# Chapter 2
# Networked Control Systems as Stochastic Team Decision Problems: A General Introduction

## 2.1 Introduction

*Networked control systems* can be viewed as stochastic decision problems with dynamic decentralized information structures or as stochastic dynamic teams, with each subcontroller viewed as an *agent* in a dynamic team. The goal of this introductory chapter is accordingly to introduce the reader to a general mathematical formulation of stochastic teams, first with static and then with dynamic information structures, and to discuss some salient features of these decision problems and associated solution concepts through some simple but illustrative examples.

The chapter discusses both *static* stochastic teams (i.e., team decision problems where the information signals received by the decision makers are not affected by actions) and *dynamic* stochastic teams (where the information of at least one decision maker is affected by action). Sections 2.2, 2.3, and 2.6 deal with static teams, whereas Sects. 2.4 and 2.5 discuss dynamic teams. Section 2.2 provides a general formulation for static teams, which is followed by a complete analysis of a finite stochastic team problem under various information patterns, in Sect. 2.3. Section 2.6 provides some general explicit results on existence, uniqueness, and characterization of optimal solutions first for general static teams and then for special classes of teams with Gaussian statistics: those with quadratic and exponentiated quadratic costs.

Sections 2.4 and 2.5 can be viewed as the counterparts of Sects. 2.2 and 2.3 for dynamic teams. First a precise mathematical formulation for dynamic team decision problems is given, in Sect. 2.4, along with various dynamic information structures and appropriate solution concepts, and then an illustrative example of a finite dynamic team is provided in Sect. 2.5, within the framework of which some important features of optimal solutions in teams are discussed. The chapter concludes with Sect. 2.7 which provides some bibliographical notes and guidelines for further reading on the topics covered herein.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_2,

## 2.2 A Mathematical Framework For Static Decision Problems

Multiple person stochastic decision problems could be formulated with varying degrees of generality, abstraction, and rigor, depending on the types of problems to be solved (*i.e.*, the scope of coverage) and the level of mathematical sophistication to be expected from the reader. Common to all possible formulations, however, is the specification of **five** basic **ingredients** which are essential for a well-founded mathematical treatment of *decision making under uncertainty*. These are:

1. The number of **decision makers** (synonymously, *agents* or *controllers*) and the sets of alternative **actions** (synonymously, *decisions* or *controls*) available to them
2. The **uncertainty** and its probabilistic description
3. The **information** acquired by each decision maker on the uncertainty and the previous actions
4. The **payoff** (or *loss*) that accrues to each decision maker as a result of joint actions (over the decision period) and realization of uncertainty
5. A **solution concept** whereby "best" or "satisfactory" decision rules can be chosen

Before going into further specification of these entities, let us pause to introduce some terminology and notation which will be needed in the sequel. We will refer to a decision problem as **static** if the information available to each decision maker is independent of the actions of other decision makers (this statement will be made precise later in the section as well as in Sect. 3.8); otherwise, the decision problem is said to be **dynamic**. We will refer to decision makers interchangeably as *agents* or *controllers*, with the *i*th one denoted $\mathbf{A}i$, where $i$ takes values in the set $\mathcal{N} := \{1, \ldots, N\}$ which is called the *agent* (*decision maker*) *set*. The variable under the control of each decision maker will be called the *action* (synonymously, *decision* or *control*) *variable* and will be denoted by $u^i$ for $\mathbf{A}i$. Each $u^i$ will take values in a given *action set* to be denoted by $U^i$. Finally, the $N$-tuple $(u^1, \ldots, u^N)$ will be denoted by $\mathbf{u}$ and the product action space $U^1 \times \cdots \times U^N$ by $\mathbf{U}$.

### *Basic Ingredients of Static Decision Problems*

In the static framework we will initially study the class of problems where the action sets, $U^i, i \in \mathcal{N}$, are either (finitely or infinitely) countable or uncountable but finite dimensional. In the latter case, we take the action set (space) to be isomorphic to the Euclidean[1] space $\mathbb{R}^{m_i}$, for some integer $m_i, i \in \mathcal{N}$; furthermore, if there are any

[1]Some background material on sets and topological notions can be found in Appendix A.

constraints imposed on the action variable $u^i$, we introduce the *action constraint set* $S^i$, for $\mathbf{A}i$, as a proper subset of $U^i$.

The uncertainty in the decision problem is captured in the so-called random state of **nature**, $\xi$, which is a random variable (or vector) defined on a given probability space $(\Omega, \mathcal{F}, P_\Omega)$[2] and taking values in the Borel space $(\Xi, \mathcal{B}(\Xi))$ where either $\Xi \equiv \mathbb{R}^m$ for some positive integer $m$ or $\Xi$ is a countable set. Let $P$ be the probability measure induced by $\xi$ on $(\Xi, \mathcal{B}(\Xi))$, corresponding to $P_\Omega$. To save from notation, the corresponding probability distribution function will also be denoted by $P$.

The decision makers do not, in general, have direct access to the true state of nature but instead observe the value of some other variable, known as the *measurement* (or *information*) *signal*. To define this quantity in precise mathematical terms, let us first introduce, for each $i \in \mathcal{N}$, the *information field*, $\mathbf{Y}^i$, for agent $\mathbf{A}i$ as a given sub $\sigma$-field of $\mathcal{B}(\Xi)$, generated by a measurable function $\eta^i$ mapping $(\Xi, \mathcal{B}(\Xi))$ onto $(Y^i, \mathbf{B}^i)$. This is known as the *information function* for $\mathbf{A}i$, and the $N$-tuple $\underline{\eta} := (\eta^1, \ldots, \eta^N)$ is called the *information structure* (or *information pattern*) of the decision problem. The information function $\eta^i$ induces a $\sigma$-field, $\mathbf{Y}^i$, of $\Xi$, and the information (measurement) signal $y^i$ of $\mathbf{A}i$ (which lies in the *measurement set* $Y^i$) is generated according to $\eta^i$, which is symbolically written as

$$y^i = \eta^i(\xi) \equiv \tilde{\eta}^i(\omega)\,, \tag{2.1}$$

where the latter relates the measurement signal directly to the original probability space $\Omega$, with elements $\omega$. This is sometimes a more convenient representation to work with, especially if $\Omega$ is finite or countable. In that case one can consider $\Omega$ and $\Xi$ to be essentially the same set and thereby view $\mathbf{Y}^i$ also as a partition of $\Omega$, which is a convention we henceforth adopt. In the case of finite probability spaces we will also adopt the convention, perhaps by a slight abuse of notation and terminology, that the measurement signal $y^i$ can be considered as an element of the partition set $\mathbf{Y}^i$.

The decision makers determine their actions using the measurement signals that they receive, under the *strategies* that they adopt for transforming measurements into actions. The *strategy* (synonymously, *decision rule* (function) or *control law*) of $\mathbf{A}i$ will be denoted by $\gamma^i$ and is formally defined as a measurable mapping from $(\Xi, \mathbf{Y}^i)$ into the space $(U^i, \mathbf{B}_{U^i})$. This can also be written as a measurable mapping from $(Y^i, \mathbf{B}^i)$ to $(U^i, \mathbf{B}_{U^i})$, as we state explicitly below. We denote the set of all such mappings, which also satisfy the additional constraints that may have been imposed on $u^i$, by $\Gamma^i$, to be called the *strategy space* of $\mathbf{A}i$, and note the relationship

$$u^i = \gamma^i(y^i) = \gamma^i(\eta^i[\xi]),$$

[2]Necessary background material on probability theory, along with an explanation of the terminology and notation used here, can be found in Appendix B.

where the latter relates the action variable to the state of nature, $\xi$. We will denote the $N$-tuple $(\gamma^1, \ldots, \gamma^N)$ by $\underline{\gamma}$, and the product strategy space $\Gamma^1 \times \cdots \times \Gamma^N$ by $\mathbf{\Gamma}$. The individual strategy spaces $\Gamma^i$ may also include the additional structural constraints that may have been imposed on the policies, such as linearity. What is not allowed in general, however, is for $\mathbf{\Gamma}$ to be *nonrectangular*, that is, for the choice out of $\Gamma^i$ (for some $i$) to restrict the choice out of $\Gamma^j (j \neq i)$. For example, our formulation (at this point) does not cover "cross-constraints" of the type $f(u^i, u^j) \leq 0$, $i \neq j$, for some functional $f$.

Given an ($N$+1)-tuple $(\xi, \mathbf{u}) \in \Xi \times \mathbf{U}$, the loss incurred to the decision makers viewed collectively as a *team* will be denoted by $L(\xi, \mathbf{u})$, where the function $L : \Xi \times \mathbf{U} \to \mathbb{R}$ is known as the *loss function* for the team. Its negative, $-L(\xi, \mathbf{u}) =: \mathtt{U}(\xi, \mathbf{u})$, is known as the *payoff function*, which all agents collectively want to "maximize," in a sense to be defined shortly. Implicit here is the assumption that for the team there exists a unique (up to equivalence) utility function which numerically orders different outcomes corresponding to joint actions and realization of the state of nature, in a way consistent with the team's preference ordering among different alternatives.

The loss incurred is generally a random quantity, the randomness appearing through both $\xi$ and $\mathbf{u}$, where the latter depends on $\xi$ through the measurement signals and the strategies adopted by the decision makers. Therefore, one rather works with the expected value of this quantity, which we will be referring to as the *cost function*.[3] Other possible terminology would be *expected loss function*, *average risk*, or *expected cost*, all of which have been used in the literature, which we will also use interchangeably. The cost function, $J : \mathbf{\Gamma} \to \mathbb{R}$, is defined on the product strategy space $\mathbf{\Gamma}$ as[4]

$$J(\underline{\gamma}) = \int_{\Xi} L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))P(d\xi) = E[L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))] =: E_\xi L(\xi, \underline{\gamma}(\underline{\eta}[\xi])), \quad (2.2)$$

where

$$\underline{\gamma}(\underline{\eta}[\xi]) := (\gamma^1(\eta^1[\xi]), \ldots, \gamma^N(\eta^N[\xi])) \quad (2.3)$$

and $E_\xi$ is the operator that takes the expected value of the quantity it precedes, over $\xi$. To show the explicit dependence of $J$ on also the information structure $\underline{\eta}$, we will sometimes use the notation $J(\underline{\gamma}, \underline{\eta})$ and occasionally use $J(\underline{\gamma}, \underline{\eta}; L, P)$ to also indicate the dependence on the loss function $L$ and the probability distribution $P$.

The specification of $J$, along with the product strategy space $\mathbf{\Gamma}$, provides a complete characterization (aside from the solution concept) of a stochastic multiple person decision problem and is known as the *normal form description*. Note that in

[3]There would be other ways of making the objective function deterministic, such as defining the cost function as the probability of the loss exceeding a given ceiling or taking it as the supremum of the loss function over $\omega \in \Omega$. We will not be devoting much discussion to such formulations in the book.

[4]See Appendix B for an explanation of the notation used here.

this description, the information structure is suppressed and it enters the problem formulation only through the strategy spaces $\Gamma^i$, $i \in \mathcal{N}$. The description which lays out explicitly the dependence of the measurement signals on the unknown state of nature is known as the *extensive form description* of the underlying (static) stochastic decision problem. The distinction between these two forms should be more transparent when we introduce dynamic decision problems, later in this chapter. We should note, however, that the two forms are in fact equivalent in the sense that they both uniquely characterize a given stochastic decision problem; the essential difference is that sometimes it is more convenient to work with one form than the other.

## *Notion of Optimality*

In the framework laid out above, it would have been possible to endow each decision maker (agent) with a different loss function and also possibly a different subjective probability measure regarding the unknown state of nature. Either of these departures would take us outside the realm of team problems and necessitate consideration of the more general framework of stochastic (zero-sum or nonzero-sum) games, with associated solution concepts, such as *saddle-point equilibrium* or *Nash equilibrium* [32]. Covering this more general framework is outside the scope and the goals of this book, as here our interest is in problems originating in networked control systems, where decision makers have common objectives and act as a *team*, even though the information may not be centralized. More precisely:

> A **team** is a collection of individual decision makers who strive for the same goal, using the same (probabilistic) model of the underlying decision process, but not necessarily sharing the same online information (such as measurements) on the uncertainty.

For an $N$-person stochastic team problem, since all agents will be striving toward the same goal, with team preferences quantified in the given loss functional, the only reasonable solution that leads to optimal behavior is the global minimization of the team cost over the product strategy space. Hence, we have

**Definition 2.2.1.** For a given stochastic team problem with a fixed information structure, $\{J; \Gamma^i, i \in \mathcal{N}\}$, a strategy $N$-tuple $\underline{\gamma}^* := ({\gamma^1}^*, \ldots, {\gamma^N}^*) \in \mathbf{\Gamma}$ is an *optimal team decision rule* (synonymously, *team-optimal decision rule* or simply *team-optimal solution*) if

$$J(\underline{\gamma}^*) = \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}) =: J^*, \tag{2.4}$$

provided that such a strategy exists. The cost level achieved by this strategy, $J^*$, is the *minimum (or optimal ) team cost.* ◇

In the above definition of "optimality in a team," we have taken the information structure as fixed and given *a priori*. Even though the class of systems one typically encounters are primarily of this type, it is worth mentioning that it is possible to consider the information structure of the problem as a variable, alongside the strategies of the agents. In fact, in the theory of organizations (as well as to a large extent in the design of networked systems), the prime goal is to obtain an optimal design for the pair $(\underline{\gamma}; \underline{\eta})$ which is known as the *organizational form* (Marschak and Radner [255]). Of course, to make the problem meaningful, we have to impose some restrictions on $\underline{\eta}$ (such as belonging to some prescribed class of comparable information structures, say $\underline{N}$) or attach some cost to it which would be directly proportional with its *value*.[5] In the absence of such realistic restraints on $\underline{\eta}$, the problem will admit the trivial solution where $\underline{\eta}^*$ (the optimal $\underline{\eta}$) allows the agents to acquire perfect information on the state of nature, $\xi$, and thereby $\underline{\gamma}^*$ to depend directly on $\xi$. Under realistic organizational constraints, however, say with $\underline{\eta} \in \underline{N}$, an optimal design $(\underline{\gamma}^*; \underline{\eta}^*) \in \mathbf{\Gamma} \times \underline{N}$ will have the property that there exists no $\underline{\eta} \in \underline{N}$ such that

$$\inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}; \underline{\eta}) < J(\underline{\gamma}^*; \underline{\eta}^*) \tag{2.5}$$

where the cost function $J$ may also include some additional (possibly additive) terms reflecting the costs associated with various $\underline{\eta}$'s. Furthermore, the policy space $\mathbf{\Gamma}$ implicitly depends on the choice out of $\underline{N}$, so that the product $\mathbf{\Gamma} \times \underline{N}$ is actually not *rectangular*. Note that a natural way of obtaining an optimal organizational form would be to minimize the function $J(\underline{\gamma}^*_{\underline{\eta}}; \underline{\eta})$ over $\underline{\eta} \in \underline{N}$, where $\underline{\gamma}^*_{\underline{\eta}}$ is the team-optimal solution corresponding to the fixed information structure $\underline{\eta}$. We use a subscript on $\underline{\gamma}^*$ here to explicitly point out the fact that the team-optimal solution depends on $\underline{\eta}$ structurally and in general in a fairly complicated manner, which makes the further optimization of $J$, as $\underline{\eta}$ varies over $\underline{N}$, a rather complex problem (not of the standard type), unless the cardinality of $\underline{N}$ is finite.

One important feature of the team-optimal solution that is worth mentioning at this point (perhaps as a cautionary remark) is that multiple solutions are *not* necessarily *interchangeable*. For a two-person team problem, for example, if the pairs of policies $(\gamma^1, \gamma^2)$ and $(\beta^1, \beta^2)$ are two team-optimal policy pairs, then it is not necessarily true that the pair $(\gamma^1, \beta^2)$ will also constitute a team-optimal solution. Hence, in case of nonuniqueness of the solution, the agents need to have a common consistent rule as to which one of the possible solutions to adopt, in order to arrive at the optimum. This may require some pre-communication and pre-commitment to some protocols among the agents.

A weaker solution concept than that of team-optimality introduced in Definition 2.2.1 is that of *person-by-person optimality*, equivalently *Nash equilibrium*, introduced next.

---

[5]At this point this is a rather imprecise statement. The precise meaning of **value** of a given information structure and the notion of one information structure being more **valuable** (or **better**) than another one will be introduced and studied in the next chapter, particularly Sect. 3.2.

**Definition 2.2.2.** For a given $N$-person stochastic team with a fixed information structure, $\{J; \Gamma^i, i \in \mathcal{N}\}$, an $N$-tuple of strategies $\underline{\gamma}^* := (\gamma^{1^*}, \ldots, \gamma^{N^*})$ constitutes a *Nash equilibrium* (synonymously, a *person-by-person optimal* (pbp optimal) solution) if, for all $\beta \in \Gamma^i$ and all $i \in \mathcal{N}$, the following inequalities hold:

$$J^* := J(\underline{\gamma}^*) \leq J(\underline{\gamma}^{-i}*, \beta), \tag{2.6}$$

where we have adopted the notation

$$(\underline{\gamma}^{-i}*, \beta) := (\gamma^1*, \ldots, \gamma^{i-1}*, \beta, \gamma^{i+1}*, \ldots, \gamma^N*). \tag{2.7}$$

$\diamond$

*Remark 2.2.1.* Nash equilibrium is a weaker solution concept than team-optimality (*cf.* Definition 2.2.1), since satisfaction of the $N$ inequalities (2.6) is clearly necessary but not sufficient for $\underline{\gamma}^*$ to be an optimal team decision rule. But, since every team-optimal solution is necessarily a *pbp* optimal solution, the latter plays an important role in the derivation of the former, as we will see later in the book, with the first demonstration being in Sect. 2.6. $\diamond$

In the next section, we depart from the abstract formulation of the present section and provide an illustrative example which will aid in better understanding of the concepts introduced above.

## 2.3 An Illustrative Example of a Finite Stochastic Team

A stochastic team problem is said to be *finite* if both the action and the uncertainty sets are finite. In this case (as we have indicated earlier) there is no need to make any distinction between $\Omega$ and $\Xi$,[6] and one may as well work in the original probability space $(\Omega, \mathcal{F}, P_\Omega)$ where the probability measure will be replaced with the probability masses $\{p_j = P_\Omega(\{\omega_j\})\}_{j=1}^{\#}$ where $\omega_j$ is an element of $\Omega$ with positive probability, and $\# := |\Omega|$, the cardinality of $\Omega$, with those elements of $\Omega$ receiving *zero* probability from $P_\Omega$ being irrelevant to the decision problem and therefore deleted. By a possible abuse of terminology, we will call the $\#$ elements of $\Omega$ the states of nature. We note that $\mathcal{F}$ is a collection of all subsets of $\Omega$ (hence it has $2^{\#}$ elements), and $\mathbf{Y}^i$ can be taken, without any loss of generality, as a partition of $\Omega$, for each $i \in \mathcal{N}$.

Every two-person finite stochastic static team can be represented by a family of matrices, each matrix (and there will be $\#$ of them) corresponding to a different state of nature, $\omega$. The rows of these matrices would correspond to action choices of one agent, say **A**1, the columns would correspond to action choices of the other agent, **A**2, and each entry would be the corresponding loss to the team for that particular $\omega$.

[6] Actually here the only requirement is that the uncertainty set be finite.

This, together with a specification of the class of all possible information signals, $(\mathbf{Y}^1, \mathbf{Y}^2)$, would constitute the *extensive form* description for the team. Such a set-up can also naturally be extended to $N$-person finite static teams, where now the matrices are replaced by $N$-dimensional hypercubes.

One approach (and a universally applicable one) toward obtaining the team-optimal solution(s) of such finite static teams is to convert the above *extensive form* into a *normal form* by relating the strategies of the agents directly to the (expected) costs that accrue to the team. As we have indicated earlier, such a formulation would suppress the information signals as well as the role of nature in the decision problem, and it would involve only a single finite, albeit larger dimensional, matrix (or hypercube, if there are more than two agents) whose columns and rows are strategy choices of the agents and whose lowest entry (or entries) would yield the team-optimal solution. Note that for $\mathbf{A}i$ the number of alternative strategies (*i.e.,* $|\Gamma^i|$) would be $|U^i|^{|\mathbf{Y}^i|}$, and hence a derivation based solely on the normal form could easily get intractable if either the number of information signals or the cardinality of the action set for at least one agent is large. It is therefore necessary to look for alternative ways of obtaining the solution, by also exploiting the nature of the information available to the agents. Note that the solution to a finite static stochastic team problem always exists (but it may be nonunique), since it involves optimization over a finite set.

With this prelude, we consider in this section a two-person static stochastic team problem where $U^1 = \{U(\text{up}), D(\text{down})\}$, $U^2 = \{L(\text{left}), R(\text{right})\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $p_1 = p_2 = 0.3$, $p_3 = 0.4$, and the loss matrices are given by

| | | A2 | | | A2 | | | A2 | |
|---|---|---|---|---|---|---|---|---|---|
| | | L | R | | L | R | | L | R |
| A1 | U | 1 | 0 | U | 2 | 3 | U | 1 | 2 |
| | D | 3 | 1 | D | 2 | 1 | D | 0 | 2 |
| $\omega$ : | $\omega_1 \leftrightarrow 0.3$ | | | $\omega_2 \leftrightarrow 0.3$ | | | $\omega_3 \leftrightarrow 0.4$ | | |

Under various information structures for the team, we now study the derivation of team-optimal decision rules and some of their properties.

1. *Perfect measurements*

Here both agents have access to the true state of nature, and hence $\mathbf{Y}^1 = \mathbf{Y}^2 = \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$, the $\sigma-$field generated by the singletons. The cardinality of the strategy spaces ($\Gamma^1$ and $\Gamma^2$) is $2^3 = 8$ each, and hence the normal form is an $8 \times 8$ matrix, requiring a comparison of $64$ entries. The normal form of a decision problem could also be called a *pre-commitment model*, since the strategies of the agents tell them what to do under all possible realizations of the information signal, even before the actual state of nature is realized. If, however, the agents wait to make their decisions until after they receive the measurements (which we may call the *post-commitment scenario*), then the dimension of the problem could be reduced significantly. This is particularly true when the agents' measurements are identical, as in the present case, where intuition tells us that we may, without any loss of generality, obtain the minimum value of $L(\omega; u^1, u^2)$ for each $\omega \in \Omega$ and then construct the optimal decision rules from the solutions of these individual (deterministic) teams. A mathematical justification for this intuitively appealing approach follows from the identity

$$\begin{aligned} J^* := \min_{\gamma^1,\gamma^2} J(\gamma^1, \gamma^2) &= \min_{\gamma^1,\gamma^2} E_\omega L(\omega; \gamma^1(\omega), \gamma^2(\omega)) \\ &\equiv E_\omega \{ \min_{u^1,u^2} L(\omega; u^1, u^2) \}, \end{aligned} \tag{2.8}$$

which is true since the agents have perfect measurements on $\omega$. Note that the inner minimization in (2.8) involves the minimization of 3 loss matrices with four elements each, while the normal form required the minimization of a cost matrix with $64$ elements.

The individual minima of $L(\omega; u^1, u^2)$ are

$$\min L(\omega_1; u^1, u^2) = L(\omega_1; U, R) = 0,$$

$$\min L(\omega_2; u^1, u^2) = L(\omega_2; D, R) = 1,$$

$$\min L(\omega_3; u^1, u^2) = L(\omega_3; D, L) = 0,$$

which lead [from (2.8)] to $J^* = 0.3$ and the unique team-optimal decision rules:

$$\gamma^{1*}(\omega) = \begin{cases} U, \ \omega = \omega_1, \\ D, \text{ else}, \end{cases} \qquad \gamma^{2*}(\omega) = \begin{cases} L, \ \omega = \omega_3, \\ R, \text{ else}, \end{cases}$$

which we rewrite symbolically as

$$\underline{\gamma}^* = (UDD, RRL),$$

a convention we adopt (and will henceforth use) for representing strategies in finite spaces.

As a final note we point to the observation that even though the policy pair $(UDD, RRL)$ is unique as a team-optimal solution (which is also, by definition, *pbp* optimal), it is not the unique *pbp* optimal solution. The policy pair $(UUD, RLL)$ is also *pbp* optimal, but it carries the unfavorable cost of $0.6$ which is significantly higher than $J^*$.

2. *Imperfect identical measurements*

Here we consider the situation where the agents can distinguish only between the pair $(\omega_1, \omega_2)$ and the singleton $\omega_3$, and hence $\mathbf{Y}^1 = \mathbf{Y}^2 =: \mathbf{Y} = \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$. The strategy spaces have four elements each, leading to a $4 \times 4$

matrix as the normal form. We write out this matrix, for instructional purposes, with the notation $\gamma^i(y^i) = (a, b)$ (with $a, b$ denoting the possible actions of the agents) standing for

$$\gamma^i(y^i) = \begin{cases} a, \; y^i = \{\omega_1, \omega_2\}, \\ b, \; \text{else}, \end{cases}$$

**A**2

| | | LL | LR | RL | RR |
|---|---|---|---|---|---|
| **A**1 | UU | 1.3 | 1.7 | 1.3 | 1.7 |
| | UD | 0.9 | 1.7 | 0.9 | 1.7 |
| | DU | 1.9 | 2.3 | 1.0 | 1.4 |
| | DD | 1.5 | 2.3 | 0.6* | 1.4 |

The matrix has a unique minimum entry, as indicated, and hence the team problem under the given information pattern admits the unique optimal solution $\underline{\gamma}^* = (DD, RL)$, yielding a cost level of $J^* = 0.6$. Note that this is twice the optimal cost level attained under the perfect state measurements, and we can refer to the difference between the two (informally) as the "value" of the additional measurement which enables the agents to distinguish between the two states $\omega_1$ and $\omega_2$. Note also that in addition to the team-optimal solution given above, the problem admits one other *pbp* optimal solution, which is $(UD, LL)$, with a corresponding (unfavorable) cost level of 0.9.

An alternative derivation for the team-optimal solution, which would involve lower-dimensional matrices, follows from a reasoning similar to the one used for the perfect information case. Here the counterpart of (2.8) would be

$$\begin{aligned} J^* := \min_{\gamma^1,\gamma^2} J(\gamma^1, \gamma^2) &= \min_{\gamma^1,\gamma^2} E_\omega L(\omega; \gamma^1(y), \gamma^2(y)) \\ &\equiv E_y\{\min_{u^1,u^2} E_{\omega|y} L(\omega; u^1, u^2)\}, \end{aligned} \tag{2.9}$$

where we have used the "iterated property" of the conditional expectation: $E_\omega = E_y E_{\omega|y}$ where $E_{\omega|y}$ is the conditional expectation of the random variable it precedes, given that $y \in \mathbf{Y}$ has been observed.[7] Also, since we are operating in finite spaces, expression (2.9) is well defined and thus we are allowed to interchange the operations of outer expectation (over $y \in \mathbf{Y}$) and minimization (over $\underline{\gamma} \in \Gamma$). Now, the inner minimization in (2.9) involves two matrices, corresponding to two different (and exhaustive) choices for $y : y_1 = \{\omega_1, \omega_2\}$ and $y_2 = \{\omega_3\}$. These

[7] For this and other properties of conditional expectation the reader is referred to Appendix B.

matrices, which we may call *conditional cost matrices*, are as follows, with the unique optimal solution indicated in each case[8]:

**A2** ($y$: $y_1 \leftrightarrow 0.6$)

| **A1** | L | R |
|---|---|---|
| U | 1.5 | 1.5 |
| D | 2.5 | 1.0* |

**A2** ($y_2 \leftrightarrow 0.4$)

| **A1** | L | R |
|---|---|---|
| U | 1 | 2 |
| D | 0* | 2 |

Since $y_1$ occurs with probability $0.6$ and $y_2$ with probability $0.4$, the (average) optimal team cost is $J^* = (0.6)(1) + (0.4)(0) = 0.6$, attained by the unique pair of decision rules $(DD, RL)$. Note that the first matrix admits one other *pbp* optimal solution $(U, L)$ which, together with the team-optimal solution of the second matrix, leads to a *pbp* optimal solution for the original team, $(UD, LL)$, which is the one found earlier using the $4 \times 4$ normal form.

3. *No measurements*

When neither agent makes any measurements, $\mathbf{Y}^1$ and $\mathbf{Y}^2$ are trivial $\sigma-$fields $\{\emptyset, \Omega\}$, and hence all permissible decision rules are *constant* maps. The normal form is the $2 \times 2$ matrix

**A2**

| **A1** | L | R |
|---|---|---|
| U | 1.3* | 1.7 |
| D | 1.5 | 1.4 |

from which we immediately read: $J^* = 1.3$ and $\underline{\gamma}^* = (U, L)$.

4. *Nonidentical measurements: Perfect for* **A**2 *and none for* **A**1

This is the first nonsymmetric information structure that we will be studying. The information sets are $\mathbf{Y}^2 = \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$, $\mathbf{Y}^1 = \sigma(\{\omega_1, \omega_2, \omega_3\})$, leading to *eight* elements for $\Gamma^2$ and *two* for $\Gamma^1$. The normal form is given by the *two-by-eight* matrix

**A2**

| **A1** | LLL | LLR | LRL | LRR | RLL | RLR | RRL | RRR |
|---|---|---|---|---|---|---|---|---|
| U | 1.3 | 1.7 | 1.6 | 2.0 | 1.0 | 1.4 | 1.3 | 1.7 |
| D | 1.5 | 2.3 | 1.2 | 2.0 | 0.9 | 1.7 | 0.6* | 1.4 |

[8]The entries of the first matrix are obtained from the relationship $E_{\omega|y_1} L(\omega; u^1, u^2) = L(\omega; u^1, u^2)p_{1|1} + L(\omega; u^1, u^2)p_{2|1}$, where $p_{1|1}$ and $p_{2|1}$ are the conditional probabilities, each equal to $1/2$.

and the unique team-optimal solution is, as indicated, $\underline{\gamma}^* = (D, RRL)$, and the value is $J^* = 0.6$. Note that the optimal cost is the same here as in case 2, even though the information structures are incomparable. (As compared with case 2, here $\mathbf{A}1$ has **worse** and $\mathbf{A}2$ has **better** information, in the sense that $\mathbf{Y}^1_{(4)} \subset \mathbf{Y}^1_{(2)}$ and $\mathbf{Y}^2_{(4)} \supset \mathbf{Y}^2_{(2)}$, where the subscripts on $\mathbf{Y}$ refer to the two different cases and inclusion is a strict one.[9])

The question arises now as to whether a procedure similar to those used in cases 1 and 2 could also be used here to simplify the derivation (i.e., to avoid working with a large dimensional matrix, each entry of which has to be computed). Clearly, an identity such as (2.9) cannot be used since the agents do not make identical measurements. However, for each fixed decision rule $\gamma^1$ of $\mathbf{A}1$ (and there are only *two*), one can obtain the best response (minimizing solution) $T(\gamma^1)$ for $\mathbf{A}2$ by using the original matrices, since $\mathbf{A}2$ has perfect information:

$$\gamma^1 = U \Rightarrow \gamma^2 = T(U) = (RLL) \Rightarrow J(U, T(U)) = 1.0,$$
$$\gamma^1 = D \Rightarrow \gamma^2 = T(D) = (RRL) \Rightarrow J(D, T(U)) = 0.6.$$

The best choice for $\mathbf{A}1$, then, is ${\gamma^1}^* = D$, and the corresponding best response for $\mathbf{A}2$ is ${\gamma^2}^* = T(D) = (RRL)$, thus agreeing with what we had obtained earlier.

The above is yet another procedure for obtaining the team-optimal solution in two-person stochastic teams: Fix the policies of one of the agents (preferably the one whose strategy space has fewer elements), obtain the best response of the other agent to each such policy, and compute the corresponding (average) team cost in each case. The lowest such cost is the optimal team cost, and the corresponding policies are the team-optimal decision rules. Such a procedure is always justified because of the following sequence of identities (where we have taken $\mathbf{A}1$ as the starting agent):

$$\begin{aligned} J^* &= \min_{\gamma^1} \min_{\gamma^2} E_\omega L(\omega; \gamma^1(y^1), \gamma^2(y^2)) \\ &\equiv \min_{\gamma^1} E_{y^2} \left\{ \min_{u^2} E_{\omega|y^2} L(\omega; \gamma^1(y^1), u^2) \right\} \\ &\equiv \min_{\gamma^1} E_\omega L(\omega; \gamma^1(y^1), T(\gamma^1)(y^2)). \end{aligned}$$

This would be applicable even if $\mathbf{Y}^1$ and $\mathbf{Y}^2$ do not satisfy an inclusion relationship (in the particular case above we had $\mathbf{Y}^1 \subset \mathbf{Y}^2$), but then one has to construct new *conditional cost matrices* ($|\mathbf{Y}^2|$ of them, each of dimension $|\Gamma^1|$-by-$|U^2|$) in order to obtain the optimal response of $\mathbf{A}2$. [See the next case for an information structure of the type where the inclusion does not hold.]

[9] An equivalent statement would be $\Gamma^1_{(4)} \subset \Gamma^1_{(2)}$ and $\Gamma^2_{(4)} \supset \Gamma^2_{(2)}$. A more formal treatment of comparison of two information structures will be done in the next chapter, in Sect. 3.2.

5. *Nonidentical imperfect measurements*

This case will serve to illustrate a point which is sometimes very useful in the derivation of team-optimal solutions. Consider the information structure given by $\mathbf{Y}^1 = \sigma(\{\{\omega_1\}, \{\omega_2, \omega_3\}\})$ and $\mathbf{Y}^2 = \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$, where an inclusion property does not hold between $\mathbf{Y}^1$ and $\mathbf{Y}^2$. This, therefore, does not fall into any of the categories of information structures considered so far in this section (for the specific example). The two methods of derivation here would be:

(a) The direct solution based on the normal form (which is a $4 \times 4$ matrix)
(b) The sequential approach (which involves *two* $4 \times 2$ matrices and hence does not offer any savings (and thereby advantage) over the normal form)

These are the two general methods which would be applicable to this class of problems; however, in the present case a simple (but useful) observation yields the solution immediately: The team-optimal decision rules $\gamma^{1^*}$ and $\gamma^{2^*}$ for case 1 (i.e., under perfect measurements) are also well-defined functions on the signal spaces $\mathbf{Y}^1$ and $\mathbf{Y}^2$ above, and hence under the right kind of interpretation, they belong to the strategy spaces $\Gamma^1$ and $\Gamma^2$ of the present problem. The information structure in case 1 being *richer* (in fact, the *richest* possible),[10] this observation directly implies that the pair $\{\gamma^{1^*} = UD, \gamma^{2^*} = RL\}$ is the unique team-optimal solution of the new problem with "coarser" information. Note that the pair $(UD, RL)$ here is indeed the pair $(UDD, RRL)$ of case 1, simply rewritten using the adopted convention, on the restricted information space. If we write them out, they both correspond to

$$\gamma^{1^*}(y^1) = \begin{cases} U, \ y^1 = \{\omega_1\}, \\ \\ D, \text{ else}, \end{cases} \qquad \gamma^{2^*}(y^2) = \begin{cases} R, \ y^2 = \{\omega_1, \omega_2\}, \\ \\ L, \text{ else}. \end{cases}$$

A mathematically precise statement of the property (of the team-optimal solution) used here will be given later in the chapter.

6. *"Noisy" measurements*

For reasons which will become clear later, it is useful to distinguish between "imperfect" and "noisy" measurements. The information signals of cases 2 and 5, studied above, belong to the former category because they do not bring in additional uncertainty into the problem formulation, other than what exists already in the complete description of the cost matrices. In a sense, an imperfect measurement brings in a refinement on the information available to an agent under the

[10] At this point, this statement should be interpreted as saying "there is no other information structure which provides the agents with more information on the state of nature." The underlying notion will be made precise later.

*no-measurement* scenario (such as case 3) **without** bringing in additional elements of uncertainty. In the "noisy measurement" case, however, the sample space has some additional elements which are not needed in the complete description of the loss (payoff) functions. To further elaborate on this point, consider the scenario depicted below, which uses essentially the same team problem as before, but with a different type of information.

Agent **A**2 makes no measurements, while **A**1 observes the value of a random variable $z$, taking two possible (distinct) values, $y_1$ and $y_2$. The loss matrices are the same as before, where we now adopt a different symbol, $x$, to replace $\omega$, with $x_i = \omega_i$. To complete the description of the team problem, we now specify, in the following table, the joint probability mass function (*pmf*) of the pair $(x, z)$, which has to be consistent with the marginal *pmf* of $x$:

| | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| $y_1$ | 0.12 | 0.21 | 0.12 |
| $y_2$ | 0.18 | 0.09 | 0.28 |

Note that

$$\mathit{Prob}(x_i \mid y_1) = \begin{cases} 4/15, \, i = 1, 3, \\ 7/15, \, i = 2, \end{cases} \qquad \mathit{Prob}(x_i \mid y_2) = \begin{cases} 18/55, \, i = 1, \\ 9/55, \quad i = 2, \\ 28/55, \, i = 3, \end{cases}$$

and hence after observing $y_1$ or $y_2$ it is not possible for **A**1 to tell, with certainty, the true value of $x$. We refer to the measurement signal as "noisy" because

(a) It does not transmit the true value of $x$ (which, along with the action variables, completely determines the loss).
(b) It introduces additional elements of uncertainty into the problem.

The problem can now be cast in the framework of the general formulation of Sect. 2.2 by constructing an appropriate sample space. Toward this end, let $\Omega$ be a set of cardinality 6, with elements $\omega_{ij}$ $(i = 1, 2, 3; j = 1, 2)$, where $\omega_{ij}$ corresponds to the pair $(x_i, y_j)$ and hence $\mathit{Prob}(\omega = \omega_{ij}) = \mathit{Prob}(x = x_i, y = y_j)$. The two possible measurement signals of **A**1 are $y_1^1 = \{\omega_{11}, \omega_{21}, \omega_{31}\}$ and $y_2^1 = \{\omega_{12}, \omega_{22}, \omega_{32}\}$ which together determine the partition $\mathbf{Y}^1$ introduced in Sect. 2.2.[11] We thus have a team problem of the standard type, for which the normal form is

| | | **A1** | | | |
|---|---|---|---|---|---|
| | | UU | UD | DU | DD |
| **A2** | L | 1.30* | 1.38 | 1.42 | 1.50 |
| | R | 1.70 | 1.70 | 1.40 | 1.40 |

[11]Here, since we have a finite decision problem, we do not distinguish between $\Omega$ and $\Xi$, and hence consider $\mathbf{Y}^1$ as a partition of the sample space $\Omega$.

which admits the unique team-optimal solution $\underline{\gamma}^* = (UU, L)$, with a corresponding value of $J^* = 1.30$. An immediate observation here is that this is the same value as that obtained in case 3, and hence the additional (noisy) information to **A**1 is of no value to the team. We leave it to the reader to verify that if, instead, agent **A**2 had received this measurement signal, then the team-optimal solution would again be unique and be given by $\underline{\gamma}^* = (D, RL)$, yielding this time a value of $J^* = 1.29$. Hence the same measurement is of some (positive) value to the team, if received by the second agent. As a final scenario, let us consider the information structure under which **both** agents have access to the realization of $z$ (i.e., they have a complete sharing of information, which makes the problem essentially no different from a single agent stochastic decision problem). In view of the discussion for case 2, and especially the relation (2.9), we first form the *conditional cost matrices* corresponding to the two realizations of the measurement signal, $y_1$ and $y_2$:

| | | **A**2 | | | **A**2 | |
|---|---|---|---|---|---|---|
| | | L | R | | L | R |
| **A**1 | U | 22/15 | 29/15 | U | 64/55* | 83/55 |
| | D | 26/15 | 19/15* | D | 72/55 | 83/55 |
| $z$ : | | $y_1 \leftrightarrow 0.45$ | | | $y_2 \leftrightarrow 0.55$ | |

Then we can readily read, from the above matrices, the unique team-optimal strategy pair: $(DU, RL)$, with a corresponding cost value of $1.21$. Note that here, to determine the optimal strategies, all we need are the six conditional probabilities, *Prob*$(x_i \mid y_j)$, $i = 1, 2, 3;\ j = 1, 2$, and not the individual probabilities for $y_1$ and $y_2$.[12] The latter are, of course, needed in the computation of the corresponding cost value.

It is worth noting that the main feature of this last case, which distinguishes it from the earlier ones, is that the random quantity $\omega$ (or, equivalently here, the state of nature $\xi$) has two identifiable components: the "payoff relevant" part, $x$, and the information signal, $y$, with some correlation between them. The role of $y$ is to carry information regarding the true value of $x$, and it affects the value of the loss function **only** through the strategy of the agent who receives this information. The advantage of splitting $\xi$ into two components, as above, may not be that obvious at this point, but we will later observe the versatility of such a formulation, especially in the context of infinite decision problems.

[12] This would not have been true if the agents had made nonidentical measurements.

## 2.4 A Mathematical Framework for Dynamic Decision Problems

As mentioned earlier in Sect. 2.2, a decision problem is said to be *dynamic* if the measurements of at least one of the agents involve past actions (of that particular agent or some other agent(s)). In the literature, the connotation "dynamic" is also used to characterize decision problems where an agent acts more than once, even if the measurements do not depend on past actions (the case of *open-loop* information structure). In principle such problems can be converted into static decision problems by essentially working in higher-dimensional spaces, but it is generally found convenient to treat them also in the context of truly dynamic problems because of the similarities in the derivation of the optimal solutions. We will have occasions to use both approaches in this book. We describe below an appropriate setup for the study of truly dynamic decision problems, restricting the exposition to discrete time.

For a truly dynamic problem, we follow the formulation of Sect. 2.2, prior to (2.1) but now replace the static relationship (2.1) with the dynamic equation

$$y^i = \eta^i(\xi; \mathbf{u}), \quad i \in \mathcal{N}, \tag{2.10}$$

where the dependence on $\mathbf{u}$ is assumed to be *strictly causal*, which means that under a given *fixed clock* the information received by each agent can depend only on actions taken in the past. To give this statement a more precise mathematical meaning, let us consider a *discrete-time* framework where actions are taken at discrete instants of time, $1, 2, \ldots, T$. Let $t$ stand for the generic time variable and $\mathcal{T}$ denote the (discrete) time set

$$\mathcal{T} := \{1, \ldots, T\}. \tag{2.11}$$

Let $u_t^i$ and $y_t^i$ denote, respectively, the action (decision) variable and the information variable of agent $\mathbf{A}i$ at the time instant $t \in \mathcal{T}$. Furthermore, introduce the notation:

$$\mathbf{u}_t := \{u_t^1, \ldots, u_t^N\}, \quad \mathbf{y}_t := \{y_t^1, \ldots, y_t^N\}, \tag{2.12}$$

$$\mathbf{u}_{[t_0,t_1)} \equiv \mathbf{u}_{[t_0,t_1-1]} := \{\mathbf{u}_{t_0}, \mathbf{u}_{t_0+1}, \ldots, \mathbf{u}_{t_1-1}\} \equiv \{u^1_{[t_0,t_1)}, \ldots, u^N_{[t_0,t_1)}\}. \tag{2.13}$$

Then, under the strict causality assumption, (2.10) becomes equivalent to

$$y_t^i = \eta_t^i(\xi; \mathbf{u}_{[1,t)}), \qquad t \in \mathcal{T}, i \in \mathcal{N} \tag{2.14}$$

for some "information functions" $\eta_t^i$, $t \in \mathcal{T}$, $i \in \mathcal{N}$. The stochastic variable $y_t^i$, taking values in $Y_t^i$, is the *online* information available to $\mathbf{A}i$ which he can use in the construction of the decision $u_t^i$ at time $t$, through an appropriate policy variable $\gamma_t^i : Y_t^i \to U_t^i$

$$u_t^i = \gamma_t^i(y_t^i) \equiv \gamma_t^i(\eta_t^i[\xi; \mathbf{u}_{[1,t)}]), \qquad t \in \mathcal{T}, i \in \mathcal{N}. \tag{2.15}$$

A permissible policy $\gamma_t^i$ is one under which $u_t^i$ becomes a well-defined random variable, defined on the original probability space, and taking values in $S_t^i \subset U_t^i$, where $S_t^i$ is the *action constraint* set for $\mathbf{A}i$ at time $t$. Let us denote the set of all such maps by $\Gamma_t^i$, which is the *policy space* of $\mathbf{A}i$ at time $t$. The construction of such a policy space will depend on the problem at hand, and we will see several such constructions throughout the book. At this point let us simply assume that such a construction is given, and rewrite (2.15) in the following compact form:

$$\mathbf{u} = \underline{\gamma}(\underline{\eta}[\xi; \mathbf{u}]), \quad \underline{\gamma} \in \mathbf{\Gamma} := \Gamma^1 \times \cdots \times \Gamma^N, \tag{2.16}$$

where $\Gamma^i$ is the *composite* (over-time) *policy space* of $\mathbf{A}i$:

$$\Gamma^i := (\Gamma_1^i, \Gamma_2^i, \ldots, \Gamma_T^i), \qquad i \in \mathcal{N}.$$

Note that the right-hand side of (2.16) also depends on $\mathbf{u}$, which is what makes dynamic decision problems intrinsically different from the static ones introduced in Sect. 2.2. Equation (2.16) is called a *loop equation*, and a dynamic decision problem is well defined only if this loop equation has a unique solution for every $\xi$, that is, for some $\underline{\tilde{\gamma}} : \Xi \to \mathbf{U}$,

$$\mathbf{u} = \underline{\tilde{\gamma}}(\xi) \tag{2.17}$$

uniquely solves (2.16). The strict causality condition, or equivalently the structural assumption (2.15), is precisely the condition that guarantees this.

It is possible to relax the strict causality condition and the fixed clock assumption and replace them by some other conditions under which the loop equations (2.16) still admit a unique solution. A precise study of these conditions is beyond the level of our treatment here; for this more general treatment the reader is referred to Witsenhausen [393] and Teneketzis [360] (see also Sect. 3.7 for further discussions). To just provide a flavor of these extensions here, let us note that in (2.15) we can allow $u_t^i$ to depend on $u_t^j, j \neq i$, provided that $u_t^j$ is not allowed to depend on $u_t^i$ either directly or through the actions of other agents. In Fig. 2.1 we have depicted two such scenarios for a four-agent problem with time step $t$ isolated. A pointed arrow indicates that information flows at this stage in the direction of the arrow. The first (upper) configuration of Fig. 2.1 does not lead to a well-defined (physically realizable) decision problem because $u_t^4$ depends on $u_t^1$ while at the same time $u_t^1$ depends on $u_t^4$; hence there is a deadlock. (Clearly closed directed graphs should not be allowed for unique solvability of the loop equations.) The second (lower) configuration of Fig. 2.1, on the other hand, depicts an acceptable information exchange[13] since it does not contain any closed direct graph. Hence,

[13]Here, and throughout, we are using a finer partition of the time interval in between the two (discrete) time points $t$ and $t+1$, so that agents can select $u_t^i$, $i \in \mathcal{N}$, in a (partially) sequential order. A possible strict time order of the configuration of Fig. 2.1 (lower) is $(u_t^1, u_t^4, u_t^2, u_t^3)$, but not all configurations have to have a strict time order.

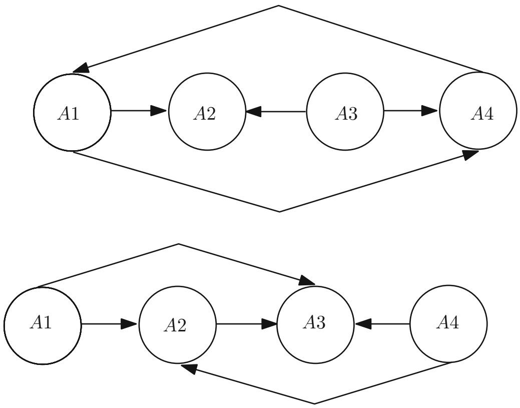


**Fig. 2.1** Two scenarios on available action information at stage $t$

the strictly causal relation (2.15) can accommodate such permissible informational relationships among the $u_t^i$'s, and the configuration could be different for each $t$.

In such a team, if there is a prespecified order in which the agents act, then such a team is said to be a *sequential team*. However, one can allow these "permissible configurations" to be sample-path dependent (i.e., dependent on $\xi$), provided that certain measurability conditions are satisfied. This is the situation where the order in which the agents act is determined (partially) by a chance mechanism. Such a dynamic team model is said to be *a nonsequential* team.

### *Sequential Dynamic Teams in State-Space Form*

Now, coming back to (2.14) and (2.15), it is generally convenient to introduce an intermediate variable, called the *state variable*, recursively defined by

$$x_{t+1} = f_t(x_t, u_t^1, \ldots, u_t^N; w_t^0), \quad t \in \mathcal{T}, \tag{2.18}$$

where $x_1$ and $w^0 := \{w_1^0, \ldots, w_T^0\}$ are random exogenous variables with given probability distributions (with the latter being the *system noise*); furthermore, $x_{t+1}$ takes values in a given topological space $X_{t+1}$. We further introduce the equation

$$y_t^i = g_t^i(x_t, u_{t-1}^1, \ldots, u_{t-1}^N; w_t^i), \qquad i \in \mathcal{N}, \qquad t \in \mathcal{T}, \tag{2.19}$$

where $y_t^i$ is again the measurement of $\mathbf{A}i$ at stage $t$ and $\{w_t^i, t \in \mathcal{T}\}$ is known as the measurement noise of $\mathbf{A}i$, which has a given probability distribution, possibly different for different agents as to be elucidated below. Note that for $t = 1$, $g_t^i$ would have as argument only $x_t$ and $w_t$, as controls have not yet been applied. Further note that the state variable $x_t$, $t \in \mathcal{T}$, can easily be eliminated (for $t > 1$) in (2.19) by recursive substitution, so that $y_t^i$ can be expressed solely in terms of the action variables and the primitive random variables:

$$y_t^i = \tilde{g}_t^i(x_1, w^0_{[1,t-1]}, \mathbf{u}_{[1,t-1]}, w_t^i). \tag{2.20}$$

Here, the primitive random variables, or the "states of nature," are

$$\xi := \{x_1, w^0_{[1,T]}, \mathbf{w}_{[1,T]}\}, \tag{2.21}$$

where $\mathbf{w}_{[1,T]}$ is defined similar to $\mathbf{u}_{[1,T]}$, with $w$ replacing $u$.

Now, the operation of a decision process described by (2.18) and (2.19) would proceed chronologically as follows:

- Generation of an initial random state $x_1$ with distribution $P_{x_1}$
- Observation of measurements $\mathbf{y}_1 := \{y_1^1, \ldots, y_1^N\}$, where the composite measurement noise $\mathbf{w}_1$ has a given conditional distribution $P_{\mathbf{w}_1|x_1}, i \in \mathcal{N}$
- Application of controls $\mathbf{u}_1$
- Generation of the "system noise" $w_1^0$, with conditional distribution $P_{w_1^0|x_1,\mathbf{u}_1}$, and transition to the next state $x_2$

. . . . . . . . . . . . . . . . . . . . .

- $\cdots$ transition to state $x_t$
- Observation of measurements $\mathbf{y}_t$, with $\mathbf{w}_t \sim P_{\mathbf{w}_t|x_t}, u_{t-1}$
- Application of controls $\mathbf{u}_t$
- Generation of the "system noise" $w_t^0 \sim P_{w_t^0|x_t,\mathbf{u}_t}$

. . . . . . . . . . . . . . . . . . . . .

- $\cdots$ transition to state $x_{T+1}$

The above evolution does not completely describe the dynamic decision process, because the construction of the controls and the allowable dependence of the controls on the past measurements and/or actions have not yet been specified. Toward this end, let $\tilde{y}_t^i$ denote some prespecified subset of the collection $\{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}$, possibly a different subcollection for different $i \in \mathcal{N}$. Note that it is possible to find an $\eta_t^i$, so that

$$\tilde{y}_t^i = \eta_t^i(\xi; \mathbf{u}_{[1,t-1]}), \quad t \in \mathcal{T}, i \in \mathcal{N},$$

where $\xi$ was introduced earlier by (2.21). The $\tilde{y}_t^i$ defined above can definitely be viewed as a (high-dimensional) vector, and it is precisely the information variable (2.14) where we have not used *tilde* on $y$ simply not to clutter the notation. But the distinction should be clear from context. In static decision problems, there is, of course, no difference between the information and measurement variables, and indeed in Sect. 2.2 we have called $y^i$ as both measurement variable and information variable. In dynamic problems, however, there is a distinction between the two, and this has to be recognized in the derivation of optimal solutions to dynamic teams, as we will see later.

The collection of individual information functions $\{\eta_t^i, t \in \mathcal{T}, i \in \mathcal{N}\}$ constitutes the *information structure* of the dynamic decision problem. Perhaps by a slight abuse of notation and terminology, we introduce, for each $t \in \mathcal{T}$ and $i \in \mathcal{N}$, a finite set $\mathcal{I}_t^i$ which specifies precisely which elements of the set of vectors $\{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}$ will be used in the construction of the control $u_t^i$, and we call the collection $\mathcal{I} := \{\mathcal{I}_t^i, i \in \mathcal{N}, t \in \mathcal{T}\}$ again as the *information structure* of the decision problem.

We list below some important information structures which will be used throughout the book.

1. *Sole prior information* (SPI): $\mathbf{A}i$ is said to have *SP* information if she makes no measurements and the only information she works with is the prior statistics on the random variables. A decision problem has *SP* information if all agents have *SP* information.
2. *Open-loop* (*OL*) *information*: $\mathbf{A}i$ is said to have *OL* information if $\mathcal{I}_t^i = \mathcal{I}_1^i$ for all $t \in \mathcal{T}$. A decision problem has *OL information structure* if all agents have OL information (which are not necessarily the same). Note that OL information is different from SPI.
3. *Complete information sharing* (*CIS*):
$$\mathcal{I}_t^i = \{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$
Here there is a complete exchange of present and past measurements as well as past actions.
4. *Complete measurement sharing* (*CMS*):
$$\mathcal{I}_t^i = \{\mathbf{y}_{[1,t]}\}, \quad i \in \mathcal{N}.$$
Here the past actions are not shared.
5. $n$*-step delayed information sharing* ($n$*DIS*):
$$\mathcal{I}_t^i = \begin{cases} \{y^i_{[t-n+1,t]}, \mathbf{y}_{[1,t-n]}, \mathbf{u}_{[1,t-n]}\}, & t > n, \\ \{y^i_{[1,t]}\}, & t \le n, \end{cases} \qquad i \in \mathcal{N}.$$
6. $n$*-step delayed measurement sharing* ($n$*DMS*):
$$\mathcal{I}_t^i = \begin{cases} \{y^i_{[t-n+1,t]}, \mathbf{y}_{[1,t-n]}\}, & t > n, \\ \{y^i_{[1,t]}\}, & t \le n, \end{cases} \qquad i \in \mathcal{N}.$$

7. *$n$-step delayed control sharing* (*$n$DCS*):

$$\mathcal{I}_t^i = \begin{cases} \{y_{[1,t]}^i, \mathbf{u}_{[1,t-n]}\}, & t > n, \\ \{y_{[1,t]}^i\}, & t \leq n, \end{cases} \qquad i \in \mathcal{N}.$$

8. *$k$-step periodic information sharing* (*$k$PIS*):

$$\mathcal{I}_t^i = \begin{cases} \{y_{[\lfloor t/k \rfloor k,t]}^i, \mathbf{y}_{[1,\lfloor t/k \rfloor k]}, \mathbf{u}_{[1,\lfloor t/k \rfloor k]}\}, & t \geq k, \\ \{y_{[1,t]}^i\}, & t < k, \end{cases} \qquad i \in \mathcal{N}.$$

9. *Completely decentralized information* (*CDI*):

$$\mathcal{I}_t^i = \{y_{[1,t]}^i\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

Note that this corresponds to *$n$DMS* with $n = T$.

All the information structures given above are of the *perfect recall* (*PR*) type, in the sense that the agents have full memory of their information in the past. An example of an information structure (*IS*) which is not of the *PR* type is

$$\mathcal{I}_t^i = \{y_t^i\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

Stochastic decision problems whose *IS*s are not of the *PR* type are relatively more difficult to analyze than those with *PR* type *IS*, as we shall see later. Another class of challenging decision problems are those with so-called nonclassical ISs. Under such *IS*s an agent sees the action variable of another agent in her information set, or her information is indirectly affected by it, but she does not have access to the measurements/information based on which that action was taken; *$n$DCS* introduced above is one such *IS*, so could *$n$DMS*, *CDI*, or information structures which are not *PR*. We will say more on such nonclassical *IS*s in the next chapter, Sect. 3.2. Furthermore, we will observe that not all nonclassical information structures lead to computational difficulties. Examples will be considered further in the next chapter, as well as in Chap. 12.

Now, fixing the *IS* of a dynamic decision problem also fixes the strategy (policy) spaces of the agents, as in (2.16). To complete the description as a team problem, we have to specify the cost structure, which we do as follows:

Adopting the description (2.18) and (2.19), also known as the *state-space model*, we associate with the team the loss function[14]

$$L(x_{[1,T+1]}, \mathbf{u}_{[1,T]}) = \sum_{t \in \mathcal{T}} c_t(x_{t+1}, \mathbf{u}_t), \tag{2.22}$$

[14] Here, an alternative form can be $L(x_{[1,T+1]}, \mathbf{u}_{[1,T]}) = \sum_{t \in \mathcal{T}} c_t(x_t, \mathbf{u}_t) + c_{T+1}(x_{T+1})$.

where each term in the summation is known as the *incremental* (*stagewise*) *loss*. Since $x_{[1,t+1]}$ can be expressed in terms of the primitive random variables and the action variables, replacing $u_t^i$ in (2.22) by $\gamma_t^i(y_t^i)$, $\gamma_t^i \in \Gamma_t^i$, constructed under the given IS, $L$ becomes a function of only $\xi$ (for each fixed $\underline{\gamma} \in \mathbf{\Gamma}$), whose expectation with respect to the subjective probability distribution function of $\xi$ leads as in (2.2) to the cost function:

$$J(\underline{\gamma}) = E_\xi L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))\,. \tag{2.23}$$

Here $L$ is given by (2.22), with the intermediate variables eliminated by using (2.18) and (2.19).

The function $J$, along with the product strategy space $\mathbf{\Gamma}$, constitutes the *normal* (*strategic*) form of the dynamic decision problem, and as such is no different (in abstract form) from the normal form introduced in Sect. 2.2 for static multiple person decision problems. Hence, all the solution concepts introduced there, viz., *team-optimality* and *person-by-person optimality* (or *Nash equilibrium*), are equally valid (and relevant) here, which we do not give to avoid repetition. In addition, however, some new features emerge due to the dynamic nature of the information pattern, which use particularly the sequential (extensive form) description of the decision problem. We introduce below two such general features associated with the team-optimal or *pbp* optimal solutions of dynamic team problems.

**Definition 2.4.1.** Let $D := \{J, \mathbf{\Gamma}, \mathcal{T}\}$ be a dynamic team problem which admits a solution $\underline{\gamma}^* \in \mathbf{\Gamma}$. Let $t > 1$ be an arbitrary point in $\mathcal{T}$ and consider the decision problem $D^{\beta}_{[t,T]}$ which is derived from $D$ by setting $\underline{\gamma}_{[1,t-1]} = \underline{\beta}_{[1,t-1]}$, for an arbitrary $\underline{\beta}_{[1,t-1]} \in \mathbf{\Gamma}_{[1,t-1]}$. Then:

(i) The solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ is *strongly time consistent* (STC) if the subpolicy $\underline{\gamma}^*_{[t,T]}$ constitutes a *solution* to the dynamic team $D^{\beta}_{[t,T]}$, this being so for every $t \in \mathcal{T}, t > 1$, and every permissible $\underline{\beta}_{[1,t-1]} \in \mathbf{\Gamma}_{[1,t-1]}$.

(ii) The solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ is *weakly time consistent* (WTC) if the subpolicy $\underline{\gamma}^*_{[t,T]}$ constitutes a *solution* to the dynamic team $D^{\beta}_{[t,T]}$ when $\underline{\beta}_{[1,t-1]} = \underline{\gamma}^*_{[1,t-1]}$.

$\diamond$

Note that if an equilibrium solution is *STC*, then the past actions do not rein in the present and future actions of the agents under the same solution concept, i.e., the agents have no reason to renege (and deviate from the equilibrium policy or the course of action) even if some inadvertent deviations have taken place in the past. With the WTC solution, however, there is no incentive to renege only if the declared course of action has been followed in the past.

### *The Intrinsic Model and the Markov Transition Model*

Before concluding this section, we should mention that in our general formulation of a dynamic team decision problem, we have allowed an agent to act multiple times, at different time instants, using possibly different information, that is, $\mathbf{A}i$ has $u^i_{[1,T]}$ as her action variable. An alternative (but equivalent) formulation would be to have an agent $\mathbf{A}i$ be split into $T$ agents, with $\mathbf{A}i(t)$ ($t$'th agent in this split, where $t = 1, \ldots, T$) controlling only $u^i_t$. This would then transform the original $N$-agent team to an $NT$-agent team problem, but other than a difference in semantics, the two formulations are essentially the same. More details on these different viewpoints to dynamic decision problems can be found in Witsenhausen [399–401] (see Sect. 3.7, where Witsenhausen's *intrinsic model* as well as other models for dynamic teams are reviewed).

Another point worth mentioning is that an alternative to the state-space model (2.18) and (2.19) exists, especially if the probability and action spaces are finite. This so-called Markov transition model involves $N+1$ conditional probability laws at each stage $t \in \mathcal{T}$, to replace (2.18) and (2.19). The state equation (2.18) is replaced by a controlled probability transition:

$$P_{x_{t+1|x_t, w^0_t}}(\mathbf{u}_t), \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

If all the variables belong to finite spaces, then the model is completely described by a finite number of finite-dimensional (probability) matrices.

## 2.5 An Illustrative Example of a Finite Dynamic Team

To illustrate some salient aspects of the formulation of dynamic decision problems, we consider in this section a finite dynamic team problem with two agents and two stages and with the agents having the same subjective prior probabilities on the random variables. We will study the derivation of the team-optimal solution under several different *IS*s of the type introduced in the previous section. Since the underlying team is finite (with a finite probability space), a team-optimal solution will exist under all *IS*s.

Now, the description of the stochastic dynamic team follows: At each stage, the control (decision) spaces of the agents have two elements, as in the static team of Sect. 2.3.

$$U^1_1 = U^1_2 = \{U(\text{up}), D(\text{down})\}, \quad U^2_1 = U^2_2 = \{L(\text{left},)R(\text{right })\}.$$

The initial state, $x_1$, is a discrete random variable, taking two values, $x_{11}$ and $x_{12}$, with respective probabilities 0.4 and 0.6. If $x_1 = x_{1i}$ and $u^1_1 = U$ or $D$, and $u^2_1 = L$ or $R$, the loss to the team (i.e., the stagewise loss, $c_1(x_1, u^1_1, u^2_1)$) is given by the "loss matrices"

$x_1 = x_{11}$ (rows: **A**1, columns: **A**2)

| | L | R |
|---|---|---|
| U | 1 | 0 |
| D | 3 | 1 |

$x_1 = x_{12}$ (rows: **A**1, columns: **A**2)

| | L | R |
|---|---|---|
| U | 1 | 2 |
| D | 0 | 2 |

We denote the first matrix above by LM1 and the second matrix by LM2. The transition to the second stage and the associated cost is now described as follows: Let $x_2$ take three distinct values, $x_{21}$, $x_{22}$ and $x_{23}$, with the rule

$$x_2 = \begin{cases} x_{21}, \text{ if } (u_1^1, u_1^2) = (U, L) \text{ and } x_1 = x_{11}, \\ \\ x_{22}, \text{ if } (u_1^1, u_1^2) = (U, L) \text{ and } x_1 = x_{12}, \\ \\ x_{23}, \text{ otherwise.} \end{cases}$$

The corresponding cost is determined by $x_2$, $u_2^1$, $u_2^2$, and an independent random variable, in terms of the loss matrices LM1 and LM2 at stage 1, according to the following table.

| $x_2$ | Loss at stage 2 ($c_2$) |
|---|---|
| $x_{21}$ | LM1 *w.p.* 0.4, LM2 *w.p.* 0.6 |
| $x_{22}$ | LM1 *w.p.* 0.5, LM2 *w.p.* 0.5 |
| $x_{23}$ | LM1 *w.p.* 0.3, LM2 *w.p.* 0.7 |

All random mechanisms are assumed to be independent and the total loss to the team is the arithmetic sum of the stagewise losses, as in (2.22). Note that here we basically have two random variables, $x_1$ and $w_2$, say, with the statistics of the latter governing the loss structure in the table above, i.e.,

$$\mathit{Prob}(w_2 = w_{21} \mid x_2) = 1 - \mathit{Prob}\ (w_2 = w_{22} \mid x_2) = \begin{cases} 0.4,\ x_2 = x_{21}, \\ 0.5,\ x_2 = x_{22}, \\ 0.3,\ x_2 = x_{23}, \end{cases}$$

where $w_{21}$ corresponds to LM1 and $w_{22}$ to LM2.

Now we specify the measurements available to the agents: It is assumed that **A**1 knows exactly the value of $x_1$ at stage 1, and does not make any further measurements (at stage 2). **A**2, on the other hand, makes no measurements at stage 1 but knows precisely the value of $w_2$ at stage 2; hence

$$y_1^1 = x_1, \quad y_2^1 : \text{void}; \quad y_1^2 : \text{void}, \quad y_2^2 = w_2.$$

For any given permissible policy, the (expected) cost to the team is given by (2.23), and with $\xi := (x_1, w_2)$. The precise form, of course, will depend on the *IS* to be adopted, as delineated below:

1. *SPI*. This will lead to the worst performance for the team, because both agents work under only the given prior information. The policy spaces of the agents are

$$\Gamma^1 = \{Uu, Ud, Du, Dd\}; \quad \Gamma^2 = \{Ll, Lr, Rl, Rr\}$$

where we have used “lower case letters” for the second-stage decisions. The only way to solve this team problem is to convert it to *normal form*, which is a $4 \times 4$ matrix

| | | **A2** | | | |
|---|---|---|---|---|---|
| | | Ll | Lr | Rl | Rr |
| | Uu | 2 | 2.08 | 2.2 | 1.8 |
| **A1** | Ud | 2.38 | 3.5 | 3.3 | 2.5 |
| | Du | 2.2 | 2.2 | 2.6 | 2.2 |
| | Dd | 3.3 | 2.9 | 3.7 | 3.2 |

The unique team-optimal solution is $(Uu, Rr)$, with a cost level of $J_{SP} = 1.8$.

2. *CMS*. This is the other extreme case, when the agents know exactly what loss matrix is being optimized at each stage. Since the lowest entries of LM1 and LM2 are both *zero*, the optimal team cost is $J_{CMS} = 0$. The solution would have been the same if, instead, we had the *CIS IS*.
3. *1-step-delayed measurement sharing* (*1DMS*). Here the information available to the agents at each stage are

$$\mathcal{I}_2^1 = \mathcal{I}_1^1 = \{x_1\}, \quad \mathcal{I}_1^2 = \phi, \mathcal{I}_2^2 = \{x_1, w_2\},$$

and hence $|\Gamma_1^1| = |\Gamma_2^1| = 4$, $|\Gamma_1^2| = 2$, and $|\Gamma_2^2| = 8$. The cardinality of the composite policy spaces are $|\Gamma^1| = 8$,[15] $|\Gamma^2| = 16$, which means that the normal form would be an $8 \times 16$ dimensional matrix. One possible approach to the problem would be to compute the entries of this matrix and choose the smallest one as the solution. An alternative approach is a sequential derivation, which makes use of the fact that measurements are shared with a delay of *one* time unit, which is what we discuss below.

Suppose that the actions at stage 1 have been taken, and the agents are facing the decision problem (at the second stage) where the value of $x_1$ is now *common knowledge*. **A**1 has no other information, and hence his possible actions (policies) are $u$ and $d$. **A**2, on the other hand, has the additional information, the precise value of $w_2$, and hence he has *four* possible policies: ll, lr, rl, and rr, where lr stands for $u_1^2 = $ l, if $w_2 = w_{21}$, and $u_1^2 = r$, otherwise. Now, conditioned on the value of $x_1$ (which is common knowledge) and the actions taken by the agents at stage 1 (which can also be considered to be common knowledge since we have a (cooperative) team problem and the information based on which these actions were taken is common knowledge), we have the following eight *total cost matrices*.

[15]This one is not 16 because for each value of $x_1$, **A**1 has four choices, which makes the total 8. Hence, one can replace $\Gamma^1 := \Gamma_1^1 \times \Gamma_2^1$ with a smaller set, without any loss in performance.

$x_1 = x_{11}$ :

$(u_1^1, u_1^2) = (U, L)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 2 | 2.6 | 1.6 | 2.2 |
| d | 2.2 | 3.4 | 1.4* | 2.6 |

$(u_1^1, u_1^2) = (U, R)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 1 | 1.3 | 0.3* | 0.6 |
| d | 2.1 | 3.3 | 0.7 | 1.3 |

$(u_1^1, u_1^2) = (D, L)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 4 | 4.3 | 3.3* | 3.6 |
| d | 5.1 | 6.3 | 3.7 | 4.3 |

$(u_1^1, u_1^2) = (D, R)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 2 | 2.3 | 1.3* | 1.6 |
| d | 3.1 | 4.3 | 1.7 | 2.3 |

$x_1 = x_{12}$ :

$(u_1^1, u_1^2) = (U, L)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 2 | 2.5 | 1.5* | 2 |
| d | 2.5 | 3.5 | 1.5* | 2.5 |

$(u_1^1, u_1^2) = (U, R)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 3 | 3.3 | 2.3* | 2.6 |
| d | 4.1 | 5.3 | 2.7 | 3.3 |

$(u_1^1, u_1^2) = (U, L)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 1 | 1.3 | 0.3* | 0.6 |
| d | 2.1 | 3.3 | 0.7 | 1.3 |

$(u_1^1, u_1^2) = (U, R)$

| | ll | lr | rl | rr |
|---|---|---|---|---|
| u | 3 | 3.3 | 2.3* | 2.6 |
| d | 4.1 | 5.3 | 2.7 | 3.3 |

In each case, the "starred" entry(ies) denotes the minimum entries of the corresponding matrices, which will be carried over to the *first* stage to determine the optimal policies there. Now, at stage 1 **A**1 knows the value of $x_1$, but **A**2 does not, so that possible policies for **A**1 are UU, UD, DU, and DD, while the permissible policies for **A**2 are L and R. Using the optimum entries above, we can construct an equivalent cost matrix at stage 1 through an appropriate averaging process:

| | | **A**2 L | **A**2 R |
|---|---|---|---|
| **A**1 | UU | 1.46 | 1.5 |
| **A**1 | UD | 0.74* | 1.5 |
| **A**1 | DU | 2.22 | 1.9 |
| **A**1 | DD | 1.5 | 1.9 |

This is known as the *optimum cost-to-go matrix* at stage 1, because of the following interpretation that the entries admit. Consider, for example, the entry with the numerical value 2.22: If **A**1 chooses D when $x_1 = x_{11}$ and U when $x_1 = x_{12}$, and **A**2 chooses L, all at stage 1, then whatever choices are made at stage 2 (under the given information) the total (expected) cost can never be lower than 2.22. To arrive at this numerical value, we note that if $x_1 = x_{11}$, $u_1^1 = D$ and $u_1^2 = L$, the cost resulting from an optimum choice of policies at stage 2 would be 3.3 (the lowest entry of the *third* conditional total cost matrix), whereas if $x_1 = x_{12}$, $u_1^1 = U$, $u_1^2 = L$, the optimum total cost would be 1.5 (the lowest entry of the *fifth* matrix). Since the value of $x_1$ is not available to **A**2 at stage 1, we average these values of $x_1$ to obtain

$$(0.4)(3.3) + (0.6)(1.5) = 2.22.$$

All other entries of the $4 \times 2$ *cost-to-go matrix* can be computed analogously.

Clearly, the minimum cost is $J_{1DMS} = 0.74$, with the unique team-optimal policy being

$$(\gamma_1^{1*}(x_1), \gamma_2^{1*}(x_1)) = \begin{cases} Ud, \; x_1 = x_{11}, \\ Du, \; x_1 = x_{12}, \end{cases}$$

$$\gamma_1^{2*} = L, \quad \gamma_2^{2*}(x_1, w_2) = \begin{cases} r, \; w_2 = w_{21}, \\ l, \; w_1 = w_{22}, \end{cases}$$

Note that $\gamma_2^{2*}$ is actually independent of the value of $x_1$ (which means that even if the measurement $y_1^1$ had not been shared, the team-optimal solution would still be the same) and that $\gamma_2^{1*}$ does depend on $x_1$ (which means that if **A**1 were not allowed to recall the value of $x_1$ at stage 2, the optimal team cost would have been higher—see the next case).

4. *No sharing, no recall* (*NSR*). Here we have

$$\mathcal{I}_1^1 = \{x_1\}, \mathcal{I}_2^1 = \phi = \mathcal{I}_1^2, \quad \mathcal{I}_2^2 = \{w_2\}.$$

If we had allowed perfect recall (i.e., $\mathcal{I}_2^1 = \mathcal{I}_1^1$), then the solution would be the one obtained in case 3, as discussed there[16]; however, without perfect recall the solution does not follow from the one in case 3. This information structure is *nonclassical* and hence a recursive derivation as in case 3 is also not possible. The only possibility is to construct the normal form for the team, which is characterized by the $8 \times 8$ matrix given below.

**A**2

| | | Lll | Llr | Lrl | Lrr | Rll | Rlr | Rrl | Rrr |
|---|---|---|---|---|---|---|---|---|---|
| | UUu | 2 | 2.54 | 1.54 | 2.08 | 2.2 | 2.5 | 1.5 | 1.8 |
| | UDu | 1.4 | 1.82 | 0.82* | 1.24 | 2.2 | 2.5 | 1.5 | 1.8 |
| | DUu | 2.8 | 3.22 | 2.22 | 2.64 | 2.6 | 2.9 | 1.9 | 2.2 |
| **A**1 | DDu | 2.2 | 2.5 | 1.5 | 1.8 | 2.6 | 2.9 | 1.9 | 2.2 |
| | UUd | 2.38 | 3.46 | 1.46 | 2.54 | 2.86 | 3.44 | 1.9 | 2.5 |
| | UDd | 2.14 | 3.34 | 0.98 | 1.82 | 3.3 | 4.5 | 1.9 | 2.5 |
| | DUd | 3.54 | 4.62 | 2.38 | 3.22 | 3.7 | 4.9 | 2.3 | 2.9 |
| | DDd | 3.3 | 4.5 | 1.9 | 2.50 | 3.7 | 4.9 | 2.3 | 2.9 |

[16] We should note that this is specific to the problem at hand and is not a general rule. In general, optimal team cost in case 4 and with perfect recall will be higher than the one at case 3 where some sharing of measurements is allowed.

The unique solution is (UDu, Lrl), with a cost level of $J_{NSR} = 0.82$. Note that UDu stands for

$$\gamma_1^{1*}(x_1) = \begin{cases} U, & x_1 = x_{11}, \\ D, & x_1 = x_{12}, \end{cases} \qquad \gamma_2^{1*} = u,$$

and Lrl denotes the policy

$$\gamma_1^{2*} = L, \quad \gamma_2^{2*}(w_2) \begin{cases} r, & w_1 = w_{21}, \\ l, & w_1 = w_{22}, \end{cases}$$

5. *Open-loop information* (*OLI*). Here the agents use only the measurements they have obtained at stage 1, which is $x_1$ for **A**1 and no measurement for **A**2. The normal form here is, in fact, a submatrix that can be obtained from the $8 \times 8$ matrix of case 4. For **A**1 the permissible policies are still the same. For **A**2, however, the permissible ones are Ll, Lr, Rl, Rr, which correspond in case 4 to Lll, Lrr, Rll, Rrr. Hence, we retain only the *first*, *fourth*, *fifth*, and *eighth* columns of the matrix of case 4, and the result is the unique team-optimal solution (UDu, Lr), with a cost level of $J_{OL} = 1.24$. We should note in passing that the normal form of case 1 can also be recovered from the normal form of case 4, this time by also eliminating the *second*, *third*, *sixth*, and *seventh* rows of the matrix of case 4.

## *Cost Comparisons*

Clearly, more information to any one agent in a team will never result in higher optimal team cost and in fact could lead to a strictly lower value. In the latter case, we say that the extra information is *useful* (or *worth receiving*). In the context of this specific example, we have the optimum cost comparisons

$$\begin{aligned} J_{SP} = 1.8 > J_{OL} = 1.24 > J_{NSR} = 0.82 > J_{1DMS} = J_{1DIS} = 0.74 \\ > J_{CMS} = J_{CIS} = 0, \end{aligned}$$

and hence in each case the extra information to one or more agents has been worth receiving (with the exception of pure action information in the cases of *CIS* and *1DIS IS*s). The equalities $J_{1DIS} = J_{1DMS}$, $J_{CMS} = J_{CIS}$ hold not only for this specific example but also for the general stochastic team problems.

## 2.6 Team-Optimal Solutions for Static Teams

We present, in this section, a theory for static $N$-person stochastic teams, by focusing on fundamental issues such as the existence, uniqueness and derivation of team-optimal solutions, establishing conditions under which person-by-person (*pbp*) optimal solutions are also team-optimal, and studying the relationships between achievable optimal team costs and information structures in static teams.

Using the terminology and notation introduced in Sect. 2.2, we represent a general team by the $(N + 1)$-tuple $\{J; \Gamma^i, i \in \mathcal{N}\}$, where the cost $J$ is derived from a loss function, using a probability measure $P$ on the states of nature, common to all agents. We consider the cases where the action spaces $(U^i, i \in \mathcal{N})$ are either finite or infinite but finite dimensional, and for the latter class we also include the possibility that some hard constraints may be imposed on the decision variables, in which case the action constraint sets $(S^i, i \in \mathcal{N})$ are taken as appropriate closed subsets of the corresponding action spaces.

In the first subsection (Sect. 2.6.1), we consider the class of teams which are either finite or have finite measurement spaces for all agents. For this class, we provide general existence and uniqueness results for team-optimality and discuss the relationship with *pbp* optimality and the notion of *stationarity* (which is to be defined shortly). In Sect. 2.6.2, we extend this study to teams where the measurement spaces are infinite (but finite dimensional) and develop conditions under which stationarity implies team-optimality. Section 2.6.3 discusses two special, but important, classes of teams: (1) those with quadratic loss functions, first under general probability distributions and then under the Gaussian distribution, and (2) static teams with exponentiated quadratic loss functions. We also discuss recursive algorithms for the computation of the team-optimal solution in each case.

### *2.6.1 Teams with Finite Measurement Spaces*

We have already identified, in Sect. 2.3, one class of team problems for which an optimal solution always exists, namely, *static finite teams*, *i.e.,* teams where both the action and the measurement spaces are finite or, equivalently, the product strategy space ($\mathbf{\Gamma}$) is finite. This conclusion would also be valid for dynamic teams where the product strategy space ($\mathbf{\Gamma}$) is finite, since we would be doing a comparison among only a finite number of choices. It would be appropriate first to present this trivial, but useful, result as a **fact**.

**Fact 2.6.1.** *Every finite stochastic team admits at least one team-optimal solution.* $\diamond$

Two other key observations we made in Sect. 2.3 were that multiple team-optimal solutions are not necessarily interchangeable (respecting the order) and that a *pbp* optimal solution is not necessarily team-optimal, both of which we summarize below for future reference. These two *facts* are naturally valid not only for finite teams but for infinite teams as well and further not only for static teams but also for dynamic teams in normal (strategic) form.

**Fact 2.6.2.** *Neither multiple team-optimal solutions nor multiple pbp optimal solutions are necessarily interchangeable.* ⋄

**Fact 2.6.3.** *Every team-optimal solution is pbp optimal, but not vice versa.* ⋄

If a team problem is not finite, then one has to bring additional structure into the formulation in order to guarantee the existence, as well as the uniqueness, of the solution. Viewing a stochastic team in normal form as one of minimization[17] of a functional, $J$, over a set, $\mathbf{\Gamma}$, an optimum may fail to exist (in infinite teams) for basically one of two reasons:

1. The (cost) functional $J$ is unbounded below.
2. There exists an infimizing sequence in $\mathbf{\Gamma}$, without any limit in $\mathbf{\Gamma}$.

The former basically says that $J^*$, defined by the *RHS* of (2.4), is $-\infty$, implying that a sequence can be found in $\mathbf{\Gamma}$ which makes the value of $J$ arbitrarily small (negative). The only way to avoid this difficulty is to formulate, from the beginning, a well-defined team problem whose cost is bounded away from $-\infty$. The latter reason, however, cannot be dispensed with that easily since it places some nontrivial restrictions on the topology of the product policy space $\mathbf{\Gamma}$ as well as on the structure of $J$. In this case $J^*$, defined by the *RHS* of (2.4), is a finite quantity, but one can only achieve values arbitrarily close to (but larger than) $J^*$, and never equal to it. This could arise if, for example, the function $J$ has some discontinuities on $\mathbf{\Gamma}$ or $\mathbf{\Gamma}$ has some "holes" in it so that the infimizing sequence cannot have a limit in $\mathbf{\Gamma}$. The most general condition that ensures that these two things do not happen is the celebrated *Weierstrass theorem* (see Appendix A, Sect. A.5), rephrased below as a fact using the team framework.

**Fact 2.6.4.** *The team problem* $\{J; \Gamma^i, i \in \mathcal{N}\}$ *admits a team-optimal solution if the product policy space* $\mathbf{\Gamma}$ *is a compact subset of a normed linear vector space, and the cost function* $J$ *is lower semicontinuous (lsc) on* $\mathbf{\Gamma}$. ⋄

As one useful application of the above result, consider the class of stochastic team problems which satisfy the following four hypotheses:

$(c.1)$ Each action constraint set $S^i$ $(i \in \mathcal{N})$ is a closed and bounded subset of the action space $U^i$ $(i \in \mathcal{N})$ which is itself a finite-dimensional vector space.

$(c.2)$ $L(\xi; u^1, \ldots, u^N)$ is *almost surely* (*a.s.*) jointly *lsc* in $(u^1, \ldots, u^N) =: \mathbf{u}$, on $\mathbf{U} := U^1 \times \cdots \times U^N$.

$(c.3)$ Each measurement set $Y^i$ $(i \in \mathcal{N})$ is finite, with no element receiving *zero* probability from the probability measure $P$, or equivalently, for each $i \in \mathcal{N}$, the

[17] For some background material on the optimization of functionals, see Sect. A.5 of Appendix A.

partition set $\mathbf{Y}^i$ has a finite number of elements, with each element receiving positive probability from $P$.

$(c.4)$ $E_{\xi|y^i} L(\xi; u^1, \ldots, u^N)$ is finite for every $y^i \in Y^i$, $u^j \in U^j$, $i, j \in \mathcal{N}$.

Then we have the following theorem:

**Theorem 2.6.1.** *For an $N$-agent static stochastic team problem satisfying $(c.1)$–$(c.4)$ above, there exists at least one team-optimal solution.* ◇

*Proof.* The result follows from Fact 2.6.4, once we observe that, under the given specifications, the normal form has a *lsc* cost function $J$ on a compact policy space $\mathbf{\Gamma}$. We first show the latter, which is equivalent to showing that, for each $i \in \mathcal{N}$, $\Gamma^i$ is a closed and bounded subset of a finite-dimensional space. Toward this end, let $Y^i$ be generated (without any loss of generality) by the $n_i$-tuple $\{y_1^i, \ldots, y_{n_i}^i\}$, where $n_i := |Y^i|$ is finite by $(c.3)$. Then, every permissible strategy $\gamma^i$ for $\mathbf{A}i$ (*i.e.,* every element of $\Gamma^i$) can be written as

$$\gamma^i(y^i) = u_j^i, \qquad \text{if} \quad y^i = y_j^i, \quad j = 1, \ldots, n_i,$$

where each $u_j^i$ lies in $S^i$. Hence, each strategy can be viewed as an $n_i$-tuple of vectors $(u_1^i, \ldots, u_{n_i}^i)$ belonging to

$$\mathbf{S}^i := \underbrace{S^i \times \cdots \times S^i}_{n_i \text{times}} \subset \underbrace{U^i \times \cdots \times U^i}_{n_i \text{times}} =: \mathbf{U}^i,$$

which makes $\Gamma^i$ isomorphic to $\mathbf{S}^i$ which is closed and bounded, and finite dimensional, since it is a finite product of $S^i$ which itself is closed and bounded [by $(c.1)$].

We now show that $J$ is *lsc* on $\mathbf{U} := \mathbf{U}^1 \times \cdots \times \mathbf{U}^N$. To obtain a description of $J$ on $\mathbf{U}$, let us first introduce the notation $\mathbf{y}$ to denote an $N$-tuple of scalars

$$\mathbf{y} := (y_{t_1}^1, \ldots, y_{t_N}^N), \quad t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N},$$

where $y_{t_i}^i$ denotes one possible (generic) measurement of $\mathbf{A}i$. By a possible abuse of terminology, we will consider $\mathbf{y}$ as a random quantity, which has $\mathbf{N} := \prod_{i \in \mathcal{N}} n_i$ different realizations. Then, we have the following sequence of equalities:

$$\begin{aligned} J(\underline{\gamma}) &= E_\xi L(\xi; \gamma^1(\eta^1(\xi)), \ldots, \gamma^N(\eta^N(\xi))) \\ &= E_{\mathbf{y}} \underbrace{E_{\xi|\mathbf{y}} L(\xi; u_{t_1}^1, \ldots, u_{t_N}^N)}_{L_{\text{av}}(\mathbf{y}; u_{t_1}^1, \ldots, u_{t_N}^N)} \\ &\equiv \sum_{t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N}} L_{\text{av}}(\mathbf{y}; u_{t_1}^1, \ldots, u_{t_N}^N) \mathit{Prob}(\mathbf{y}), \end{aligned} \qquad (\star)$$

where the second line follows from the "iterated property" of conditional expectations (see Appendix B) and the adopted convention that $\gamma^i(y^i) = u^i_{t_i}$, when $y^i = y^i_{t_i}$.

Now, under $(c.2)$ and $(c.4)$, $L_{\text{av}}(\mathbf{y}_t; u^1_{t_1}, \dots, u^N_{t_N})$ is a *lsc* function on $\mathbf{u}$ (as well as on $U$) for each $\mathbf{y}_t$, since it is the integral of a *lsc* function ($L$) under the conditional measure *Prob*$(\xi|\mathbf{y})$, which is finite by $(c.4)$. In view of this, the last line of $(\star)$ (which is a finite weighted sum of individual *lsc* functions) provides a representation for $J$ on $U$, which is *lsc*. This then completes the proof of the theorem. It is worth noting at this point that the result would be true even if condition $(c.2)$ is relaxed somewhat, requiring instead that the function $L_{\text{av}}(\mathbf{y}; \mathbf{u})$ be *lsc* on $\mathbf{u}$, where $L_{\text{av}}(\mathbf{y}; \mathbf{u}) := E_{\xi|\mathbf{y},\mathbf{u}} L(\xi; \mathbf{u})$. □

For the general result of Theorem 2.6.1 to be valid, conditions $(c.2)$ and $(c.3)$ cannot be relaxed any further, because the relaxation of $(c.2)$ (with the provision above) would lead to violation of the *lsc* part of Fact 2.6.4, and the relaxation of $(c.3)$ (meaning that some of the $Y^i$'s might be infinite sets) leads to policy spaces that are no longer finite dimensional, in which case $(c.1)$ does not imply the compactness of $\mathbf{\Gamma}$. Relaxation of $(c.4)$, on the other hand, would lead to a $J$ that is not *lsc* at those points of $\mathbf{u}$ where it is unbounded. Of course, this condition is automatically satisfied if the underlying probability space is finite. The only condition that can be relaxed, without affecting the basic result of the theorem is $(c.1)$. To accommodate in our formulation the situation where some (or all) of the decision variables do not have hard constraints imposed on them, we have the following substitute condition:

$(c.1')$ Let $\mathcal{N}_h$ and $\mathcal{N}_s$ be two complementary subsets of $\mathcal{N}$ (*i.e.*, $\mathcal{N}_h \cup \mathcal{N}_s = \mathcal{N}$, and $\mathcal{N}_h \cap \mathcal{N}_s = \emptyset$) such that $S^i$ is compact for all $i \in \mathcal{N}_h$, and $S^j \equiv U^j$ for all $j \in \mathcal{N}_s$. Then, as $\sum_{j\in\mathcal{N}_s} |u^j| \to \infty$, $L(\xi; u^1, \dots, u^{\mathcal{N}}) \to \infty$ *a.s.*, for every fixed $u^i \in S^i, i \in \mathcal{N}_h$.

This condition ensures that $u^j, j \in \mathcal{N}_s$, can be restricted to a (possibly sufficiently large) compact set, thus making the result of Theorem 2.6.1 still valid. Hence we have

**Corollary 2.6.1.** *An $N$-agent static stochastic team problem satisfying* $(c.1')$, $(c.2)$, $(c.3)$, *and* $(c.4)$ *admits at least one team-optimal solution.* ◇

### Uniqueness

In view of Fact 2.6.2, it may be important to determine the conditions under which a team-optimal solution is unique, since as we have discussed earlier in Sect. 2.3, multiple optima may lead to an inferior outcome if the agents do not have a consistent protocol to resolve the dilemma. Once the existence of an optimum has been established, there would be two ways to verify uniqueness of the solution. One would be to write down a set of necessary conditions to be satisfied by the team-optimal solution and show that these conditions admit at most one solution—as to be discussed later in this subsection. A second way to verify unicity would be to use the normal form for the team and show strict convexity[18] of $J$ over the product policy space $\mathbf{\Gamma}$, which has to be a convex set. The following theorem now does precisely that, by relating the (strict) convexity of $L$ to the convexity of $J$ under the hypotheses of Theorem 2.6.1.

**Theorem 2.6.2.** *In addition to the four hypotheses of Theorem 2.6.1 or of Corollary 2.6.1, let $S^i$ be a convex set for each $i \in \mathcal{N}$ and $L(\xi;\cdot)$ be strictly convex on $\mathbf{U}$ a.s.*[19] *Then, the stochastic team problem admits a unique team-optimal solution.* $\diamond$

*Proof of Theorem 2.6.2.* First note that for $X$ a finite-dimensional vector space and $I$ a finite index set, if $f_i$, $i \in I$, is a convex (respectively, strictly convex) functional defined on $X$, then the functional $f$: $f = \sum_{i \in I} f_i$ is also convex (respectively, strictly convex) on $X$. Now, the construction given for $J$, in the proof of Theorem 2.6.1, satisfies the hypotheses of this result with $f_t(\cdot) = L_{\text{av}}(\mathbf{y}_t;\cdot)\mathit{Prob}(\mathbf{y}_t)$, since for each $\mathbf{y}_t$, $L_{\text{av}}(\mathbf{y}_t;\cdot)$ is strictly convex on $\mathbf{u}$ (being the conditional average of an *a.s.* strictly convex functional), and every $u^i_{t_i}$, $t_i \in \{1,\ldots,n_i\}, i \in \mathcal{N}$, appears in at least one of the additive terms in the representation $(\star)$ for $J$. Note, in passing, that $\mathit{Prob}(\mathbf{y}_t)$ may not be positive for every possible $N$-tuple $(t_1,\ldots,t_N)$, $t_i \in \{1,\ldots,n_i\}, i \in \mathcal{N}$, but for any $j \in \mathbf{N}$, and $t_j \in \{1,\ldots,n_j\}$, $y^j_{t_j}$ will receive positive probability in at least one such sequence, since otherwise this would imply that $\mathit{Prob}(y^j = y^j_{t_j}) = 0$, a contradiction to our initial hypothesis. $\square$

The following example serves to illustrate some of the fine points of the results of Theorems 2.6.1 and 2.6.2 and the analyses that led to these results, including the construction $(\star)$ used in the proof of Theorem 2.6.1.

*Example 2.6.1.* Let $N = 2$, $\Xi = U^1 = U^2 \equiv \mathbb{R}$, $\xi$ be a (continuous) random variable *uniformly distributed* on the open interval $(0,2)$, and the loss functional $L$ be given by

$$L(\xi;u^1,u^2) = (u^1)^2 + (u^2)^2 + \xi u^1 u^2 - u^1 - 2u^2.$$

Suppose that $\mathbf{A}1$ can tell (through his measurements and with certainty) whether the realized value of $\xi$ belongs to the open interval $(0,1)$ or not and $\mathbf{A}2$ can similarly tell whether it belongs to the subinterval $(\frac{1}{2},\frac{3}{2})$ or not. The question is whether this (static) stochastic team problem admits a team-optimal solution or not and, if it does, whether it is unique and how it can be computed.

Let us first check the conditions of Theorem 2.6.1 and Corollary 2.6.1. Clearly $(c.2)$ and $(c.3)$ are satisfied, where in the latter we choose $\mathbf{y}^1 = \{(0,1),\,[1,2)\}$, $\mathbf{y}^2 = \{(\frac{1}{2},\frac{3}{2}),\,(0,\frac{1}{2}] \cup [\frac{3}{2},2)\}$. We associate the measurement $y^i_1$ with the first subinterval (in each corresponding partition) and the measurement $y^i_2$ with the

[18] See Appendix A, Sect. A.4, for a definition.

[19] A random function $L(\xi;\mathbf{u})$ is *a.s. strictly convex* in $\mathbf{u}$ if the set of $\xi$ for which $L$ is not strictly convex in $\mathbf{u}$ is of $P_\xi$-measure zero.

complement (*i.e.*, the second) set. Condition $(c.1)$ is not satisfied, but $(c.1')$ is (with $\mathcal{N}_s = \mathcal{N}$), because, for each $\xi \in (0,2)$, the *Hessian matrix* of $L$ (see Appendix A, Sect. A.4),

$$\bigtriangledown^2 L(\xi, \mathbf{u}) = \begin{pmatrix} 2 & \xi \\ \xi & 2 \end{pmatrix},$$

is *positive definite* (*p.d.*), implying that $L(\xi, \mathbf{u}) \to \infty$ as $|u^1| + |u^2| \to \infty$, for every fixed $\xi \in (0,2)$. Note that even if the open interval $(0,2)$ is replaced with the closed interval $[0,2]$ (the distribution still being the same), $(c.1')$ would still be satisfied because, even though $\bigtriangledown^2 L(\xi, \mathbf{u})$ is no longer *p.d.* at $\xi = 2$ (in fact, then choosing $u^1 = -u^2$ and letting $u^2 \to \infty$, one can drive $L$ to $-\infty$), the singleton event $\{\xi = 2\}$ receives *zero* probability under the given continuous distribution, and hence the condition holds in the *a.s.* sense. If, however, we had a probability distribution with a *jump* at the point $\xi = 2$, assigning, say, a weight of $\frac{1}{3}$ to that single value, then $(c.1')$ would have been violated.

Now, in the course of the discussion above, we have also established the validity of the two additional hypotheses of Theorem 2.6.2 (the first one trivially and the second one because of the reason that the Hessian matrix of $L$ is *p.d. a.s.*), from which it follows that the problem indeed admits a unique team-optimal solution.

To obtain a characterization of the solution, let us first construct the normal form, following the steps outlined in the proof of Theorem 2.6.1. Noting that the summation in $(\star)$ has *four* terms, with *Prob*$((y_i^1, y_j^2)) = \frac{1}{4}, i, j = 1, 2$, some algebra leads to the expression

$$J(\underline{\gamma}) = \frac{1}{2} \sum_{i,j=1}^{2} (u_j^i)^2 - \frac{1}{2} \sum_{j=1}^{2} u_j^1 - \sum_{j=1}^{2} u_j^2 + \frac{1}{16}[u_1^1 u_2^2 + 3u_1^1 u_1^2 + 5u_2^1 u_1^2 + 7u_2^1 u_2^2],$$

which is to be minimized with respect to $(u_1^1, u_2^1, u_1^2, u_2^2)$ over $\mathbb{R}^4$. This is a strictly convex functional and is differentiable, which means that the unique solution should satisfy (uniquely) the *stationarity* conditions:

$$\partial J / \partial u_j^i = 0, \quad i, j = 1, 2.$$

These conditions reduce to the set of *four* linear equations:

$$\left.\begin{aligned} 16u_1^1 + 3u_1^2 + u_2^2 &= 8, \\ 16u_2^1 + 5u_1^2 + 7u_2^2 &= 8, \\ 3u_1^1 + 5u_2^1 + 16u_1^2 &= 16, \\ u_1^1 + 7u_2^1 + 16u_2^2 &= 16, \end{aligned}\right\} \qquad (\circ)$$

which admits the unique solution (to the nearest 6 decimal places):

$$u^{1*} := (u_1^{1^*}, u_2^{1^*}) = (0.231214, -0.323699),$$

$$u^{2^*} := (u_1^{2^*}, u_2^{2^*}) = (1.057803, 1.127168),$$

with the minimum value being

$$J^* \approx -1.141618.$$

It would be instructive to compare this value for the team cost with what would have been achieved if the agents had not made any measurements (*i.e.*, operated in an "open-loop" fashion with no measurements). In such a case, the normal form for the team would be given by the cost functional $J_{\mathrm{OL}}$ (where the subscript $_{\mathrm{OL}}$ stands for "open-loop");

$$J_{\mathrm{OL}} = (u^1)^2 + (u^2)^2 - u^1 - 2u^2 + u^1 u^2,$$

since $E[\xi] = 1$. A straightforward minimization of this quadratic (and strictly convex) functional leads to the unique solution

$$u^{1^*} = 0, u^{2^*} = 1 \quad \Rightarrow \quad J_{\mathrm{OL}}^* = -1.$$

Hence, we observe that the presence of the measurements (which bring the uncertainty in the true value of $\xi$ to intervals of length 1, instead of the original interval of length 2) leads to an improvement of (approximately) $14\,\%$ in the performance attained by the team.

Another extreme case to consider would be the information structure that provides the agents with the "maximum" information regarding the true value of $\xi$, which, unquestionably, is the measurement signal $y^i = \xi, i = 1, 2$ (*i.e.*, perfect measurement), unless some restrictions are imposed on the information structure. Our results, so far, as embodied in Theorems 2.6.1 and 2.6.2 and Corollary 2.6.1, are not (strictly speaking) applicable to problems of this type, since the measurements belong to infinite sets—this is the topic of the next subsection. However, because of the fact that both agents make perfect measurements here, the problem would be easy to analyze, since it is no different from a deterministic (quadratic) optimization problem. As such, this particular (team) problem is not well defined, since the objective functional is not strictly convex at $\xi = 2$, meaning (in this case) that for $\xi = 2$ the loss functional can be driven to $-\infty$, which in turn implies (in that case) that under the uniform distribution on $(0, 2)$ the cost (average loss) can be made arbitrarily small (negative).[20] The message here is that not only the characterization

[20]One can determine the optimal decision rules for each value of $\xi \in (0, 2)$ (this can be done analytically), substitute these unique rules into the given loss functional, and see that its integral over the interval $(0, 2)$ does not exist—which shows that $J^*$ is unbounded under the perfect measurement information structure.

but also the existence of a solution in a stochastic team problem could very much depend on the underlying information structure. $\diamond$

We have thus seen, in the preceding example, a constructive procedure for obtaining closed-form solutions to a stochastic static team problem with finite measurement spaces. The question now is whether there are other (alternative) ways of obtaining the solution and also of verifying team-optimality of a given candidate solution without going through the derivation. Such tools would be provided by the *necessary conditions* satisfied by a team-optimal solution, one of which is *person-by-person* (*pbp*) optimality, the necessity of which has already been given in Fact 2.6.3. Recalling Remark 2.2.1, and particularly inequality (2.7), a *pbp* solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ for a static team problem $(J, \mathbf{\Gamma})$ would be given by

$$\min_{\beta \in \Gamma^i} J(\underline{\gamma}^{-i}*, \beta) = J(\underline{\gamma}^*), \quad i \in \mathcal{N}, \tag{2.24}$$

which can equivalently be written as

$$\min_{u \in S^i} E_{\xi|y^i} L(\xi; \underline{\gamma}^{-i} * (\mathbf{y}^{-i}), u) = E_{\xi|y^i} L(\xi; \underline{\gamma} * (\mathbf{y})), \quad i \in \mathcal{N}. \tag{2.25}$$

In other words, we have $N$ separate optimization problems, one for each agent, and in each case the remaining agents' policies frozen at their *pbp* optimal choices. Note that (2.24) is an optimization (of total expectation) in the policy space, whereas (2) is optimization (of conditional expectation) in the action space, for every value of the conditioning variable. If (2.25) admits a unique solution and if the original problem is known to have a team-optimal solution (as in the case of Theorem 2.6.1 or Corollary 2.6.1), then (2.25) provides an alternative way of obtaining that solution. If, however, (2.25) admits more than one solution, then one would like to determine whether all or some of these are team-optimal. Hence, derivation of conditions under which *pbp* optimality implies team-optimality is of natural interest. At the outset, one would expect *a.s.* convexity of $L(\xi; \mathbf{u})$ over $\mathbf{u} \in \mathbf{U}$ to play a role here. This is indeed the case, but convexity in itself is not a sufficient condition, as the following example demonstrates:

Consider the purely deterministic loss function $L : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$L(u^1, u^2) = \begin{cases} (u^1)^2 + (1 - u^2)^2, \; u^1 \geq u^2, \\ \\ (u^2)^2 + (1 - u^1)^2, \; u^1 < u^2, \end{cases}$$

which is strictly convex on $\mathbb{R}^2$. For any fixed $u^2 \in \mathbb{R}$, $\arg\min_{u^1} L(u^1, u^2) = u^2$, and likewise for fixed $u^1 \in \mathbb{R}$, $\arg\min_{u^2} L(u^1, u^2) = u^1$. Hence, there exist infinitely many *pbp* optimal solutions $(u^1 = u^2 = u, u \in \mathbb{R})$, but only one of these, namely, $(u^1 = u^2 = \frac{1}{2})$ is team-optimal.

The function $L$ above is nondifferentiable at the *pbp* optimal points $(u^1 = u^2)$, and in view of this observation one might wonder whether a similar "negative" result can be obtained if the loss function were continuously differentiable. The following lemma outrules this possibility for deterministic problems and provides a set of (tight) sufficient conditions for a *pbp* optimal solution to be globally optimal.

**Lemma 2.6.1.** *Let* $L : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N} \to \mathbb{R}$ *be a convex (deterministic) loss function, with a pbp optimal solution* $\mathbf{u}^\circ := (u^{1^\circ}, \ldots, u^{N^\circ})$*. If* $L$ *is continuously differentiable*[21] *at* $\mathbf{u}^\circ$*, then* $\mathbf{u}^\circ$ *is globally (team) optimal.* $\diamond$

*Proof.* From the definition of convexity, we have the inequality

$$L(\alpha\mathbf{v} + (1-\alpha)\mathbf{u}^\circ) \leq \alpha L(\mathbf{v}) + (1-\alpha)L(\mathbf{u}^\circ)$$

for any $\mathbf{v} = (v^1, \ldots, v^N) \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$ and every $\alpha \in [0,1]$. Rearranging this inequality, we obtain, for $0 < \alpha \leq 1$,

$$\frac{1}{\alpha}[L(\mathbf{u}^\circ + \alpha(\mathbf{v} - \mathbf{u}^\circ)) - L(\mathbf{u}^\circ)] \leq L(\mathbf{v}) - L(\mathbf{u}^\circ),$$

and letting $\alpha \downarrow 0$, we arrive at

$$\sum_{i=1}^{N} \nabla_{u^i} L(\mathbf{u}^\circ)(v^i - u^{oi}) \leq L(\mathbf{v}) - L(\mathbf{u}^\circ), \quad \forall \mathbf{v} \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N},$$

where the required derivatives exist and the chain rule applies since $L$ is continuously differentiable at the given point. Furthermore, by the *pbp* optimality of $\mathbf{u}^\circ$, all these partial derivatives vanish, leading to

$$L(\mathbf{u}^\circ) \leq L(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N},$$

which proves global optimality of $\mathbf{u}^\circ$. $\square$

This lemma now finds a natural generalization to static stochastic team problems with finite measurement spaces. First, we formally introduce the notion of a "stationary policy."

**Definition 2.6.1.** Given a static stochastic team problem $\{J; \Gamma^i, i \in \mathcal{N}\}$, a policy $N$-tuple $\underline{\gamma} \in \mathbf{\Gamma}$ is *stationary* if (i) $J(\underline{\gamma})$ is finite, (ii) the $N$ partial derivatives in the following equations are well defined (locally), and (iii) $\underline{\gamma}$ satisfies these equations:

$$\left[\nabla_{u^i} E_{\xi|y^i} L(\xi; \underline{\gamma}^{-i}(\mathbf{y}^{-i}), u^i)\right]|_{u^i=\gamma^i(y^i)} = 0, a.s. \quad i \in \mathcal{N}. \tag{2.26}$$

$\diamond$

[21] Here, if one is to generalize the space on which $L$ is defined, such as an infinite-dimensional space, Fréchet differentiability would be a sufficient condition. In fact continuous differentiability, and thus continuity of partial derivatives for a finite-dimensional function (see Appendix A.4), implies Fréchet differentiability [140]. The key aspect required is that the chain rule in differentiation applies, which is the case for Fréchet differentiable functions, and not necessarily the case for weaker forms of differentiability. See also Radner [316] for a related discussion.

Clearly, the *stationarity condition* (2.26) is a necessary condition for (2.25) if $L(\xi;\mathbf{u})$ is continuously differentiable in each agent's action variable (not necessarily jointly) for every $\xi \in \Xi$, and $S^i, i \in \mathcal{N}$, are open subsets of finite-dimensional vector spaces. It is equivalent to (2.25) if furthermore $S^i, i \in \mathcal{N}$, are convex sets, and $L(\xi;\mathbf{u})$ is convex in $u^i, i \in \mathcal{N}$, for every $\xi \in \Xi$. The following theorem now basically says that if the convexity and continuous differentiability of $L$ is jointly in all the agents' action variables, then a stationary policy is necessarily team-optimal.

**Theorem 2.6.3.** *For an $N$-agent static stochastic team problem, let the hypotheses $(c.3)$ and $(c.4)$ be satisfied, $S^i$ be an open convex subset of a finite-dimensional vector space, for each $i \in \mathcal{N}$, and $L(\xi,\cdot)$ be convex and continuously differentiable on $\mathbf{S} := S^1 \times \cdots \times S^N$. Under these conditions, if the policy $\underline{\gamma}^{\circ}$, taking values in $\mathbf{S}$, is stationary, it is team-optimal.* $\diamond$

*Proof.* Using the construction given in the proof of Theorem 2.6.2, $J$ admits a representation on the space $U$, which is convex and continuously differentiable (this last property follows because by $(c.4)$ the function $L_{av}(\mathbf{y}_t;\cdot)$ is continuously differentiable, and the representation for $J$ is a finite weighted sum of such functions). Then, the result follows by a direct application of Lemma 2.6.1. $\square$

*Example 2.6.1. continued.* Returning to the static team of Example 2.6.1, so as to apply Theorem 2.6.3, first by the "monotone convergence theorem" (see the proof of Theorem 2.6.4), the conditional expectation and differentiation in (2.26) can be interchanged, leading to the equivalent stationarity conditions

$$E_{\xi|y^i}\{(\partial/\partial u^i)L(\xi;\gamma^j(y^j),u^i)\}|_{u^i=\gamma^i(y^i)} = 0, \quad i \neq j, i, j = 1,2$$

$\Leftrightarrow$

$$2\gamma^1(y^1) + E[\xi\gamma^2(y^2)|y^1] - 1 = 0,$$

$$2\gamma^2(y^2) + E[\xi\gamma^1(y^1)|y^2] - 2 = 0.$$

Since $y^1$ and $y^2$ each take two different values, this pair of equations is in fact a set of *four* linear equations, identical with the equations (∘) encountered earlier—as we would have expected. Note that to further simplify the pair of equations above, we can substitute for $\gamma^2$ from the second into the first, to arrive at a single equation in terms of $\gamma^1$,

$$4\gamma^1(y^1) + 2E[\xi|y^1] - 1 - E[\xi E[\gamma^1(y^1)|y^2]|y^1] = 0,$$

which can be solved uniquely for $u_1^1 = \gamma^1(y_1^1)$ and $u_2^1 = \gamma^1(y_2^1)$, to yield $u_1^{1^*} = 0.231214, u_2^{1^*} = -0.323699$. The stationary policies of agent 2 can likewise be obtained. Since the loss function $L$ satisfies all the hypotheses of Theorem 2.6.3, these stationary policies are indeed team-optimal. $\diamond$

### 2.6.2 Teams on Finite-Dimensional Spaces

We now extend the theory of the previous subsection from finite spaces to a class of uncountable (but finite-dimensional) measurement spaces, with the action spaces again taken as finite-dimensional vector spaces. Given such a team problem $\{J; \Gamma^i, i \in \mathcal{N}\}$, at least one of the policy spaces $(\Gamma^i)$ will be infinite dimensional, which means that condition $(c.1)$ will no longer imply that the policy space $\mathbf{\Gamma}$ is compact. Hence, even though Fact 2.6.1 would still be applicable in this case, a counterpart of Theorem 2.6.1 (on the existence of a team solution) will not follow from the given conditions. If $\underline{\gamma}^* \in \mathbf{\Gamma}$ is a team-optimal solution, then it will necessarily satisfy the *pbp* optimality condition (2.25) where now $Y^i = \mathbb{R}^{r_i}, i \in \mathcal{N}$. If furthermore, the action constraint sets are open, and the function to be minimized in (2.25) is continuously differentiable in the minimizing argument, this being so for all $i \in \mathcal{N}$, then the team solution should satisfy the stationarity conditions (2.26). The question now is whether there exists a counterpart of Theorem 2.6.3, to ensure that every stationary solution is also team-optimal. We first have the following theorem which provides a set of sufficient conditions for a policy $N$-tuple to be team-optimal.

**Theorem 2.6.4.** *Let $\{J; \Gamma^i, i \in \mathcal{N}\}$ be a static stochastic team problem where $U^i \equiv \mathbb{R}^{m_i}, i \in \mathcal{N}$, the loss function $L(\xi, \mathbf{u})$ is convex and continuously differentiable in $\mathbf{u}$ a.s., and $J(\underline{\gamma})$ is bounded from below on $\mathbf{\Gamma}$. Let $\underline{\gamma}^* \in \mathbf{\Gamma}$ be a policy $N$-tuple with a finite cost $(J(\underline{\gamma}^*) < \infty)$, and suppose that for every $\underline{\gamma} \in \mathbf{\Gamma}$ such that $J(\underline{\gamma}) < \infty$, the following $N$ inequalities hold:*

$$E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i*}(y^i)]\} \geq 0, \quad i \in \mathcal{N}, \tag{2.27}$$

*where $E\{\cdot\}$ denotes the total expectation. Then, $\underline{\gamma}^*$ is a team-optimal policy, and it is unique if $L$ is strictly convex in $\mathbf{u}$.* $\diamond$

*Proof.* First, by the convexity of $L$, we obtain (as in the proof of Lemma 2.6.1)

$$\frac{1}{\alpha}\left[L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})]) - L(\xi; \underline{\gamma}^*(\mathbf{y}))\right] \leq L(\xi; \underline{\gamma}(\mathbf{y})) - L(\xi; \underline{\gamma}^*(\mathbf{y})),$$

for all $\alpha \in (0, 1]$. Using the definition of $J$, this inequality can equivalently be written as (by taking the total expectation):

$$h(\alpha) := \frac{1}{\alpha}[E\{L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})])\} - J(\underline{\gamma}^*)] \quad \leq J(\underline{\gamma}) - J(\underline{\gamma}^*),$$

where $\alpha \in (0, 1]$. Note that both $J(\underline{\gamma})$ and $J(\underline{\gamma}^*)$ are finite, by hypothesis, and the first random variable (i.e., the first loss function) also has a finite expectation for every $\alpha \in (0, 1]$ because of the bound provided by the inequality. Now, due to the convexity of $L$, its finite integral, $E\{L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})])\}$ is also convex in $\alpha$. This leads to the conclusion that (by a property of convex functionals, given in Appendix A, Sect. A.4) $h(\alpha)$ is a monotonically nonincreasing function as $\alpha \downarrow 0$, and furthermore $h(1) \equiv J(\underline{\gamma}) - J(\underline{\gamma}^*)$ is bounded (by hypothesis). It then follows from the *monotone convergence theorem* (see Appendix B) that $\lim_{\alpha \downarrow 0} h(\alpha)$ exists, and the limit and expectation operations can be interchanged. As a consequence of continuous differentiability, this then leads to the inequality

$$\sum_{i=1}^{N} E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i^*}(y^i)]\} \quad \leq J(\underline{\gamma}) - J(\underline{\gamma}^*)$$

from which team-optimality of $\underline{\gamma}^*$ follows, since the left-hand side is nonnegative, by (2.27).

If $L$ were strictly convex in $\mathbf{u}$, a.s., then all the inequalities above would be strict, for $\underline{\gamma} \neq \underline{\gamma}^*$, thus leading to

$$J(\underline{\gamma}^*) < J(\underline{\gamma}),$$

which says that $\underline{\gamma}^*$ is the unique team-optimal solution. □

Note that the conditions of Theorem 2.6.4 above do not include the stationarity of $\underline{\gamma}^*$, and furthermore inequalities (2.27) may not generally be easy to check, since they involve all permissible policies $\underline{\gamma}$ (with finite cost)—generally an uncountable set. It is therefore important to obtain more readily checkable conditions to replace (2.27) and to relate team-optimality to stationarity. Either one of the following two conditions will accomplish this goal:

$(c.5)$ For all $\underline{\gamma} \in \mathit{\Gamma}$ such that $J(\underline{\gamma}) < \infty$, the following random variables have well-defined (finite) expectations (i.e., mean values):

$$\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i*}(y^i)], \quad i \in \mathcal{N}$$

$(c.6)$ $\mathit{\Gamma}^i$ is a Hilbert space for each $i \in \mathcal{N}$ and $J(\underline{\gamma}) < \infty$ for all $\underline{\gamma} \in \mathit{\Gamma}$. Furthermore,

$$E_{\xi|y^i}\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y})\} \in \mathit{\Gamma}^i, \quad i \in \mathcal{N}.$$

Of course, $(c.6)$ can be obtained from $(c.5)$ if $\mathit{\Gamma}^i$, $i \in \mathcal{N}$, are taken as Hilbert spaces. Here we give it as a separate condition because in some problems (such as linear quadratic—as we shall see shortly) $(c.6)$ follows quite readily from the problem formulations.

**Theorem 2.6.5.** *Let $\{J; \mathit{\Gamma}^i, i \in \mathcal{N}\}$ be a static stochastic team problem which satisfies all the hypotheses of Theorem 2.6.4, with the exception of the set of inequalities (2.27). Instead of (2.27), let either (c.5) or (c.6) be satisfied. Then, if $\underline{\gamma}^* \in \mathbf{\Gamma}$ is a stationary policy, it is also team-optimal. Such a policy is unique if $L(\xi; \mathbf{u})$ is strictly convex in $\mathbf{u}$, a.s.* ◇

*Proof.* We prove the result under condition $(c.6)$ and leave its verification under $(c.5)$ as an exercise. Clearly, what we need to show is that stationarity of $\underline{\gamma}^*$ implies [under $(c.6)$] the set of inequalities (2.27). Firstly note that since $\mathit{\Gamma}^i$ is a vector space, $\gamma^i - \gamma^{i*} \in \mathit{\Gamma}^i$ for every $\gamma^i \in \mathit{\Gamma}^i$, and for every $\beta^i \in \mathit{\Gamma}^i$, there exists a $\gamma^i \in \mathit{\Gamma}^i$

such that $\beta^i = \gamma^i - \gamma^{i*}$. Since $\beta^i \in \Gamma^i \Rightarrow -\beta^i \in \Gamma^i$, the set of inequalities (2.27) become equivalent to

$$E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))\beta^i(y^i)\} = 0, \quad \forall \beta^i \in \Gamma^i, \quad i \in \mathcal{N}$$

$\Leftrightarrow$

$$E_{y^i}\{E_{\xi|y^i}[\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))]\beta^i(y^i)\} = 0, \quad \forall \beta^i \in \Gamma^i, \quad i \in \mathcal{N},$$

where the second line follows from the iterated property of conditional expectation, under condition ($c.6$). Since both product terms above belong to $\Gamma^i$ which is a Hilbert space, and the equality is required to hold for every element of $\Gamma^i$, $i \in \mathcal{N}$, the last line becomes equivalent to

$$E_{\xi|y^i}[\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))] = 0, \text{a.s. } y^i, \quad i \in \mathcal{N}.$$

To complete the proof, we now have to show that the stationarity condition (2.26) implies the above, which would be true if we were able to interchange the derivative (which is a limit) and conditional expectation operations. This however is justified (using again the monotone convergence theorem, as in the proof of Theorem 2.6.4), since $J(\underline{\gamma})$ is finite for all $\underline{\gamma} \in \mathbf{\Gamma}$ and the conditional expectation above is well defined (as an element of a Hilbert space). □

Theorem 2.6.5 above thus provides an extension of the result of Theorem 2.6.3 from finite to infinite measurement sets. To appreciate some of the fine points of Theorems 2.6.4 and 2.6.5, let us now consider the following example, which was discussed by Radner [316] and Krainak et al. [218].

*Example 2.6.2.* Let $N = 2$, $\Xi = U^1 = U^2 = \mathbb{R}$, $\xi = x$ be a Gaussian random variable with zero mean and unit variance ($\sim N(0,1)$), and the loss functional be given by

$$L(x; u^1, u^2) = (u^1 - u^2)^2 e^{x^2} + 2u^1 u^2.$$

Note that $L$ is strictly convex and continuously differentiable in $(u^1, u^2)$ for every value of $x$. Hence, if the true value of $x$ were known to both agents, the problem would admit a unique team-optimal solution: $u^1 = u^2 = 0$, which is also stationary. Since this team-optimal solution does not use the precise value of $x$, it is certainly optimal also under "no-measurement" information (the other extreme scenario). Note, however, that in this case the only pairs that make $J(\gamma)$ finite are $u^1 = u^2 = u \in \mathbb{R}$, since

$$E[e^{x^2}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{+\frac{x^2}{2}} dx = \infty.$$

With the set of permissible policies not being an open set, clearly we cannot talk about stationarity in this case. Theorem 2.6.4 (which does not involve stationarity) is applicable here, where inequality (2.27) is satisfied trivially. Note also that for every $u \in \mathbb{R}$, $u^1 = u^2 = u$ is a *pbp optimal* solution, but only one of these is team-optimal.

Now, as a more interesting case, consider the measurement scheme:

$$y^1 = x + w^1; \qquad y^2 = x + w^2,$$

where $w^1$ and $w^2$ are independent random variables uniformly distributed on the interval $[-1, 1]$, which are also independent of $x$.[22] Clearly, $u^1 = u^2 = 0$ is team-optimal for this case also, but it is not obvious at the outset whether it is stationary or not. Toward this end, let us evaluate (2.26) for $i = 1$ and with $\gamma^2(y^2) = 0$:

$$(\partial/\partial u^1)E_{x,y^2|y^1}\{(u^1)^2 e^{\xi^2}\} = (\partial/\partial u^1)[(u^1)^2 E_{x|y^1}\{e^{\xi^2}\}] = 2u^1 E_{x|y^1}\{e^{\xi^2}\}$$

where the last step follows because the conditional probability density of $x$ given $y^1$ is nonzero only in a finite interval (thus making the conditional expectation finite). By symmetry, it follows that both derivatives in (2.26) vanish at $u^1 = u^2 = 0$, and hence the team-optimal solution is stationary. It is not difficult to see that in fact this is the only pair of stationary policies. Note that all the hypotheses of Theorem 2.6.5 are satisfied here, under condition $(c.5)$. $\diamond$

### 2.6.3 Two Special Cost Structures

We now specialize the above general result to two classes of teams with special cost structures, namely, quadratic and exponentiated quadratic loss functions. In both cases the team loss function will be strictly convex and continuously differentiable, so that (2.27) provides a sufficient condition for a policy $\underline{\gamma}^* \in \boldsymbol{\Gamma}$ to be team-optimal. We will further observe that the conditions of Theorem 2.6.5 are satisfied, so that stationary policies are also team-optimal.

#### Static Teams with Quadratic Loss

Given a probability space $(\Omega, \mathbf{F}, P_\Omega)$, and an associated vector-valued random variable $\xi$, let $\{J; \Gamma^i, i \in \mathcal{N}\}$ be a static stochastic team problem with the following specifications:

(i) $U^i \equiv \mathbb{R}^{m_i}$, $i \in \mathcal{N}$, i.e., the action spaces are unconstrained Euclidean spaces.
(ii) The loss function is a quadratic function of $\mathbf{u}$ for every $\xi$:

$$L(\xi; \mathbf{u}) = \sum_{i,j \in \mathcal{N}} u^{i'} R_{ij}(\xi) u^j + 2 \sum_{i \in \mathcal{N}} u^{i'} r_i(\xi) + c(\xi), \qquad (2.28)$$

[22] Note that here the random state of nature, $\xi$, is chosen as $(x, w^1, w^2)'$.

where $R_{ij}(\xi)$ is a matrix-valued random variable (with $R_{ij} \equiv R'_{ji}$), $r_i(\xi)$ is a vector-valued random variable, and $c(\xi)$ is a random variable, all generated by measurable mappings on the random state of nature, $\xi$.

(iii) $L(\xi; \mathbf{u})$ is strictly (and uniformly) convex in $\mathbf{u}$ a.s., i.e., there exists a positive scalar $\alpha$ such that, with $R(\xi)$ defined as a matrix comprised of $N$ blocks, with the *ij*'th block given by $R_{ij}(\xi)$, the matrix $R(\xi) - \alpha I$ is positive definite a.s., where $I$ is the appropriate dimensional identity matrix.

(iv) $R(\xi)$ is uniformly bounded above, i.e., there exists a positive scalar $\beta$ such that the matrix $\beta I - R(\xi)$ is positive definite a.s.

(v) $Y^i \equiv \mathbb{R}^{r_i}$, $i \in \mathcal{N}$, i.e., the measurement spaces are unconstrained Euclidean spaces.

(vi) $y^i = \eta^i(\xi), i \in \mathcal{N}$, for some appropriate Borel measurable functions $\eta^i$, $i \in \mathcal{N}$.

(vii) $\Gamma^i$ is the (Hilbert) space of all Borel measurable mappings of $\gamma^i : \mathbb{R}^{r_i} \to \mathbb{R}^{m_i}$, which have bounded second moments, i.e., $E_{y^i}\{\gamma^{i'}(y^i)\gamma^i(y^i)\} < \infty$.

(viii) $E_\xi[r'_i(\xi)r_i(\xi)] < \infty, i \in \mathcal{N}; \quad E_\xi[c(\xi)] < \infty$.

**Definition 2.6.2.** A static stochastic team is *quadratic* if it satisfies (i)–(viii) above. It is a *standard quadratic team* if furthermore the matrix $R$ is constant for all $\xi$ (i.e., it is deterministic). If, in addition, $\xi$ is a Gaussian distributed random vector, and $r_i(\xi) = Q_i\xi, \eta^i(\xi) = H^i\xi$, $i \in \mathcal{N}$, for some deterministic matrices $Q_i, H^i$, $i \in \mathcal{N}$, the decision problem is a *quadratic-Gaussian team* (more widely known as a linear-quadratic-Gaussian (LQG) team under some further structure on $Q_i$ and $H^i$). ◇

We now first show that the cost function of a quadratic team is bounded and strictly convex on $\mathbf{\Gamma}$.

**Proposition 2.6.1.** *For a quadratic team,*

*(i)* $|J(\underline{\gamma})| < \infty$ *for all* $\underline{\gamma} \in \mathbf{\Gamma}$.
*(ii)* $J(\underline{\gamma})$ *is strictly convex on* $\mathbf{\Gamma}$. ◇

*Proof.* For each $\underline{\gamma} \in \mathbf{\Gamma}$, each component of $u^i = \gamma^i(y^i)$ is a random variable on $(\Omega, \mathcal{F}, P_\Omega)$ with a bounded second moment (i.e., it is a *second-order* random variable), this being true for all $i \in \mathcal{N}$. Now using the fact that the product of any two second-order random variables defined on the same probability space is a well-defined random variable (on the same probability space) with a finite mean value (see Appendix B), it follows that the expected value of the second term of (2.28) is finite. Furthermore, since $R(\xi)$ is uniformly bounded, the expected value of the first term satisfies the bound

$$0 \le E\{\sum_{i,j} u^{i'} R_{ij}(\xi)u^j\} \equiv E\{\mathbf{u}'R(\xi)\mathbf{u}\} \le \beta E\{\mathbf{u}'\mathbf{u}\},$$

where $E\{\mathbf{u}'\mathbf{u}\}$ is finite by the same reasoning as above. Then, it readily follows that $L(\xi; \mathbf{u})$, with $u^i = \gamma^i(y^i)$, $i \in \mathcal{N}$, is a well-defined random variable with a *finite* expectation. Now, since $L(\xi, \mathbf{u})$ is strictly convex in $\mathbf{u}$ for every $\xi$, we have the strict inequality

$$L(\xi; \tilde{\alpha}\underline{\gamma}(\mathbf{y}) + (1-\tilde{\alpha})\underline{\hat{\gamma}}(\mathbf{y})) < \tilde{\alpha}L(\xi; \underline{\gamma}(\mathbf{y})) + (1-\tilde{\alpha})L(\xi; \underline{\hat{\gamma}}(\mathbf{y}))$$

for all $\tilde{\alpha} \in (0,1)$, and every $\underline{\gamma}$, $\hat{\underline{\gamma}} \in \mathbf{\Gamma}$, $\underline{\gamma} \neq \hat{\underline{\gamma}}$. Taking the expected values of both sides, which are finite as shown above, we arrive at

$$J(\tilde{\alpha}\underline{\gamma} + (1-\tilde{\alpha})\hat{\underline{\gamma}}) < \tilde{\alpha} J(\underline{\gamma}) + (1-\tilde{\alpha})J(\hat{\underline{\gamma}}),$$

which shows that $J$ is strictly convex. □

Now, the stationarity conditions (2.26) associated with the loss functional (2.28) can be evaluated:

$$[\nabla_{u^i}\left\{E_{\xi|y^i}\sum_{k,j\in\mathcal{N}} u^{k'}R_{kj}(\xi)u^j + 2u^{i'}E_{\xi|y^i}r_i(\xi) + 2E_{\xi|y^i}\sum_{j\in\mathcal{N},j\neq i} u^{j'}r_j(\xi) + E_{\xi|y^i}c(\xi)\right\}]|_{u^i=\gamma^i(y^i)} = 0, \quad i\in\mathcal{N}$$

$$\Leftrightarrow [E_{\xi|y^i}[R_{ii}(\xi)]u^i + \sum_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}R_{ij}(\xi)u^j + E_{\xi|y^i}r_i(\xi)]|_{u^i=\gamma^i(y^i)} = 0, \quad i\in\mathcal{N}$$

$$\Leftrightarrow E_{\xi|y^i}[R_{ii}(\xi)]\gamma^i(y^i) + \sum_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j(y^j)] + E_{\xi|y^i}r_i(\xi) = 0, \quad i\in\mathcal{N}, \tag{2.29}$$

where in going from the first to the second line of the equation we have simply performed vector differentiation with respect to $u^i$ which is outside the conditional expectation and have also used the fact that $R_{ij} \equiv R'_{ji}$.

Hence, (2.29) constitutes the set of stationarity conditions for the quadratic team. The following theorem, due to Radner [316], now says that the solution is unique and is team-optimal.

**Theorem 2.6.6.** *A quadratic static team (á la Definition 2.6.2) admits a unique team-optimal solution $\underline{\gamma}^* \in \mathbf{\Gamma}$, which is also the unique stationary solution satisfying (2.29).* ◇

*Proof.* Assuming that there exists a stationary solution [i.e., a solution to (2.29)], the uniqueness and team-optimality follow from Theorem 2.6.5, since all its hypotheses are satisfied along with condition ($c$.6). Hence the proof will be completed if we can show that there exists at least one solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ to (2.29). Here the verification is somewhat technical and requires some results from functional analysis and particularly Hilbert spaces (which are summarized in Appendix A, Sect. A.2). We outline here the crucial steps in this verification; the approach is essentially due to Radner [316].

Let us first note that the quadratic loss function (2.28) can equivalently be written as

$$L(\xi;\mathbf{u}) = \mathbf{u}'R(\xi)\mathbf{u} + 2\mathbf{u}'\mathbf{r}(\xi) + c(\xi)$$

$$\equiv [\mathbf{u} + R^{-1}(\xi)\mathbf{r}(\xi)]'R(\xi)[\mathbf{u} + R^{-1}(\xi)\mathbf{r}(\xi)] + c(\xi) - \mathbf{r}(\xi)'R^{-1}(\xi)\mathbf{r}(\xi),$$

where $R(\xi)$ is the matrix whose $ij$th block is $R_{ij}(\xi)$, and has an inverse (pointwise) by the a.s. strict convexity of $L$. Hence, if all agents had perfect access to the precise value of $\xi$, the minimum value of $J$ would be given by the expected value of the last two terms above. Since this is not the case, the actual minimum value of $J$ will be higher, the difference being due to the error in "approximating the vector $-R^{-1}(\xi)\mathbf{r}(\xi)$ using policies out of $\mathbf{\Gamma}$." An equivalent problem, therefore, is

$$\min_{\underline{\gamma}\in\mathbf{\Gamma}} \tilde{J}(\underline{\gamma}), \quad \tilde{J}(\underline{\gamma}) := E\{\|\underline{\gamma}(\mathbf{y}) + R^{-1}(\xi)\mathbf{r}(\xi)\|^2_{R(\xi)}\}, \tag{2.30}$$

and the statement just made (in "inverted commas") can be given a precise mathematical meaning as follows.

First note that the policy space $\mathbf{\Gamma}$ is the product space $\Gamma^1 \times \cdots \times \Gamma^N$, where each $\Gamma^i$ is in fact a Hilbert space, with the inner product

$$\langle \alpha^i, \beta^i \rangle_i := E[\alpha^i(y^i)'\beta^i(y^i)]$$

(see Appendix A, Sect. A.2, and Appendix B, Sect. B.1). This makes $\mathbf{\Gamma}$ also a Hilbert space, with the inner product

$$\langle \underline{\alpha}, \underline{\beta} \rangle := E\left\{\sum_{i\in\mathcal{N}} \alpha^i(y^i)'\beta^i(y^i)\right\} \equiv E[\underline{\alpha}(\mathbf{y})'\underline{\beta}(\mathbf{y})]. \tag{$\circ$}$$

Note the important restriction that $\underline{\alpha}$ and $\underline{\beta}$ are not allowed to depend on all components of $\mathbf{y}$, because different agents do not have access to the same set of measurements. Now, in order to be able to use (2.30) as a norm compatible with the given inner product, we have to change the ($\circ$) somewhat by weighting it with $R(\xi)$:

$$\langle \underline{\alpha}, \underline{\beta} \rangle = E[\underline{\alpha}(\mathbf{y})'R(\xi)\underline{\beta}(\mathbf{y})]. \tag{$\circ\circ$}$$

This actually changes $\mathbf{\Gamma}$, the space where $\underline{\alpha}$ and $\underline{\beta}$ belong, but because of the given properties of $R(\xi)$, we have an isometry between the two spaces and therefore can denote the one under the new inner product ($\circ\circ$) also by $\mathbf{\Gamma}$. If every component of $\underline{\gamma}$ were allowed to depend on the entire measurement vector $\mathbf{y}$ (which would be the case if all the agents were to share their measurements), then the set of all permissible $\underline{\gamma}$'s bounded under the norm induced by ($\circ\circ$) would be a much larger (than $\mathbf{\Gamma}$) space. Let us denote this space by $\mathbf{H}$, and note that it is also a Hilbert space, under the inner product ($\circ\circ$). An important observation now is that $\mathbf{\Gamma}$ is a closed linear subspace of $\mathbf{H}$, *closed* because every convergent sequence in $\mathbf{\Gamma}$ with a limit point will have the limit point in $\mathbf{\Gamma}$. Hence, the team-minimization problem (2.30) is in fact an orthogonal projection problem, one of orthogonally projecting the random vector $\mathbf{x}(\xi) := R^{-1}(\xi)\mathbf{r}(\xi)$ from $\mathbf{H}$ onto $\mathbf{\Gamma}$. The conditions of the *orthogonal projection theorem* given in Appendix A, Sect. A.2, are satisfied, and therefore there exists a *unique* element of $\mathbf{\Gamma}$ that solves (2.30). Furthermore, this unique element, say $\underline{\gamma}^*$, has the property that

$$\underline{\gamma}^* + \mathbf{x} \perp \gamma, \qquad \forall \gamma \in \mathbf{\Gamma}$$

(see Appendix A, Sect. A.2, for notation and terminology). Using the inner product (∘∘), this orthogonality relationship can be written as

$$\begin{aligned}
&\langle \underline{\gamma}^* + \mathbf{x}, \underline{\gamma} \rangle = E\left\{ [\underline{\gamma}^*(\mathbf{y}) + \mathbf{x}(\xi)]' R(\xi) \underline{\gamma}(\mathbf{y}) \right\} = 0 \\
\Leftrightarrow \\
&E\left\{ \sum_{i \in \mathcal{N}} \gamma^i(y^i)' \left[ \sum_{j \in \mathcal{N}} R_{ij}(\xi) \gamma^{*j}(y^j) + r_i(\xi) \right] \right\} = 0 \\
\Leftrightarrow \\
&E\Bigg\{ \sum_{i \in \mathcal{N}} \gamma^i(y^i)' \Bigg[ E_{\xi|y^i}\{R_{ii}(\xi)\} \gamma^{*i}(y^i) + \sum_{j \in \mathcal{N}, j \neq i} E_{\xi|y^i}\left\{ R_{ij}(\xi) \gamma^{*j}(y^j) \right\} \\
&\qquad\qquad + E_{\xi|y^i} r_i(\xi) \Bigg] \Bigg\} = 0,
\end{aligned}$$

where in arriving at the last line we have used the iterative property of conditional expectations. Now, since this equality has to hold for all $\underline{\gamma} \in \mathbf{\Gamma}$ and since $\mathbf{\Gamma}$ is a Hilbert space, it follows that the expression in brackets should vanish for every $i \in \mathcal{N}$[23] which is precisely (2.29). □

The proof of the theorem, as presented above, provides us with a new interpretation to the stationarity conditions (2.29). Note that they can be rewritten (compactly) as

$$PR\underline{\gamma} + P\mathbf{r} = 0, \tag{2.31}$$

where $P$ is a linear operator, block diagonal, with the *ii*'th block defined through

$$P_{ii}\beta^i(\xi) = E_{\xi|y^i}\beta^i(\xi), \quad i \in \mathcal{N},$$

where $\beta^i(\xi)$ is an $m_i$-dimensional measurable function of $\xi$, satisfying the boundedness condition $E\{\beta^i(\xi)'\beta^i(\xi)\} < \infty$. As such, the linear operator $P$ is a *projection operator* defined on a Hilbert space, whose operator norm is one (see Appendix A, Sect. A.2). Note that if the agents had full access to the value of $\xi$, then the stationarity condition would be

$$R(\xi)\underline{\gamma}(\xi) + \mathbf{r}(\xi) = 0, \tag{2.32}$$

[23] Here we have used the following property of Hilbert spaces: if $\langle \alpha, \beta \rangle = 0$ for all $\beta \in \mathbf{H}$, then $\alpha \equiv 0$.

which of course admits a unique solution since the matrix $R(\xi)$ is invertible for all $\xi$. Hence, the stationarity equation in the decentralized measurement case is a "projected" version of the one in the centralized full information case, but note, however, that the unique (decentralized) team-optimal solution is *not* a projected version of the centralized one $(-R^{-1}\mathbf{r})$.

The unique team-optimal solution can be obtained using some approximation schemes. Viewing (2.29) (or equivalently (2.31) as a fixed-point equation (see Appendix A, Sect. A.6, for details), one approach would be to use successive approximations:

$$\begin{cases} \gamma^i_{(k+1)}(y^i) = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}\left\{\sum\limits_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k)}(y^j)] + E_{\xi|y^i}r_i(\xi)\right\}, \\ \gamma^i_{(0)}(y^i) \quad\; \equiv 0, \quad i\in\mathcal{N}, \quad k=0,1,\ldots \end{cases} \tag{2.33}$$

which is called the *parallel update scheme*, where we have taken the starting points of the iteration as the *zero* function, as an arbitrary choice. This iteration models a dynamic decision process where the agents exchange policy information at every (discrete) point in time, and at the $k+1$'th instant agent $i$ solves the (stochastic) optimization problem

$$\begin{aligned} \min\nolimits_{\gamma^i\in\Gamma^i} J(\gamma^1_{(k)},\ldots,\gamma^{i-1}_{(k)},\gamma^i,\gamma^{i+1}_{(k)},\ldots,\gamma^N_{(k)}) \\ = J(\gamma^1_{(k)},\ldots,\gamma^{i-1}_{(k)},\gamma^i_{(k+1)},\gamma^{i+1}_{(k)},\ldots,\gamma^N_{(k)}),\; i\in\mathcal{N}. \end{aligned}$$

Clearly, if the parallel scheme converges, it will yield the (unique) team-optimal solution in the limit. However, there is generally no guarantee that it will converge, unless some conditions are imposed on the matrix $R$ and the probabilistic structure of the problem. To state two such conditions, let us first write (2.33) in compact form:

$$\underline{\gamma}_{(k+1)} = \mathbf{F}\underline{\gamma}_{(k)} + \hat{\mathbf{r}}, \tag{2.34}$$

where $\mathbf{F}$ is a linear operator mapping $\mathbf{\Gamma}$ into itself and composed of block operators with the diagonal blocks being zero and off-diagonal blocks given by

$$[\mathbf{F}_{ij}\gamma^j](y^i) = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}E_{\xi|y^i}[R_{ij}(\xi)\gamma^j(y^j)], \quad j\neq i, j\in\mathcal{N}. \tag{2.35}$$

Furthermore, $\hat{\mathbf{r}}\in\mathbf{\Gamma}$, with the $i$th block vector given by

$$[\hat{\mathbf{r}}(\mathbf{y})]_i = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}E_{\xi|y^i}r_i(\xi), \quad i\in\mathcal{N}. \tag{2.36}$$

Using the notation introduced in Appendix A, Sect. A.2, let $\ll \mathbf{F} \gg$ denote the *operator norm* of $\mathbf{F}$ and $\rho(\mathbf{F})$ denote its *spectral radius*; furthermore, note the inequality

$$\rho(\mathbf{F}) \leq \;\; \ll \mathbf{F} \gg .$$

The *Banach* and *successive approximation* theorems of Appendix A, Sect. A.6, now readily lead to the following result.

**Proposition 2.6.2.** *Consider the parallel update scheme (2.33) [equivalently (2.34)] for the general stochastic static team problem:*

*(i) The iteration converges for all starting points* $\underline{\gamma}_{(0)} \in \mathbf{\Gamma}$ *if, and only if,*

$$\rho(\mathbf{F}) < 1. \tag{2.37}$$

*(ii) The iteration converges for all starting points* $\underline{\gamma}_{(0)} \in \mathbf{\Gamma}$ *if*

$$\ll \mathbf{F} \gg \quad < 1, \tag{2.38}$$

*which is therefore a sufficient condition for (2.37).*

$\diamond$

It is important to note that nonsatisfaction of (2.37) does not necessarily imply that there is no recursive scheme which would compute $\underline{\gamma}^*$; in fact, there may exist nonparallel schemes or schemes that use relaxation (i.e., higher-order memory), which will have better convergence properties. As an example of a nonparallel scheme consider the so-called sequential scheme where the agents take their turns, one at a time and in strict order, to re-optimize their policies, i.e., $\gamma^i_{(k+1)}$ is determined through the minimization of

$$\min_{\gamma^i \in \Gamma^i} J(\gamma^1_{(k+1)}, \ldots, \gamma^{i-1}_{(k+1)}, \gamma^i, \gamma^{i+1}_{(k)}, \ldots, \gamma^N_{(k)}).$$

This then leads to the following counterpart of (2.33):

$$\begin{cases} \gamma^i_{(k+1)} = [E_{\xi|y^i} R_{ii}(\xi)]^{-1} \Bigg\{ \displaystyle\sum_{j \in \mathcal{N}, j<i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k+1)}(y^i)] \\ \qquad\qquad + \displaystyle\sum_{j \in \mathcal{N}, j>i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k)}(y^j)] + E_{\xi|y^i} r_i(\xi) \Bigg\}, i \in \mathcal{N}, k = 0, 1 \ldots, \\ \gamma^i_{(0)} \quad \equiv 0, \quad i \in \mathcal{N}, i \neq 1. \end{cases} \tag{2.39}$$

Note that this recursion cannot be written in a compact form as in (2.34). However, for such convex team problems, sequential schemes have more desirable convergence properties since the sequence of minimizations above leads to a monotone non-increasing sequence of positive real numbers (associated with the team cost) which has a limit (unlike the general parallel scheme in (2.33)).

### Standard Quadratic Teams

We now study the class of quadratic teams where the matrix $R(\xi)$ is a constant in $(\xi)$, i.e., $R$ is deterministic. The basic equation of stationarity, (2.29), simplifies to

$$\gamma^i(y^i) + \sum_{j\in\mathcal{N}, j\neq i} \tilde{R}_{ij} E_{\xi|y^i}[\gamma^j(y^j)] + E_{\xi|y^i}\tilde{r}_i(\xi) = 0, \quad i \in \mathcal{N}, \tag{2.40}$$

where

$$\tilde{R}_{ij} := R_{ii}^{-1} R_{ij}; \quad \tilde{r}_i(\xi) = R_{ii}^{-1} r_i(\xi). \tag{2.41}$$

Clearly, by Theorem 2.6.6, this equation admits a unique solution $\underline{\gamma}^* \in \mathbf{\Gamma}$, whenever the loss function is strictly convex (equivalently, if the constant matrix $R$ is positive definite). The counterpart of the parallel update scheme (2.33) is

$$\gamma^i_{(k+1)}(y^i) = -\sum_{j\in\mathcal{N}, j\neq i} \tilde{R}_{ij} E_{y^j|y^i}[\gamma^j_{(k)}(y^j)] - E_{\xi|y^i}\tilde{r}_i(\xi) \quad i \in \mathcal{N}, k = 0, 1, \ldots, \tag{2.42}$$

which we now further study for the case $N = 2$ (i.e., with only two agents). Substituting $\gamma^2_{(k+1)}$ obtained from (2.42) into the same for $i = 1$, we obtain

$$\gamma^1_{(k+2)}(y^1) = \tilde{R}_{12}\tilde{R}_{21} E_{y^2|y^1} E_{y^1|y^2}[\gamma^1_{(k)}(y^1)] + c^1(y^1), \tag{2.43}$$

where

$$c^1(y^1) := -E_{\xi|y^1}\tilde{r}_1(\xi) + \tilde{R}_{12} E_{y^2|y^1} E_{\xi|y^2}\tilde{r}_2(\xi). \tag{2.44}$$

Note that if we instead had the sequential update, (2.39), the resulting equation for $i = 1$ would be exactly (2.43) with simply $\gamma^1_{(k)}$ replaced by $\gamma^1_{(k+1)}$. Hence the parallel and sequential update schemes are essentially identical in the case of a two-agent team problem. The following proposition states this result, along with two other useful observations.

**Proposition 2.6.3.** *For the standard quadratic team with $N = 2$:*

*(i) The parallel update schemes (2.42) converge (to a limiting policy pair $\underline{\gamma}^* \in \mathbf{\Gamma}$, which is a team-optimal solution) if, and only if, the single update scheme (2.43) converges.*

*(ii) If (2.43) converges to a limiting policy $\underline{\gamma}^{1*} \in \Gamma^1$, then $\underline{\gamma}^{2*}$ is the unique team-optimal policy of agent $\mathbf{A}1$, and*

$$\gamma^{2*}(y^2) = -\tilde{R}_{21} E_{y^1|y^2}[\gamma^{1*}(y^1)] - E_{\xi|y^2}\tilde{r}_2(\xi)$$

*is the unique team-optimal policy of agent $\mathbf{A}2$.*

*(iii) The parallel and sequential update schemes require the same condition of convergence, which is*

$$\rho(\tilde{R}_{12}\tilde{R}_{21} E_{\xi|y^1} E_{\xi|y^2}) < 1\,,$$

*where $\rho(\cdot)$ is the spectral radius of its linear operator argument mapping $\Gamma^1$ into itself.*

$\diamond$

*Proof.* Parts (i) and (ii) are mere observations and require no proof. Part (iii) follows from the *successive approximation theorem* of Appendix A, Sect. A.6, since in (2.43) $\tilde{R}_{12}\tilde{R}_{21}E_{y^2|y^1}E_{y^1|y^2}$, which can equivalently be written as $\tilde{R}_{12}\tilde{R}_{21}E_{\xi|y^1}E_{\xi|y^2}$ is a linear bounded operator mapping $\Gamma^1$ (a Hilbert space) into itself. See also Proposition 2.6.2, and compare (2.43) with (2.31). $\square$

The following lemma now paves the way toward showing that the condition of Proposition 2.6.3(iii) is satisfied for all standard quadratic teams.

**Lemma 2.6.2.** *The loss function (2.28), with $N = 2$ and $R_{ij}$ constant matrices, is strictly convex if, and only if, $R_{22}$ is positive definite and*

$$\rho(\tilde{R}_{12}\tilde{R}_{21}) < 1.$$

$\diamond$

*Proof.* Strict convexity of $L$ is equivalent to the positive definiteness of the matrix

$$R := \begin{pmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{pmatrix},$$

which is further equivalent to (by definition)

$$\begin{pmatrix} x \\ y \end{pmatrix}' \begin{pmatrix} R_{11} & R_{12} \\ R_{12}' & R_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} > 0, \qquad \forall \begin{pmatrix} x \\ y \end{pmatrix} \neq \underline{o},$$

where $\underline{o}$ is the zero vector in $\mathbb{R}^m$, $m := m_1 + m_2$, and $x$, $y$ have dimensions compatible with the dimensions of the blocks of $R$. The above can be rewritten as

$$x'R_{11}x + 2x'R_{12}y + y'R_{22}y > 0$$

from which it follows that $R_{11} > 0$, $R_{22} > 0$ are necessary conditions for strict convexity. Now, minimizing this expression with respect to $y$, we have, by differentiation,

$$y = -R_{22}^{-1}R_{21}x$$

as the unique solution, substitution of which into the original expression leads to

$$\begin{aligned}\min_y \begin{pmatrix} x \\ y \end{pmatrix}' R \begin{pmatrix} x \\ y \end{pmatrix} &= x'(R_{11} - R_{12}R_{22}^{-1}R_{12}')x \\ &\equiv (R_{11}^{\frac{1}{2}}x)'(I - (R_{11}^{\frac{1}{2}})^{-1}R_{12}R_{22}^{-1}R_{12}'(R_{11}^{\frac{1}{2}})^{-1})(R_{11}^{\frac{1}{2}}x) > 0,\end{aligned}$$

for $x \neq 0$, where $R_{11}^{\frac{1}{2}}$ is the unique square root of $R_{11}$.

The strict inequality holds, for all nonzero $x$, if, and only if, the matrix

$$I - (R_{11}^{\frac{1}{2}})^{-1} R_{12} R_{22}^{-1} R_{12}' (R_{11}^{\frac{1}{2}})^{-1}$$

is positive definite, which is equivalent to all eigenvalues of the second (nonnegative definite) matrix to be less than one. Hence,

$$\rho((R_{11}^{\frac{1}{2}})^{-1} R_{12} R_{22}^{-1} R_{12}' (R_{11}^{\frac{1}{2}})^{-1}) \equiv \rho(R_{11}^{-1} R_{12} R_{22}^{-1} R_{12}') < 1,$$

where we used the fact that for two square matrices $A$ and $B$, $\rho(AB) = \rho(BA)$. Since $\tilde{R}_{12} = R_{11}^{-1} R_{12}$, $\tilde{R}_{21} = R_{22}^{-1} R_{21} \equiv R_{22}^{-1} R_{12}'$, this completes the proof of the lemma. □

This brings us to the following strengthened version of Theorem 2.6.6 for standard quadratic teams with $N = 2$.

**Theorem 2.6.7.** *For the two-agent standard quadratic team,*

*(i)* *There exists a unique team-optimal solution $\underline{\gamma}^* \in \mathbf{\Gamma}$, which is also the unique solution of (2.40) with $N = 2$.*
*(ii)* *Both the parallel and sequential update schemes converge for all starting points in $\mathbf{\Gamma}$.*
*(iii)* *Agent $\mathbf{A}i$'s optimal policy is given by the infinite sum*

$$\gamma^{i*}(y^i) = \sum_{k=0}^{\infty} (\tilde{R}_{ij} \tilde{R}_{ji} E_{y^j|y^i} E_{y^i|y^j})^k c^i(y^i), \quad i,j = 1,2; j \neq i, \tag{2.45}$$

*where $c^1$ is given by (2.44) and $c^2$ is defined by the same with 1's and 2's interchanged.* ◇

*Proof.* Of course, (i) follows from Theorem 2.6.6, but since (ii) implies (i) in view of Proposition 2.6.3, the independent proof that we will give for (ii) will also provide an alternative proof to this special case of Theorem 2.6.6.

To prove part (ii), it will be sufficient to verify the condition of Proposition 2.6.3(iii). Toward this end, let us first introduce a (Hilbert) space $\hat{\Gamma}^1$ of all $m_1$-dimensional measurable functions $\hat{\gamma}(y^1, y^2)$ with bounded second moments: $E_{\mathbf{y}}\{|\hat{\gamma}(y^1, y^2)|^2\} < \infty$. Clearly, $\Gamma^1$ is a subspace of $\hat{\Gamma}^1$. Now, the conditional expectation operator $E_{\xi|y^i} =: P^i$ is a projection operator on $\hat{\Gamma}^1$ (see Appendix B) and hence has operator norm one, for both $i = 1$ and $i = 2$. Since the product $P^1 P^2$ is also a linear bounded operator on $\hat{\Gamma}^1$, its norm is bounded by

$$\ll P^1 P^2 \gg \quad \leq \quad \ll P^1 \gg \ll P^2 \gg = 1.$$

An important observation here is that for any $\hat{\gamma} \in \hat{\Gamma}'$, $P^1 P^2 \hat{\gamma} \in \Gamma^1 \subset \hat{\Gamma}^1$, and hence we can also view the product operator $P^1 P^2$ as a bounded linear operator mapping $\Gamma^1$ into itself. Since $\tilde{R}_{12} \tilde{R}_{21}$ also maps $\Gamma^1$ into itself, we have

$$\rho(\tilde{R}_{12} \tilde{R}_{21} P^1 P^2) \leq \rho(\tilde{R}_{12} \tilde{R}_{21}) \rho(P^1 P^2),$$

which is the spectral radius inequality on Hilbert spaces (see Appendix A, Sect. A.2). The first product term above is strictly less than *one* by Lemma 2.6.2, and the second term is no greater than *one*, by

$$\rho(P^1P^2) \leq \quad \ll P^1P^2 \gg \quad \leq 1.$$

This completes the proof of (ii), in view of Proposition 2.6.3(iii).

For part (iii) simply note that for $i = 1$, (2.45) is the infinite summation obtained from (2.43) by taking $\gamma^1_{(0)} \equiv 0$, with the limit being a valid (well-defined) element of $\Gamma^1$ by part (ii). Clearly the same result holds for $i = 2$. □

*Remark 2.6.1.* The iteration (2.43), or more generally (2.42), is sometimes called the *infinite second guessing* scheme. If the agents had known each other's (optimal) policies, then the iteration would halt after one step. Since this knowledge is not there, they have to estimate (or guess) each other's actions, which would also involve the estimates of each other's estimates, etc., leading in general to an infinite, albeit convergent, sequence. ◇

Even though iteration (2.42) converges for $N = 2$, it does not necessarily converge for $N > 2$. This is mainly due to the fact that strict convexity of $L$ (equivalently, positive definiteness of $R$) does not imply that[24] $\rho(\tilde{R}) < 1$, unless $N = 2$.

Also one intuitive explanation for this discrepancy is that for $N = 2$ the iteration (2.42) corresponds to the sequence of minimizations

$$J(\gamma^1_{(k+1)}, \gamma^2_{(k)}) \quad = \min_{\gamma^1 \in \Gamma^2} J(\gamma^1, \gamma^2_{(k)}) =: J_{(k+1)},$$

$$J(\gamma^1_{(k+1)}, \gamma^2_{(k+2)}) = \min_{\gamma^2 \in \Gamma^2} J(\gamma^1_{(k+1)}, \gamma^2) =: J_{(k+2)}, \quad k = 0, 1, \ldots,$$

with the property

$$J_{(1)} \geq J_{(2)} \geq \ldots \geq J_{(k)} \geq J_{(k+1)} \geq \cdots,$$

Hence, at every step of the iteration, the value of $J$ cannot increase, thus generating a nonincreasing convergent sequence (of costs). This, of course, could also converge to a *pbp*-optimal solution, but we know in this case (as already shown) that a *pbp*-optimal solution (which is also stationary because of the special quadratic structure of the loss function) is also team-optimal.

For $N > 2$, however, the iteration (2.42) does not necessarily generate a monotonic cost sequence nor a subsequence that is monotonic, which is a reason for the failure of (2.42) to converge.

---

[24] Here $\tilde{R}$ is the $(m \times m)$ matrix whose diagonal blocks are zero, and off-diagonal blocks are given by $[\tilde{R}]_{ij} = \tilde{R}_{ij}$, as defined by (2.41).

In view of the above, a natural question that arises is whether there exists some other (computational) algorithm that would yield the unique solution of (2.40) (which is known to exist by Theorem 2.6.6). Toward studying this question, let us first rewrite (2.40) as follows, using the compact notation of (2.31):

$$PR\underline{\gamma} + P\mathbf{r} = 0. \qquad (*)$$

Let us add $-\epsilon\underline{\gamma}$ to both sides (where $\epsilon > 0$), and divide throughout by $\epsilon$, to obtain

$$\underline{\gamma} = P(I - \frac{1}{\epsilon}R)\underline{\gamma} - \frac{1}{\epsilon}P\mathbf{r}, \qquad (**)$$

where we have used the fact that the projection operator $P$ and the identity operator (matrix) $I$ commute. Note that $(*)$ and $(**)$ are in fact identical equations. Now, we associate the following iteration with $(**)$:

$$\underline{\gamma}_{(k+1)} = P(I - \frac{1}{\epsilon}R)\underline{\gamma}_{(k)} - \frac{1}{\epsilon}P\mathbf{r}, \quad k = 0, 1, \qquad (2.46)$$

which, in component form, is, for $i \in \mathcal{N}$, $k = 0, 1, \ldots$,

$$\gamma^i_{(k+1)}(y^i) = (I - \frac{1}{\epsilon}R_{ii})\gamma^i_{(k)}(y^i) - \frac{1}{\epsilon}\sum_{j\in\mathcal{N}, j\neq i} R_{ij}E_{\xi|y^i}\gamma^j(y^j) - \frac{1}{\epsilon}E_{\xi|y^i}r_i(\xi). \qquad (2.47)$$

Clearly, if the sequence generated by (2.46) converges to a limit in $\mathbf{\Gamma}$, this also solves $(**)$ and equivalently $(*)$. Furthermore, (2.40) being a linear iteration, we know from Proposition 2.6.3(i) that the sequence $\{\underline{\gamma}_{(k)}\}$ converges if, and only if,

$$\rho(P(I - \frac{1}{\epsilon}R)) < 1.$$

Since both $P$ and $(I - \frac{1}{\epsilon}R)$ map $\mathbf{\Gamma}$ into itself and since $P$ has operator norm equal to *one*, this inequality will be satisfied if

$$\rho(I - \frac{1}{\epsilon}R) < 1.$$

The matrix $R$ being positive definite, this inequality can always be met by choosing $\epsilon > 0$ sufficiently large. If $\lambda_{\max}(R)$ denotes the maximum eigenvalue of $R$, choosing $\epsilon > \frac{1}{2}\lambda_{\max}(R)$ will in fact do the job. The specific choice of $\epsilon$ within this region may be dictated by other considerations, such as the speed of convergence. Since the smaller the spectral radius $\rho(I - \frac{1}{\epsilon}R)$ is, the "faster" the algorithm will (in general) converge, a reasonable choice of $\epsilon$, with this in mind, is

$$\epsilon = \frac{1}{2}[\lambda_{\max}(R) + \lambda_{\min}(R)], \qquad (2.48)$$

where $\lambda_{\min}(R)$ denotes the minimum eigenvalue of $R$. These results are now summarized in the following proposition.

**Proposition 2.6.4.** *For the standard quadratic team with $N$ agents, the parallel update scheme (31b) converges to the unique team-optimal solution whenever $\epsilon > \frac{1}{2}\lambda_{\max}(R)$. A particular value of $\epsilon$ which leads to relatively fast convergence is given by (2.48).* $\diamond$

*Proof.* The result has already been verified prior to the statement of the proposition. Note that this also provides an alternative proof for Theorem 2.6.6 for the special case of standard quadratic teams. $\square$

*Remark 2.6.2.* The algorithm (2.46) should be viewed (at this point) only as a computational tool, and not carry any significant interpretation in terms of the original team decision problem. There are also other variations of this algorithm which lead to convergence, but further discussion is beyond the scope and goal of this book. $\diamond$

### Quadratic-Gaussian Teams

One class of quadratic teams for which the team-optimal solution can be obtained in closed form are those where the random state of nature $\xi$ is a Gaussian random vector. Let us decompose $\xi$ into $N+1$ block vectors

$$\xi = (x', y^{1'}, y^{2'}, \ldots, y^{N'})' \tag{2.49}$$

of dimensions $r_0$, $r_1$, $r_2, \ldots, r_N$, respectively. Being a Gaussian random vector, $\xi$ is completely described in terms of its mean value and covariance matrix, which we specify below:

$$E[\xi] =: \bar{\xi} = (\bar{x}', \bar{y}^{1'}, \ldots, \bar{y}^{N'}), \tag{2.50}$$

$$\text{cov}\ (\xi) =: \Sigma,\ \text{with}\ [\Sigma]_{ij} =: \Sigma_{ij}, \quad i,j = 0, 1, \ldots, N, \tag{2.51}$$

$[\Sigma]_{ij}$ denotes the *ij*th block of the matrix $\Sigma$ of dimension $r_i \times r_j$, which stands for the cross-variance between the *i*th and *j*th block components of $\xi$. We further assume (in addition to the natural condition $\Sigma \geq 0$) that $\Sigma_{ii} > 0$ for $i \in \mathcal{N}$, which means that the measurement vectors $y^i$'s have nonsingular distributions. To complete the description of the quadratic-Gaussian team, we finally take the linear terms $r_i(\xi)$ in the loss function (2.28) to be linear in $x$, which makes $x$ the "payoff relevant" part of the state of nature (recall the earlier discussion in Sect. 2.4 on the use of this terminology):

$$r_i(\xi) = D_i x, \quad i \in \mathcal{N}, \tag{2.52}$$

where $D_i$ is an $(r_i \times r_0)$ dimensional constant matrix. Note that in Definition 2.6.2 we simply have $Q_i = (D_i, 0, 0)$.

In the characterization of the team-optimal solution for the quadratic-Gaussian team we will need the following important result on the conditional distributions of Gaussian random vectors, which we will have occasion to use also in other chapters in the book. A proof of this result can be found in any standard text on probability theory.

**Lemma 2.6.3.** *Let $z$ and $y$ be jointly Gaussian distributed random vectors with mean values $\bar{z}$, $\bar{y}$, and covariance*

$$cov\,(z,y) = \begin{pmatrix} \Sigma_{zz} & \Sigma_{zy} \\ \Sigma'_{zy} & \Sigma_{yy} \end{pmatrix} \geq 0, \quad \Sigma_{yy} > 0. \tag{2.53}$$

*Then, the conditional distribution of $z$ given $y$ is Gaussian, with mean*

$$E[z|y] = \bar{z} + \Sigma_{zy}\Sigma_{yy}^{-1}(y-\bar{y}) \tag{2.54}$$

*and covariance*

$$cov(z|y) = \Sigma_{zz} - \Sigma_{zy}\Sigma_{yy}^{-1}\Sigma'_{zy} \tag{2.55}$$

$\diamond$

The complete solution to the quadratic-Gaussian team is now given in the following theorem:

**Theorem 2.6.8.** *The quadratic-Gaussian team decision problem as formulated above admits a unique team-optimal solution that is affine in the measurement of each agent:*

$$\gamma^{i*}(y^i) = \Pi^i(y^i - \bar{y}^i) + M^i\bar{x}, \quad i \in \mathcal{N}. \tag{2.56}$$

*Here, $\Pi^i$ is an $(m_i \times r_i)$ matrix $(i \in \mathcal{N})$, uniquely solving the set of linear matrix equations:*

$$R_{ii}\Pi^i\Sigma_{ii} + \sum_{j\in\mathcal{N}, j\neq i} R_{ij}\Pi^j\Sigma_{ji} + D_i\Sigma_{0i} = 0, \tag{2.57}$$

*and $M^i$ is an $(m_i \times r_0)$ matrix for each $i \in \mathcal{N}$, obtained as the unique solution of*

$$\sum_{j\in\mathcal{N}} R_{ij}M^j + D_i = 0, \quad i \in \mathcal{N}. \tag{2.58}$$

$\diamond$

*Proof.* Referring back to iteration (2.47), and initializing it with $\gamma^i_{(0)} \equiv 0$, $i \in \mathcal{N}$, it follows from repeated application of Lemma 2.6.3 that $\gamma^i_{(k)}(y^i)$ that is generated by (2.47) is necessarily affine in $y^i$, for all $k = 1, 2, \ldots$, with the structure given by

$$\gamma^i_{(k)}(y^i) = \Pi^i_{(k)}(y^i - \bar{y}^i) + M^i_{(k)}\bar{x}.$$

By Proposition 2.6.4, this sequence converges, and the limiting solution is necessarily in the form (2.56). Further, by Theorem 2.6.6, this limiting policy should uniquely solve the stationarity equations (2.40). Therefore, all that remains to be done is to substitute (2.56) into (2.40), to arrive at (in view of Lemma 2.6.3):

$$[R_{ii}\Pi^i + \sum_{j\in\mathcal{N}, j\neq i} R_{ij}\Pi^j\Sigma_{ji}\Sigma_{ii}^{1} + D_i\Sigma_{oi}\Sigma_{ii}^{-1}](y^i - \bar{y}^i) + [\Sigma_{j\in\mathcal{N}} R_{ij}M^j + D_i]\bar{x} \equiv 0,$$

which is an identity for each $i \in \mathcal{N}$. Since $y^i - \bar{y}^i$ and $\bar{x}$ are independent, (2.57) and (2.58) readily follow. Clearly, in view of our reasoning above, the solutions to (2.57) and (2.58) have to be unique. This algebraic result can in fact also be proven directly. For (2.58), it trivially follows because it can be rewritten as

$$RM + D = 0,$$

where $M := (M^{1'}, M^{2'}, \ldots, M^{N'})'$; $D := (D^{1'}, D^{2'}, \ldots, D^{N'})'$ and hence the unique solution is

$$M = -R^{-1}D.$$

□

A quadratic-Gaussian team is known as a LQG team, if furthermore the measurements have the special structure

$$y^i = H^i x + w^i, \quad i \in \mathcal{N}, \tag{2.59}$$

where $w^i$, $i \in \mathcal{N}$, constitutes an independent sequence of zero-mean Gaussian random vectors, also independent of $x$. Let us denote the covariance of $w^i$, known as the measurement noise for agent $\mathbf{A}i$, by $N^i > 0$, $i \in \mathcal{N}$. Note that in this setup the state of nature is given as

$$\xi = (x', w^{1'}, \ldots, w^{N'})',$$

which is again an $r := \sum_{i=0}^{N} r_i$-dimensional Gaussian random vector.

Now, in view of (2.59), and the independence of the noise sequence, we have

$$\bar{y}^i = H^i\bar{x}, \Sigma_{0i} = \Sigma_{00}H^{i'}, \Sigma_{ij} = H^i\Sigma_{00}H^{j'}, \Sigma_{ii} = H^i\Sigma_{00}H^{i'} + N^i, i \in \mathcal{N}.$$

Clearly, by the positive definiteness of $N^i$'s, $\Sigma_{ii}$'s are positive definite, which means that all the hypotheses of Theorem 2.6.8 are satisfied. The following corollary then follows as a special case.

**Corollary 2.6.2.** *The LQG team decision problem as formulated above admits a unique team-optimal solution given by*

$$\gamma^{i*}(y^i) = \Pi^i(y^i - H^i\bar{x}) + M^i\bar{x}, \quad i \in \mathcal{N}, \tag{2.60}$$

*where $M^i$, $i \in \mathcal{N}$, is the unique solution of (2.58) and $\Pi^i$ solves uniquely the following version of (2.57):*

$$R_{ii}\Pi^i + \left( \sum_{j \in \mathcal{N}, j \neq i} R_{ij}\Pi^j H^j \Sigma_{00} H^{i'} + D_i \Sigma_{00} H^{i'} \right) (H^i \Sigma_{00} H^{i'} + N^i)^{-1} = 0\,. \tag{2.61}$$

◇

*Example 2.6.3.* To illustrate the preceding results, consider the two-agent scalar LQG team with loss function

$$L(x, \mathbf{u}) = (u^1 + u^2 + x)^2 + (u^1)^2 + (u^2)^2$$

and measurements

$$y^1 = x + w^1, \quad y^2 = x + w^2$$

under the independent statistics

$$x \sim N(1,2), w^1 \sim N(0,2), w^2 \sim N(0,1).$$

Direct application of Corollary 2.6.2 leads to the unique team-optimal solution

$$\left.\begin{aligned} \gamma^{1*}(y^1) &= -\tfrac{2}{11}(y^1 - 1) - \tfrac{1}{3}, \\ \gamma^{2*}(y^2) &= -\tfrac{3}{11}(y^2 - 1) - \tfrac{1}{3}. \end{aligned}\right\} \tag{$*$}$$

The corresponding minimum team cost can be computed to be

$$J^* := J(\gamma^{1*}, \gamma^{2*}) \simeq 1.424.$$

Note that this is a symmetric team as far as the loss function goes (i.e., $L(u^1, u^2; x) = L(u^2, u^1, x)$), but as far as the measurements go agent $\mathbf{A}1$ has "higher" measurement noise than agent $\mathbf{A}2$. This is reflected in the team-optimal policies, with the measurement of $\mathbf{A}1$ weighted less than the measurement of $\mathbf{A}2$ (compare the gain $\frac{2}{11}$ against the gain $\frac{3}{11}$).

If the agents did not have access to any measurements, and thus optimize in the class of constant policies, the unique solution can easily be read off from $(*)$ to be

$$\gamma^1_{OL} = \gamma^2_{OL} = -\frac{1}{3}$$

with the corresponding cost being

$$J_{OL} := J(\gamma^1_{OL}, \gamma^2_{OL}) = \frac{7}{3} \cong 2.3333.$$

Hence, the decentralized measurements lead to about $39\%$ improvement (reduction) in the team cost, as compared with the no-measurement (*open-loop*) case.

If, on the other hand, the agents shared their measurements, with the team's common measurement now being $\mathbf{y} = (y^1, y^2)'$, the optimum cost should be lower than $J^*$. To study this specific model, we first note that Theorem 2.6.8 is directly applicable here, with

$$\Sigma_{11} = \Sigma_{12} = \Sigma_{22} = \text{cov}(\mathbf{y}) = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}; \Sigma_{01} = \Sigma_{02} = (2, 2).$$

The unique team-optimal solution can readily be obtained to be

$$\gamma^1_{sh}(\mathbf{y}) = \gamma^2_{sh}(\mathbf{y}) = -\frac{1}{4}[\frac{1}{3}(y^1 - 1) + \frac{2}{3}(y^2 - 1)] - \frac{1}{3}$$

with the corresponding cost being

$$J_{sh} := J(\gamma^1_{sh}, \gamma^2_{sh}) = \frac{4}{3} \cong 1.3333.$$

The improvement here, over the open-loop cost, is 43 %, and over the decentralized case is about $6\%$.

Finally, if both agents had perfect access to the true value of $x$ (the case of perfect measurements), the unique optimal solution would be

$$\gamma^1_{pr}(x) = \gamma^2_{pr}(x) = -\frac{1}{3}x$$

with a cost level of

$$J_{pr} := J(\gamma^1_{pr}, \gamma^2_{pr}) = \frac{1}{3}E[x^2] = 1,$$

which is the lowest possible value for $J$, under any measurement scheme. ◇

### Positively Exponentiated Quadratic Loss

Consider again the formulation of the quadratic-Gaussian team (á la Definition 2.6.2) but with the loss function being a positively exponentiated quadratic function, i.e.,

$$L(\xi; \mathbf{u}) = \theta e^{\frac{\theta}{2}C(\xi;\mathbf{u})}, \quad \theta > 0, \tag{2.62}$$

where $C$ is a strictly convex (in $\mathbf{u}$) function:

$$C(\xi; \mathbf{u}) = \sum_{i,j \in \mathcal{N}} u^{i'} R_{ij} u^j + 2 \sum_{i \in \mathcal{N}} u^{i'} D_i x + x' Q x. \tag{2.63}$$

The state of nature, $\xi$, is a Gaussian vector, as specified earlier by (2.49)–(2.51).

A static team problem with the structure above is known as an *exponential-Gaussian team* or a *linear-exponential-Gaussian team* (LEGT), the latter used especially if the measurements are given in the form (2.59). An exponential (of a quadratic) loss function captures phenomena not obtainable from a quadratic loss function and is preferred in situations where higher (than second) order moments of the statistical quantities should also be taken into consideration. A team using an exponential quadratic loss function in the construction of policies is called *risk averse* if $\theta > 0$ and *risk preferring* if $\theta < 0$. Here we will discuss only the case $\theta > 0$ because it is only in this case that $L$ is convex in $\mathbf{u}$, which will enable us to apply some of the results of Sect. 2.6.2. The "optimistic" case $\theta < 0$ does not lead to a convex loss function, and hence it is not possible to obtain a general theory to cover this case as well. However, this should not be construed as the $\theta < 0$ case not being well defined or interesting. In fact, the stationarity conditions could hold in this case also, but one has to study each problem individually before concluding global (team) optimality.

Returning to the LEGT problem with $\theta > 0$, the first step toward studying its solution would be to obtain a characterization of the stationarity conditions (2.26). To avoid some unnecessary complexity in the analysis to follow, let us take the mean value of $\xi$ to be zero, and furthermore let us restrict ourselves at the outset to linear decision rules (policies) for the agents—the latter will actually create no loss of generality as we shall see later.

Accordingly, let the decision rules be given as

$$\gamma^j(y^j) = A^j y^j, \quad j \in \mathcal{N}. \tag{2.64}$$

Let us fix all but one (say $i$th one) as above, and substitute them into (2.63) to obtain

$$C(\xi; u, \{\gamma^j\}_{j\neq i}) = u' R_{ii} u + 2u' T_i' \xi_i + \xi_i' S_i \xi_i =: C_i(\xi_i, u),$$

where $u$ stands for $u^i$ and $T_i$, $S_i$ are defined as follows:

$$\xi_i := (x', y^{1'} \dots y^{i-1'} y^{i+1'} \dots y^{N'})',$$

$$T_i'\xi_i = D_i x + \sum_{j\in\mathcal{N}, j\neq i} R_{ij} A^j y^j, \tag{2.65}$$

$$\xi_i' S_i \xi_i = x'Qx + \sum_{j,k\in\mathcal{N}; j,k\neq i} y^{j'} A^{j'} R_{jk} A^k y^k + 2 \sum_{j\in\mathcal{N}, j\neq i} y^{j'} A^{j'} D_j x. \tag{2.66}$$

The important point here is that $T_i$ and $S_i$ are constant matrices (not dependent on $\xi$), but they depend on the policy gain matrices $A^j$, for all $j \in \mathcal{N}$, except $j = i$.

We now evaluate

$$E_{\xi|y^i} L(\xi; u, \{\gamma^j\}_{j\neq i}) =: J_i(u; y),$$

where, for convenience, we have dropped the superscript from $y^i$. Using Lemma 2.6.3, the distribution of $\xi$ conditioned on $y^i = y$ is Gaussian, with mean and covariance given by

$$E[\xi|y] = (\Sigma'_{0i}\Sigma'_{1i}\dots\Sigma'_{Ni})'\Sigma_{ii}^{-1}y =: \hat{\xi}_i, \tag{2.67}$$

$$\mathrm{cov}(\xi_i|y) = \Sigma^{(i)} - (\Sigma'_{0i}\dots\Sigma'_{Ni})'\Sigma_{ii}^{-1}(\Sigma'_{0i}\dots\Sigma'_{Ni}) =: \hat{\Sigma}_i, \tag{2.68}$$

which we assume to be positive definite. In (2.67) the matrix $\Sigma_{ii}$ does not appear in $(\dots)$, and likewise in (2.68). Furthermore $\Sigma^{(i)}$ in (2.68) is the covariance of $\xi_i$, which is the $\Sigma$ of (2.51) with the $(i+1)$th row and column block deleted.

Now, aside from a *pdf* normalization constant,

$$J_i(u;y) = \int \theta e^{\frac{\theta}{2}C_i(\xi_i;u)}e^{-\frac{1}{2}(\xi_i-\hat{\xi}_i)'\hat{\Sigma}_i^{-1}(\xi_i-\hat{\xi}_i)}d\xi_i,$$

where the integration is over the vector $\xi_i$ belonging to an appropriate dimensional Euclidean space. This integral will have a finite value if, and only if, the quadratic term in $\xi_i$ is negative definite, which brings in the condition

$$M_i := \hat{\Sigma}_i > 0. \tag{2.69}$$

Under this condition, the integral can be evaluated (using a property of Gaussian *pdf*'s) to yield (again aside from a positive multiplying constant)

$$J_i(u;y) = \theta e^{\tilde{C}_i(\hat{\xi}_i,u)},$$

where

$$\tilde{C}_i(\hat{\xi}_i,u) := \frac{\theta}{2}u'R_{ii}u + \frac{1}{2}(\theta T_i u + \hat{\Sigma}_i^{-1}\hat{\xi}_i)'M_i^{-1}(\theta T_i u + \hat{\Sigma}_i^{-1}\hat{\xi}_i) - \frac{1}{2}\hat{\xi}_i^{-1}\hat{\Sigma}_i^{-1}\hat{\xi}_i.$$

Note that for $\theta > 0$, $\tilde{C}_i$ is strictly convex in $u$, which implies that $J_i(u;y)$ is also strictly convex in $u$. Minimization of $J_i$ is equivalent to minimization of $\tilde{C}_i$ which, being quadratic, immediately leads to

$$u = \gamma^i(y^i) = -(R_{ii} + \theta T_i' M_i^{-1} T_i)^{-1} T_i M_i^{-1} \hat{\Sigma}_i^{-1}\hat{\xi}_i =: A^i y^i. \tag{2.70}$$

Clearly this solution also satisfies the stationarity condition (2.26), with all other (than $i$th) agents' policies fixed as in (2.64).

Note that $A^i$ determined as in (2.70) depends on the fixed gain matrices $A^j$'s, for $j \neq i$, this dependence being through $M_i$ and $T_i$. Let us denote this relationship by

$$A^i = f^i(A^i,\dots,A^{i-1},A^{i+1},\dots,A^N) \tag{2.71}$$

where $f^i$ is a nonlinear but continuous function, determined uniquely by (2.70). Since agent $\mathbf{A}i$ was selected arbitrarily in the preceding analysis, a similar function will exist for each agent, so that (2.71) will hold for all $i \in \mathcal{N}$. This readily brings us to the following proposition.

**Proposition 2.6.5.** *The positively exponentiated Gaussian team problem admits a linear stationary solution if, and only if, there exists a set of matrices $A^i$, $i \in \mathcal{N}$, mutually satisfying (2.71), and under which (2.69) holds for all $i \in \mathcal{N}$, and $J(\gamma)$ remains finite.* ◇

*Proof.* In view of Definition 2.6.1, the result follows from the analysis that led to the proposition. □

*Remark 2.6.3.* If the random state of nature, $\xi$, has nonzero mean, as in (2.50), then the policies (2.64) will have to be replaced by the affine structure

$$\gamma^j(y^j) = A^j(y^j - \bar{y}^j) + b^j, \quad i \in \mathcal{N}.$$

Within this structure, one can again proceed through the preceding analysis and arrive at a counterpart of Proposition 2.6.5. ◇

*Remark 2.6.4.* The boundedness of the cost corresponding to the linear stationary solution can be checked by evaluating the quantity

$$E_{y^i} J_i(A^i y^i; y^i), \tag{$*$}$$

where $J_i$ was defined in the analysis leading to the proposition. We first substitute (2.70) in $C_i(\hat{\xi}_i, u)$ to obtain

$$\tilde{C}_i(\hat{\xi}_i, \gamma^i(y^i)) = -\frac{1}{2}\hat{\xi}_i' N_i \hat{\xi}_i,$$

where

$$N_i := \hat{\Sigma}_i^{-1} + \hat{\Sigma}_i^{-1} M_i^{-1} T_i (\frac{1}{\theta} R_{ii} + T_i' M_i^{-1} T_i)^{-1} T_i' M_i^{-1} \hat{\Sigma}_i^{-1} - \hat{\Sigma}_i^{-1} M_i^{-1} \hat{\Sigma}_i^{-1}.$$

Then, we observe that $(*)$ is finite if, and only if, the integral

$$\int \theta e^{-\frac{1}{2}\hat{\xi}_i' N_i \hat{\xi}_i} e^{-\frac{1}{2} y' \Sigma_{ii}^{-1} y} dy$$

is finite, where $\hat{\xi}_i$ is related to $y$ through (2.67). This condition is equivalent to the exponent being negative definite, that is,

$$\hat{\xi}_i' N_i \hat{\xi}_i + y' \Sigma_{ii}^{-1} y > 0 \qquad \forall y \in \mathbb{R}^{m_i}, y \neq 0. \tag{$**$}$$

◇

Proposition 2.6.5 above leaves a number of questions unanswered. First, we would like to know whether a linear stationary solution, whenever it exists, is team-optimal and secondly whether there would be other team-optimal solutions if the linear stationary solution (*lss*) ceases to exist. Clearly, we would not generally expect the *lss* to exist for all (especially arbitrarily large) values of $\theta$, because of condition (2.69). The theorem below now provides an answer to the first question raised above; the second question is a most difficult one for which no general answer is known as yet.

**Theorem 2.6.9.** *Let $\underline{\gamma}^* \in \mathbf{\Gamma}$ be the linear stationary solution of Proposition 2.6.5, and let there exist some other linear policy $\underline{\beta} \in \mathbf{\Gamma}$ such that $J(\underline{\beta}) < \infty$. Then, $\underline{\gamma}^*$ is the unique team-optimal solution of the (positively) exponentiated-Gaussian team problem.* $\diamond$

*Proof.* Here we resort to Theorem 2.6.5, which delineates the conditions under which stationarity implies team-optimality. Clearly $L(\xi; \mathbf{u})$ is strictly convex and continuously differentiable (in $\mathbf{u}$), and $J(\underline{\gamma})$ is bounded from below (by zero) for all $\underline{\gamma} \in \mathbf{\Gamma}$. The stationary solution $\underline{\gamma}^*$ has finite cost by hypothesis, and the subset (say, $\hat{\mathbf{\Gamma}}$) of $\mathbf{\Gamma}$ on which $J$ is finite is not a singleton, again by hypothesis. Hence, to apply Theorem 2.6.5, one has to show that condition (*c.5*) holds for this problem. This is indeed the case and follows from the fact that the subset $\hat{\mathbf{\Gamma}}$ referred to above is not a singleton. The proof of this result is quite technical and will not be given here; it can be found in [218]. $\square$

*Remark 2.6.5.* A sufficient condition for the second hypothesis of Theorem 2.6.9 is the following: Choose $\beta \equiv 0$, which is clearly a linear policy. Then,

$$J(\beta) = E\{\theta e^{\frac{\theta}{2} x' Q x}\},$$

which is finite if, and only if,

$$\Sigma_{00}^{-1} - \theta Q > 0.$$

Hence, if $\theta$ is chosen to be smaller than $1/[\lambda_{\max}(\Sigma_{00})\lambda_{\max}(Q)]$, the second hypothesis is satisfied. Of course, this condition (on $\theta$) can be made less stringent by choosing some other (nonzero) $\beta$. $\diamond$

*Remark 2.6.6.* For the negatively exponentiated Gaussian team problem, Proposition 2.6.5 remains equally valid (now in fact condition (2.69) would be satisfied with a bigger margin on $\theta$), but we do not have the counterpart of Theorem 2.6.9 because of lack of convexity. $\diamond$

*Example 2.6.4.* To illustrate the main result of Theorem 2.6.9 and to study the restrictions imposed on the parameters of the problem by the various conditions stated there, let us reconsider the static two-agent team problem of Example 2.6.3, with two differences: Now, the loss functional is a positive exponential of the one given there, i.e.,

$$L(x,u) = e^{\theta C(x,\mathbf{u})},$$
$$C(x,\mathbf{u}) = (u^1 + u^2 + x)^2 + (u^1)^2 + (u^2)^2,$$

and the random variable $x$ has zero mean, i.e.,

$$x \sim N(0,2), \quad w^1 \sim N(0,2), \quad w^2 \sim N(0,1).$$

The measurements are still given by

$$y^1 = x + w^1, \quad y^2 = x + w^2.$$

Writing out the stationarity conditions (2.71), we obtain, after some algebra,

$$A^1 = -\frac{1}{2}[1+A^2-\theta(A^2)^2]/c_1(\theta,A^2), \quad A^2 = -[1+A^1-\theta(A^1)^2]/c_2(\theta,A^1) \quad (*)$$

where

$$c_1(\theta,A^2) := 2 - \theta[1 + 2A^2 + 6(A^2)^2] + \theta^2(A^2)^2,$$
$$c_2(\theta,A^1) := 3 - \theta[1 + 12(A^1)^2 + 2A^1] + 2\theta^2(A^1)^2.$$

The matrices $M_1$ and $M_2$, defined by (2.69), are given by

$$M_1 = \begin{pmatrix} 2-\theta & -1-\theta A^2 \\ -1-\theta A^2 & 1-2\theta(A^2)^2 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 2-\theta & -\frac{1}{2}-\theta A^1 \\ -\frac{1}{2}-\theta A^1 & \frac{1}{2}-2\theta(A^1)^2 \end{pmatrix},$$

so that condition (2.69) reads

$$0 < \theta < 2, \quad c_1(\theta,A^2) > 0, \quad c_2(\theta,A^1) > 0. \quad (**)$$

Trying out two different values of $\theta$, namely, $\theta = 1$ and $\theta = \frac{1}{3}$, we find that for the former there is no solution to $(*)$ that also satisfies $(**)$; for the latter, however, there exists a unique solution $(*)$ that also meets $(**)$, which is

$$A^{1^*} = -0.236375, \quad A^{2*} = -0.345398 \quad (\theta = \frac{1}{3}).$$

This solution and the associated value of $\theta$ also satisfy the conditions of Remarks 2.6.4 and 2.6.5, and hence by Theorem 2.6.9 there exists a unique team-optimal solution to the LEGT problem:

$$\gamma^{1*}(y^1) = -0.236375y^1, \quad \gamma^{2*}(y^2) = -0.345398y^2.$$

It is important to note that, as $\theta \to 0$ in $(*)$, the nonlinear equations reduce to linear equations:

$$A^1 = -\frac{1}{4}(1 + A^2), \quad A^2 = -\frac{1}{3}(1 + A^1),$$

whose unique solution $(-\frac{2}{11}, -\frac{3}{11})$ is precisely the gain coefficients in the team-optimal solution of Example 2.6.3. Hence in the limit as $\theta \to 0$ we recover the solution of the corresponding LQG team (with loss function $C(x, \mathbf{u})$). This is to be expected because for any positive function $C$,

$$\lim_{\theta \to 0} \frac{1}{\theta}(e^{\theta C} - 1) = C.$$

*Remark 2.6.7.* A method for solving the set of coupled nonlinear equations $(*)$ of Example 2.6.4 for a particular value of $\theta$, or more generally equations (2.71), is provided by the parallel update scheme

$$A^i_{(k+1)} = f^i(A^1_{(k)}, \ldots, A^{i-1}_{(k)}, A^{i+1}_{(k)}, \ldots, A^N_{(k)}), \quad i \in \mathcal{N}, \quad k = 0, 1, \ldots,$$

where the starting point is arbitrary. This set of recursive equations admits exactly the same interpretation as in the LQ team case, and as was the case there, this recursion may not converge (even if the LEG team problem admits a solution) for $N \geq 3$. For $N = 2$, however, the recursion will converge whenever the LEQ team problem admits a solution; this is because at each step this corresponds to an agent's minimization of $J$ by fixing the other agent's policy at its most recently updated value. Since one is basically minimizing a strictly convex functional (in the LEGT problem), the unique minimum, whenever it exists, should be reachable by such a unilaterally cost-minimizing update scheme. $\diamond$

## 2.7 Concluding Remarks

This chapter has provided a general introduction to stochastic team decision problems and associated solution concepts. Static and dynamic teams have been identified, and in the context of static teams conditions for existence of team-optimal solutions and for person-by-person optimality to imply team-optimality have been obtained. The chapter has also discussed iterative methods for obtaining team-optimal solutions and illustrated the theory presented with numerical examples.

## 2.8 Bibliographic Notes

Team decision theory has its roots in both control theory and economics. Jacob Marschak [254] was perhaps the first to introduce the basic elements of teams and to provide the first steps toward the development of a *team theory*. Roy Radner [316] followed with a precise mathematical formulation and provided conclusive results to some classes of static teams, establishing precise connections between person-by-person optimality, stationarity, and team-optimality. Marschak's and Radner's collaborative work culminated in the publication of their influential 1972 book [255]. At the time when such developments were being made, significant progress in the theory of statistical decision theory was also taking place: Bahadur's characterization of information fields and sufficient statistics [35, 36]; Blackwell's sufficient statistics and comparison of experiments results [61]; and Wald's [383], Savage's [333], and Chernoff's [94] contributions to statistical decision theory, among other major developments in probability theory, contributed to the rapid development of team decision theory.

Contributions of Hans Witsenhausen [393, 394, 399–401] on dynamic teams and characterization of information structures have been crucial in the progress of our understanding of dynamic teams; see Sect. 3.7, where Witsenhausen's *intrinsic model* as well as other models for dynamic teams are discussed in detail. This section also includes a brief discussion for nonsequential teams where important contributions in the literature have been due to Andersland and Teneketzis [9, 10] and Teneketzis [360], in addition to Witsenhausen [393].

Considerations of risk sensitivity motivated researchers to look into team problems with exponentiated loss function, with substantial results in this domain obtained (for teams) by Krainak et al. [219]. De Waal and van Schuppen [114] considered extensions to discrete action spaces. Bagchi and Başar [34] studied teams in continuous time as well as non-Gaussian settings.

Başar [24] studied team problems and more general nonzero-sum stochastic games when agents do not agree on a common *a priori* probability measure on the primitive random variable and work under their own subjective views of the environment, with team models in this context necessarily leading to game formulations. A more detailed discussion in this context of inconsistent probability models among a group of decision makers is presented in Chap. 12.

Further discussion on design of information structures in the context of team theory and economics applications is available in [15, 372], among a rich collection of other contributions.

In the next chapter, Chap. 3, we will see extensions of the static team theory of this chapter to dynamic teams, where information structures are of paramount importance. We will also consider Witsenhausen's intrinsic model more explicitly in Sect. 3.7. We refer the reader to also Teneketzis [360], in addition to [400], in this context.

Part of the chapter uses results from [219,316], however, with somewhat different proofs for some of the key results. The update schemes considered in Sect. 2.6 are based on [24].

http://www.springer.com/978-1-4614-7084-7

Stochastic Networked Control Systems
Stabilization and Optimization under Information Constraints
Yüksel, S.; Başar, T.
2013, XVIII, 482 p., Hardcover
ISBN: 978-1-4614-7084-7
A product of Birkhäuser Basel