow \mathcal{L}(\mathcal{H})$  from the  $\sigma$ -algebra  $\mathcal{X}$  of  $\mathbb{X}$  to the Hermitian operators  $\mathcal{L}(\mathcal{H})$  on  $\mathcal{H}$  such that

- (1)  $M(\emptyset) = 0$ ,  $M(\mathbb{X}) = \text{Id}$ ;
- (2)  $M(B) \geq 0 \forall B \in \mathcal{X}$ ;
- (3)  $M(\bigcup_j B_j) = \sum_j M(B_j)$  for any disjoint collection  $\{B_j\}$ , where the series is weakly convergent.

Here, weak convergence means the following:

$$\lim_{n \rightarrow \infty} \left\langle \psi, \sum_{j=1}^n M(B_j) \psi \right\rangle = \left\langle \psi, \sum_{j=1}^{\infty} M(B_j) \psi \right\rangle,$$

for all  $\psi \in \mathcal{H}$ .

If the measurable outcome space  $(\mathbb{X}, \mathcal{X})$  is finite in POVM  $M : \mathcal{X} \rightarrow \mathcal{L}(\mathcal{H})$ , it is sufficient to define POVM  $M : \mathbb{X} \rightarrow \mathcal{L}(\mathcal{H})$  on  $\mathbb{X}$  and extend the definition to  $\mathcal{X}$  as follows: for any  $A \in \mathcal{X}$ , define  $M(A) = \sum_{x \in A} M(x)$ . For instance, if  $\mathbb{X} = \{0, 1\}$ , then we will only need to define  $M(0)$  and  $M(1)$ . In the finite case,  $M : \mathbb{X} \rightarrow \mathcal{L}(\mathcal{H})$  should satisfy the following conditions:  $M(x) \geq 0 \forall x \in \mathbb{X}$  and  $\sum_x M(x) = \text{Id}$ .

When one applies the measurement  $M : \mathcal{X} \rightarrow \mathcal{L}(\mathcal{H})$  to the quantum system in the state  $\rho$ , the probability of obtaining the event  $B \in \mathcal{X}$  is given by

$$P(B | \rho) = \text{Tr}\{M(B) \rho\}.$$

Note that since  $\text{Tr}\{\rho\} = 1$  and  $M(B) \geq 0$  for all  $B \in \mathcal{X}$ , we have  $P(B | \rho) \geq 0$  for all  $B \in \mathcal{X}$  and  $P(\mathbb{X} | \rho) = 1$ . Moreover, for any disjoint collection  $\{B_j\}$  of events, by (3) we have

$$P(B | \rho) = \sum_j P(B_j | \rho),$$

where  $B := \bigcup_j B_j$ . Hence,  $P(\cdot | \rho)$  is indeed a probability measure on  $\mathbb{X}$ . We note that if  $\rho$  is a pure state given by a unit vector  $u$ , then

$$P(B | \rho) = \langle u, M(A) u \rangle.$$

In quantum physics, a compound of  $N$  quantum systems with the underlying Hilbert spaces  $\{\mathcal{H}_i, i = 1, \dots, N\}$  is represented by the tensor product  $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_N$  of the individual Hilbert spaces. Therefore, any state  $\rho$  (called

the *compound state*) of this compound quantum system is an element of  $\mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \dots \otimes \mathcal{H}_N)$ . If the state  $\rho$  is in product form

$$\rho = \rho_1 \otimes \dots \otimes \rho_N,$$

then it is called *product state*. In this case, the individual quantum systems are said to be independent. If  $\rho$  is a mixture of product states, it is called *separable*. If the state is not separable, then it is called *entangled*.

**Quantum-correlated strategic measures.** With these definitions, we can now define quantum-correlated strategic measures. An element  $P$  is a *quantum-correlated* strategic measure if agents have access to a part of a quantum compound state  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \dots \otimes \mathcal{H}_N)$ , where  $\{\mathcal{H}_i, i = 1, \dots, N\}$  is a collection of arbitrary Hilbert spaces, and, for each  $i = 1, \dots, N$ , DM  $i$  makes measurements  $M^{i,y^i} : \mathbb{U}^i \rightarrow \mathcal{L}(\mathcal{H}_i)$  on  $i^{\text{th}}$  part of the state  $\rho$  depending on its observations  $y^i$  to generate its action  $u^i$  as the output of the measurement (recall that  $\mathcal{U}^i$  is the Borel  $\sigma$ -algebra on the action space  $\mathbb{U}^i$  of DM  $i$ ); that is, the conditional distribution  $P(d\mathbf{u}|\omega_0, \mathbf{y})$  is of the following form:

$$P(d\mathbf{u}|\omega_0, \mathbf{y}) = \text{Tr} \left\{ \left( M^{1,y^1}(du^1) \otimes \dots \otimes M^{N,y^N}(du^N) \right) \rho \right\}.$$

Let the state  $\rho$  be separable; that is,

$$\rho = \int \nu(dz) \rho_1^z \otimes \dots \otimes \rho_N^z.$$

Suppose that on each individual quantum system, we are performing the following measurement  $M^{i,y^i} : \mathbb{U}^i \rightarrow \mathcal{L}(\mathcal{H}_i)$ . Then, the corresponding joint probability of obtaining the event  $du^1 \times \dots \times du^N$  given observations  $(y^1, \dots, y^N)$  and state  $\omega_0$  is the following:

$$\begin{aligned} P(du^1, \dots, du^N | \omega_0, \mathbf{y}) &= \int \nu(dz) \text{Tr}\{M^{1,y^1}(du^1)\rho_1^z\} \dots \text{Tr}\{M^{N,y^N}(du^N)\rho_N^z\} \\ &= \int \nu(dz) P(du^1 | y^1, z) \dots P(du^N | y^N, z), \end{aligned}$$

where  $P(du^i | y^i, z) = \text{Tr}\{M^{i,y^i}(du^i)\rho_i^z\}$ . This follows from  $\text{Tr}(A \otimes B) = \text{Tr}(A)\text{Tr}(B)$  and that  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ . Therefore,  $P \in L_C(\nu)$ . Hence, separable states can only generate correlations that can be classically generated; that is, if individual quantum systems share a common randomness  $\nu(dz)$ , then they can also realize  $P(du^1, \dots, du^N | \omega_0, \mathbf{y})$  independently using classical devices given the common randomness. Hence, to generate correlations that cannot be realized classically, one must use entangled states.

Let  $L_Q(\mu)$  denote the set of quantum-correlated strategic measures. In the next section, we will show that  $L_C(\mu)$  is a strict subset of  $L_Q(\mu)$ . In view of the discussion in the previous paragraph, the strategic measure  $P \in L_Q(\mu) \setminus L_C(\mu)$  if the shared state  $\rho$  is entangled.

As an example, consider the two-agent team problem with  $\mathbb{Y}^1 = \mathbb{Y}^2 = \mathbb{U}^1 = \mathbb{U}^2 = \{0, 1\}$ . Suppose that  $\Omega_0$  is degenerate. Let  $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$ . Consider the following states on  $\mathbb{C}^2 \times \mathbb{C}^2$ :

$$\begin{aligned}\eta &= \frac{1}{2}E_{1,1} \otimes E_{1,1} + \frac{1}{2}E_{2,2} \otimes E_{2,2} \\ \tau &= \frac{1}{2}E_{1,1} \otimes E_{1,1} + \frac{1}{2}E_{1,2} \otimes E_{1,2} + \frac{1}{2}E_{2,1} \otimes E_{2,1} + \frac{1}{2}E_{2,2} \otimes E_{2,2},\end{aligned}$$

where  $E_{a,b}$  is a matrix on  $\mathbb{C}^2$  such that  $E_{a,b}(i, j) = 1_{\{(i,j)=(a,b)\}}$ , for  $a, b \in \{1, 2\}$ . Note that  $E_{1,1}, E_{2,2} \in \mathcal{D}(\mathbb{C}^2)$ . Hence,  $\eta$  is a mixture of product states, and so, it is separable. However, note that  $E_{1,2}, E_{2,1} \notin \mathcal{D}(\mathbb{C}^2)$ . Therefore,  $\tau$  is an entangled state (indeed it is maximally entangled state). For any angle  $\theta$ , let us define

$$\Pi_\theta = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{pmatrix} = (\cos(\theta) \ \sin(\theta)) \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

Using  $\Pi_\theta$ , we now define two collection of measurements  $\{M^{1,y^1} : \mathbb{U}^1 \rightarrow \mathcal{L}(\mathbb{C}^2); y^1 \in \mathbb{Y}^1\}$ ,  $\{M^{2,y^2} : \mathbb{U}^2 \rightarrow \mathcal{L}(\mathbb{C}^2); y^2 \in \mathbb{Y}^2\}$  for DM 1 and DM 2, respectively, as follows:

$$M^{1,0}(0) = \Pi_0, M^{1,0}(1) = \Pi_{\pi/2}; \quad M^{1,1}(0) = \Pi_{\pi/4}, M^{1,1}(1) = \Pi_{3\pi/4}$$

and

$$M^{2,0}(0) = \Pi_{\pi/8}, M^{2,0}(1) = \Pi_{5\pi/8}; \quad M^{2,1}(0) = \Pi_{7\pi/8}, M^{2,1}(1) = \Pi_{3\pi/8}.$$

Namely, if the observation of DM 1 is  $y^1$ , then DM 1 applies the measurement  $M^{1,y^1} : u^1 \mapsto M^{1,y^1}(u^1)$  to the quantum state. Similarly, if the observation of DM 2 is  $y^2$ , then DM 2 applies the measurement  $M^{2,y^2} : u^2 \mapsto M^{2,y^2}(u^2)$  to the quantum state.

Now, if we apply these measurements to the quantum states  $\eta$  and  $\tau$ , we obtain two quantum-correlated strategic measures  $P_\eta$  and  $P_\tau$ , respectively. The conditional distributions of these quantum-correlated strategic measures on actions given observations can be computed as follows:

$$\begin{aligned}P_\eta(u^1, u^2 | y^1, y^2) &= \frac{1}{2} \left( \text{Tr}\{M^{1,y^1}(u^1)E_{1,1}\} \text{Tr}\{M^{2,y^2}(u^2)E_{1,1}\} + \text{Tr}\{M^{1,y^1}(u^1)E_{2,2}\} \right. \\ &\quad \left. \times \text{Tr}\{M^{2,y^2}(u^2)E_{2,2}\} \right) \\ P_\tau(u^1, u^2 | y^1, y^2) &= \frac{1}{2} \left( \text{Tr}\{M^{1,y^1}(u^1)E_{1,1}\} \text{Tr}\{M^{2,y^2}(u^2)E_{1,1}\} + \text{Tr}\{M^{1,y^1}(u^1)E_{1,2}\} \right. \\ &\quad \left. \times \text{Tr}\{M^{2,y^2}(u^2)E_{1,2}\} \right) \\ &\quad + \frac{1}{2} \left( \text{Tr}\{M^{1,y^1}(u^1)E_{2,1}\} \text{Tr}\{M^{2,y^2}(u^2)E_{2,1}\} + \text{Tr}\{M^{1,y^1}(u^1)E_{2,2}\} \right)\end{aligned}$$

$$\times \text{Tr}\{M^{2,y^2}(u^2)E_{2,2}\}.$$

Since  $E_{1,1}, E_{2,2} \in \mathcal{D}(\mathbb{C}^2)$ , the conditional distribution  $P_\eta(u^1, u^2|y^1, y^2)$  is a mixture of the following independent conditional probability measures

$$P_{1,\eta}(u^1|y^1)P_{1,\eta}(u^2|y^2) := \text{Tr}\{M^{1,y^1}(u^1)E_{1,1}\} \text{Tr}\{M^{2,y^2}(u^2)E_{1,1}\}$$

and

$$P_{2,\eta}(u^1|y^1)P_{2,\eta}(u^2|y^2) := \text{Tr}\{M^{1,y^1}(u^1)E_{2,2}\} \text{Tr}\{M^{2,y^2}(u^2)E_{2,2}\}.$$

Hence  $P_\eta$  can be classically realized via common randomness; that is,  $P_\eta \in L_C(\mu)$ . However, this is not true for  $P_\tau$  since  $E_{1,2}, E_{2,1} \notin \mathcal{D}(\mathbb{C}^2)$ , and so,

$$\text{Tr}\{M^{1,y^1}(u^1)E_{1,2}\} \text{Tr}\{M^{2,y^2}(u^2)E_{1,2}\}$$

and

$$\text{Tr}\{M^{1,y^1}(u^1)E_{2,1}\} \text{Tr}\{M^{2,y^2}(u^2)E_{2,1}\}$$

are not probability measures. Therefore,  $P_\tau$  cannot be realized classically; that is,  $P_\tau \in L_Q(\mu) \setminus L_C(\mu)$ .

Here, one can view the following collection as the joint strategy of the decision makers:

$$\left\{ \mathcal{H}_i (i \in \mathcal{N}), \rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \dots \otimes \mathcal{H}_N), \{M^{i,y^i} : \mathcal{U}^i \rightarrow \mathcal{L}(\mathcal{H}_i); y^i \in \mathbb{Y}^i\} (i \in \mathcal{N}) \right\}.$$

Namely, at the beginning of the problem, decision makers agree on the Hilbert spaces  $\mathcal{H}_i (i \in \mathcal{N})$  and share a compound state  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \dots \otimes \mathcal{H}_N)$  (i.e., entangled state) to be used in the measurements. Then, they choose decentralized measurements  $\{M^{i,y^i} : \mathcal{U}^i \rightarrow \mathcal{L}(\mathcal{H}_i); y^i \in \mathbb{Y}^i\} (i \in \mathcal{N})$  without communicating with each other.

Given the Hilbert spaces and the compound state, it is possible to view decentralized measurements as the decentralized strategies of the decision makers. However, when designing optimal quantum-correlated strategic measures, decision makers must also find the optimal Hilbert spaces and the corresponding optimal compound state. Note that the maximally entangled state is in general not the optimal choice [42]. Therefore, the optimization variable in the decision problem corresponding to the quantum-correlated strategic measures is in general quite large even if the measurement and the action spaces are finite. However, note that [48, 84] (see also Example 5.1) give evidence that quantum-correlated teams with finite measurement and action spaces can be computationally tractable as opposed to their classical randomized counterparts even though the optimization space is quite large compared to the classical case. Namely, in these papers, optimization problems associated with quantum-correlated strategic measures can be written as (or can be approximated by) semi-definite programs whose sizes scale with the cardinality of the measurement and action spaces. As a result, they can be solved exactly or approximately in polynomial

time. In particular, [48] computes the optimal value of XOR team and [84] approximates the optimal value of the unique teams via semi-definite programs.

However, there are other instances of finite team problems [83, 131, 107, 81], where exact or approximate computation of the optimal cost of quantum-correlated teams is NP-hard, and therefore, cannot be cast as a semi-definite program. One of the reasons for such negative complexity results might be the fact that the optimal value of quantum-correlated teams can be attained via quantum systems living in infinite-dimensional Hilbert spaces (see, e.g., [58]). Therefore, in order to obtain a computationally tractable problem, we may need to put a constraint on the dimensions of the Hilbert spaces where quantum systems live in.

To that end, let  $L_{Q(d)}(\mu) \subset L_Q(\mu)$  denote the set of quantum-correlated strategic measures where each agent has access to a part of quantum compound state  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \dots \otimes \mathcal{H}_N)$  with  $\dim(\mathcal{H}_i) \leq d$  for all  $i \in \mathcal{N}$ .

In the following, we state convexity and topological properties of the sets  $L_Q(\mu)$  and  $L_{Q(d)}(\mu)$ .

**Theorem 5.2.** (i)  $L_Q(\mu)$  is a convex set but it is not closed under the weak convergence topology.

(ii)  $L_{Q(d)}(\mu)$  is not convex if  $L_C(\mu) \not\subset L_{Q(d)}(\mu)$ . In particular,  $L_{Q(1)}(\mu)$  is not convex and  $L_{Q(2)}(\mu)$  is not convex for certain models.

(iii) For each  $d$ ,  $L_{Q(d)}(\mu)$  is closed under the weak convergence topology if the observation spaces are finite and action spaces are compact.

(iv)  $L_Q(\mu)$  is a strict super-set of  $\bigcup_d L_{Q(d)}(\mu)$ .

*Proof.* Proof of (i): The convexity of  $L_Q(\mu)$  can be proved easily and so we omit the details. Non-closedness of  $L_Q(\mu)$  has been a longstanding open problem and is proved recently in [128] who showed that the weak closure of the set is strictly larger than the set itself by explicit construction.

Proof of (ii): By [53, Corollary 2],  $L_{Q(d)}(\mu)$  is not convex if  $L_C(\mu) \not\subset L_{Q(d)}(\mu)$ . In particular,  $L_{Q(1)}(\mu)$  is not convex by [53, Proposition 1] and  $L_{Q(2)}(\mu)$  is not convex for certain models by [53, Proposition 2].

Proof of (iii): Since both the set of states and the set of POVMs of fixed (local) dimension are compact in finite measurement space and compact action space case and the trace is a continuous map,  $L_{Q(d)}(\mu)$  is closed (see [67, Appendix B] where the finite action case is considered but the same argument applies for the compact case also).

Proof of (iv): The fact that  $L_Q(\mu)$  is a strict super-set of  $\bigcup_d L_{Q(d)}(\mu)$  is proved in [49]; that is, the authors show that there exists a joint strategy which is not attainable by quantum systems of any arbitrary finite dimension, but is attained exclusively by infinite-dimensional quantum systems  $\square$

The following result is a corollary of Theorem 5.2-(iii).

**Corollary 5.1.** Consider a static team decision problem with finite measurement spaces (and hence independent measurements reduction) and compact action spaces. Then, the following team problem  $\inf_{P \in L_{Q(d)}(\mu)} J(P)$  has an optimal team policy.

To show the effectiveness (both computationally and optimally) of quantum-correlated strategic measures over classical randomized ones, we consider the XOR team problem and a variation below.

**Example 5.1.** *In the XOR team, we have two agents with binary action spaces  $\{0, 1\}$ ; that is,  $\mathbb{U}^1 = \mathbb{U}^2 = \{0, 1\}$ . Measurements are generated independently and uniformly over some finite sets  $\mathbb{Y}^1$  and  $\mathbb{Y}^2$  via probability measure  $\mu(y^1, y^2)$ . Hence, there is no  $\omega_0$  variable in the problem, and so the problem is automatically static. The reward<sup>1</sup> function is defined as*

$$r(y^1, y^2, u^1, u^2) = \begin{cases} 1, & \text{if } u^1 \oplus u^2 = h(y^1, y^2) \\ -1, & \text{otherwise,} \end{cases}$$

where  $h : \mathbb{Y}^1 \times \mathbb{Y}^2 \rightarrow \{0, 1\}$  is some arbitrary binary-valued function. This team problem with quantum-correlated policies can be written as a semi-definite program due to Tsirelson's Theorem [135, Theorem 6.62].

Indeed, let us define

$$g(y^1, y^2) = \mu(y^1, y^2) (-1)^{h(y^1, y^2)}.$$

Given any  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$  for some finite-dimensional Hilbert spaces  $\mathcal{H}_1, \mathcal{H}_2$  (finite dimensional Hilbert spaces are sufficient for this problem) and given any two collection of POVMs  $\{M^{1,y^1}, y^1 \in \mathbb{Y}^1\}, \{M^{2,y^2}, y^2 \in \mathbb{Y}^2\}$ , the corresponding strategic measure  $P$  is  $P(y^1, y^2) = \mu(y^1, y^2)$  and

$$P(u^1, u^2 | y^1, y^2) = \text{Tr} \left\{ \left( M^{1,y^1}(u^1) \otimes M^{2,y^2}(u^2) \right) \rho \right\}$$

and its expected reward function can be written as

$$J(P) = \sum_{y^1, y^2} g(y^1, y^2) \text{Tr} \{ (M^{1,y^1}(0) - M^{1,y^1}(1)) \otimes (M^{2,y^2}(0) - M^{2,y^2}(1)) \rho \}.$$

Note that an operator  $H$  is Hermitian with operator norm  $\|H\| \leq 1$  if and only if it can be written as  $H = M(0) - M(1)$ , where  $M(i)$  ( $i = 0, 1$ ) are positive semi-definite operators with  $M(0) + M(1) = \text{Id}$ . Therefore, for any pair  $(y^1, y^2)$ ,  $H^{y^1} = M^{1,y^1}(0) - M^{1,y^1}(1)$  and  $H^{y^2} = M^{2,y^2}(0) - M^{2,y^2}(1)$  are Hermitian with operator norms less than 1. Conversely, for any pair  $(y^1, y^2)$ , any Hermitian operators  $H^{y^1}, H^{y^2}$  with operator norms less than 1 can be decomposed as above.

Let  $X$  be a  $|\mathbb{Y}^1| \times |\mathbb{Y}^2|$  real matrix. Tsirelson's Theorem states that the following assertions are equivalent:

1. There exist Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , a state  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ , and two collections of Hermitian operators

$$\{H^{y^1}, y^1 \in \mathbb{Y}^1\} \text{ and } \{H^{y^2}, y^2 \in \mathbb{Y}^2\}$$

---

<sup>1</sup>All results in this paper are also true for maximization of a reward function.



whose operator norms are less than 1, and

$$X(y^1, y^2) = \text{Tr} \left\{ \left( H^{y^1} \otimes H^{y^2} \right) \rho \right\},$$

for all  $y^1 \in \mathbb{Y}^1$ ,  $y^2 \in \mathbb{Y}^2$ .

2. There exists two collections  $\{u_{y^1}, y^1 \in \mathbb{Y}^1\}$ ,  $\{v_{y^2}, y^2 \in \mathbb{Y}^2\} \subset \mathbb{R}^{|\mathbb{Y}^1| \times |\mathbb{Y}^2|}$  of unit vectors such that

$$X(y^1, y^2) = u_{y^1}^T v_{y^2}$$

for all  $y^1 \in \mathbb{Y}^1$ ,  $y^2 \in \mathbb{Y}^2$ .

Therefore, Tsirelson's Theorem and the above fact about Hermitian operators imply that

$$\sup_{P \in L_Q(\mu)} J(P) = \sup_{\substack{u_{y^1}, v_{y^2} \in \mathbb{R}^{|\mathbb{Y}^1| \times |\mathbb{Y}^2|} \\ y^1 \in \mathbb{Y}^1, y^2 \in \mathbb{Y}^2}} \sum_{y^1, y^2} g(y^1, y^2) u_{y^1}^T v_{y^2},$$

subject to

$$u_{y^1}^T u_{y^1} = 1 \text{ and } v_{y^2}^T v_{y^2} = 1, \text{ for } y^1 \in \mathbb{Y}^1, y^2 \in \mathbb{Y}^2.$$

This optimization problem is indeed a semi-definite program. Therefore, the optimal value of the XOR team with quantum-correlated policies can be computed in polynomial time as opposed to its classical counterpart.

**Example 5.2.** A special case for XOR team is the celebrated CHSH (Clauser-Horne-Shimony-Holt) team [135]. In CHSH team, we have binary observation and action spaces  $\{0, 1\}$  and the reward function is defined as

$$r(y^1, y^2, u^1, u^2) = \begin{cases} 1, & \text{if } u^1 \oplus u^2 = y^1 \cdot y^2 \\ -1, & \text{otherwise.} \end{cases}$$

For this problem, the optimal value of randomized policies is 0.5 [135, Section 6.3.2], which can be verified by an inspection of 16 measurable strategic measures as they are sufficient for optimality. However, quantum-correlated policies can achieve the maximum reward of  $2\sqrt{2}/4$ , that is obtained by solving the corresponding semi-definite program. Therefore, for CHSH team, we have

$$\sup_{P \in L_Q(\mu)} J(P) > J_*,$$

that is, quantum-correlated policies improve the optimal value  $J^*$  of the original team as opposed to randomized policies.

Indeed, Let  $\mathcal{H}_1 = \mathcal{H}_2 = \mathbb{C}^2$  and let  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2)$  be the maximally entangled state

$$\rho = \frac{1}{2} \sum_{a,b \in \{1,2\}} E_{a,b} \otimes E_{a,b}.$$

Recall the two collection of POVMs  $\{M^{1,y^1}, y^1 \in \mathbb{Y}^1\}, \{M^{2,y^2}, y^2 \in \mathbb{Y}^2\}$ :

$$M^{1,0}(0) = \Pi_0, M^{1,0}(1) = \Pi_{\pi/2}; M^{1,1}(0) = \Pi_{\pi/4}, M^{1,1}(1) = \Pi_{3\pi/4}$$

and

$$M^{2,0}(0) = \Pi_{\pi/8}, M^{2,0}(1) = \Pi_{5\pi/8}; M^{2,1}(0) = \Pi_{7\pi/8}, M^{2,1}(1) = \Pi_{3\pi/8},$$

where

$$\Pi_\theta = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{pmatrix}.$$

Then, the corresponding quantum-correlated strategic measure  $P$  has the following conditional distribution  $P(u^1, u^2|y^1, y^2)$ :

$$\begin{aligned} P(\cdot, \cdot|0, 0) &= \left( \frac{2 + \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8} \right), \\ P(\cdot, \cdot|0, 1) &= \left( \frac{2 - \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8} \right) \\ P(\cdot, \cdot|1, 0) &= \left( \frac{2 + \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8} \right) \\ P(\cdot, \cdot|1, 1) &= \left( \frac{2 - \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8}, \frac{2 + \sqrt{2}}{8}, \frac{2 - \sqrt{2}}{8} \right) \end{aligned} \tag{17}$$

Now, it is straightforward to compute that  $J(P) = 2\sqrt{2}/4$ . Hence,  $P$  is the optimal quantum-correlated strategic measure.

### 5.3. Relations between convex relaxations

Up to this point we have introduced a number of convex strategic measures, we now present a comparison. First, we note that the joint strategies in the set  $L_C(\mu)$  indeed correspond to the *local hidden variable* correlations in quantum mechanics. In [56], Einstein, Podolsky, and Rosen have made the point that due to the probabilistic nature of quantum mechanics, the theory cannot be complete and should be supplemented by a local hidden variable to describe the probabilistic nature, where the local hidden variable describes the uncertainties in the experimental setup. In other words, they claimed that  $L_C(\mu) = L_Q(\mu)$ . However, John Bell, in his influential paper [25], constructed a quantum mechanical system whose statistical behavior cannot be explained via the local hidden variable theory; that is,  $L_Q(\mu)$  is a strict super-set of  $L_C(\mu)$ . This result implies that quantum mechanical systems are intrinsically probabilistic and cannot be explained by the lack of precision in experimental measurements. Experimental verification of Bell's prediction has been established by Alain Aspect in [10]. Additionally, in [116], Popescu and Rohrlich also discussed that correlations

(strategic measures in our setting) achieved by quantum mechanical systems can be as large as  $L_{NS}(\mu)$  which itself is larger than  $L_Q(\mu)$ . However, no experimental verification has been provided for this claim as yet, to our knowledge.

We also note that the lower bounds on the optimal team costs achieved by classical strategies  $L_C(\mu)$  is known to be *Bell inequalities* in the quantum information theory literature. Bell inequality violations correspond to quantum-correlated strategies that improve the classical optimal team cost [48]. Identifying Bell inequality violations is an active area of research in quantum information theory. One such violation is established in Example 5.2 for CHSH team. Since the considered sets are convex, by the Hahn-Banach Separation Theorem [99], identifying Bell inequality violations is equivalent to finding a *separating* function in an inner-product form. Accordingly identifying Bell inequality violations would be implied by finding a cost function for which the optimal quantum-correlated team cost is strictly smaller than the optimal classical team cost; that is;

$$\inf_{P \in L_Q(\mu)} \int P(ds) c(s) < \inf_{P \in L_C(\mu)} \int P(ds) c(s).$$

Since both  $L_Q(\mu)$  and  $L_C(\mu)$  are convex sets, the above inequality states that the hyperplane  $\{P : \int P(ds) c(s) = J^*\}$  separates the convex sets  $L_Q(\mu)$  and  $L_C(\mu)$ . It is in general very hard to find all Bell inequality violations (i.e., separating hyperplanes) for a certain setup.

Before we establish the relation between convex relaxations, we introduce a further information structure, denoted by  $L_{CJ}$ , which allows for the common random variable  $z$  to be correlated with  $\omega_0, \mathbf{y}$ ; in this case, the information structure is essentially centralized, since the dependence between  $z$  and the exogenous variables are arbitrary:

$$L_{CJ}(\mu) := \left\{ P \in \mathcal{P} \left( \Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k) \right) : \right. \\ \left. P(B) = \int_{[0,1]^N \times B^0 \times \prod_{k=1}^N A^k} \mu(dz, d\omega_0, d\mathbf{y}) \prod_k \mathbf{1}_{\{u^k = \gamma^k(y^k, z) \in B^k\}}, \right. \\ \left. \gamma^k(\cdot, z) \in \Gamma^k, B \in \mathcal{B}(\Omega_0 \times \prod_k (\mathbb{Y}^k \times \mathbb{U}^k)) \right\},$$

where the marginal of  $\omega_0, \mathbf{y}$  is fixed as  $\mu$ .

**Theorem 5.3.** *We have the following relation between convex relaxations.*

- (i)  $L_C(\mu) \subset L_Q(\mu) \subset L_{NS}(\mu) \subset L_M(\mu) \subset L_{CJ}(\mu)$ .
- (ii) *The sets  $L_C(\mu), L_Q(\mu), L_{NS}(\mu), L_M(\mu), L_{CJ}(\mu)$  are convex.*
- (iii) *The inclusions among the convex sets above are strict.*
- (iv) *There exist problems for which*

$$\inf_{P \in L_Q(\mu)} \int P(ds) c(s) < \inf_{P \in L_C(\mu)} \int P(ds) c(s)$$

$$\inf_{P \in L_{NS}(\mu)} \int P(ds) c(s) < \inf_{P \in L_Q(\mu)} \int P(ds) c(s)$$

and in particular

$$\inf_{P \in L_{NS}(\mu)} \int P(ds) c(s) < \inf_{P \in L_C(\mu)} \int P(ds) c(s).$$

One implication of the above is that, the convex programs:

$$\inf_{P \in L_Q(\mu)} \int P(ds) c(s) \text{ and } \inf_{P \in L_{NS}(\mu)} \int P(ds) c(s)$$

or  $\inf_{P \in L_M(\mu)} \int P(ds) c(s)$  provide lower bounds to the original problem. In particular, since the optimization problem associated with non-signaling strategic measures can be written as a linear program, it can be solved in a polynomial time if the measurement and action spaces are finite.

*Proof.* Proof of (ii): We have already proved the convexity of  $L_C(\mu)$ ,  $L_Q(\mu)$ ,  $L_{NS}(\mu)$ , and  $L_M(\mu)$ .

Proof of (i): Let  $P \in L_R(\mu)$ ; that is,

$$P(d\mathbf{u}|\omega_0, \mathbf{y}) = \prod_{i=1}^N \Pi^i(du^i|y^i).$$

Let  $\mathcal{H}$  be a Hilbert space with dimension 1. Let  $e_w \in \mathcal{H}$  with unit norm. We define

$$\begin{aligned} \rho &= e_w e_w^* \otimes \dots \otimes e_w e_w^* \\ M^{y^i}(A) &= \Pi^i(A|y^i) e_w e_w^* \end{aligned}$$

for all  $y^i \in \mathbb{Y}^i$ ,  $A \in \mathcal{B}(\mathbb{U}^i)$ , and  $i = 1, \dots, N$ . Then, we have

$$P(d\mathbf{u}|\omega_0, \mathbf{y}) = \text{Tr}\left\{ \left( M^{y^1}(du^1) \otimes \dots \otimes M^{y^N}(du^N) \right) \rho \right\}.$$

This implies that  $L_R(\mu) \subset L_Q(\mu)$ . Since the convex hull of  $L_R(\mu)$  is  $L_C(\mu)$  and  $L_Q(\mu)$  is convex, we have  $L_C(\mu) \subset L_Q(\mu)$ .

Let  $P \in L_Q(\mu)$ ; that is, agents have access to a part of a compound quantum state  $\rho \in \mathcal{D}(\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_N)$ , where  $\{\mathcal{H}_i, i = 1, \dots, N\}$  is a collection of arbitrary Hilbert spaces, and, for each  $i = 1, \dots, N$ , DM  $i$  makes measurements  $M^{y^i} : \mathcal{U}^i \rightarrow \mathcal{L}(\mathcal{H}_i)$  on the  $i^{\text{th}}$  part of the state  $\rho$  depending on its observations  $y^i$  to generate its action as the output of the following measurement:

$$P(d\mathbf{u}|\omega_0, \mathbf{y}) = \text{Tr}\left\{ \left( M^{y^1}(du^1) \otimes \dots \otimes M^{y^N}(du^N) \right) \rho \right\}.$$

We prove that  $P$  satisfies the non-signaling condition. Fix any  $k \in \{1, \dots, N\}$ . Then, we have

$$\int_{\mathbb{U}^k} P(d\mathbf{u}|\omega_0, y^1, \dots, y^k, \dots, y^N)$$

$$\begin{aligned}
&= \int_{\mathbb{U}^k} \text{Tr}\left\{\left(M^{y^1}(du^1) \otimes \dots \otimes M^{y^k}(du^k) \otimes \dots \otimes M^{y^N}(du^N)\right) \rho\right\} \\
&= \text{Tr}\left\{\left(\int_{\mathbb{U}^k} M^{y^1}(du^1) \otimes \dots \otimes M^{y^k}(du^k) \otimes \dots \otimes M^{y^N}(du^N)\right) \rho\right\} \\
&= \text{Tr}\left\{\left(M^{y^1}(du^1) \otimes \dots \otimes \text{Id} \otimes \dots \otimes M^{y^N}(du^N)\right) \rho\right\} \\
&= \int_{\mathbb{U}^k} \text{Tr}\left\{\left(M^{y^1}(du^1) \otimes \dots \otimes M^{\hat{y}^k}(du^k) \otimes \dots \otimes M^{y^N}(du^N)\right) \rho\right\} \\
&= \int_{\mathbb{U}^k} P(\mathbf{d}\mathbf{u} \mid \omega_0, y^1, \dots, \hat{y}^k, \dots, y^N)
\end{aligned}$$

for all  $y^1, \dots, y^{k-1}, y^k, \hat{y}^k, y^{k+1}, \dots, y^N$  and  $\omega_0$ . Therefore,  $P$  is in  $L_{NS}(\mu)$ .

Proof of (iii) and (iv): We prove (iii) and (iv) by providing an example. This example is CHSH team. Recall that in CHSH team, we have two agents with binary measurement and action spaces  $\{0, 1\}$ . Observations are generated independently and uniformly. Hence, there is no state variable in the problem, and so the problem is automatically static. The reward function is defined as

$$r(y^1, y^2, u^1, u^2) = \begin{cases} 1, & \text{if } u^1 \oplus u^2 = y^1 \cdot y^2 \\ 0, & \text{otherwise.} \end{cases}$$

For this problem, randomized strategic measures  $L_C(\mu)$  can achieve 0.5 [48]. Quantum-correlated strategic measures can achieve the maximum reward of  $2\sqrt{2}/4$  [135, Section 6.3.2] (see also Example 5.2). However, non-signaling policies can achieve the maximum reward of 1 using the following policy, which is called Popescu-Rohrlich (PR) box in the literature [116]:

$$P(u^1, u^2 \mid y^1, y^2) = \begin{cases} 1/2, & \text{if } u^1 \oplus u^2 = y^1 \cdot y^2 \\ 0, & \text{otherwise} \end{cases}$$

It is easy to show that  $P$  is non-signaling; that is,  $u^i$  is independent of  $y^j$  given  $y^i$ , for  $i, j = 0, 1$  and  $i \neq j$ . The reward of  $P$  is 1, which is the maximum achievable reward by any policy as  $0 \leq r \leq 1$ . Hence,  $P$  is the optimal non-signaling policy. Therefore, for CHSH team, we have

$$\sup_{P \in L_{NS}(\mu)} \int P(ds) r(s) > \sup_{P \in L_Q(\mu)} \int P(ds) r(s) > \sup_{P \in L_C(\mu)} \int P(ds) r(s)$$

that is, non-signaling policies improve the optimal team cost of quantum-correlated policies and quantum correlated policies improve the optimal team cost of randomized policies.  $\square$

**Remark 5.1** (On comparing quantum and non-signaling policies). *Note that to implement the optimal non-signaling policy in the CHSH team, agents should communicate their observations to a mediator, and then, mediator directs them to apply either the same actions or different actions based on the product of*

their observations. This kind of communication is, in general, infeasible for team decision problems. Therefore, although allowing non-signaling correlations among actions of agents enables us to formulate the team problem as a linear program (solvable in a polynomial time for finite case), it is in general not realistic to assume that agents can apply such policies in real-life applications due to communication constraints dictated by decentralized information structure. Therefore, non-signaling relaxation can only be used to provide a lower bound to the original team problem.

However, quantum correlated relaxation is indeed an admissible extension or relaxation of classical team problem since it does not require any communication between agents and a mediator (assuming that we live in the quantum world equipped with appropriate instrumentation!). Hence, the solution or approximation (as explained in the previous section) of quantum correlated team problems will be a significant contribution to team decision theory.

**Remark 5.2** (Information theoretic relaxations). *At the heart of information theory's success is the arrival at single-letter characterizations of optimal information transmission problems which are, in operational formulations, strictly non-convex optimization problems. As detailed out in [149, Section 5.4] in the context of optimal quantization problems, information theory convexifies these problem by first relaxing the constraints (such as conditional independence) with mutual information constraints and randomized codes and then showing the attainability of such bounds as the dimension of the problem reaches infinity. Building on this insight, various efforts have been presented for team theoretic problems. Notably, [89] studied strategic measures in the context of product measures involving an encoder and a decoder, and a class of team problems, and have considered convexification of the strategic measures by characterizing information channels with information theoretic inequalities. Here, the idea is to abstractly view a channel by its mutual information properties, and thus avoiding the product-structure which ultimately makes  $L_R(\mu)$  non-convex.*

#### 5.4. Linear programming formulation and its dual

Since the non-signaling constraints are linear in  $P$ , the optimization problem associated with non-signaling strategic measures can be written as a linear program over an appropriate vector space. The dual of this linear program and its approximation will provide a lower bound to the original problem. This lower bound, serving as a benchmark, can be quite useful in evaluating how a sub-optimal policy performs in the team problem.

One such linear programming formulation can be done as follows. For any metric space  $\mathbb{E}$ , let  $\mathcal{M}(\mathbb{E})$  denote the set of finite signed measures on  $\mathbb{E}$  and  $C(\mathbb{E})$  denotes the vector space of some real measurable functions on  $\mathbb{E}$ . Depending on the topological properties of the cost function  $c$ , one should choose  $C(\mathbb{E})$  appropriately. For instance, if  $c$  is continuous, then  $C(\mathbb{E})$  must be chosen as the set of all continuous functions.

Consider the vector spaces

$$\mathcal{M} \left( \Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k) \right), \quad C \left( \Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k) \right).$$

For each  $i \in \mathcal{N}$ , we also define

$$\mathcal{M} \left( \Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k \times \prod_{k \neq i} \mathbb{U}^k \right), \quad C \left( \Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k \times \prod_{k \neq i} \mathbb{U}^k \right).$$

To ease the notation, we will denote these sets as  $\mathcal{M}, C, \mathcal{M}_i, C_i$  for  $i \in \mathcal{N}$ . Let us define bilinear forms on  $(\mathcal{M}, C)$  and on  $(\prod_{i=1}^N \mathcal{M}_i \times \mathbb{R}, \prod_{i=1}^N C_i \times \mathbb{R})$  by letting

$$\begin{aligned} \langle \rho, v \rangle_1 &:= \int v(\omega_0, \mathbf{y}, \mathbf{u}) \rho(d\omega_0, d\mathbf{y}, d\mathbf{u}), \quad (18) \\ \left\langle \left( \prod_{i=1}^N \rho_i, a \right), \left( \prod_{i=1}^N v_i, b \right) \right\rangle_2 &:= \sum_{i=1}^N \int v_i(\omega_0, \mathbf{y}, \mathbf{u}^{-i}) \rho_i(d\omega_0, d\mathbf{y}, d\mathbf{u}^{-i}) + ab. \quad (19) \end{aligned}$$

The bilinear forms in (18) and (19) constitute duality between spaces [22, Chapter IV.3]. Hence, the topologies on these spaces should be understood as the weak topology of the duality induced by these bilinear forms. One should not confuse these topologies with the topologies induced by dual vector spaces.

We define the linear map  $L : \mathcal{M} \rightarrow \prod_{i=1}^N \mathcal{M}_i \times \mathbb{R}$  by  $L(\rho) = (\prod_{i=1}^N L_i(\rho), T(\rho))$ , where

$$\begin{aligned} L_i : \rho(d\omega_0, d\mathbf{y}, d\mathbf{u}) &\mapsto \rho(d\omega_0, d\mathbf{y}, d\mathbf{u}^{-i}) - \rho(d\mathbf{y}^{-i}, d\mathbf{u}^{-i}) \mu(d\omega_0, dy^i), \quad i \in \mathcal{N} \\ T : \rho(d\omega_0, d\mathbf{y}, d\mathbf{u}) &\mapsto \langle \rho, 1 \rangle_1. \end{aligned}$$

Using  $L$ , the optimal team cost with non-signaling strategic measures can be written as a linear program as follows:

$$\begin{aligned} (\mathbf{NS}) \quad &\text{minimize}_{\rho \in \mathcal{M}_+} \langle \rho, c \rangle_1 \\ &\text{subject to } L(\rho) = (0, \dots, 0, 1). \quad (20) \end{aligned}$$

With this linear programming formulation, the optimal team cost with non-signaling strategic measures can be found in polynomial time if the measurement and action spaces are finite sets. This is not possible for teams with randomized policies [112] and also class of teams with quantum-correlated relaxations [83, 131, 107, 81]. Note that if the measurement and action spaces are continuous, the linear program (NS) is infinite dimensional, and so, should be approximated. But, this approximation may not provide a lower bound to the original team problem. To achieve this, we should first formulate the dual of

(NS) and approximate the resulting infinite dimensional dual linear program to obtain a lower bound to the original team problem.

Note that the dual of  $L$  is given by  $L^* : \prod_{i=1}^N C_i \times \mathbb{R} \rightarrow C$ , where

$$L^*(v_1, \dots, v_N, b) := \sum_{i=1}^N \left( v_i(\omega_0, \mathbf{y}, \mathbf{u}^{-i}) - \int_{\Omega_0 \times \mathbb{Y}^i} \mu(d\omega_0, dy^i) v_i(\omega_0, \mathbf{y}, \mathbf{u}^{-i}) \right) + b.$$

Then the dual program of (NS) can then be written as [22, Chapter IV.6]

$$\begin{aligned} (\mathbf{NS}^*) \text{ maximize}_{(v_1, \dots, v_N, b) \in \prod_{i=1}^N C_i \times \mathbb{R}} \quad & \langle (0, \dots, 0, 1), (v_1, \dots, v_N, b) \rangle_2 = b \\ \text{subject to } & L^*(v_1, \dots, v_N, b) \leq c. \end{aligned} \tag{21}$$

Note that this dual linear program (NS\*) is in infinite dimensional spaces that are often computationally intractable. By weak duality, the maximum value of (NS\*) is a lower bound to the original team decision problem. Therefore, its approximation is also a lower bound to the original team problem as we are maximizing the objective function as opposed to the linear program (NS). To approximate the infinite dimensional linear program (NS\*), we can use the techniques developed in [106] in which approximations to infinite dimensional linear programming problems were introduced. Indeed, in this work, authors applied the findings of the paper to approximate the Markov decision processes (MDPs) using the linear programming formulation of the MDPs. Similar to the (NS\*), this linear programming problem is defined on a function space. In [106], the authors first approximate the infinite dimensional function space with a finite dimensional subspace spanned by finitely many independent functions. For instance, one can use finitely many Fourier basis functions to generate this finite dimensional subspace. With this reduction, the problem becomes semi-infinite since the functions should still satisfy the inequality constraint (i.e.,  $L^*(v_1, \dots, v_N, b) \leq c$  in (NS\*)) for uncountably many variables (i.e., for all  $(\omega_0, \mathbf{y}, \mathbf{u}) \in \Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  in (NS\*)). The next step is the randomization step; that is, simulate i.i.d. samples using some distribution and let functions satisfy the inequality constraint only for these sampled points. As a result, we obtain a (random) finite dimensional approximation to the original linear program with some probabilistic convergence guarantee. In our setup, by solving this finite dimensional approximation, we can then obtain a lower bound to the original team problem.

In the following, we will apply a similar reasoning, but to the primal problem to arrive at (as tight as desired) upper bounds, instead.

## 6. Finite approximations of information structures via quantization

In this section, we consider the finite approximation of static team problems. Since results of this section can also be applied to static reduction of dynamic teams, we suppose that the cost function  $c$  also depends on the measurements  $\mathbf{y}$  (which is not the case in the original problem formulation). Recall that, in the



independent static reduction of a dynamic team, the reduced cost function  $c_s$  is a function of  $\omega_0$ ,  $\mathbf{u}$ , and  $\mathbf{y}$ . To obtain finite approximation result, the following assumptions are imposed on the components of the model.

- Assumption 6.1.** (a) *The cost function  $c$  is continuous in  $(\mathbf{u}, \mathbf{y})$  for any fixed  $\omega_0$ . In addition, it is bounded on any compact subset of  $\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$ .*
- (b) *For each  $k$ ,  $\mathbb{U}^k$  is a closed and convex subset of a completely metrizable locally convex vector space.*
- (c) *For each  $k$ ,  $\mathbb{Y}^k$  is locally compact.*
- (d) *For any subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$ , the function  $w_G(\omega_0, \mathbf{y}) := \sup_{\mathbf{u} \in G} c(\omega_0, \mathbf{y}, \mathbf{u})$  is integrable with respect to  $\mu(d\omega_0, d\mathbf{y})$ , for any compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$  of the form  $G = \prod_{k=1}^N G^k$ .*
- (e) *For any  $\underline{\gamma} \in \mathbf{\Gamma}$  with  $J(\underline{\gamma}) < \infty$  and each  $k$ , there exists  $u^{k,*} \in \mathbb{U}^k$  such that  $J(\underline{\gamma}^{-k}, \gamma_{u^{k,*}}^k) < \infty$ , where  $\gamma_{u^{k,*}}^k \equiv u^{k,*}$ .*

Note that Assumption 6.1-(d),(e) hold if the cost function is bounded. Indeed, conditions in Assumption 6.1 are quite mild and hold for the counterexample of Witsenhausen [137].

In what follows, for any subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$ , we let

$$\mathbf{\Gamma}_G := \left\{ \underline{\gamma} \in \mathbf{\Gamma} : \underline{\gamma} \left( \prod_{k=1}^N \mathbb{Y}^k \right) \subset G \right\}$$

and  $\mathbf{\Gamma}_{c,G} := \mathbf{\Gamma}_c \cap \mathbf{\Gamma}_G$ , where  $\mathbf{\Gamma}_c$  denotes the set of continuous strategies. Using these definitions, let us define the following set of strategic measures for any subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$ :

$$L_A^G(\mu) := \left\{ P \in L_A(\mu) : P(B) = \int_{B^0 \times \prod_{k=1}^N A^k} \mu(d\omega_0, d\mathbf{y}) \prod_{k=1}^N 1_{\{u^k = \gamma^k(y^k) \in B^k\}}, \underline{\gamma} \in \mathbf{\Gamma}_G \right\}.$$

Let  $L_A^{G,c}(\mu)$  denote the set of strategic measures in  $L_A^G(\mu)$  induced by continuous policies.

The following result states that, there exists a near optimal strategic measure whose support on the product of action spaces  $\prod_{k=1}^N \mathbb{U}^k$  is convex and compact (and thus bounded) subset  $G$  of it, and conditional distributions of actions given measurements are induced by continuous policies.

**Proposition 6.1.** *Suppose Assumption 6.1 holds. Then, for any  $\varepsilon > 0$  there exists a compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$  of the form  $G = \prod_{i=1}^N G^i$ , where each  $G^i$  is convex and compact, such that*

$$\inf_{P \in L_A^{G,c}(\mu)} \int P(ds) c(s) < J^* + \varepsilon.$$

*Sketch of Proof.* Given any strategic measure, using Assumption 6.1-(e) and the fact that every measure on a Borel space is tight [113, Theorem 3.2], one can construct a strategic measure in  $L_A(\mu)$  whose support on the product of action spaces is convex and compact and whose cost is  $\varepsilon/2$ -close to the cost of the given strategic measure.

For the new strategic measure, since it has a convex and compact support on the product of action spaces, using Lusin's theorem [54, Theorem 7.5.2], we can construct a strategic measure induced by continuous policies whose cost function is  $\varepsilon/2$ -close to the cost of bounded support strategic measure.

We can complete the proof by combining these two results.  $\square$

Since each  $\mathbb{Y}^i$  is a locally compact separable metric space, there exists an increasing sequence of compact subsets  $\{K_l^i\}$  such that  $K_l^i \subset \text{int } K_{l+1}^i$  and  $\mathbb{Y}^i = \bigcup_{l=1}^{\infty} K_l^i$  [3, Lemma 2.76], where  $\text{int } D$  denotes the interior of the set  $D$ .

Let  $d_i$  denote the metric on  $\mathbb{Y}^i$ . For each  $l \geq 1$ , let  $\mathbb{Y}_{l,n}^i := \{y_{i,1}, \dots, y_{i,i_{l,n}}\} \subset K_l^i$  be a  $1/n$ -net in  $K_l^i$ . Recall that if  $\mathbb{Y}_{l,n}^i$  is a  $1/n$ -net in  $K_l^i$ , then for any  $y \in K_l^i$  we have

$$\min_{z \in \mathbb{Y}_{l,n}^i} d_i(y, z) < \frac{1}{n}.$$

For each  $l$  and  $n$ , let  $q_{l,n}^i : K_l^i \rightarrow \mathbb{Y}_{l,n}^i$  be a nearest neighborhood quantizer given by

$$q_{l,n}^i(y) = \arg \min_{z \in \mathbb{Y}_{l,n}^i} d_i(y, z),$$

where ties are broken so that  $q_{l,n}^i$  is measurable. If  $K_l^i = [-M, M] \subset \mathbb{Y}^i = \mathbb{R}$  for some  $M \in \mathbb{R}_+$ , the finite set  $\mathbb{Y}_{l,n}^i$  can be chosen such that  $q_{l,n}^i$  becomes a uniform quantizer. We let  $Q_{l,n}^i : \mathbb{Y}^i \rightarrow \mathbb{Y}_{l,n}^i$  denote the extension of  $q_{l,n}^i$  to  $\mathbb{Y}^i$  given by

$$Q_{l,n}^i(y) := \begin{cases} q_{l,n}^i(y), & \text{if } y \in K_l^i, \\ y_{i,0}, & \text{otherwise,} \end{cases}$$

where  $y_{i,0} \in \mathbb{Y}^i$  is some auxiliary element.

Define  $\Gamma_{l,n}^i = \Gamma^i \circ Q_{l,n}^i \subset \Gamma^i$ ; that is,  $\Gamma_{l,n}^i$  is defined to be the set of all strategies  $\tilde{\gamma}^i \in \Gamma^i$  of the form  $\tilde{\gamma}^i = \gamma^i \circ Q_{l,n}^i$ , where  $\gamma^i \in \Gamma^i$ . Define also  $\mathbf{\Gamma}_{l,n} := \prod_{i=1}^N \Gamma_{l,n}^i \subset \mathbf{\Gamma}$ . Note that, for any  $i = 1, \dots, N$ ,  $\Gamma_{l,n}^i$  is the set of policies for DM  $i$  which can only use the output levels of the quantizer  $Q_{l,n}^i$ . In other words, in addition to the measurement channel  $g^i(dy^i|\omega_0)$  between DM  $i$  and the state of nature, there is also an analog-to-digital converter (quantizer) between them.

Using these definitions, let us define the following set of strategic measures for any  $l$  and  $n$ :

$$L_A^{l,n}(\mu)$$

$$\begin{aligned} &:= \left\{ P \in L_A(\mu) : P(B) \right. \\ &= \left. \int_{B^0 \times \prod_{k=1}^N A^k} \mu(d\omega_0, d\mathbf{y}) \prod_{k=1}^N 1_{\{u^k = \gamma^k(y^k) \in B^k\}}, \underline{\gamma} \in \Gamma_{l,n} \right\}. \end{aligned}$$

The following theorem states that an optimal (or almost optimal) strategic measure can be approximated with arbitrarily small approximation error for the induced costs by strategic measures in  $L_A^{l,n}(\mu)$  for sufficiently large  $l$  and  $n$ .

**Theorem 6.1.** [120] *For any  $\varepsilon > 0$ , there exist  $(l, n(l))$ , a compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$  of the form  $G = \prod_{i=1}^N G^i$ , where each  $G^i$  is convex and compact, and  $P \in L_A^{l,n(l)}(\mu) \cap L_A^G(\mu)$  such that*

$$\int P(ds) c(s) < J^* + \varepsilon$$

*Sketch of Proof.* Fix any strategic measure  $P \in L_A(\mu)$ . Then, by Proposition 6.1, there exists a strategic measure  $P_G \in L_A^{G,c}(\mu)$  for some compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$  of the form  $G = \prod_{i=1}^N G^i$ , where each  $G^i$  is convex and compact, such that

$$\int P_G(ds) c(s) \leq \int P(ds) c(s) + \varepsilon/2.$$

Let  $\underline{\gamma} \in \Gamma_{G,c}$  be the strategy that induces  $P_G$ . Then define  $\gamma_{l,n}^i = \gamma^i \circ Q_{l,n}^i$  for each  $i = 1, \dots, N$ . Note that on any compact subset of  $\mathbb{Y}^i$ , we have  $\gamma_{l,n}^i \rightarrow \gamma^i$  as  $\gamma^i$  is continuous. Let  $P_{l,n}$  the strategic measure induced by the strategy  $(\gamma_{l,n}^1, \dots, \gamma_{l,n}^N)$ .

Note that since  $\mu(d\omega_0, d\mathbf{y})$  is tight, for any  $\delta > 0$ , there exists a compact subset  $K$  of  $\prod_{k=1}^N \mathbb{Y}^k$  of the form  $K = \prod_{i=1}^N K^i$ , where each  $K^i$  is compact such that  $\mu(\Omega_0 \times K) \geq 1 - \delta$ . By choosing a sequence of  $\{\delta(l, n)\}$  and the corresponding compact sets  $\{K_{l,n}\}$  appropriately, we can prove that

$$\lim_{(l,n) \rightarrow \infty} \int P_{l,n}(ds) c(s) = \int P_G(ds) c(s),$$

as  $\gamma_{l,n}^i \rightarrow \gamma^i$  on each compact set  $K_{l,n}$  and  $\mu(\Omega_0 \times K_{l,n}) \geq 1 - \delta(l, n)$ . Hence, there exists  $(l, n(l))$  such that

$$\int P_{l,n(l)}(ds) c(s) \leq \int P_G(ds) c(s) + \varepsilon/2.$$

This completes the proof. □

The above result implies that to compute a near optimal strategic measure for the team problem it is sufficient to consider quantized measurements  $(Q_{l,n}^1(y^1), \dots, Q_{l,n}^N(y^N))$  for sufficiently large  $l$  and  $n$ . Furthermore, this nearly optimal strategic measure can have a compact support of the form  $G = \prod_{i=1}^N G^i$  on  $\prod_{k=1}^N \mathbb{U}^k$ , where  $G^i$  is convex and compact for each  $i = 1, \dots, N$ .

For ease of reference, we define

$$L_{A(q)}(\mu) = \bigcup_{(l,n)} L_A^{l,n}(\mu)$$

and we define  $L_{A(c)}(\mu)$  to be the subset of  $L_A(\mu)$  defined in (7) where the policies are restricted to be in  $\Gamma_c$ .

**6.1. Finite measurement approximate models**

In this section, for each  $(l, n)$ , we define a team model with finite measurement spaces. We prove that, for sufficiently large  $l$  and  $n$ , an optimal strategic measure of the team model corresponding to  $(l, n)$  will provide a strategic measure to the original team model which is nearly optimal.

To this end, fix any  $(l, n)$ . For the pair  $(l, n)$ , the corresponding finite measurement team model has the following measurement spaces:  $\mathbb{Z}_{l,n}^i := \{y_{i,0}, y_{i,1}, \dots, y_{i,i_{l,n}}\}$  (i.e., the output levels of  $Q_{l,n}^i$ ),  $i \in \mathcal{N}$ . The stochastic kernels  $g_{l,n}^i(\cdot | \omega_0)$  from  $\Omega_0$  to  $\mathbb{Z}_{l,n}^i$  denote the measurement constraints and are given by:

$$g_{l,n}^i(\cdot | \omega_0) := \sum_{j=0}^{i_{l,n}} g(S_{i,j}^{l,n} | \omega_0) \delta_{y_{i,j}}(\cdot),$$

where  $S_{i,j}^{l,n} := \{y \in \mathbb{Y}^i : Q_{l,n}^i(y) = y_{i,j}\}$ . Indeed,  $g_{l,n}^i(\cdot | \omega_0)$  is the push-forward of the measure  $g^i(\cdot | \omega_0)$  with respect to the quantizer  $Q_{l,n}^i$ .

Let  $\Phi_{n,l}^i := \{\phi^i : \mathbb{Z}_{l,n}^i \rightarrow \mathbb{U}^i, \phi^i \text{ measurable}\}$  denote the set of measurable policies for DM  $i$  and let  $\Phi_{l,n} := \prod_{i=1}^N \Phi_{l,n}^i$ . The cost of this team model is  $J_{l,n} : \Phi_{l,n} \rightarrow \mathbb{R}_+$  and defined as

$$J_{l,n}(\underline{\phi}) := \int_{\Omega_0 \times \prod_{i=1}^N \mathbb{Z}_{l,n}^i} c(\omega_0, \mathbf{y}, \mathbf{u}) P_{l,n}(d\omega_0, d\mathbf{y}),$$

where  $\underline{\phi} = (\phi^1, \dots, \phi^N)$ ,  $\mathbf{u} = \underline{\phi}(\mathbf{y})$ , and

$$P_{l,n}(d\omega_0, d\mathbf{y}) = P(d\omega_0) \prod_{i=1}^N g_{l,n}^i(dy^i | \omega_0) =: \mu_{l,n}(d\omega_0, d\mathbf{y}).$$

For any compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$ , we also define  $\Phi_{l,n}^G := \{\underline{\phi} \in \Phi_{l,n} : \underline{\phi}(\prod_{i=1}^N \mathbb{Z}_{l,n}^i) \subset G\}$ .

In order to obtain the approximation result, we need to impose the following additional assumption.

**Assumption 6.2.** For any compact subset  $G$  of  $\prod_{k=1}^N \mathbb{U}^k$  of the form  $G = \prod_{i=1}^N G^i$ , we assume that the function  $w_G$  is uniformly integrable with respect to the measures  $\{\mu_{l,n}\}$ ; that is,

$$\lim_{R \rightarrow \infty} \sup_{l,n} \int_{\{w_G > R\}} w_G(\omega_0, \mathbf{y}) d\mu_{l,n} = 0.$$

This assumption is quite mild and satisfied by the celebrated counterexample of Witsenhausen.

Let  $T_A(\mu_{l,n})$  denote the set of strategic measures induced by measurable policies for this team model. Let  $T_A^G(\mu_{l,n})$  be the strategic measures in  $T_A(\mu_{l,n})$  whose support on  $\prod_{k=1}^N \mathbb{U}^k$  is a subset  $G$ .

The following theorem is the main result of this section. It states that to compute a near optimal strategic measure for the original team problem, it is sufficient to compute an optimal strategic measure for the finite measurement team problem corresponding to sufficiently large  $l$  and  $n$ .

**Theorem 6.2.** [120] *Suppose Assumptions 6.1 and 6.2 hold. Then, for any  $\varepsilon > 0$ , there exists a pair  $(l, n(l))$  and a compact subset  $G = \prod_{i=1}^N G^i$  of  $\prod_{k=1}^N \mathbb{U}^k$  such that an optimal (or almost optimal) strategic measure  $P^{l,n(l)}$  in the set  $T_A^G(\mu_{l,n(l)})$  for the  $(l, n(l))$  team is  $\varepsilon$ -optimal for the original team problem when  $P^{l,n(l)}$  is extended to  $\Omega \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  via quantizers  $Q_{l,n(l)}^i$ ; that is,*

$$P_{\text{ex}}^{l,n(l)}(\cdot) = \int \mu(d\omega_0, d\mathbf{y}) \prod_{k=1}^N 1_{\{u^k = \gamma^k \circ Q_{l,n(l)}^k(y^k) \in \cdot\}}$$

where

$$P^{l,n(l)}(\cdot) = \int \mu_{l,n(l)}(d\omega_0, d\mathbf{y}) \prod_{k=1}^N 1_{\{u^k = \gamma^k(y^k) \in \cdot\}}$$

*Proof.* We first prove the following fact. Let  $\{\underline{\phi}_{l,n}\}$  be a sequence of strategies such that  $\underline{\phi}_{l,n} \in \Phi_{l,n}^G$ , where  $G = \prod_{i=1}^N G^i$  and each  $G^i$  is convex and compact. For each  $l$  and  $n$ , define  $\underline{\gamma}_{l,n} := \underline{\phi}_{l,n} \circ Q_{l,n}$ , where  $Q_{l,n} := (Q_{l,n}^1, \dots, Q_{l,n}^N)$ . Then, we have

$$\lim_{l,n \rightarrow \infty} |J_{l,n}(\underline{\phi}_{l,n}) - J(\underline{\gamma}_{l,n})| = 0.$$

Indeed, let us introduce the following finite measures on  $\Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k$ :

$$\begin{aligned} \mu_G(S) &:= \int_S w_G(\omega, \mathbf{y}) \, d\mu, \\ \mu_G^{l,n}(S) &:= \int_S w_G(\omega, \mathbf{y}) \, d\mu_{l,n}. \end{aligned}$$

Since  $\mu_{l,n}$  converges to  $\mu$  weakly, by [127, Theorem 3.5] and Assumption 6.2 we have  $\mu_G^{l,n} \rightarrow \mu_G$  weakly as  $l, n \rightarrow \infty$ . Hence, the sequence  $\{\mu_G^{l,n}\}$  is tight. Therefore, there exists a compact subset  $K$  of  $\Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k$  such that  $\mu_G(K^c) < \varepsilon/2$  and  $\mu_G^{l,n}(K^c) < \varepsilon/2$  for all  $l, n$ . Then, we have

$$\begin{aligned} &|J_{l,n}(\underline{\phi}_{l,n}) - J(\underline{\gamma}_{l,n})| \\ &= \left| \int_{\Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k} c(\omega_0, Q_{l,n}(\mathbf{y}), \underline{\gamma}_{l,n}(\mathbf{y})) \, d\mu - \int_{\Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k} c(\omega_0, \mathbf{y}, \underline{\gamma}_{l,n}(\mathbf{y})) \, d\mu \right| \end{aligned}$$

$$\begin{aligned} &\leq \int_K |c(\omega_0, Q_{l,n}(\mathbf{y}), \underline{\gamma}_{l,n}(\mathbf{y})) - c(\omega_0, \mathbf{y}, \underline{\gamma}_{l,n}(\mathbf{y}))| d\mu \\ &\quad + \int_{K^c} w_G(\omega_0, \mathbf{y}) d\mu + \int_{K^c} w_G(\omega_0, \mathbf{y}) d\mu_{l,n}. \end{aligned}$$

The first term in the last expression goes to zero as  $l, n \rightarrow \infty$  by the dominated convergence theorem and the fact that  $c$  is bounded and continuous in  $\mathbf{y}$ . The second and third terms are less than  $\varepsilon$  by tightness. Since  $\varepsilon$  is arbitrary, this completes the proof of the fact. Note that this fact implies that for any  $P^{l,n} \in T_A^G(\mu_{l,n})$ , we have

$$\lim_{l,n \rightarrow \infty} \left| \int P^{l,n}(ds) c(s) - \int P_{\text{ex}}^{l,n}(ds) c(s) \right| = 0.$$

Note that using this fact, it is straightforward to prove that

$$\liminf_{l,n \rightarrow \infty} \inf_{P \in T_A^G(\mu_{l,n})} \int P(ds) c(s) \geq J^*.$$

Since the extension of the strategic measures  $T_A^G(\mu_{l,n})$  to  $\Omega_0 \times \prod_{k=1}^N (\mathbb{Y}^k \times \mathbb{U}^k)$  forms the set of strategic measures  $L_A^{l,n}(\mu) \cap L_A^G(\mu)$ , by Theorem 6.1, we can complete the proof. Indeed, let  $P_{\text{ex}}^{l,n(l)} \in L_A^{l,n}(\mu) \cap L_A^G(\mu)$  be the  $\varepsilon$  optimal policy for the original problem. Then we have

$$\begin{aligned} J^* + \varepsilon &\geq \limsup_{l,n \rightarrow \infty} \int P_{\text{ex}}^{l,n(l)}(ds) c(s) \\ &= \limsup_{l,n \rightarrow \infty} \int P^{l,n(l)}(ds) c(s) \\ &\geq \limsup_{l,n \rightarrow \infty} \inf_{P \in T_A^G(\mu_{l,n})} \int P(ds) c(s). \end{aligned}$$

Hence

$$J^* + \varepsilon \geq \limsup_{l,n \rightarrow \infty} \inf_{P \in T_A^G(\mu_{l,n})} \int P(ds) c(s) \geq \liminf_{l,n \rightarrow \infty} \inf_{P \in T_A^G(\mu_{l,n})} \int P(ds) c(s) \geq J^*.$$

This completes the proof. □

We interpret this result as saying that the space of information structures obtained by quantizing the original information structure is approximately optimal: denseness under weak convergence implies denseness of strategic measures as far as optimality is concerned.

**Remark 6.1.** *Results proved in this section about approximation of static team problems can be directly applied to static reductions of dynamic teams if static reductions satisfy Assumption 6.1 and Assumption 6.2. In particular, the above approximation results hold for Witsenhausen’s counterexample.*

## 7. Product topology induced by individual policies and relations with information structures

Instead of strategic measures, one may choose to directly induce topologies on control policies or strategies alone. We will see that such an approach will let us arrive at complementary conditions compared with what we have studied earlier. In the following, we first revisit some classical results in optimal control theory via this approach.

### 7.1. Some remarks on classical deterministic and stochastic control and Young measures

It is instructive to revisit various control topologies that are already well-known in classical control theory (when there is a single controller who has access to the state variable). In deterministic nonlinear, geometric, and continuous-time control, properties on stabilizability, controllability, and reachability are drastically impacted by the restrictions on the classes of allowed controls (e.g., continuous, Lipschitz, finitely differentiable, or smooth control functions in the state or time when control is open-loop [118, 40, 129, 80]) and naturally the control topology induced is dictated by the class of admissible controls.

For optimal control, to allow for continuity/compactness arguments, a priori imposing compactness over spaces of measurable functions would be an artificial restriction, and the use of powerful theorems such as the Arzela-Ascoli theorem which necessarily entail (usually very restrictive and suboptimal) conditions on continuity properties of the considered policies. In deterministic optimal control theory, relaxed controls [146] [134] allow for this machinery to be applied with no artificial restrictions on the classes of control policies considered; these are known as Young measures.

Let us consider an open-loop controller, where the control is only a function of the time variable. We let  $\nu(dt, du)$  be a measure on  $[0, T] \times \mathbb{U}$  where the first marginal  $\lambda(dt)$  is the normalized Lebesgue measure on time interval  $[0, T]$  and let  $\nu(du|t) = 1_{\{\gamma(t) \in du\}}$  be the conditional measure induced by deterministic open loop control. So, any deterministic open-loop control is embedded via:

$$\nu(dt, du) = \lambda(dt) 1_{\{\gamma(t) \in du\}}.$$

If we allow for randomized policies, we obtain the set  $\mathcal{P}_\lambda([0, T] \times \mathbb{U})$  of all probability measures with fixed marginal on  $[0, T]$ . This set is weakly closed, whose extreme points are those induced by deterministic policies (as also was discussed in the context of (4)). Thus, any deterministic optimal control problem, which can be written in an integral form and has a lower semi-continuous cost functions in actions, will have an optimal solution, which will then be deterministic as these form the extreme points of randomized controls. It can also in fact be shown that such policies are dense in the space of randomized policies, in addition to these policies forming the extreme points in the set of randomized policies (see e.g., [24, Proposition 2.2] [92], [105, 19, Theorem 3], but also many

texts in optimal stochastic control where denseness of deterministic controls has been established inside the set of relaxed controls [33]). We refer the reader to [103] (see also the review paper [92]) for further discussion.

The following example builds on these, with somewhat different arguments. Let  $\mathbb{X} = \mathbb{R}, \mathbb{U} = [0, 1]$ , and let  $f : \mathbb{X} \times \mathbb{U} \rightarrow [0, 1]$  and  $c : \mathbb{X} \times \mathbb{U} \rightarrow [0, 1]$  be measurable functions continuous in the control action variable. Consider the following optimal control problem:

$$\inf_{\substack{\gamma: \mathbb{X} \rightarrow \mathbb{U} \\ u_t = \gamma(x_t)}} \int_0^1 c(x_t, u_t) \lambda(dt) \tag{22}$$

subject to

$$\frac{dx}{dt} = f(x_t, u_t) \tag{23}$$

The natural space to consider is the set of all control functions which depends on the current state, where the only restriction is measurability. However, allowing for measurability only does not facilitate continuity/compactness arguments since, as noted above, imposing compactness on a space of functions is an unnecessarily restrictive condition. Accordingly, one often cites appropriate but tedious measurable selection theorems building on optimality equations through dynamic programming.

On the other hand, every deterministic function of the state can be expressed as a deterministic function of time, and so, be considered open-loop. Accordingly, we consider open loop controls and those which are relaxed. Let  $\mathcal{P}_\lambda([0, T] \times \mathbb{U})$  be the set of relaxed open loop policies (Young measures). Now consider the space  $C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$ , where  $C([0, 1]; \mathbb{X})$  is the space of continuous functions from  $[0, 1]$  to  $\mathbb{X}$ . We endow this space with the product topology with the first component being under the supremum norm and the second under the Prohorov metric (or any weak convergence inducing metric). Note now that the cost (22) is continuous on  $C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$ . Note that since  $f$  is uniformly bounded, we have that the set  $\mathcal{A}$  of all admissible sample paths of the state  $x : [0, 1] \rightarrow \mathbb{X}$  is equicontinuous, and so, by the Arzela-Ascoli theorem,  $\mathcal{A}$  is relatively compact in  $C([0, 1]; \mathbb{X})$ . Accordingly, our space of interest  $\mathcal{A} \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$  is a relatively compact subset of  $C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$ .

Define now

$$H = \left\{ (x, m) \in C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U}) : x_t - \int_0^t f(x_s, u) m_s(du) \lambda(ds) = 0 \right\},$$

where  $m_s(du) = m(du|s)$ . This set is closed under the topology defined on  $C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$  and is a subset of  $C([0, 1]; \mathbb{X}) \times \mathcal{P}_\lambda([0, T] \times \mathbb{U})$ . Hence,  $H$  is compact. Now, the problem then is to find an optimal  $(x, m) \in H$  which minimizes (22), reformulated as:

$$\inf_{(x, m) \in H} \int_0^1 c(x_t, u) m(dt, du)$$



This is continuous in  $(x, m)$  by an application of the generalized weak convergence theorem under continuous convergence [126, Theorem 3.5] or [95, Theorem 3.5]. Therefore, there exists an optimal solution to the problem.

Note that as has been reported in various literatures (e.g. in optimal quantization [150, p. 878]) that the set of deterministic controls is not weakly closed (or setwise) under Young's relaxation. In fact, there exist problems where an optimal solution exists among relaxed controls but not in deterministic controls [92].

On the other hand, in the continuous-time stochastic context, the analysis can be quite subtle due to the fact that the control policy (only restricted to be measurable in general) may violate conditions needed for the existence of strong solutions for a given stochastic differential equation since the control policy may couple the state dynamics with the past in an arbitrarily complicated, though measurable, way and hence violating the existence conditions for strong solutions to stochastic differential equations [98] [90] (e.g., in [143] Lipschitz continuity conditions are imposed, where Lipschitz property holds in the control when viewed as a map from the normed linear space of continuous functions on measurements to control actions). To avoid such technical issues on strong solutions, relaxed solution concepts were introduced and studied in the literature based on the measure transformation technique due to Girsanov [26, 50, 51] which allows the control to be a function of an independent Brownian innovations process. These approaches require strong absolute continuity conditions on the measurement process which may not be always applicable. See [92] for a detailed analysis for controlled stochastic differential systems. As we will see later in the paper, in the partially observed setup, a further interpretation of relaxed controls, called *wide sense admissible controls* [62, 36, 38] has been utilized to arrive at existence results on optimal control policies.

As we will see in the following, adding more general, non-classical, information structures will entail further intricacies but also facilitate additional creativity for the analysis in optimal stochastic control.

### 7.2. Trading-off continuity with compactness: a product space approach for individual control policies

To make the ideas in this section more explicit, consider the set of individually randomized strategic measures  $L_R(\mu)$  studied in Section 3.2. In this set, a policy of a DM  $k$  is a stochastic kernel  $\Pi^k(du^k|y^k)$  from  $\mathbb{Y}^k$  to  $\mathbb{U}^k$ . Instead of studying continuity and compactness properties regarding strategic measures (joint probability measures on the state, measurement and action spaces), we can focus on individual control policies and measurements, and their Cartesian products. Let us denote, as in Section 3.1,

$$\Theta^i = \{v \in \mathcal{P}(\mathbb{Y}^i \times \mathbb{U}^i) : v(dy^i \times \mathbb{U}^i) = \xi^i(dy^i)\}$$

to be the distributional strategy for DM  $i$ , which is the joint probability measure with a fixed marginal  $\xi^i$  on the measurement variable  $y^i$ . Note that if the problem

is an independent static reduction of a dynamic team, then one can choose  $\xi^i$  as the reference measure  $Q^i$  in the static reduction.

By our earlier analysis, we know that  $\Theta^i$  (a subset of  $\Theta^i$  that is sufficient for optimality) is compact under the weak topology or the  $w$ - $s$  topology if  $\mathbb{U}^i$  is compact (if this subset satisfies some moment condition). Hence, since the marginal on measurements are fixed, we can view the weak topology or  $w$ - $s$  topology on  $\Theta^i$  as a topology on the set of policies  $\Gamma^i$ . In particular, any sequence of policies:

$$\underline{\gamma}_n := (\gamma_n^1, \dots, \gamma_n^N)$$

will have a converging subsequence,  $\underline{\gamma}_{n_k}$  to a limit  $\underline{\gamma}$ , and therefore, closedness (in particular compactness) is not an issue under this topology. In other words, the IS is always preserved under this topology, which is not the case in the strategic measure approach as it has been shown in Theorem 4.2.

Note that if the measurements are independent, then strong existence results can be obtained using this approach. This is precisely the same condition obtained via the strategic measures approach studied in Theorem 4.5.

On the other hand, what is not clear is, in general, whether we have

$$\lim_{n \rightarrow \infty} J(\underline{\gamma}_n) = J(\underline{\gamma})$$

under this topology (this is always true when we place a topology on the set of strategic measures). In particular, if we have a general  $\mu(d\omega_0, d\mathbf{y})$ , it is not clear if we have this *joint continuity* condition. The following example will demonstrate this subtlety through a negative implication. Therefore, there is a trade-off between continuity and compactness in these approaches: The strategic measure approach provides continuity but may not preserve the IS (so compactness); and conversely, the approach via topology on individual policies leads to a compact strategy space but loses continuity.

**Example 7.1.** Consider the counterexample presented in Theorem 4.2. By considering each individual decision makers' policy separately, we have that each will converge individually; that is,

$$P_n(dy^1, du^1) \rightarrow P(dy^1, du^1), \quad P_n(dy^2, du^2) \rightarrow P(dy^2, du^2).$$

Now, consider a cost function  $c(\omega_0, u^1, u^2) = (u^1 - u^2)^2$ . Note that, for each  $n$ , under  $P_n$ , actions  $u^1$  and  $u^2$  are the same (i.e.,  $u^1 = u^2$ ) given  $y^1 = y^2 = y$ , and so,  $J(\underline{\gamma}_n) = \int P_n(ds) c(s) = 0$  for every  $n$ . However, in the limit where  $u^i$  and  $y = y^i$  ( $i = 1, 2$ ) are independent under  $P$ , we have  $J(\underline{\gamma}_n) = \int P(ds) c(s) = 1/2$ . This is a consequence of viewing individual policies separately without their joint convergence properties. Thus, we do not even have lower semi-continuity. Hence, while we have established compactness, we have lost lower semi-continuity by placing the topologies directly on control policies and not on the strategic measures.

**Remark 7.1** (Hilbert space construction for static quadratic Gaussian teams). For static Linear Quadratic Gaussian (LQG) team problems, it is instructive to

note that [149, Theorem 2.6.6] (building on [117]) utilizes a Hilbert space theoretic construction of control topologies on individual control policies (and on the team policy via a product space construction). Via this topology on the controls, the optimality of affine solutions was established via the Projection theorem. In view of the discussion presented above (on the trade-off between continuity and compactness), it is interesting to note that the Hilbert space topology simultaneously leads to both closedness and continuity, though the setup is restricted only to quadratic cost criteria.

### 7.3. Policies defined by conditional independence given measurements

[148] considered the following topology on control policies, while developing a universal dynamic programming algorithm applicable to any sequential decentralized stochastic control problem, generalizing Witsenhausen's program [140] which was tailored primarily for countable probability spaces.

Define

- (i) State:  $x_t = \{\omega_0, u^1, \dots, u^{t-1}, y^1, \dots, y^t\}$ ,  $1 \leq t \leq N$ .
- (i') Extended State:  $\pi_t \in \mathcal{P}(\Omega_0 \times \prod_{i=1}^t \mathbb{Y}^i \times \prod_{i=1}^{t-1} \mathbb{U}^i)$  where, for Borel  $B \in \Omega_0 \times \prod_{i=1}^t \mathbb{Y}^i \times \prod_{i=1}^{t-1} \mathbb{U}^i$ ,

$$\pi_t(B) := E_{\pi_t}[1_{\{(\omega_0, y^1, \dots, y^t; u^1, \dots, u^{t-1}) \in B\}}].$$

Thus,  $\pi_t \in \mathcal{P}(\Omega_0 \times \prod_{i=1}^t \mathbb{Y}^i \times \prod_{i=1}^{t-1} \mathbb{U}^i)$  where the space of probability measures is endowed with the weak convergence topology.

- (ii) Control Action: Given  $\pi_t$ ,  $\hat{\gamma}^t$  is a probability measure in  $\mathcal{P}(\Omega_0 \times \prod_{k=1}^t \mathbb{Y}^k \times \prod_{k=1}^t \mathbb{U}^k)$  that satisfies the conditional independence relation:

$$u^t \leftrightarrow y^t \leftrightarrow x_t = (\omega_0, y^1, \dots, y^t; u^1, \dots, u^{t-1})$$

(that is, for every Borel  $B \in \mathbb{U}^t$ , almost surely under  $\hat{\gamma}^t$ , the following holds:

$$P(u^t \in B | y^t, (\omega_0, y^1, \dots, y^t; u^1, \dots, u^{t-1})) = P(u^t \in B | y^t)$$

with the restriction

$$x_t \sim \pi_t.$$

Denote by  $\Gamma^t(\pi_t)$  the set of all such probability measures. Any  $\hat{\gamma}^t \in \Gamma^t(\pi_t)$  defines, for almost every realization  $y^t$ , a conditional probability measure on  $\mathbb{U}^t$ . When the notation does not lead to confusion, we will denote the action at time  $t$  by  $\gamma^t(du^t | y^t)$ , which is understood to be consistent with  $\hat{\gamma}^t$ .

- (ii') Alternative Control Action for Static Teams with Independent Measurements: Given  $\pi_t$ ,  $\hat{\gamma}^t$  is a probability measure on  $\mathbb{Y}^t \times \mathbb{U}^t$  with a fixed marginal  $P(dy^t)$  on  $\mathbb{Y}^t$ , that is  $\pi_t^{\mathbb{Y}^t}(dy^t) = P(dy^t)$ . Denote by  $\Gamma^t(\pi_t^{\mathbb{Y}^t})$  the

set of all such probability measures. As above, when the notation does not lead to confusion, we will denote the action at time  $t$  by  $\gamma^t(du^t|y^t)$ , which is understood to be consistent with  $\hat{\gamma}^t$ . In particular,  $(y^t, u^t)$  is independent of  $(y^k, u^k)$  for  $k \neq t$ .

With the control actions defined as in the above [148] developed a universal dynamic program for any sequential decentralized stochastic control and established, as a corollary of the program, further existence results, one of which is essentially identical to that presented in Theorem 4.5, but slightly more restrictive in that the cost function is assumed to be continuous in all of its arguments.

**Theorem 7.1.** [148]

- (i) Under the kernel (8) and controlled Markov construction presented, the optimal team problem admits a well-defined backwards-induction (dynamic programming) recursion.
- (ii) In particular, if the problem is independent static-reducible, actions are compact-valued and the cost function is continuous, an optimal policy exists and the value function is continuous in the prior (that is, in the distribution of primitive noise variables) under weak convergence.

**Remark 7.2.** The above construction is related to an interpretation put forward by Witsenhausen in his standard form [140] where all the uncertainty is embedded into the initial state and the controlled system evolves deterministically. Witsenhausen had considered only countable probability spaces for an optimality analysis.

**7.4. Weak-\* topology on randomized policies**

In this section, inspired by the topology on policies for classical stochastic control problems constructed in Borkar [34] (see also [9]), we study a topology on the set of randomized policies for static teams or dynamic teams that admit static reduction introduced in [119]. With an abuse of notation, we denote the set of randomized policies for DM  $i$  by  $\Gamma^i$ , which can be written as

$$\Gamma^i = \left\{ \gamma^i : \gamma^i \text{ is a measurable function from } \mathbb{Y}^i \text{ to } \mathcal{P}(\mathbb{U}^i) \right\},$$

where  $\mathcal{P}(\mathbb{U}^i)$  is endowed with the Borel  $\sigma$ -algebra generated by the weak convergence topology. We will use the latter characterization of  $\Gamma^i$  when introducing the topology.

For a metric space  $\mathbb{E}$ , let  $C_0(\mathbb{E})$  be the Banach space of all continuous real functions on  $\mathbb{E}$  vanishing at infinity endowed with the norm

$$\|g\|_\infty = \sup_{e \in \mathbb{E}} |g(e)|.$$

Recall that  $\mathcal{M}(\mathbb{E})$  and  $\mathcal{P}(\mathbb{E})$  denote the set of all finite signed measures and probability measures on  $\mathbb{E}$ . Let  $\mathbb{E}_1$  and  $\mathbb{E}_2$  be two metric spaces. For any  $\mu \in$

$\mathcal{M}(\mathbb{E}_1 \times \mathbb{E}_2)$ ,  $\text{Proj}_{\mathbb{E}_1}(\mu)(\cdot) = \mu(\cdot \times \mathbb{E}_2)$  be the marginal of  $\mu$  on  $\mathbb{E}_1$ . For a Banach space  $\mathbb{K}$ , let  $\mathbb{K}^*$  denote its topological dual. For any  $i$ , we now introduce a topology on  $\Gamma^i$ . To this end, we impose the following assumption.

**Assumption 7.1.** (a)  $\Omega_0$ ,  $\mathbb{Y}^i$ , and  $\mathbb{U}^i$  ( $i = 1, \dots, N$ ) are locally compact.  
 (b) For all  $i$ , the measurement channel  $g^i(dy_i|\omega_0) = g^i(y_i, \omega_0) \mu^i(dy^i)$  for some probability measure  $\mu^i$  on  $\mathbb{Y}^i$ ; that is,  $g^i(dy_i|\omega_0)$  is absolutely continuous with respect to  $\mu^i(dy^i)$  for all  $\omega_0$ .

Note that by Riesz representation theorem [63, Theorem 7.17], the topological dual of  $(C_0(\mathbb{U}^i), \|\cdot\|_\infty)$  is  $(\mathcal{M}(\mathbb{U}^i), \|\cdot\|_{TV})$ , where  $\|\cdot\|_{TV}$  is the total variation norm on  $\mathcal{M}(\mathbb{U}^i)$ .

We now define the set of  $w^*$ -measurable functions from  $\mathbb{Y}^i$  to  $\mathcal{M}(\mathbb{U}^i)$ . Later, we will prove that the set of randomized policies  $\Gamma^i$  for DM  $i$  is a bounded subset of this set. A function  $\gamma : \mathbb{Y}^i \rightarrow \mathcal{M}(\mathbb{U}^i)$  is called  $w^*$ -measurable if the mapping

$$\mathbb{Y}^i \ni y \mapsto \int d\gamma(y)(du) g(u) \in \mathbb{R}$$

is measurable for all  $g \in C_0(\mathbb{U}^i)$  [43, p. 18]. Let  $\mathcal{L}(\mu^i, \mathcal{M}(\mathbb{U}^i))$  denote the set of all such functions. With this definition, we now define the following set

$$\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i)) := \left\{ \gamma \in \mathcal{L}(\mu^i, \mathcal{M}(\mathbb{U}^i)) : \|\gamma\|_\infty := \text{ess sup}_{y \in \mathbb{Y}^i} \|\gamma(y)\|_{TV} < \infty \right\},$$

where  $\text{ess sup}$  is taken with respect to the measure  $\mu^i$ . Recall that  $\mu_i$  is the reference probability measure in Assumption 7.1-(b) for the measurement channel  $g^i$ .

Let  $L_1(\mu^i, C_0(\mathbb{U}^i))$  denote the set of all Bochner-integrable [52] functions from  $\mathbb{Y}^i$  to  $C_0(\mathbb{U}^i)$  endowed with the norm

$$\|f\|_1 := \int_{\mathbb{Y}^i} \|f(y)\|_\infty \mu^i(dy^i).$$

By using the fact that  $C_0(\mathbb{U}^i)^* = \mathcal{M}(\mathbb{U}^i)$ , one can prove that the topological dual of  $(L_1(\mu^i, C_0(\mathbb{U}^i)), \|\cdot\|_1)$  can be identified with  $(\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i)), \|\cdot\|_\infty)$  [43, Theorem 1.5.5, p. 27]; that is,

$$L_1(\mu^i, C_0(\mathbb{U}^i))^* = \mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i)).$$

Using this identification, we equip  $\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i))$  with the  $w^*$ -topology induced by  $L_1(\mu^i, C_0(\mathbb{U}^i))$ . We write  $\gamma_n \xrightarrow{*} \gamma$ , if  $\gamma_n$  converges to  $\gamma$  in  $\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i))$  with respect to the  $w^*$ -topology. As noted earlier, for this topology, we have been inspired by the topology introduced in [9, 34], where in these works, a similar topology is introduced for randomized Markov policies to study continuous-time stochastic control problems with average cost optimality criterion.

Note that we can identify the set of randomized policies  $\Gamma^i$  as a bounded subset of  $\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i))$ :

$$\Gamma^i = \left\{ \gamma \in \mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i)) : \gamma(y) \in \mathcal{P}(\mathbb{U}^i) \mu^i - \text{a.e.} \right\}.$$

Hence, we can equip  $\Gamma^i$  with the relative  $w^*$ -topology inherited by the  $w^*$ -topology on  $\mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i))$ .

Now, we derive some properties of this topology on  $\Gamma^i$ . Firstly, since  $\mathcal{P}(\mathbb{U}^i)$  is bounded in  $\mathcal{M}(\mathbb{U}^i)$ , by the Banach-Alaoglu Theorem [63, Theorem 5.18],  $\Gamma^i$  is relatively compact with respect to  $w^*$ -topology. Since  $C_0(\mathbb{U}^i)$  is separable, then by [71, Lemma 1.3.2],  $\Gamma^i$  is also relatively sequentially compact. However,  $\Gamma^i$  is not closed with respect to the  $w^*$ -topology. Indeed, let  $\mathbb{Y}^i = \mathbb{U}^i = \mathbb{R}$ . Define  $\gamma_n(y^i)(\cdot) := \delta_n(\cdot)$  and  $\gamma(y^i)(\cdot) := 0(\cdot)$ , where  $\delta_n$  denotes Dirac-delta measure at point  $n$  and  $0(\cdot)$  denotes the degenerate measure on  $\mathbb{U}^i$ ; that is,  $0(D) = 0$  for all  $D \in \mathcal{B}(\mathbb{R})$ . Let  $g \in L_1(\mu^i, C_0(\mathbb{U}^i))$ . Then we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\mathbb{Y}^i} \left( \int_{\mathbb{U}^i} \gamma_n(y)(du) g(y)(u) \right) \mu^i(dy) &= \lim_{n \rightarrow \infty} \int_{\mathbb{Y}^i} g(y)(n) \mu^i(dy) \\ &= \int_{\mathbb{Y}^i} \lim_{n \rightarrow \infty} g(y)(n) \mu^i(dy) \text{ (as } \|g(y)\|_\infty \text{ is } \mu^i\text{-integrable)} \\ &= 0 \text{ (as } g(y) \in C_0(\mathbb{U}^i)\text{)}. \end{aligned}$$

Hence,  $\gamma_n \rightharpoonup^* \gamma$ . But,  $\gamma \notin \Gamma^i$ , and so,  $\Gamma^i$  is not closed.

Thus, as opposed to the topology introduced in Section 7 on randomized policies, here the set of policies is not closed under the  $w^*$ -topology; that is, the IS may not be preserved. However, as we will see in the sequel, under this topology, the cost  $J$  is lower semi-continuous, which is in general not the case for the topology introduced in Section 7. However, by imposing additional assumptions on the components of the problem, we may first ensure the closedness of the set of policies or subset of it that is sufficient for optimality under this topology using lower semi-continuity of the cost. Then, we can prove the existence of an optimal policy. This result will be complementary to the existence results presented in Section 4.

In the remainder of this section,  $\Gamma^i$  is equipped with this topology. In addition,  $\mathbf{\Gamma}$  has the product topology induced by these  $w^*$ -topologies; that is,  $\underline{\gamma}_n$  converges to  $\underline{\gamma}$  in  $\mathbf{\Gamma}$  with respect to the product topology if and only if  $\gamma_n^i \rightharpoonup^* \gamma^i$  for all  $i = 1, \dots, N$ . In this case we write  $\underline{\gamma}_n \rightharpoonup^* \underline{\gamma}$ . Note that  $\mathbf{\Gamma}$  is sequentially relatively compact under this topology.

To this end, for any  $L > 0$ , we define

$$\mathbf{\Gamma}_L := \left\{ \underline{\gamma} \in \mathbf{\Gamma} : J(\underline{\gamma}) < J^* + L \right\}$$

and

$$\begin{aligned} S_L := \left\{ \lambda \in \mathcal{P} \left( \Omega_0 \times \prod_{i=1}^N (\mathbb{Y}^i \times \mathbb{U}^i) \right) : \right. \\ \left. \lambda(d\omega_0, d\mathbf{y}, d\mathbf{u}) = \mu(d\omega_0) \prod_{i=1}^N \gamma^i(du^i | y^i) \mu^i(dy^i), \underline{\gamma} \in \mathbf{\Gamma}_L \right\}. \end{aligned}$$

For each  $i = 1, \dots, N$ , we define  $S_L^i := \text{Proj}_{\mathbb{Y}^i \times \mathbb{U}^i}(S_L)$ . In order to prove that  $\Gamma_L$  is closed with respect to the  $w^*$ -topology, we should impose the following assumption.

**Assumption 7.2.** *For some  $L > 0$ ,  $S_L^i$  is tight for  $i = 1, \dots, N$ .*

We provide several conditions that imply Assumption 7.2.

**Theorem 7.2** ([119, Theorem 4]). *Suppose either of the following conditions hold:*

- (i)  $\mathbb{U}^i$  is compact for all  $i$ .
- (ii) For the non-compact case, we assume
  - (a) The cost function  $c$  satisfies the following condition: for each  $j$ , for any  $M \geq 0$ , and for any compact  $K \subset \Omega_0 \times \prod_{k=1}^N \mathbb{Y}^k \times \prod_{k=1}^{j-1} \mathbb{U}^k$ , there exists a compact  $L \subset \mathbb{U}^j$  such that

$$\inf_{K \times L^c \times \prod_{k=j+1}^N \mathbb{U}^k} c(\omega_0, \mathbf{y}, \mathbf{u}) \geq M.$$

- (b) For all  $j$ ,  $g^j > 0$  and  $g^j$  is lower semi-continuous.

Then, Assumption 7.2 holds.

The following theorem establishes the existence of an optimal team decision rule using the  $w^*$ -topology.

**Theorem 7.3** ([119, Theorem 6]). *Suppose Assumptions 7.1 and 7.2 hold. Moreover, we assume that  $c$  is lower semi-continuous and measurement channels  $g^i(dy^i|\omega_0)$  are continuous with respect to the total variation distance. Then, there exists  $\underline{\gamma}^* \in \Gamma_L$  which is optimal.*

*Sketch of Proof.* Note that when the cost function  $c$  is a compactly supported continuous function, by the Stone-Weierstrass Theorem [94, Lemma 6.1],  $c$  can be uniformly approximated by functions of the form

$$\sum_{j=1}^k r_j \prod_{i=1}^N f_{j,i} g_{j,i}, \tag{24}$$

where  $r_j \in C_0(\Omega_0)$ ,  $f_{j,i} \in C_0(\mathbb{Y}^i)$ , and  $g_{j,i} \in C_0(\mathbb{U}^i)$  for each  $j = 1, \dots, k$  and  $i = 1, \dots, N$ , are compactly supported. One can prove that the cost  $J$  is continuous with respect to the  $w^*$ -topology if  $c$  is of the form in (24) [119, Theorem 5]. Since any compactly supported function can be uniformly approximated by such functions,  $J$  is also continuous if  $c$  is compactly supported. Note that if  $c$  is lower semi-continuous, then  $c$  can be approximated pointwise from below by compactly supported functions (see the proof of [71, Proposition 1.4.18]). Hence, if  $c$  is lower semi-continuous, then  $J$  is also lower semi-continuous with respect to the  $w^*$ -topology by the monotone convergence theorem.

Now, we prove that  $\Gamma_L$  is closed under the  $w^*$ -topology, which will complete the proof since  $\Gamma_L$  is relatively compact. To this end, let

$$\underline{\gamma}_n \rightharpoonup^* \underline{\gamma},$$

for some  $\underline{\gamma} \in \prod_{i=1}^N \mathcal{L}_\infty(\mu^i, \mathcal{M}(\mathbb{U}^i))$  (recall that  $\Gamma$  is not closed with respect to the  $w^*$ -topology). If  $\gamma^i \in \mathcal{P}(\mathbb{U}^i)$   $\mu^i$ -a.e. for all  $i$ , then  $\underline{\gamma}$  is in  $\Gamma_L$ , and so,  $\Gamma_L$  is closed.

Fix any  $i$ . Note that the sequence  $\{\gamma_n^i \otimes \mu^i\}$  is tight as it is a subset of  $S_L^i$ . Thus, there exists a further subsequence, denoted for simplicity by  $\{\gamma_i^i \otimes \mu^i\}$ , that converges weakly to some  $\lambda \in \mathcal{P}(\mathbb{Y}^i \times \mathbb{U}^i)$ . This implies that  $\gamma^i \otimes \mu^i = \lambda$ , and so,  $\gamma^i \otimes \mu^i(\mathbb{Y}^i \times \mathbb{U}^i) = 1$ . Hence,  $\gamma^i \in \mathcal{P}(\mathbb{U}^i)$   $\mu^i$ -a.e. Thus,  $\Gamma_L$  is closed. This completes the proof.  $\square$

**Remark 7.3.** *Weak topology has been used in a similar context in [87, 88]; [87, Lemma 4.9] and [88, Lemma A.1] can be exploited to prove the existence of optimal policies. Indeed, if one can prove that actions are square integrable under any strategy, the set of actions is weakly relatively compact, and the cost function is lower semi-continuous, then the existence of optimal policies can be deduced by the Weierstrass Extreme Value Theorem since [87, Lemma 4.9] and [88, Lemma A.1] guarantees that the limiting actions preserve the information structure of the problem.*

**7.5. Exchangeability, infinite products of individual policies and optimality of symmetric randomized policies for mean-field teams**

As a final example on the utility of placing a product topology on individual policies, we consider stochastic team problems with infinitely many decision makers. Such problems have seen a significant activity in the context of mean field theory [78, 77, 96] (see also more recent papers [61, 29, 8, 91]) and in mean-field team problems [79, 133][7] [101] [122]. In the context of mean-field team problems [101] and [122] have shown that, under sufficient convexity conditions, a sequence of optimal policies for teams with  $N$  number of decision makers as  $N \rightarrow \infty$  converges to a team optimal policy for the static team with a countably infinite number of decision makers, where the latter establishes the optimality of symmetric (i.e., identical for each DM) policies as well as existence of optimal team policies for both finite and infinite DM setups. In the following, we demonstrate how the control topology approach can be utilized to relax such strong convexity conditions.

Consider the following:

**Problem ( $\mathcal{P}_N$ ):** Let  $\mathcal{N} = \{1, \dots, N\}$ . Let  $\underline{\gamma}_N = (\gamma^1, \dots, \gamma^N)$  and  $\Gamma_N = \prod_{i=1}^N \Gamma^i$ . Define an expected cost function of  $\underline{\gamma}_N$  as

$$J_N(\underline{\gamma}_N) = \mathbb{E}^{\underline{\gamma}_N} [c(\omega_0, \underline{y}_N)] := \mathbb{E}[c(\omega_0, \gamma^1(y^1), \dots, \gamma^N(y^N))], \quad (25)$$



for some Borel measurable cost function  $c : \Omega_0 \times \prod_{k=1}^N \mathbb{U}^k \rightarrow \mathbb{R}_+$ . We define  $\omega_0$  to be the  $\Omega_0$ -valued cost function relevant exogenous random variable as  $\omega_0 : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\Omega_0, \mathcal{F}_0)$ , where  $\Omega_0$  is a Borel space with its Borel  $\sigma$ -field  $\mathcal{F}_0$ . Here, we have the notation  $\underline{u}_N := \{u^i, i \in \mathcal{N}\}$ .

**Problem ( $\mathcal{P}_\infty$ ):** Consider a stochastic team with a countably infinite number of decision makers, that is,  $\mathcal{N} = \mathbb{N}$ . Let  $\Gamma = \prod_{i \in \mathbb{N}} \Gamma^i$  and  $\underline{\gamma} = (\gamma^1, \gamma^2, \dots)$ . Let  $c : \Omega_0 \times \mathbb{U} \times \mathbb{U} \rightarrow \mathbb{R}_+$ . Define the expected cost of  $\underline{\gamma}$  as

$$J(\underline{\gamma}) = \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}^\gamma \left[ \sum_{i=1}^N c \left( \omega_0, u^i, \frac{1}{N} \sum_{p=1}^N u^p \right) \right]. \quad (26)$$

With a slight abuse of notation, we use the same notation for the cost function  $c$  as in (25).

**Assumption 7.3.** Assume for any  $DM^i$ , there exists a probability measure  $Q^i$  on  $\mathbb{Y}^i$  and a function  $f^i$  such that for all Borel sets  $S = S^1 \times \dots \times S^N$ , we have

$$\mathbb{P}((y^1, \dots, y^N) \in S | \omega_0) = \prod_{i=1}^N \int_{S^i} f^i(y^i, \omega_0, y^1, \dots, y^{i-1}) Q^i(dy^i). \quad (27)$$

As observed in Section 2.2, Assumption 7.3 allows us to reduce the problem to a static team problem where the observation of each DM is independent of observations of other DMs and also independent of  $\omega_0$ . Hence, under Assumption 7.3, we can focus on each  $DM^i$  separately and identify  $\Gamma^i$  via the set of probability measures

$$\Theta^i := \left\{ P \in \mathcal{P}(\mathbb{U}^i \times \mathbb{Y}^i) \mid P(du^i, dy^i) = 1_{\{\gamma^i(y^i) \in du^i\}} Q^i(dy^i), \gamma^i \in \Gamma^i \right\}. \quad (28)$$

As noted earlier, the above set is the set of extreme points of the set of probability measures on  $(\mathbb{U}^i \times \mathbb{Y}^i)$  with fixed marginals  $Q^i$  on  $\mathbb{Y}^i$ . Hence it inherits Borel measurability and topological properties of that Borel measurable set [35]. As before, we define convergence of policies as  $\gamma_n^i \rightarrow \gamma^i$  iff  $1_{\{\gamma_n^i(y^i) \in du^i\}} Q^i(dy^i) \rightarrow 1_{\{\gamma^i(y^i) \in du^i\}} Q^i(dy^i)$  (in the weak convergence topology) as  $n \rightarrow \infty$  for each DM. We will also allow for randomized (relaxed) policies. Accordingly, each individual control policy  $\gamma^i \in \Gamma^i$  is an element in the set of probability measures  $\mathcal{P}(\mathbb{U}^i \times \mathbb{Y}^i)$  with a fixed marginal,  $Q^i$ , on  $\mathbb{Y}^i$ .

Now that we have a standard Borel space formulation for policies, we can define the set of probability measures on policies with product topology on  $\Gamma_N = \prod_{i=1}^N \Gamma^i$ . We define the following set of Borel probability measures on admissible relaxed policies  $\Gamma_N$  as follows:

$$L^N := \mathcal{P}(\Gamma_N), \quad (29)$$

where Borel  $\sigma$ -field  $\mathcal{B}(\Gamma^i)$  is induced by the topology defined above.

Recall the definition of *exchangeability* for random variables.

**Definition 7.1.** Random variables  $x^1, x^2, \dots, x^N$  defined on a common probability space are  $N$ -exchangeable if for any permutation  $\sigma$  of the set  $\{1, \dots, N\}$ ,

$$\begin{aligned} & P\left(x^{\sigma(1)} \in A^1, x^{\sigma(2)} \in A^2, \dots, x^{\sigma(N)} \in A^N\right) \\ &= P\left(x^1 \in A^1, x^2 \in A^2, \dots, x^N \in A^N\right) \end{aligned}$$

for any measurable  $\{A^1, \dots, A^N\}$ , and  $(x^1, x^2, \dots)$  is infinitely-exchangeable if it is  $N$ -exchangeable for all  $N \in \mathbb{N}$ .

Now, we define the set of exchangeable probability measures on policies as:

$$\begin{aligned} L_{\text{EX}}^N := & \left\{ P_\pi \in L^N \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) \text{ and for all } \sigma \in S_N : \right. \\ & \left. P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = P_\pi(\gamma^{\sigma(1)} \in A_1, \dots, \gamma^{\sigma(N)} \in A_N) \right\}, \end{aligned}$$

where  $S_N$  is the space of permutations of  $\{1, \dots, N\}$ . We note that  $L_{\text{EX}}^N$  is a convex subset of  $L^N$ . Define the set of probability measures on policies induced by a common randomness as:

$$\begin{aligned} L_{\text{CO}}^N := & \left\{ P_\pi \in L^N \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : \right. \\ & \left. P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = \int_{z \in [0,1]} \prod_{i=1}^N P_\pi^i(\gamma^i \in A_i | z) \eta(dz), \eta \in \mathcal{P}([0,1]) \right\}, \end{aligned}$$

where  $\eta$  is the distribution of common, but independent (from intrinsic exogenous system variables), randomness. Note that conditioned on  $z$ , policies are independent. We also define the set  $L_{\text{CO,SYM}}^N$  as the set of identical probability measures on policies induced by a common randomness:

$$\begin{aligned} L_{\text{CO,SYM}}^N := & \left\{ P_\pi \in L^N \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : \right. \\ & \left. P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = \int_{z \in [0,1]} \prod_{i=1}^N P_\pi(\gamma^i \in A_i | z) \eta(dz), \eta \in \mathcal{P}([0,1]) \right\}, \end{aligned}$$

where we drop the index  $i$  in  $P_\pi$  to indicate that the independent randomization is identical through DMs. Also, define the set of probability measures on policies with only private independent randomness as:

$$\begin{aligned} & L_{\text{PR}}^N \\ := & \left\{ P_\pi \in L^N \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = \prod_{i=1}^N P_\pi^i(\gamma^i \in A_i) \right\}. \end{aligned}$$

Finally, define the set of probability measures on policies with identical and independent randomness:

$$L_{\text{PR,SYM}}^N := \left\{ P_\pi \in L^N \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = \prod_{i=1}^N P_\pi(\gamma^i \in A_i) \right\}.$$

For a team with a countably infinite number of decision makers, we define sets of probability measures  $L, L_{\text{EX}}, L_{\text{CO}}, L_{\text{CO,SYM}}, L_{\text{PR}}, L_{\text{PR,SYM}}$  similarly using the Ionescu Tulcea extension theorem by iteratively adding new coordinates for our probability measure (see e.g., [3, 72]). We define the set of probability measures  $L$  on the infinite product Borel spaces  $\Gamma = \prod_{i \in \mathbb{N}} \Gamma^i$  as:

$$L := \mathcal{P}(\Gamma). \quad (30)$$

Now, we define the set of infinitely exchangeable probability measures on policies as:

$$L_{\text{EX}} := \left\{ P_\pi \in L \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) \text{ and for all } N \in \mathbb{N}, \text{ and for all } \sigma \in S_N: \right. \\ \left. P_\pi(\gamma^1 \in A_1, \dots, \gamma^N \in A_N) = P_\pi(\gamma^{\sigma(1)} \in A_1, \dots, \gamma^{\sigma(N)} \in A_N) \right\},$$

and we define

$$L_{\text{CO}} := \left\{ P_\pi \in L \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : \right. \\ \left. P_\pi(\gamma^1 \in A_1, \gamma^2 \in A_2, \dots) = \int_{z \in [0,1]} \prod_{i \in \mathbb{N}} P_\pi^i(\gamma^i \in A_i | z) \eta(dz), \eta \in \mathcal{P}([0,1]) \right\}.$$

Note that  $L_{\text{CO}}$  is a convex subset of  $L$  and its extreme points are in the set of probability measures on policies with private independent randomness:

$$L_{\text{PR}} := \left\{ P_\pi \in L \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : P_\pi(\gamma^1 \in A_1, \gamma^2 \in A_2, \dots) = \prod_{i \in \mathbb{N}} P_\pi^i(\gamma^i \in A_i) \right\}.$$

Also, we define

$$L_{\text{CO,SYM}} := \left\{ P_\pi \in L \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : \right. \\ \left. P_\pi(\gamma^1 \in A_1, \gamma^2 \in A_2, \dots) = \int_{z \in [0,1]} \prod_{i \in \mathbb{N}} P_\pi(\gamma^i \in A_i | z) \eta(dz), \eta \in \mathcal{P}([0,1]) \right\},$$

and we define

$$L_{\text{PR,SYM}} := \left\{ P_\pi \in L \mid \text{for all } A_i \in \mathcal{B}(\Gamma^i) : P_\pi(\gamma^1 \in A_1, \gamma^2 \in A_2, \dots) = \prod_{i \in \mathbb{N}} P_\pi(\gamma^i \in A_i) \right\}.$$

**Theorem 7.4.** [121] *Suppose that Assumption 7.3 holds. Then, any  $P_\pi \in L_{EX}$  satisfying the following condition:*

- *For every  $i \in \mathbb{N}$ ,  $\mathbb{E}(\phi_i(u^i)) \leq K$  for some finite  $K$ , where  $\phi_i : \mathbb{U}^i \rightarrow \mathbb{R}_+$  is a lower semi-continuous moment function.*

*is in  $L_{CO,SYM}$ , i.e., for any  $P_\pi \in L_{EX}$  satisfying the above moment condition, there exists a  $[0, 1]$ -valued random variable  $z \sim \eta$  such that for any  $A_i \in \mathcal{B}(\Gamma^i)$*

$$P_\pi(\gamma^1 \in A_1, \gamma^2 \in A_2, \dots) = \int_{z \in [0,1]} \prod_{i \in \mathbb{N}} P_\pi(\gamma^i \in A_i | z) \eta(dz) \quad (31)$$

**Assumption 7.4.** *The cost function  $c$  in problem  $(\mathcal{P}_N)$  is exchangeable with respect to actions for all  $\omega_0$ , i.e., for any permutation  $\sigma$  of  $\{1, \dots, N\}$*

$$c(\omega_0, u^1, \dots, u^N) = c(\omega_0, u^{\sigma(1)}, \dots, u^{\sigma(N)})$$

*for all  $\omega_0$ .*

**Lemma 7.1.** *For a fixed  $N$ , consider an  $N$ -DM static team. Assume  $\bar{L}^N$  is an arbitrary convex subset of  $L^N$ . Under Assumption 7.3 and Assumption 7.4, if observations of DMs are exchangeable conditioned on  $\omega_0$ , then*

$$\begin{aligned} & \inf_{P_\pi \in \bar{L}^N} \int P_\pi(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0) \\ &= \inf_{P_\pi \in \bar{L}^N \cap L_{EX}^N} \int P_\pi(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0), \end{aligned}$$

*where  $c^N(\underline{\gamma}, \underline{y}, \omega_0) := c(\omega_0, \gamma^1(y^1), \dots, \gamma^N(y^N))$ .*

**Assumption 7.5.** *The cost function  $c : \Omega_0 \times \mathbb{U} \times \mathbb{U} \rightarrow \mathbb{R}_+$  in problem  $(\mathcal{P}_\infty)$  is continuous in its second and third arguments for all  $\omega_0$ .*

Under mild conditions, we can show that the optimal expected cost function induced by  $L_{EX}^N$  and  $L_{EX}$  are equal as  $N$  goes to infinity. Hence, by Lemma 7.1, under symmetry, this allows us to show that without loss of global optimality, optimal policies of static mean-field teams with a countably infinite number of DMs can be considered to be an infinitely exchangeable type.

**Lemma 7.2.** *Suppose that Assumption 7.3 and Assumption 7.5 hold. Assume further that  $\mathbb{U}$  is compact and the cost function is bounded. If observations of DMs are i.i.d. random vectors conditioned on  $\omega_0$ , then*

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \inf_{P_\pi \in L_{EX}^N} \int P_\pi^N(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0) \\ &= \limsup_{N \rightarrow \infty} \inf_{P_\pi \in L_{EX}} \int P_{\pi,N}(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0), \end{aligned} \quad (32)$$

*where  $c^N(\underline{\gamma}, \underline{y}, \omega_0) := \frac{1}{N} \sum_{i=1}^N c\left(\omega_0, \gamma^i(y^i), \frac{1}{N} \sum_{p=1}^N \gamma^p(y^p)\right)$ ,  $P_{\pi,N}$  is the marginal of the  $P_\pi \in L_{EX}$  to the first  $N$  components and  $\mu^N$  is the marginal of the fixed probability measure on  $(\omega_0, y^1, y^2, \dots)$  to the first  $N+1$  components.*

We now establish an existence of a randomized optimal policy for  $(\mathcal{P}_\infty)$ , which is symmetric.

**Theorem 7.5.** [121] *Consider a static team problem  $(\mathcal{P}_\infty)$  where Assumption 7.3 and Assumption 7.5 hold. Assume further that  $\mathbb{U}$  is compact. If observations of DMs are i.i.d. random vectors conditioned on  $\omega_0$ , then there exists a randomized optimal policy  $P_\pi^*$  for  $(\mathcal{P}_\infty)$  which is in  $L_{PR,SYM}$ :*

$$\min_{P_\pi \in L_{PR,SYM}} \limsup_{N \rightarrow \infty} \int P_{\pi,N}(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0) \quad (33)$$

$$= \inf_{P_\pi \in L_{PR}} \limsup_{N \rightarrow \infty} \int P_{\pi,N}(d\underline{\gamma}) \mu^N(d\underline{\omega}_0, d\underline{y}) c^N(\underline{\gamma}, \underline{y}, \omega_0). \quad (34)$$

Following from Lemma 7.1, Lemma 7.2, and our analysis in the proof of Theorem 7.5, thanks to Theorem 7.4, we can show that without losing global optimality, optimal policies for mean-field teams can be considered to be symmetric and privately randomized ( $L_{PR,SYM}$ ).

We note that if one also has convexity in the cost as well as action sets  $\mathbb{U}^i$ , then one can also establish that for every finite  $N$ , the optimal policies are symmetric and deterministic, but in the infinite limit, randomization may be required [122]. Similar results also hold for dynamic team problems [123]. We emphasize that a *strategic measures* approach would not be feasible for arriving at this solution since exchangeability in the actions is not sufficient to ensure that the *dominating* random variable (in the de Finetti representation) is independent of the intrinsic randomness in the system.

### 7.6. Extended weak convergence, topology of information, and adapted topologies

A versatile topology, which has evidently been used in a variety of contexts in stochastic analysis, is essentially given by the following convergence notion. Consider a stochastic process converging to another one in the following sense: all finite dimensional marginals converge weakly and the conditional kernels on the future random variables given the past converge weakly as well when conditional kernels are viewed as measure valued stochastic processes.

This notion has been applied for different problems: Aldous has termed it *extended weak convergence* [2] and Hellwig has named it *the information topology* [70]; these have recently been shown to be equivalent in discrete-time [14, Theorem 1.1]. In addition to the conditional independence preservation [70] [68], applications in robust stochastic control also follow from this discussion, with details being very context-specific: [82], [95], [23, 15].

In relation to our context, such a convergence notion again requires strong continuity conditions under all admissible policies. In the same spirit of the discussion in Section 7.2, though continuity and closedness hold under convergence with this notion, compactness will require more restrictive conditions.

With this approach, the recent work [68] (see also [20]) has established existence results for a setup where either the measurements are countable or there is

a common information among decision makers which is countable space-valued with the private information satisfying an absolute continuity condition. As noted earlier in the paper, static reduction applies in both such setups and the results presented in this paper (notably Theorem 4.5) generalize those reported in [68]. We note also that the use of the  $w$ - $s$  topology in Theorem 4.5 significantly relaxes the requirements of continuity.

## 8. Revisiting (and avoiding a subtle potential error in) relaxed policies for partially observed stochastic control

In this section, we will revisit the concept of *relaxed* control policies for classical stochastic control problems, with a further relaxation known as *wide sense admissible policies* introduced by Fleming and Pardoux [62] and prominently used to establish the existence of optimal solutions for partially observed stochastic control problems. Borkar [36, 38, 37] (see also [39]) has utilized these policies for a coupling/simulation method to arrive at optimality results for average cost partially observed stochastic control problems.

The main goal of this section is to show that if wide sense admissible control policies are not defined in a correct form, this can lead to a significant error in reasoning: the controllers may be allowed to have access to information that they should not.

Relaxed control policies are extremely useful concepts as shown in Section 3.2 and allows one to use topologies on the sets of probability measures to study existence, optimality, and structural results. A key aspect of such relaxations is that, the relaxation should not allow for optimal expected cost values to be improved; they should only be means to facilitate stochastic analysis. Our goal here is, building on the insights developed in Section 5.1 (in particular by the analysis on the CHSH team [47] reviewed in Theorem 5.3), to highlight a subtlety which may lead to incorrect conclusions if the relaxation is not cautiously constructed. Accordingly, we first present a brief overview of relaxed control policies in continuous-time or discrete-time stochastic control.

Consider a continuous-time Markov decision process  $\{x_t\}$  on an Euclidean space  $\mathbb{R}^N$ , controlled by a control process  $\{u_t\}$  taking values in a convex and compact Borel action space  $\mathbb{U} \subset \mathbb{R}^L$ , and with an associated observation process  $\{y_t\}$  taking values in  $\mathbb{R}^M$ , where  $0 \leq t \leq T$ . The evolution of  $\{x_t, y_t\}$  is given by stochastic differential equations

$$dx_t = b(x_t, y_t, u_t)dt + \sigma(x_t, y_t)dW_t, \quad (35)$$

$$dy_t = h(x_t)dt + dB_t. \quad (36)$$

Here,  $W$  and  $B$  are independent standard Wiener processes with values in  $\mathbb{R}^D$  and  $\mathbb{R}^M$ , respectively (hence,  $\sigma$  is a  $N \times D$ -matrix). Typically the functions  $b, \sigma, h$  above are assumed to have some regularity and growth conditions to ensure the existence of strong or weak solutions. The objective is to minimize the following

cost function

$$E \left[ \int_0^T F(x_t, u_t) dt + G(x_T) \right],$$

where  $F : \mathbb{R}^N \times \mathbb{U} \rightarrow [0, \infty)$  and  $G : \mathbb{R}^N \rightarrow [0, \infty)$ . In the literature, it is customary to require that control process  $\{u_t\}$  be adapted to the filtration generated by the observation process  $\{y_t\}$ ; that is, for each  $t \in [0, T]$ ,  $u_t$  is  $\sigma(y_s, 0 \leq s \leq t)$ -measurable. We will call such policies (strict-sense or precise) admissible policies. In [62], Fleming and Pardoux introduced another class of policies which they called *wide-sense admissible policies*. Using this relaxed class of policies, they study the existence of optimal policies to the above problem.

To define wide sense admissible policies, we first reproduce the above processes on a canonical probability space

$$\Omega = \Omega_0 \times \Omega_1 \times \Omega_2 \times \Omega_3,$$

where  $\Omega_0$ ,  $\Omega_1$ , and  $\Omega_2$  are  $C([0, T]; \mathbb{R}^m)$  with  $m = D, N, M$ , respectively, and  $\Omega_3 = L^2([0, T]; \mathbb{U})$ . We identify the Wiener process  $W$ , state process  $x$ , observation process  $y$ , and action process  $u$  as follows. If  $\omega = (W, x, y, u) \in \Omega$ , then

$$w(t) = (w_t(w), x_t(w), y_t(w), u_t(w)), \quad 0 \leq t \leq T.$$

Here,  $\Omega_1$ ,  $\Omega_2$ , and  $\Omega_3$  are endowed with the usual sup-norm topology and  $\Omega_3$  is endowed with the weak topology. Let  $\Omega^2 = \Omega_2 \times \Omega_3$  and define

$$\mathcal{F}_t(y) = \sigma(y_s, 0 \leq s \leq t), \quad \mathcal{F}_t(u) = \sigma(u_s, 0 \leq s \leq t), \quad \mathcal{G}_t^2 = \mathcal{F}_t(y) \times \mathcal{F}_t(u).$$

Here,  $\mathcal{G}_t^2$  is the product  $\sigma$ -field generated by  $\mathcal{F}_t(y)$  and  $\mathcal{F}_t(u)$ . Note that  $\{\mathcal{G}_t^2\}$  is a filtration on  $\Omega^2$ .

We can now define the class of wide-sense admissible policies. A *wide-sense admissible policy*  $\pi$  is a probability measure on  $(\Omega^2, \mathcal{G}_T^2)$  such that  $y$  is a  $(\pi, \{\mathcal{G}_t^2\})$  Wiener process. Note that this definition requires that the projection  $(y, u) \rightarrow y$  maps  $\pi$  onto Wiener measure and  $\{u_s, 0 \leq s \leq t\}$  is independent of the increment  $y_r - y_t$  for all  $t \leq r \leq T$ . The latter condition states that actions up to time  $t$  are independent of the observations after time  $t$  given past observations and actions. In other words, instead of saying that actions should be dependent on current and past observations, this condition states that actions should be independent of future observations given past observations and actions.

Given a distribution  $\mu$  of  $x_0$ , each wide-sense admissible policy induces a joint measure  $P^\pi$  of  $(w, x, y, u)$  as follows. Let  $P^{y,u}$  denote the conditional probability of  $(w, x)$  given  $(y, u)$  induced by stochastic differential equations (35), (36), which is independent of the policy as the action process  $u$  is given. Then define  $P_0^\pi$  as follows

$$P_0^\pi(dw, dx, dy, du) = P^{y,u}(dw, dx) \otimes \pi(dy, du).$$

Let us define the following density function

$$Z_T = \exp \left[ \int_0^T h(x_s) dy_s - \frac{1}{2} \int_0^T |h(x_s)|^2 ds \right].$$

Then,  $P^\pi$  is given by

$$\frac{dP^\pi}{dP_0^\pi} = Z_T.$$

Here, we a priori assume that this is integrable under the new measure, and so, we are indeed applying Girsanov's transformation. Let  $B_t = y_t - \int_0^t h(x_s) ds$ . Then, under  $P^\pi$ ,  $W$  and  $B$  are independent standard Wiener processes and stochastic differential equations (35) and (36) hold. Hence, the cost function of the wide-sense admissible policy  $\pi$  is given by

$$J(\pi) = E^\pi \left[ \int_0^T F(x_t, u_t) dt + G(x_T) \right],$$

where  $E^\pi$  denotes the expectation with respect to  $P^\pi$ .

Note that in this setup, a policy  $\pi$  is admissible in the classical sense if  $\pi$  is wide-sense admissible and there exists  $\mathcal{F}_T(y)/\mathcal{F}_T(u)$ -measurable  $f : \Omega_2 \rightarrow \Omega_3$  such that  $\pi(dy, du) = \delta_{f(y)}(du) \otimes w(dy)$ , where  $w$  is a Wiener measure. Likewise, we can say that the policy is relaxed if  $f : y \mapsto \mathcal{P}(\mathbb{U})$  is probability measure valued. We refer the reader to [62] for a more explicit construction of wide-sense admissible policies in the continuous time setup.

In [62, Theorem 7.2], Fleming and Pardoux proved the existence of an optimal wide-sense policy by converting the original problem to a fully-observed continuous-time Markov decision process on the belief-space  $P(\mathbb{R}^N)$ ; that is, the state of the belief-space MDP is the conditional distribution of the state  $x_t$  given the past observations and actions  $\mathcal{G}_t^2$ . Then, by adding some mild conditions on the stage-wise cost functions  $F$  and  $G$ , they also proved that the infimum achieved by classical admissible policies is the same as the infimum achieved by wide-sense admissible policies [62, Theorem 6.1]. Therefore, without loss of generality, one can work with wide-sense admissible policies instead of classical policies in order to further analyze such problems.

**Remark 8.1.** *It may be important to note that Bismut [28] arrived at further existence results, through an approach which avoids separation (and the construction of a belief-MDP), in discrete-time a similar approach is given in [148, Section 5.4.2].*

### 8.1. Discrete-time case

Inspired by the work of Fleming and Pardoux [62], Borkar introduced wide-sense control policies to study discrete-time partially-observed finite state-observation Markov decision processes with average cost criterion in [36, p.675], [38, item 1].



Using coupling methods, Borkar proved that fully-observed belief-space Markov decision processes under wide-sense admissible policies admit the solution of the average cost optimality equation, and any stationary policy for the belief-space MDP that solves this equation is optimal. Borkar later extended this result to continuous state-observation MDPs in [37, 39]. While it is evident that the discussion in [36, p.675], [38, item 1] is just an oversight (since this was corrected later, as we will discuss below), nonetheless we wish to note that if a reader applies this as written, this may lead to a consequential error, as we note in the following and as this was not explicitly noted.

To this end, we first review the construction of Borkar, which is very similar to the above construction, and we wish to recognize also that Borkar achieves what is in essence equivalent to Witsenhausen's static reduction reviewed in Section 2.2. For simplicity, we only consider here the case where state and observation spaces are finite. We consider a discrete-time Markov decision process  $\{x_n\}$  on a finite state space  $\mathbb{X}$ , controlled by a control process  $\{u_n\}$  taking values in a compact Borel action space  $\mathbb{U}$ , and with an associated observation process  $\{y_n\}$  taking values in a finite observation space  $\mathbb{Y}$ , where  $n = 0, 1, 2, \dots$ . The evolution of  $\{x_n, y_n\}$  is given by

$$P(x_{n+1}, y_{n+1} \in \cdot | x_m, x_m, u_m, m \leq n) = \rho(x_{n+1}, y_{n+1} \in \cdot | x_n, u_n),$$

where  $\rho : \mathbb{X} \times \mathbb{U} \rightarrow \mathcal{P}(\mathbb{X}) \times \mathcal{P}(\mathbb{Y})$  is some transition kernel. To ease the exposition, we assume that  $\rho$  is of the following form:

$$\rho(x_{n+1}, y_{n+1} | x_n, u_n) = r(y_{n+1} | x_{n+1}) \otimes p(x_{n+1} | x_n, u_n),$$

where  $p$  is the state transition kernel and  $r$  is the observation kernel. The initial distribution of  $x_0$  is  $\mu$ .

A control process  $\{u_n\}$  is admissible in the classical sense if it is adapted to the filtration  $\{\sigma(y_m, m \leq n)\}$  generated by observations  $\{y_n\}$ . In this case, one can write

$$u_n = \pi_n(y_0, \dots, y_n), n \geq 0, \tag{37}$$

for some  $\pi_n : \prod_{k=0}^n \mathbb{Y} \rightarrow \mathbb{U}$ . Let us denote  $\pi = \{\pi_n\}$ .

Note that one can always write the evolution of the state process  $\{x_n\}$  as a noise-driven dynamical system

$$x_{n+1} = F(x_n, u_n, w_n), \tag{38}$$

where  $F : \mathbb{X} \times \mathbb{U} \times [0, 1] \rightarrow \mathbb{X}$  is measurable and  $\{w_n\}$  are independently and identically distributed uniformly on  $[0, 1]$ . Using this dynamical system, we now reproduce the above process on a more convenient probability space. This will then enable us to define *wide-sense admissible policies*.

In the following, we reduce the problem to an independent static one via Witsenhausen/Girsanov/Borkar, see Borkar's [36, 38] explicit analysis or Witsenhausen's method presented in Section 2.2.

Under this reduction, we obtain a new probability measure  $P_0^\pi$  under which:

- (a)  $\{y_n\}$  is i.i.d. uniform on  $\mathbb{Y}$  and independent of  $x_0$  and  $\{w_n\}$ ,
- (b)  $\{u_n, y_0, \dots, y_n\}$  is independent of  $\{w_n\}$ ,  $x_0$ , and  $\{y_m, m > n\}$ , for all  $n$ .

Using these properties, Borkar defined wide sense admissible policies in [36, 38] as follows. A policy  $P_0$  is *wide sense admissible* if  $P_0$  satisfies (a) and (b). Note that condition (b) is very similar to the non-signaling condition introduced in Section 5.1. It assumes that the action at time  $n$  is independent of the observations after time  $n$  given past observations (but **not** necessarily past actions). In other words, instead of saying that action  $u_n$  should be dependent on current and past observations  $\{y_0, \dots, y_n\}$ , this condition states that action  $u_n$  should be independent of future observations  $\{y_m, m > n\}$  given past observations. Borkar and Budhiraja thankfully realized this seemingly simple, but consequential as we will see later, typo as this was corrected in further publications: in [37, 39], when Borkar and Budhiraja extend this definition to the continuous space case, they have slightly changed the condition: This condition is denoted by (b') and stated as follows:

- (b')  $\{u_0, \dots, u_n, y_0, \dots, y_n\}$  is independent of  $\{w_n\}$ ,  $x_0$ , and  $\{y_m, m > n\}$ , for all  $n$ .

In (b'), in addition to  $\{u_n, y_0, \dots, y_n\}$ , we suppose that past actions  $u_0, \dots, u_{n-1}$  are independent of  $\{w_n\}$ ,  $x_0$ , and  $\{y_m, m > n\}$ . In other words,  $u_n$  is independent of the observations after time  $n$  given past observations and *actions* (note that in condition (b) past actions are missing). This is indeed the right relaxation since in the next section, we establish via a counterexample that the optimal value achieved by wide sense admissible policies (with condition (b)) is strictly better than the optimal value achieved by classically admissible policies. Moreover, we show that an optimal wide-sense admissible policy under (b) evidently has access to a control policy that it should not have; that is, it violates the causality of the problem, which is prohibited in general.

## 8.2. A counterexample

We have a partially observed MDP with the components  $\mathbb{X} = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$ ,  $\mathbb{U} = \{0, 1\}$ , and  $\mathbb{Y} = \{y^*\}$  contains only one element. Let  $x_0 \sim \pi_0 \otimes \gamma_0$ , where  $\pi_0$  is a distribution on  $\{0, 1\} \times \{0, 1\}$  and  $\gamma_0$  is a distribution on  $\{0, 1\}$ . A typical element of  $\mathbb{X}$  is denoted by  $x = [x^1, x^2, x^3]$ . The transition and the observation kernels are defined as follows:

$$p([x_{n+1}^1, x_{n+1}^2, x_{n+1}^3] | [x_n^1, x_n^2, x_n^3], u_n) = \lambda(x_{n+1}^1, x_{n+1}^2 | x_n^1, x_n^2) \otimes \delta_{u_n}(x_{n+1}^3)$$

$$r(y_n | x_n) = \delta_{y^*}(y_n),$$

where  $\lambda : \{0, 1\} \times \{0, 1\} \rightarrow P(\{0, 1\} \times \{0, 1\})$  is a stochastic kernel such that  $\pi_0$  is an invariant probability measure of  $\lambda$ . The reward<sup>2</sup> function is given by

$$c(x, u) = \begin{cases} 1, & \text{if } x^3 \oplus u = x^1 \cdot x^2 \\ 0, & \text{otherwise.} \end{cases}$$

To construct this counterexample, we have been inspired by the CHSH team [47] reviewed earlier in Theorem 5.3, which establishes that non-signaling policies are not admissible relaxations as they strictly improve the performance. A similar conclusion will be obtained in this counterexample.

In this problem, the observation process  $\{y_n\}$  is non-informative as, for each  $n$ ,  $y_n \sim \delta_{y^*}$ . Since any admissible control process  $\{u_n\}$  is adapted to the filtration generated by the observation process; that is, one can write

$$u_n = \pi_n(y_0, \dots, y_n), n \geq 0, \quad (39)$$

for some  $\pi_n : \prod_{k=0}^n \mathbb{Y} \rightarrow \mathbb{U}$ , any admissible control process  $\{u_n\}$  can be represented as a deterministic sequence  $\{a_n\}_{n=0}^\infty \subset \{0, 1\}^\infty$ , i.e.,  $u_n \sim \delta_{a_n}$  for all  $n$ . Note that, by the definition of the state transition kernel, the distribution  $\pi_0$  of the first two components of the state  $(x_n^1, x_n^2)$  remains as it is during the evolution of the state. With these observations, the maximum reward an admissible policy can attain is

$$\max \{ \pi_0(x_0^1 \cdot x_0^2 = 0), \pi_0(x_0^1 \cdot x_0^2 = 1) \}. \quad (40)$$

For instance, if

$$\max \{ \pi_0(x_0^1 \cdot x_0^2 = 0), \pi_0(x_0^1 \cdot x_0^2 = 1) \} = \pi_0(x_0^1 \cdot x_0^2 = 0),$$

then this reward can be obtained by picking the control sequence as follows:

$$u_0 = 0, u_1 = 1, u_2 = 0, u_3 = 1, \dots$$

We now construct an optimal wide sense admissible policy with the reward function 1, which is in general strictly larger than (40). Let the probability distribution  $P_0$  on  $\Omega$  have the following properties:

- (a)  $\{y_n\}$  is i.i.d. uniform on  $\mathbb{Y}$  and independent of  $x_0$  and  $\{w_n\}$ ,
- (b) For any  $n \geq 0$ ,

$$P_0(u_n, u_{n+1} | x_m, y_m, w_m, m \leq n) = \begin{cases} 1/2, & \text{if } u_n \oplus u_{n+1} = x_n^1 \cdot x_n^2 \\ 0, & \text{otherwise} \end{cases}$$

One can prove that  $P_0$  is wide sense admissible; that is,  $\{u_n, y_0, \dots, y_n\}$  is independent of  $\{w_n\}$ ,  $x_0$ , and  $\{y_m, m > n\}$ , for all  $n$ , under  $P_0$ . Indeed, we have  $P_0(u_n | x_m, y_m, w_m, m \leq n) = \mathbb{U}_{\{0,1\}}(u_n)$  and  $P_0(u_{n+1} | x_m, y_m, w_m, m \leq n) =$

<sup>2</sup>We note that all results in this paper apply with straightforward modifications for the case of maximizing reward instead of minimizing cost.

$U_{\{0,1\}}(u_{n+1})$ , where  $U_{\{0,1\}}$  is the uniform distribution on  $\{0, 1\}$ . The average reward of  $P_0$  is 1, which can be the maximum achievable by any policy as  $0 \leq c \leq 1$ . Hence,  $P_0$  is the optimal wide-sense policy.

Note that if

$$\max\{\pi_0(x_0^1 \cdot x_0^2 = 0), \pi_0(x_0^1 \cdot x_0^2 = 1)\} = \pi_0(x_0^1 \cdot x_0^2 = 0) = \pi_0(x_0^1 \cdot x_0^2 = 1) = 1/2,$$

then classical policies at most attain the reward of  $1/2$ , while wide sense admissible policies can get the reward of 1 which is twice as big as  $1/2$ . Hence, this proves the fact that wide sense admissible policies are much bigger than the classical policies. Furthermore,  $P_0$  cannot be implemented in real life control applications as the action  $u_n$  at any time  $n$  depends on the knowledge of the action  $u_{n+1}$  at time  $n + 1$ , which violates the causality of the problem. Hence, this shows that for certain class of problems, wide-sense admissible policies are not a legitimate class of policies to study in the discrete-time setup. In order to work with such policies, it is necessary to establish a theorem like [62, Theorem 6.1] that establishes equivalency of these two classes of policies in terms of achievable optimal value. This can indeed be done if condition (b) is replaced by condition (b'). Note that the wide sense policy  $P_0$  defined above violates the condition (b'). Therefore, it is not wide sense admissible under condition (b').

In summary, if the relaxation is done with the interpretation that  $u_n$  is conditionally independent from future observations given the past observations, an incorrect conclusion can be made. If the relaxations is so that the control action  $u_n$  is conditionally independent from future observations, given all the past, then relaxation is valid.

## 9. Conclusion, topics left out and some open problems

### 9.1. Concluding remarks

The way information is decentralized is a key attribute determining how to approach a problem in various areas of applied mathematics. In this review article, we studied information structures in a probability theoretic and topological context. We defined information structures, placed various topologies on them, and study closedness and compactness properties. We presented existence and approximation results for optimal decision/control policies. We discussed various upper and lower bounding techniques, through relaxations and convex programs ranging from classically realizable and classically non-realizable (such as quantum and non-signaling) relaxations. Figure 2 depicts a summary of our findings on strategic measures.

We later presented various topologies on decision/control strategies defined independently of information structures, but for which information structures determine whether the topologies have utility in arriving at existence, compactness, convexification or approximation results.

We showed that viewing decentralization with regard to the induced strategic measures and viewing decentralization with respect to the applied control poli-

cies, lead to different operational conclusions and tools to arrive at optimality conditions and results for optimal decentralized decision making.

We showed that externally provided randomness should be well motivated and when one defines relaxed control; which is a common solution technique in discrete-time and continuous-time classical stochastic control, this should be introduced cautiously since such randomness may indeed lead to unacceptable performance improvement.

In the paper we considered the set of correlations *given* an information structure. A further related set of results involve the problem of optimal *design* of information structures. This subject is beyond the scope of this review article with some related results in [150, Thms. 3.2, 3.3, 3.4] and [149], which also directly apply to the multi-agent setting.

## 9.2. Discussion on existence and some related results

As noted earlier, while the topological constructions on policies are quite useful, Theorem 4.5 (for static teams or dynamic teams with an independent-measurements reduction) and Theorems 4.6 and 4.7 (for sequential teams that do not allow an independent-measurements reduction) are the most general existence results, to our knowledge, for the problems considered here. However, some slightly weaker versions of these results can be arrived at through different methods, as laid out in the paper. We also note, for completeness, that existence of optimal policies for static and a class of sequential dynamic teams had been studied in [144], [137] [69, 151], [150] [148]. Conditions for optimality have been established in [45] for a class of continuous-time decentralized stochastic control problems. We also noted in the article, for (classical) partially observed models, existence results in [62], [28] with a discrete-time review in [148]. For a class of teams which are convex, one can reduce the search space to a smaller parametric class of policies, such as linear policies for quasi-classical linear quadratic Gaussian problems [117, 85, 73].

Note that standard dynamic programming is known to be a useful tool for a class of dynamic teams since the 1970s. Clearly, if all the information at any given decision maker is common knowledge between all decision makers, then the system can be viewed as a centralized system and standard dynamic programming is applicable. However, if only some of the system variables are common knowledge, the remaining unknowns may or may not lead to a computationally tractable program generating an optimal solution. A possible approach toward establishing a tractable program is through the construction of a controlled Markov chain where the controlled Markov state may now live in a larger state space (for example a space of probability measures) and the actions may be elements from appropriately defined function spaces [149]. This controlled Markov construction may lead to a computation of optimal policies. Such a *dynamic programming approach* has been adopted extensively in the literature (see for example, [11], [145], [46], [1], [147], [93] and significantly generalized and termed as the *common information approach* in [109] and [110]) through the use of a

team-policy which uses common information to generate partial functions for each DM to generate their actions using local information. This construction requires a common knowledge among decision makers, which is a unifying assumption in the aforementioned contributions in the literature. Witsenhausen [140] and [148] developed universal dynamic programming algorithms, which are conceptually useful and mathematically consequential (on existence and recursive analysis) but practically of limited algorithmic use with our current knowledge.

### 9.3. Some open problems

In the following, we present a number of open problems:

- (a) Note that one can prove, using the same argument in Theorem 5.3-(ii), that

$$\inf_{P \in L_{Q(1)}(\mu)} \int P(ds) c(s) \leq J^*,$$

since there exists an optimal individually randomized policy that achieves  $J^*$ . Therefore, for any  $d$ , the team problem  $\inf_{P \in L_{Q(d)}(\mu)} \int P(ds) c(s)$  is indeed an *admissible* extension or relaxation of classical team problem since it does not require any communication between agents and a mediator. Hence, the solution of this problem will be a significant contribution to team decision theory as the optimal quantum-correlated policy can be realizable in real life in view of recent and potential advances in quantum technology. However, it is important to note that  $L_{Q(d)}(\mu)$  cannot be convex if the dimension constraint  $d$  is small [53] (see also Theorem 5.2). Therefore, with dimensionality constraint, the optimization problem corresponding to the quantum-correlated strategic measures can be non-convex. To convexify the problem, we can either add unlimited common randomness without changing  $d$  or increase  $d$  and allow for a limited common randomness [53]. We can now state the following open problem:

- (OP1)** Under what conditions on the components of the team, there exists  $d \geq 1$  such that the optimization problem

$$\inf_{P \in L_{Q(d)}(\mu)} \int P(ds) c(s) \tag{41}$$

can be written or can be approximated by a semi-definite program?

One way to approach **(OP1)** may be to adopt the *so-called* NPA hierarchy [108], which provides an infinite hierarchy of semidefinite programming (SDP) outer approximations to the set of quantum-correlated strategic measures. In the NPA hierarchy, given a set of finitely many operators that are linear combinations of products of projectors, one forms the moment matrix, where each entry of the moment matrix is obtained by computing the trace of the product of the operators in this finite set. This moment matrix satisfies some semidefinite constraints and cost can be written as a

trace of the product of moment matrix and properly extended cost operator. Hence, for each finite set of operators, one can write the optimization problem over moment matrix as a semidefinite program. Since semidefinite constraints satisfied by moment matrix are necessary conditions for the original quantum-correlated relaxation, it gives an outer approximation to the original problem. As the number of operators in the finite set converges to infinity, it can be proved that the outer approximation converges to the original quantum-correlated setup. However, although this gives SDP outer approximations, there are no bounds on the rate of convergence quantifying how the approximation improves as the level in the hierarchy increases. On the other hand, if we can bound the rate of convergence for the NPA hierarchy, it will be possible to use it for solving **(OP1)**.

- (b) Recall the dual program introduced in Section 5.4 using non-signaling policies. In this section, we establish the dual problem for Witsenhausen's counterexample and pose several open problems. To this end, we let  $\mathbb{Y}^1 = \mathbb{Y}^2 = \mathbb{U}^1 = \mathbb{U}^2 = \mathbb{R}$ , which denote the observation and action spaces of agents. In Witsenhausen's celebrated counterexample [138], depicted in Fig. 4, there are two decision makers: Agent 1 observes a zero mean and  $\sigma$ -variance Gaussian random variable  $y^1 \in \mathbb{Y}^1$  and decides its action  $u^1 \in \mathbb{U}^1$ . Agent 2 observes  $y^2 := u^1 + v \in \mathbb{Y}^2$ , where  $v \in \mathbb{V} = \mathbb{R}$  is zero mean and unit variance Gaussian noise independent of  $y^1$ , and decides its action  $u^2 \in \mathbb{U}^2$

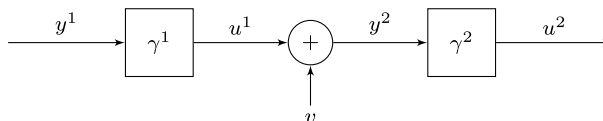


FIG 4. *Witsenhausen's counterexample.*

The cost function of the team is given by

$$c(y^1, u^1, u^2) = k^2(u^1 - y^1)^2 + (u^2 - u^1)^2,$$

where  $k > 0$ . Let  $g(y) := \frac{1}{\sqrt{2\pi}} \exp\{-y^2/2\}$  be the probability density function of a zero mean and unit variance Gaussian random variable. Then we have

$$P(y^2 \in S | u^1) = \int_S g(y^2 - u^1) m(dy^2),$$

where  $m$  denotes the Lebesgue measure on  $\mathbb{Y}^2$ . Let

$$f(u^1, y^2) := \exp\left\{-\frac{(u^1)^2 - 2y^2u^1}{2}\right\} \quad (42)$$

so that  $g(y^2 - u^1) = f(u^1, y^2) \frac{1}{\sqrt{2\pi}} \exp\{-(y^2)^2/2\} = f(u^1, y^2) g(y^2)$ . The independent static reduction of Witsenhausen's counterexample proceeds as follows: for any strategic measure  $P \in L_R(\mu)$ , the expected cost can be written as

$$\begin{aligned} J(P) &= \int c(y^1, u^1, u^2) P(du^2|y^2) P(dy^2|u^1) P(du^1|y^1) \mu_\sigma(dy^1) \\ &= \int c(y^1, u^1, u^2) f(u^1, y^2) \mu_1(dy^2) \mu_\sigma(dy^1), \end{aligned}$$

where  $\mu_\rho$  denotes zero mean and  $\rho$ -variance Gaussian distribution. Hence, by defining  $c_s(y^1, y^2, u^1, u^2) = c(y^1, u^1, u^2) f(u^1, y^2)$  and  $\mu(dy^1, dy^2) = \mu_\sigma(dy^1) \mu_1(dy^2)$ , we can write  $J(P)$  as

$$J(P) = \int c_s(y^1, y^2, u^1, u^2) P(du^2|y^2) P(du^1|y^1) \mu(dy^1, dy^2). \quad (43)$$

Therefore, in the static reduction of Witsenhausen's counterexample, the agents observe independent zero mean Gaussian random variables. In the remainder of this note, we consider the static reduction of Witsenhausen's counterexample. Note that a strategic measure  $P \in \mathcal{P}(\mathbb{Y}^1 \times \mathbb{Y}^2 \times \mathbb{U}^1 \times \mathbb{U}^2)$  is *non-signaling* if

$$\begin{aligned} P(du^1|y^1, y^2) &= P(du^1|y^1), \\ P(du^2|y^1, y^2) &= P(du^2|y^2). \end{aligned} \quad (44)$$

As noted in Section 5.4, since the constraints in (44) for  $P$  are linear, the optimal team cost with non-signaling policies can be written as a linear program over an appropriate vector spaces as follows. Recall that, for any metric space  $\mathbb{E}$ ,  $\mathcal{M}(\mathbb{E})$  denotes the set of finite signed measures on  $\mathbb{E}$  and  $C(\mathbb{E})$  denotes the set of continuous real functions. Consider the vector spaces  $\mathcal{M}(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ ,  $C(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ ,  $\mathcal{M}(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ ,  $C(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ ,  $\mathcal{M}(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ , and  $C(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ . Let us define bilinear forms on

$$(\mathcal{M}(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2), C(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2))$$

and on

$$(\mathcal{M}(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathcal{M}(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathbb{R}, C(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times C(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathbb{R})$$

by letting

$$\begin{aligned} \langle \rho, v \rangle_1 &:= \int_{\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2} v(u^1, u^2, y^1, y^2) \rho(du^1, du^2, dy^1, dy^2), \quad (45) \\ \langle (\rho_1, \rho_2, a), (v_1, v_2, b) \rangle_2 &:= \int_{\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2} v_1(u^1, y^1, y^2) \rho_1(du^1, dy^1, dy^2) \end{aligned}$$



$$+ \int_{\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2} v_2(u^2, y^1, y^2) \rho_2(du^2, dy^1, dy^2) + ab. \quad (46)$$

The bilinear forms in (45) and (46) constitute duality between spaces [22, Chapter IV.3]. Hence, the topologies on these spaces should be understood as the weak topology of the duality induced by these bilinear forms. We define the linear map  $L : \mathcal{M}(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \rightarrow \mathcal{M}(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathcal{M}(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathbb{R}$  by

$$L(\rho) = (L_1(\rho), L_2(\rho), L_3(\rho)),$$

where

$$\begin{aligned} L_1 : \rho(du^1, du^2, dy^1, dy^2) &\mapsto \rho(du^1, dy^1, dy^2) - \rho(du^1, dy^1) \mu_1(dy^2) \\ L_2 : \rho(du^1, du^2, dy^1, dy^2) &\mapsto \rho(du^2, dy^1, dy^2) - \rho(du^2, dy^2) \mu_\sigma(dy^1) \\ L_3 : \rho(du^1, du^2, dy^1, dy^2) &\mapsto \langle \rho, 1 \rangle_1. \end{aligned}$$

Using  $L$ , the optimal value of the team with non-signaling policies can be written as a linear program as follows:

$$\begin{aligned} \text{(NS)} \quad &\text{minimize}_{\rho \in \mathcal{M}_+(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)} \langle \rho, c_s \rangle_1 \\ &\text{subject to } L(\rho) = (0, 0, 1). \end{aligned} \quad (47)$$

Since  $L_C(\mu) \subset L_{NS}(\mu)$ , the solution of above linear program gives a lower bound to the original formulation of Witsenhausen's counterexample. Note that the dual of  $L$  is given by  $L^* : C(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times C(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathbb{R} \rightarrow C(\mathbb{U}^1 \times \mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2)$ , where

$$\begin{aligned} L^*(v_1, v_2, b) &= v_1(u^1, y^1, y^2) - \int_{\mathbb{Y}^2} v_1(u^1, y^1, y^2) d\mu_1(y^2) \\ &\quad + v_2(u^2, y^1, y^2) - \int_{\mathbb{Y}^1} v_2(u^2, y^1, y^2) d\mu_\sigma(y^1) + b. \end{aligned} \quad (48)$$

Then the dual program of (NS) can be written as [22, Chapter IV.6]

$$\begin{aligned} \text{(NS}^*) \quad &\text{maximize}_{(v_1, v_2, b) \in C(\mathbb{U}^1 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times C(\mathbb{U}^2 \times \mathbb{Y}^1 \times \mathbb{Y}^2) \times \mathbb{R}} b \\ &\text{subject to } L^*(v_1, v_2, b) \leq c_s. \end{aligned} \quad (49)$$

We can now state the following problems:

- (OP2) Approximate numerically the dual linear program (NS\*) to obtain a lower bound to  $J^*$ .
- (OP3) Analyze the difference

$$\inf_{P \in L_{NS}(\mu)} \int P(ds) c_s(s) - \inf_{P \in L_C(\mu)} \int P(ds) c_s(s),$$

or at least with an upper bound obtained with numerical methods (see [120] for a review of numerical results in the literature).

- (OP4) We know that when there is no  $u^1 \cdot u^2$  term in the original cost function  $c$ , there is an affine optimal policy which can be obtained analytically using information theoretic tools [17]. In this case, what is the relation between

$$\inf_{P \in L_{NS}(\mu)} \int P(ds) c_s(s)$$

and

$$\inf_{P \in L_C(\mu)} \int P(ds) c_s(s)?$$

More generally, for which cost functions, can we establish that

$$\inf_{P \in L_{NS}(\mu)} \int P(ds) c_s(s) = \inf_{P \in L_C(\mu)} \int P(ds) c_s(s)?$$

- (c) In the paper, we considered only sequential decentralized stochastic control. As noted earlier, if there is a pre-defined order in which the decision makers act, then we say that a system is *sequential*; otherwise, the system is non-sequential. Such non-sequential systems are substantially more difficult to study, since the ambiguities in the order of actions lead to challenges on the interpretation of local information. Optimal design of such non-sequential models requires the systems to be deadlock-free, that is the actions of a given DM should not depend on the actions of DMs acting in the future, for any realized random ordering. Furthermore, the optimization problem for such systems should be well posed/solvable, since for some designs the expected cost may not be well-defined. We refer the reader to Witsenhausen [139], Andersland and Teneketzis [5], [6] and Teneketzis [130] for a comprehensive study of non-sequential systems, see [149] for a brief review. We also note that with the observation that the information fields generated by local measurements lead to subtle conditions on solvability and causality, an alternative probabilistic model, based on quantum mechanics, for describing such problems has been proposed by Baras in [19] and [18]. These papers also present an accessible review of related developments on the quantum information literature prior to the publications. In summary, the study of non-sequential decentralized stochastic control systems in the context of what we studied throughout this paper is an open problem.

## Acknowledgements

The authors would like to acknowledge discussions and collaborations with Tamer Başar, Tamás Linder, Sina Sanjari, Ian Hogeboom-Burr, Ankur Kulkarni, Abhishek Gupta, Aditya Mahajan, Nuno Martins, Demos Teneketzis, Can Deha Karıksız, Maxim Raginsky, Bora Yongacoglu and Eric Chitambar.

## References

- [1] M. AICARDI, F. DAVOLI, AND R. MINCIARDI, *Decentralized optimal control of Markov chains with a common past information set*, IEEE Transactions on Automatic Control, 32 (1987), pp. 1028–1031. [MR0909973](#)
- [2] D. J. ALDOUS, *Weak convergence and the general theory of processes*, Editeur inconnu, 1981.
- [3] C. ALIPRANTIS AND K. BORDER, *Infinite Dimensional Analysis*, Berlin, Springer, 3rd ed., 2006. [MR2378491](#)
- [4] V. ANANTHARAM AND V. S. BORKAR, *Common randomness and distributed control: A counterexample*, Systems and control letters, 56 (2007), pp. 568–572. [MR2332009](#)
- [5] M. ANDERSLAND AND D. TENEKETZIS, *Information structures, causality, and non-sequential stochastic control, I: design-independent properties*, SIAM J. Control and Optimization, 30 (1992), pp. 1447–1475. [MR1185632](#)
- [6] M. ANDERSLAND AND D. TENEKETZIS, *Information structures, causality, and non-sequential stochastic control, II: design-dependent properties*, SIAM J. Control and Optimization, 32 (1994), pp. 1726–1751. [MR0472143](#)
- [7] J. ARABNEYDI AND A. MAHAJAN, *Team-optimal solution of finite number of mean-field coupled LQG subsystems*, in IEEE 54th Annual Conference on Decision and Control (CDC), 2015, pp. 5308–5313.
- [8] A. ARAPOSTATHIS, A. BISWAS, AND J. CARROLL, *On solutions of mean field games with ergodic cost*, Journal de Mathématiques Pures et Appliquées, 107 (2017), pp. 205–251. [MR3597374](#)
- [9] A. ARAPOSTATHIS, V. S. BORKAR, AND M. K. GHOSH, *Ergodic control of diffusion processes*, vol. 143, Cambridge University Press, 2012. [MR2884272](#)
- [10] A. ASPECT, *Proposed experiment to test the nonseparability of quantum mechanics*, Phys. Rev. D, 14 (1976), pp. 1944–1951.
- [11] M. ATHANS, *Survey of decentralized control methods*, Washington D.C, 1974, 3rd NBER/FRB Workshop on Stochastic Control.
- [12] V. AULETTA, D. FERRAIOLI, A. RAI, G. SCARPA, AND A. WINTER, *Belief-invariant equilibria in games with incomplete information*. arXiv:[1605.07896](#), 2016.
- [13] R. J. AUMANN, *Correlated equilibrium as an expression of bayesian rationality*, Econometrica: Journal of the Econometric Society, (1987), pp. 1–18. [MR0875515](#)
- [14] J. BACKHOFF-VERAGUAS, D. BARTL, M. BEIGLBÖCK, AND M. EDER, *All adapted topologies are equal*, arXiv preprint arXiv:[1905.00368](#), (2019). [MR4168395](#)
- [15] J. BACKHOFF-VERAGUAS, D. BARTL, M. BEIGLBÖCK, AND M. EDER, *Adapted wasserstein distances and stability in mathematical finance*, Finance and Stochastics, 24 (2020), pp. 601–632. [MR4118990](#)
- [16] E. J. BALDER, *On  $ws$ -convergence of product measures*, Mathematics of Operations Research, 26 (2001), pp. 494–518. [MR1849882](#)
- [17] R. BANSAL AND T. BASAR, *Stochastic teams with nonclassical information revisited: When is an affine law optimal?*, IEEE Transactions on Automatic

- Control, 32 (1987), pp. 554–559. [MR0887593](#)
- [18] J. S. BARAS, *Noncommutative probability models in quantum communication and multi-agent stochastic control*, *Ricerca di Automatica*, 55 (1979), pp. 217–264. [MR0614261](#)
- [19] J. S. BARAS, *Multi-agent stochastic control: Models inspired from quantum physics*, Saint Petersburg, Russia, 2003, International Conference, Physics and Control (PhysCon), pp. 747–758.
- [20] M. BARBIE AND A. GUPTA, *The topology of information on the space of probability measures over polish spaces*, *Journal of Mathematical Economics*, 52 (2014), pp. 98–111. [MR3212192](#)
- [21] J. BARRETT, N. LINDEN, S. MASSAR, S. PIRONIO, S. POPESCU, AND D. ROBERTS, *Nonlocal correlations as an information-theoretic resource*, *Phys. Rev. A*, 71 (2005), p. 022101.
- [22] A. BARVINOK, *A Course in Convexity*, American Mathematical Society, 54 of Graduate Studies in Mathematics. Providence, RI, 2002. [MR1940576](#)
- [23] E. BAYRAKTAR, Y. Y. DOLINSKY, AND J. GUO, *Continuity of utility maximization under weak convergence*, *Mathematics and Financial Economics*, (2020), pp. 1–33. [MR4135703](#)
- [24] M. BEIGLBÖCK AND D. LACKER, *Denseness of adapted processes among causal couplings*, *arXiv*, (2018), pp. arXiv-1805.
- [25] J. S. BELL, *On the einstein podolsky rosen paradox*, *Physics Physique Fizika*, 1 (1964), pp. 195–200. [MR3790629](#)
- [26] V. E. BENEŠ, *Existence of optimal stochastic control laws*, *SIAM Journal on Control*, 9 (1971), pp. 446–472. [MR0300726](#)
- [27] P. BILLINGSLEY, *Convergence of probability measures*, New York: Wiley, 2nd ed., 1999. [MR0233396](#)
- [28] J.-M. BISMUT, *Partially observed diffusions and their control*, *SIAM Journal on Control and Optimization*, 20 (1982), pp. 302–309. [MR0646955](#)
- [29] A. BISWAS, *Mean field games with ergodic cost for discrete time markov processes*, *arXiv preprint arXiv:1510.08968*, (2015).
- [30] D. BLACKWELL, *Memoryless strategies in finite-stage dynamic programming*, *Annals of Mathematical Statistics*, 35 (1964), pp. 863–865. [MR0162642](#)
- [31] D. BLACKWELL, *The stochastic processes of Borel gambling and dynamic programming*, *Annals of Statistics*, (1976), pp. 370–374.
- [32] D. BLACKWELL AND C. RYLL-NADZEWSKI, *Non-existence of everywhere proper conditional distributions*, *Annals of Mathematical Statistics*, 34 (1963), pp. 223–225. [MR0148097](#)
- [33] V. S. BORKAR, *The probabilistic structure of controlled diffusion processes*, *Acta Applicandae Mathematica*, 11 (1988), pp. 19–48. [MR0927371](#)
- [34] V. S. BORKAR, *A topology for markov controls*, *Applied Mathematics and Optimization*, 20 (1989), pp. 55–62. [MR0989430](#)
- [35] V. S. BORKAR, *White-noise representations in stochastic realization theory*, *SIAM J. on Control and Optimization*, 31 (1993), pp. 1093–1102. [MR1233993](#)
- [36] V. S. BORKAR, *Average cost dynamic programming equations for con-*

- trolled Markov chains with partial observations*, SIAM J. Control Optim., 39 (2000), pp. 673–681. [MR1786324](#)
- [37] V. S. BORKAR, *Dynamic programming for ergodic control with partial observations*, Stochastic Processes and their Applications, 103 (2003), pp. 293–310. [MR1950768](#)
- [38] V. S. BORKAR, *Dynamic programming for ergodic control of Markov chains under partial observations: A correction*, SIAM J. Control Optim., 45 (2007), pp. 2299–2304. [MR2285725](#)
- [39] V. S. BORKAR AND A. BUDHIRAJA, *A further remark on dynamic programming for partially observed Markov processes*, Stochastic Processes and their Applications, 112 (2004), pp. 79–93. [MR2062568](#)
- [40] R. W. BROCKETT, *Asymptotic stability and feedback stabilization*, Differential geometric control theory, 27 (1983), pp. 181–191. [MR0708502](#)
- [41] N. BRUNNER, D. CAVALCANTI, S. PIRONIO, V. SCARANI, AND S. WEHNER, *Bell nonlocality*, Reviews of Modern Physics, 86 (2014), p. 419.
- [42] N. BRUNNER, N. GISIN, AND V. SCARANI, *Entanglement and non-locality are different resources*, New Journal of Physics, 7 (2005), pp. 88–88.
- [43] P. CEMBRANOS AND J. MENDOZA, *Banach Spaces of Vector-Valued Functions*, Springer-Verlag, 1997. [MR1489231](#)
- [44] C. D. CHARALAMBOUS, *Decentralized optimality conditions of stochastic differential decision problems via Girsanov’s measure transformation*, Mathematics of Control, Signals, and Systems, 28 (2016), pp. 1–55. [MR3519592](#)
- [45] C. D. CHARALAMBOUS AND N. U. AHMED, *Centralized versus decentralized optimization of distributed stochastic differential decision systems with different information structures-part i: A general theory*, IEEE Transactions on Automatic Control, 62 (2017), pp. 1194–1209. [MR3622490](#)
- [46] C. Y. CHONG AND M. ATHANS, *On the periodic coordination of linear stochastic systems*, Automatica, 12 (1976), pp. 321–335. [MR0426925](#)
- [47] J. F. CLAUSER, M. A. HORNE, A. SHIMONY, AND R. A. HOLT, *Proposed experiment to test local hidden-variable theories*, Physical Review Letters, 23 (1969), p. 880.
- [48] R. CLEVE, P. HOYER, B. TONER, AND J. WATROUS, *Consequences and limits of nonlocal strategies*, in Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004., June 2004, pp. 236–249.
- [49] A. COLADANGELO AND J. STARK, *An inherently infinite-dimensional quantum correlation*, Nat Commun, 11 (2020), p. 3335.
- [50] M. H. A. DAVIS AND P. VARAIYA, *Information states for linear stochastic systems*, Journal of Mathematical Analysis and Applications, 37 (1972), pp. 384–402. [MR0484688](#)
- [51] M. H. A. DAVIS AND P. VARAIYA, *Dynamic programming conditions for partially observable stochastic systems*, SIAM Journal on Control, 11 (1973), pp. 226–261. [MR0319642](#)
- [52] J. DIESTEL AND J. UHL, *Vector Measures*, American Mathematical Society, 1977. [MR0453964](#)
- [53] J. M. DONOHUE AND E. WOLFE, *Identifying nonconvexity in the sets*

- of limited-dimension quantum correlations, *Phys. Rev. A*, 92 (2015), p. 062120.
- [54] R. M. DUDLEY, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2nd ed., 2002. [MR1932358](#)
- [55] E. B. DYNKIN AND A. A. YUSHKEVICH, *Controlled Markov processes*, vol. 235, Springer, 1979. [MR0554083](#)
- [56] A. EINSTEIN, B. PODOLSKY, AND N. ROSEN, *Can quantum-mechanical description of physical reality be considered complete?*, *Phys. Rev.*, 47 (1935), pp. 777–780.
- [57] S. N. ETHIER AND T. G. KURTZ, *Markov processes: characterization and convergence*, vol. 282, John Wiley & Sons, 2009. [MR0838085](#)
- [58] K. F. PÁL AND T. VERTESI, *Maximal violation of a bipartite three-setting, two-outcome Bell inequality using infinite-dimensional quantum systems*, *Phys. Rev. A*, 82 (2010). [MR2884097](#)
- [59] E. A. FEINBERG, *Non-randomized Markov and semi-Markov strategies in dynamic programming*, *Theory of Probability & Its Applications*, (1982), pp. 116–126. [MR0645132](#)
- [60] E. A. FEINBERG, *On measurability and representation of strategic measures in markov decision processes*, *Institute of Mathematical Statistics Lecture Notes*. Eds. T. S. Ferguson, L. S. Shapley, J. B. MacQueen, (1996), pp. 29–43.
- [61] M. FISCHER, *On the connection between symmetric  $N$ -player games and mean field games*, *The Annals of Applied Probability*, 27 (2017), pp. 757–810. [MR3655853](#)
- [62] W. FLEMING AND E. PARDOUX, *Optimal control for partially observed diffusions*, *SIAM J. Control Optim.*, 20 (1982), pp. 261–285. [MR0646953](#)
- [63] G. B. FOLLAND, *Real Analysis: Modern Techniques and Their Applications*, John Wiley and Sons, 1999. [MR1681462](#)
- [64] F. FORGES, *An approach to communication equilibria*, *Econometrica: Journal of the Econometric Society*, (1986), pp. 1375–1385. [MR0868148](#)
- [65] I. I. GIHMAN AND A. V. SKOROHOD, *Controlled stochastic processes*, Springer Science & Business Media, 2012. [MR0544839](#)
- [66] I. V. GIRSANOV, *On transforming a certain class of stochastic processes by absolutely continuous substitution of measures*, *Theory of Probability & Its Applications*, 5 (1960), pp. 285–301. [MR0133152](#)
- [67] K. T. GOH, J. KANIEWSKI, E. WOLFE, T. VÉRTESI, X. WU, Y. CAI, Y.-C. LIANG, AND V. SCARANI, *Geometry of the set of quantum correlations*, *Phys. Rev. A*, 97 (2018), p. 022104.
- [68] A. GUPTA, *Existence of team-optimal solutions in static teams with common information: A topology of information approach*, *SIAM Journal on Control and Optimization*, 58 (2020), pp. 998–1021. [MR4082276](#)
- [69] A. GUPTA, S. YÜKSEL, T. BAŞAR, AND C. LANGBORT, *On the existence of optimal policies for a class of static and sequential dynamic teams*, *SIAM Journal on Control and Optimization*, 53 (2015), pp. 1681–1712. [MR3361435](#)
- [70] M. F. HELLWIG, *Sequential decisions under uncertainty and the maxi-*

- mum theorem*, Journal of Mathematical Economics, 25 (1996), pp. 443–464. [MR1394333](#)
- [71] O. HERNÁNDEZ-LERMA AND J. LASSERRE, *Markov Chains and Invariant Probabilities*, Birkhauser, 2003. [MR1974383](#)
- [72] O. HERNÁNDEZ-LERMA AND J. B. LASSERRE, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer, 1996. [MR1363487](#)
- [73] Y. C. HO AND K. C. CHU, *Team decision theory and information structures in optimal control problems - part I*, IEEE Transactions on Automatic Control, 17 (1972), pp. 15–22. [MR0359875](#)
- [74] Y. C. HO AND K. C. CHU, *On the equivalence of information structures in static and dynamic teams*, IEEE Transactions on Automatic Control, 18 (1973), pp. 187–188. [MR0441531](#)
- [75] I. HOGEBOOM-BURR AND S. YÜKSEL, *Zero-sum games involving teams against teams: Existence of equilibria, and comparison and regularity in information*, arXiv:[2206.12048](#), (2022).
- [76] A. S. HOLEVO, *Probabilistic and Statistical Aspects of Quantum Theory*, Edizioni della Normale, 2011. [MR2797301](#)
- [77] M. HUANG, P. E. CAINES, AND R. P. MALHAMÉ, *Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle*, Communications in Information and Systems, 6 (2006), pp. 221–251. [MR2346927](#)
- [78] M. HUANG, P. E. CAINES, AND R. P. MALHAMÉ, *Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria*, IEEE Transactions on Automatic Control, 52 (2007), pp. 1560–1571. [MR2352434](#)
- [79] M. HUANG, P. E. CAINES, AND R. P. MALHAMÉ, *Social optima in mean field LQG control: centralized and decentralized strategies*, IEEE Transactions on Automatic Control, 57 (2012), pp. 1736–1751. [MR2945936](#)
- [80] S. JAFARPOUR AND A. LEWIS, *Locally convex topologies and control theory*, Mathematics of Control, Signals, and Systems, 28 (2016), p. 29. [MR3583883](#)
- [81] Z. JI, A. NATARAJAN, T. VIDICK, J. WRIGHT, AND H. YUEN, *Mip\*=re*, arXiv preprint arXiv:[2001.04383](#), (2020).
- [82] A. KARA AND S. YÜKSEL, *Robustness to incorrect system models in stochastic control*, SIAM Journal on Control and Optimization, 58 (2020), pp. 1144–1182. [MR4091166](#)
- [83] J. KEMPE, H. KOBAYASHI, K. MATSUMOTO, B. TONER, AND T. VIDICK, *Entangled games are hard to approximate*, SIAM Journal on Computing, 40 (2011), pp. 848–877. [MR2823510](#)
- [84] J. KEMPE, O. REGEV, AND B. TONER, *Unique games with entangled provers are easy*, SIAM Journal on Computing, 39 (2010), pp. 3207–3229. [MR2678072](#)
- [85] J. KRAINAK, J. SPEYER, AND S. MARCUS, *Static team problems – part I: Sufficient conditions and the exponential cost criterion*, IEEE Transactions on Automatic Control, 27 (1982), pp. 839–848. [MR0680485](#)
- [86] H. KUHN, *Extensive games and the problem of information*, In Contributions to the Theory of Games, (H. Kuhn and A. Tucker, editors), (1953),

- pp. 193–216. [MR0054924](#)
- [87] A. KULKARNI, *Near-optimality of linear strategies for static teams with 'big' non-gaussian noise*. arXiv:1603.03160, 2016. [MR4060251](#)
- [88] A. KULKARNI, *Local and networked mean-square estimation with high dimensional log-concave noise*, IEEE Trans. Inf. Theory, 64 (2018), pp. 2759–2773. [MR3782286](#)
- [89] A. A. KULKARNI AND T. P. COLEMAN, *An optimizer's approach to stochastic control problems with nonclassical information structures*, IEEE Transactions on Automatic Control, 60 (2014), pp. 937–949. [MR3340786](#)
- [90] H. KUSHNER, *Stochastic stability and control*, Academic Press, New York, 1967. [MR0216894](#)
- [91] D. LACKER, *A general characterization of the mean field limit for stochastic differential games*, Probability Theory and Related Fields, 165 (2016), pp. 581–648. [MR3520014](#)
- [92] D. LACKER, *Probabilistic compactification methods for stochastic optimal control and mean field games*, (2018).
- [93] A. LAMPERSKI AND L. LESSARD, *Optimal decentralized state-feedback control with sparsity and delays*, Automatica, 58 (2015), pp. 143–151. [MR3355640](#)
- [94] S. LANG, *Real and Functional Analysis*, Springer-Verlag, 1993. [MR1216137](#)
- [95] H. LANGEN, *Convergence of dynamic programming models*, Math. Oper. Res., 6 (1981), pp. 493–512. [MR0703092](#)
- [96] J. M. LASRY AND P. L. LIONS, *Mean field games*, Japanese J. of Mathematics, 2 (2007), pp. 229–260. [MR2295621](#)
- [97] E. LEHRER, D. ROSENBERG, AND E. SHMAYA, *Signaling and mediation in games with common interests*, Games and Economic Behavior, 68 (2010), pp. 670–682. [MR2655227](#)
- [98] A. LINDQUIST, *On feedback control of linear stochastic systems*, SIAM Journal on Control, 11 (1973), pp. 323–343. [MR0323425](#)
- [99] D. LUENBERGER, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, NY, 1969. [MR0238472](#)
- [100] A. MAHAJAN, N. MARTINS, M. ROTKOWITZ, AND S. YÜKSEL, *Information structures in optimal decentralized control*, in IEEE Conference on Decision and Control, Hawaii, USA, 2012.
- [101] A. MAHAJAN, N. C. MARTINS, AND S. YÜKSEL, *Static LQG teams with countably infinite players*, in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, IEEE, 2013, pp. 6765–6770.
- [102] L. MASANES, A. ACIN, AND N. GISIN, *General properties of nonsignaling theories*, Phys. Rev. A, 73 (2006), p. 012112.
- [103] E. MASCOLO AND L. MIGLIACCIO, *Relaxation methods in control theory*, Applied Mathematics and Optimization, 20 (1989), pp. 97–103. [MR0989434](#)
- [104] J.-F. MERTENS, S. SORIN, AND S. ZAMIR, *Repeated games*, vol. 55, Cambridge University Press, 2015. [MR3560551](#)
- [105] P. R. MILGROM AND R. J. WEBER, *Distributional strategies for games with incomplete information*, Mathematics of operations research, 10 (1985), pp. 619–632. [MR0812820](#)



- [106] P. MOHAJERIN ESFAHANI, T. SUTTER, D. KUHN, AND J. LYGEROS, *From infinite to finite programs: Explicit error bounds with applications to approximate dynamic programming*, SIAM Journal on Optimization, 28 (2018), pp. 1968–1998. [MR3820392](#)
- [107] A. NATARAJAN AND T. VIDICK, *Two-player entangled games are NP-hard*, in Proceedings of the 33rd Computational Complexity Conference, CCC '18, 2018, pp. 20:1–20:18. [MR3824471](#)
- [108] M. NAVASCUÉS, S. PIRONIO, AND A. ACÍN, *A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations*, New Journal of Physics, 10 (2008), p. 073013.
- [109] A. NAYYAR, A. MAHAJAN, AND D. TENEKETZIS, *Optimal control strategies in delayed sharing information structures*, IEEE Transactions on Automatic Control, 56 (2011), pp. 1606–1620. [MR2848272](#)
- [110] A. NAYYAR, A. MAHAJAN, AND D. TENEKETZIS, *The common-information approach to decentralized stochastic control*, in *Information and Control in Networks*, Editors: G. Como, B. Bernhardsson, A. Rantzer, Springer, 2013. [MR3183777](#)
- [111] M. NIELSEN AND I. CHUANG, *Quantum computation and quantum information*, Cambridge University Press, Cambridge, 2010. [MR1796805](#)
- [112] C. PAPADIMITRIOU AND J. TSITSIKLIS, *Intractable problems in control theory*, SIAM J. Control Optim., 24 (1986), pp. 639–654. [MR0846372](#)
- [113] K. PARTHASARATHY, *Probability Measures on Metric Spaces*, AMS Bookstore, 1967. [MR0226684](#)
- [114] R. PHELPS, *Lectures on Choquet's theorem*, Van Nostrand, New York:, 1966. [MR0193470](#)
- [115] A. B. PIUNOVSKIY, *Controlled random sequences: methods of convex analysis and problems with functional constraints*, Russian Mathematical Surveys, 53 (1998), pp. 1233–1293. [MR1702690](#)
- [116] S. POPESCU AND D. ROHRLICH, *Quantum nonlocality as an axiom*, Foundations of Physics, 24 (1994), pp. 379–385. [MR1265577](#)
- [117] R. RADNER, *Team decision problems*, Ann. Math. Statist., 33 (1962), pp. 857–881. [MR0146937](#)
- [118] E. P. RYAN, *On brockett's condition for smooth stabilizability and its necessity in a context of nonsmooth feedback*, SIAM Journal on Control and Optimization, 32 (1994), pp. 1597–1604. [MR1297100](#)
- [119] N. SALDI, *A topology for team policies and existence of optimal team policies in stochastic team theory*, IEEE Transactions on Automatic Control, 65 (2020), pp. 310–317. [MR4052878](#)
- [120] N. SALDI, S. YÜKSEL, AND T. LINDER, *Finite model approximations and asymptotic optimality of quantized policies in decentralized stochastic control*, IEEE Transactions on Automatic Control, 62 (2017), pp. 2360–2373. [MR3641451](#)
- [121] S. SANJARI, N. SALDI, AND S. YÜKSEL, *Optimality of independently randomized symmetric policies for exchangeable stochastic teams with infinitely many decision makers*, arXiv preprint [arXiv:2008.11570](#), (2020).
- [122] S. SANJARI AND S. YÜKSEL, *Optimal solutions to infinite-player stochas-*

- tic teams and mean-field teams*, IEEE Transactions on Automatic Control, (2020). [MR4226756](#)
- [123] S. SANJARI AND S. YÜKSEL, *Optimal policies for convex symmetric stochastic dynamic teams and their mean-field limit*, SIAM Journal on Control and Optimization, 59 (2021), pp. 777–804. [MR4222180](#)
- [124] M. SCHÄL, *Conditions for optimality in dynamic programming and for the limit of  $n$ -stage optimal policies to be optimal*, Z. Wahrscheinlichkeitsth, 32 (1975), pp. 179–296. [MR0378841](#)
- [125] M. SCHÄL, *On dynamic programming: compactness of the space of policies*, Stochastic Processes and their Applications, 3 (1975), pp. 345–364. [MR0386706](#)
- [126] R. SERFOZO, *Convergence of lebesgue integrals with varying measures*, Sankhyā: The Indian Journal of Statistics, Series A, (1982), pp. 380–402. [MR0705462](#)
- [127] R. SERFOZO, *Convergence of Lebesgue integrals with varying measures*, Sankhya Ser.A, (1982), pp. 380–402. [MR0705462](#)
- [128] W. SLOFSTRA, *The set of quantum correlations is not closed*, in Forum of Mathematics, Pi, vol. 7, Cambridge University Press, 2019. [MR3898717](#)
- [129] E. D. SONTAG, *Mathematical control theory: deterministic finite dimensional systems*, vol. 6, Springer Science & Business Media, 2013. [MR1640001](#)
- [130] D. TENEKETZIS, *On information structures and nonsequential stochastic control*, CWI Quarterly, 9 (1996), pp. 241–260. [MR1434810](#)
- [131] T. VIDICK, *Three-player entangled xor games are NP-hard to approximate*, SIAM Journal on Computing, 45 (2016), pp. 1007–1063. [MR3516877](#)
- [132] A. B. WAGNER AND V. ANANTHARAM, *An improved outer bound for multiterminal source coding*, IEEE Transactions on Information Theory, 54 (2008), pp. 1919–1937. [MR2450842](#)
- [133] B. C. WANG AND J. F. ZHANG, *Social optima in mean field linear-quadratic-gaussian models with markov jump parameters*, SIAM Journal on Control and Optimization, 55 (2017), pp. 429–456. [MR3609229](#)
- [134] J. WARGA, *Optimal control of differential and functional equations*, Academic press, 2014. [MR0372708](#)
- [135] J. WATROUS, *The Theory of Quantum Information*, Cambridge University Press, 2018.
- [136] R. L. WHEEDEN AND A. ZYGMUND, *Measure and Integral*, Marcel Dekker, New York, 1977. [MR0492146](#)
- [137] H. S. WITSENHAUSEN, *A counterexample in stochastic optimal control*, SIAM J. Contr., 6 (1968), pp. 131–147. [MR0231649](#)
- [138] H. S. WITSENHAUSEN, *A counterexample in stochastic optimum control*, SIAM J. on Control and Optimization, 6 (1968), pp. 131–147. [MR0231649](#)
- [139] H. S. WITSENHAUSEN,, *On information structures, feedback and causality*, SIAM J. Control, 9 (1971), pp. 149–160. [MR0386831](#)
- [140] H. S. WITSENHAUSEN, *A standard form for sequential stochastic control*, Mathematical Systems Theory, 7 (1973), pp. 5–11. [MR0396029](#)
- [141] H. S. WITSENHAUSEN, *The intrinsic model for discrete stochastic control:*

- Some open problems*, Lecture Notes in Econ. and Math. Syst., Springer-Verlag, 107 (1975), pp. 322–335.
- [142] H. S. WITSENHAUSEN, *Equivalent stochastic control problems*, Math. Control, Signals and Systems, 1 (1988), pp. 3–11. [MR0923272](#)
- [143] W. M. WONHAM, *On the separation theorem of stochastic control*, SIAM Journal on Control, 6 (1968), pp. 312–326. [MR0237219](#)
- [144] Y. WU AND S. VERDÚ, *Witsenhausen’s counterexample: a view from optimal transport theory*, in Proceedings of the IEEE Conference on Decision and Control, Florida, USA, December 2011, pp. 5732–5737. [MR1819812](#)
- [145] T. YOSHIKAWA, *Dynamic programming approach to decentralized control problems*, IEEE Transactions on Automatic Control, 20 (1975), pp. 796–797. [MR0403777](#)
- [146] L. C. YOUNG, *Generalized curves and the existence of an attained absolute minimum in the calculus of variations*, Comptes Rendus de la Societe des Sci. et des Lettres de Varsovie, 30 (1937), pp. 212–234.
- [147] S. YÜKSEL, *Stochastic nestedness and the belief sharing information pattern*, IEEE Transactions on Automatic Control, 54 (2009), pp. 2773–2786. [MR2583718](#)
- [148] S. YÜKSEL, *A universal dynamic program and refined existence results for decentralized stochastic control*, SIAM Journal on Control and Optimization, 58 (2020), pp. 2711–2739. [MR4146376](#)
- [149] S. YÜKSEL AND T. BAŞAR, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*, Springer, New York, 2013. [MR3088779](#)
- [150] S. YÜKSEL AND T. LINDER, *Optimization and convergence of observation channels in stochastic control*, SIAM J. on Control and Optimization, 50 (2012), pp. 864–887. [MR2914232](#)
- [151] S. YÜKSEL AND N. SALDI, *Convex analysis in decentralized stochastic control, strategic measures and optimal solutions*, SIAM J. on Control and Optimization, 55 (2017), pp. 1–28. [MR3590643](#)