arXiv:2304.01445v1 [eess.SY] 4 Apr 2023

战略性多智能体机器人团队的协调效率

Marcos M. Vasconcelos[1] 和 Behrouz Touri[2]

摘要—我们研究在随机环境中，一组战略决策者通过选择是否参与某项任务来实现分散式协调的问题。首先，我们定义了一类对称效用博弈，它涵盖了广泛的协调博弈，其中包括一种广为人知的框架——全球博弈。为了研究智能体参与随机协调博弈时实际实现协调的程度，我们提出了一种新的协调效率的概率度量。随后，我们给出一个信息论上界，说明协调效率与观测信道中的噪声大小之间的关系。最后，我们重新考察了一大类全球博弈，并指出，尽管纳什均衡策略能够带来更高的期望效用，其协调效率却可能低于确定性等价策略。这一反直觉结果表明，协调博弈中的协调效率与期望效用之间存在非平凡的权衡。

I. 引言

在许多分布式自主系统中，例如机器人、社会经济和生物网络，协调行为都十分重要 [1]–[5]。工程领域的大多数协调研究都假设智能体通过通信网络交换消息，以渐近地就共同的决策变量达成一致，例如观点动力学和分布式优化。然而，在经济学领域，协调问题的研究视角有所不同：智能体不交换（显式）消息，而是采取策略性行动。这种方法与协调博弈有关，其中两个或多个交互智能体受到激励而采取相同的行动。确定性协调博弈的特点是存在多个均衡，并且往往会引出对社会困境的分析。解决均衡多重性的一种方法采用了称为全球博弈 [6]的框架。

全球博弈是一种贝叶斯协调博弈，其中每个代理人在观察到有关世界状态的噪声信号后选择行动。世界状态（我们简称为状态 ）反映了银行挤兑模型中的经济实力、政治制度等特征
在机制转换模型中，或在任务分配问题中任务的难度。对于效用结构的某些假设，即使观测中存在无限趋近于零的噪声，全局博弈也能导出唯一的贝叶斯纳什均衡，从而解决具有多个均衡的博弈中的均衡选择问题 [7]。

[1]M. M. Vasconcelos 任职于美国佛罗里达州立大学 FAMU-FSU 工程学院电气与计算机工程系。电子邮箱：m.vasconcelos@fsu.edu[2]Behrouz Touri 任职于美国加州大学圣迭戈分校电气工程系。电子邮箱：btouri@eng.ucsd.edu

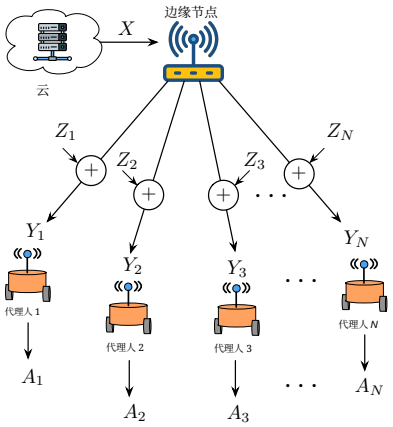


图1. 机器人团队中策略协调的系统架构：团队接收关于远程位置处可用的随机状态变量的信息。

这一类博弈的近期文献主要关注以下方面：在不同信息模式下纳什均衡的存在性 [8]，、智能体观测之间的相关性所产生的影响 [9], [10], ，以及就智能体效用函数中的外部性而言，不同局部连接模式的影响 [11]。其他近期进展则研究了针对含噪信号的非传统概率模型 [12],，以及多任务分配问题中的多维状态 [13]。

我们考虑图1所示的联邦系统架构，其中状态可在远程位置（例如云服务器）获取，并由边缘节点或网关通过并行噪声信道广播给多个智能体。代理人收到噪声信号后，会做出二元决策，例如最大化满足策略互补性的期望效用函数，从而促成协调行为 [14]。我们研究全局博弈中的协调如何随着通信信道噪声水平的增加而退化。此外，我们还希望刻画与机器人智能体通信时，在给定信噪比下协调能力的极限。

本文的主要贡献如下：

- 我们提出了一类博弈，即齐次协调博弈，其中包含了范围广泛的全局博弈。
- 我们提出了一种新的协调效率概念，用于衡量齐次协调博弈中的协调情况。
- 我们得出了基于信息论工具的任意策略在全球博弈中协调效率的基本面限制。

II. 系统模型

本节将讨论我们的全球博弈模型，特别是介绍这类博弈的一个重要子类，即同质协调博弈。

A. 效用结构

全局博弈是一种由 N参与者 $[N]\stackrel{\text{def}}{=}\{1,\dots,N\}$与自然进行的不完全信息博弈。形式上，全局博弈是一个元组（$[N],\mathcal{A},X,\mathbf{u},\mathbf{Y}$），其中：

（i） $\mathcal{A}=\mathcal{A}_1\times\mathcal{A}_2\times\cdots\times\mathcal{A}_N$是参与者的联合行动集，其中 $\mathcal{A}_i$ 是参与者$i\in[N]$的行动集；（ii） X 是一个决定自然类型的随机变量。我们将 X 称为博弈的状态或基础状态；（iii）**效用** $=(u_1,\dots,u_N)$：$\mathcal{A}\times\mathbb{R}\to\mathbb{R}^N$是参与者的效用，其中 u_i：$\mathcal{A}\times\mathbb{R}\to\mathbb{R}$是第$i$个参与者的效用，它取决于每位参与者的行动以及基础状态的取值 X；（iv）$\mathbf{Y}$ $=(Y_1,\dots,Y_N)$，其中 Y_i 是一个随机变量，表示参与者 $i\in[N]$ 对基础状态的含噪观测（由此形成其信念） X。

对于任意向量 $\mathbf{v}=(v_1,\dots,v_N)$ 以及任意 $i\in[N]$，我们使用记号 $\mathbf{v}_{-i}=(v_1,\dots,v_{i-1},v_{i+1},\dots,v_N)$；并且为方便起见，我们称 $\mathbf{v}$ $=(v_i,\mathbf{v}_{-i})$。例如，对于联合行动 $\mathbf{a}\in\mathcal{A}$，我们对所有 $i\in[N]$ 写作 $\mathbf{a}=(a_i,\mathbf{a}_{-i})$。

我们的研究受到这样一种观察的启发：关于全球博弈及其应用的研究中，绝大多数底层博弈都具有以下共同特征：

(a) 对称／置换不变：在许多情境中，个体智能体的效用函数在其他智能体行动的任意置换下保持不变。换言之，对于任意置换矩阵[1] P，$u_i(a_i,\mathbf{a}_{-i})=u_i(a_i,\mathbf{a}_{-i}P)$。

(b) 同质效用函数： N 参与者的效用函数相同，其含义是：对于任意参与者 $i\in[N]$ 和任意行动组合 $\mathbf{a}\in\mathcal{A}$，都有$u_i(a_i,\mathbf{a}_{-i})=u_1(a_i,\mathbf{a}_{-i})$。

(c) 同质行动集：在许多全局博弈中，我们讨论的是一个庞大的群体，所有参与者的行动集都相同。例如，在政治骚乱中，面对一种政治制度，所有参与者都决定采取风险行动或安全行动。在这种情况下，$\mathcal{A}_i=\{0,1\}$ 其中 0 和 1 分别对应安全行动和风险行动。

(d) 促进协调：同样，在大多数我们感兴趣的情境中，参与者的效用结构会促进协调。例如，在政治起义、银行挤兑等情况下，一个经过充分研究的效用函数是 $u_i(\mathbf{a}\ X)=a_i(\sum_{j=1}^N a_i-X)$。因此，在所有参与者都掌握关于 X 的完美信息时，即当 $Y_i=X$ 对所有 $i\in[N]$ 都成立时，均衡唯一；它取决于 $X>\frac{1}{n}$ 还是 $X<\frac{1}{n}$，而这个博弈的博弈要么是 $\mathbf{a}=\mathbf{1}$，要么是 $\mathbf{a}=\mathbf{0}$，从而使 N 参与者实现协调。

受此启发，我们引入同质协调博弈的概念，以形式化满足上述性质的一大类博弈。为此，令 $\Delta^k:=\{\mathbf{q}\in\mathbb{R}_+^k\mid\sum_{i=1}^k q_i=1\}$ 为 $\mathbb{R}^k$ 中的概率单纯形。对于有限集 $\mathcal{A}=\{0,\dots,M-1\}$ 和向量$\mathbf{v}\in\mathcal{A}^d$，其中 $d\geq 1$，定义经验质量函数 $G(\mathbf{v})=\big(q_0,\dots,q_{M-1}\big)\in\Delta^M$ 如下：

$$q_\ell=\frac{1}{\gamma}\sum_j\mathbf{1}(v_j=\ell),\quad \ell\in\mathcal{A}.$$

基本上， q_ℓ 是$\mathbf{v}$ 中等于 $\ell\in\mathcal{A}$的条目所占的比例。现在，我们可以正式定义齐次协调博弈这一类博弈。

定义 *1*（齐次协调博弈）：齐次协调博弈是指所有智能体具有相同的行动集 $\mathcal{A}$和相同的效用函数u：$\mathcal{A}^N\times\mathbb{R}\to\mathbb{R}$，并满足以下条件的博弈：

(1) 存在一个函数$\hat{u}$：$\mathcal{A}\times\Delta^{N-1}\times\mathbb{R}\to\mathbb{R}$，使得对于所有$\mathbf{a}\in\mathcal{A}^N$、所有 $i\in[N]$和所有 $x\in\mathbb{R}$，均有

$$\hat{u}(a_i,G(\mathbf{a}^{-i}),x)=u(a_i,\mathbf{a}^{-i},x).\tag{1}$$

(2) 对于所有 $i\in[N]$、所有 $x\in\mathbb{R}$和所有$\mathbf{g}\in\Delta^{N-1}$，存在一个最优行动 $a^\star(x)$以及参与者的多数 $c^\star(x)$，使得如果多数参与者正在采取 $a^\star(x)$，那么参与者 i 采取该行动会更有利。用数学形式表示，存在 $c^\star(x)\in[0,1]$，使得对于任意满足$g_{a^\star(x)}=\max_{\ell\in[M]}g_\ell\geq c^\star(x)$的$\mathbf{g}\in\Delta^M$，均有

$$\hat{u}\big(a_i^\star(x),\mathbf{g},x\big)\geq\hat{u}\big(a_i,\mathbf{g},x\big),\qquad\forall a_i\in\mathcal{A}.\tag{2}$$

请注意，性质（1）本质上意味着每个参与者的效用函数都是对称的／置换不变的。换言之，对于有限行动集 $\mathcal{A}=[M]$，任何对称／置换不变函数 f：$\mathcal{A}^{N-1}\to\mathbb{R}$（如 (a) 中所定义），都可以表示为参与者行动的经验质量函数的函数。也就是说，对于此类效用函数，哪个参与者采取哪个行动并不重要，重要的是有多少参与者或多大比例的参与者采取了各个行动。

在本文余下部分中，略去符号上的繁琐之处，我们可以将齐次协调博弈的效用函数视为仅为经验质量的函数，并用记号 $u(a_i,\ \mathbf{g}_{-i},\ x)$ 代替 $\hat{u}(a_i,\ G(\mathbf{a}_{-i})\ x)$，而不再使用 $u(a_i,\ \mathbf{a}_{-i},\ x)$，其中 $\mathbf{g}_{-i}=G(\mathbf{a}_{-i})$。

B. 信息结构与策略

这里，我们讨论关于基本面 X以及个体智能体对 Y_i 的含噪观测所作的假设。我们始终假设 X 是方差为 σ_X^2 的零均值高斯随机变量，即 $X\sim\mathcal{N}(0,\ \sigma_X^2)$。我们假设采用文献中常见的模型（参见 [6], [8], [9]），第 i 个代理人的含噪观测 Y_i 为 $Y_i=X+Z_i$。我们假设噪声序列 $\{Z_i\}_{i\in[N]}$ 在各智能体之间独立同分布，且 $Z_i\sim\mathcal{N}(0,\ \sigma_Z^2)$。此外，$\{Z_i\}$ 与 X 相互独立。

[1]如果一个矩阵的所有元素都为 0 或 1，且每一行和每一列都恰好有一个非零元素，则该矩阵称为置换矩阵。

注意，由于 X 和 Y_i 联合高斯，给定 $Y_i = y_i$ 时对 X 的最小均方误差估计是线性的，且可表示为

$$\hat{x}_{\mathrm{mmse}}(y_i) \overset{\mathrm{def}}{=} \mathbb{E}[X \mid Y_i = y_i] = \Big(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}\Big) y_i. \quad (3)$$

一般而言，在不完全信息博弈（包括全球博弈和齐次协调博弈）中，智能体根据其观测采取行动。这引出了策略这一概念。对于行动集为 $\mathcal{A} = \{1, \ldots, M\}$ 的齐次协调博弈，策略是一个映射 $\gamma_i : \mathbb{R} \to \mathcal{A}$，它将智能体 i 的观测转换为行动，即智能体 $i \in [N]$ 采取行动 $a_i = \gamma_i(Y_i)$。

C. 贝叶斯纳什均衡

设 u_i 为齐次协调博弈（定义 1）的效用函数。在该博弈中，智能体以非合作方式行动，寻求通过协调策略最大化各自的期望效用。设 $\boldsymbol{\gamma} \overset{\mathrm{def}}{=} (\gamma_1, \ldots, \gamma_N)$ 为一个策略组合，即系统中所有智能体所使用策略的集合。

给定 $\boldsymbol{\gamma}_{-i}$，第 i 个智能体的目标是求解以下随机优化问题

$$\underset{\gamma_i}{\text{maximize}} \;\; \mathcal{J}_i(\gamma_i, \boldsymbol{\gamma}_{-i}) \overset{\mathrm{def}}{=} \mathbb{E}\Big[u_i(A_i, G_{-i}, X)\Big],$$

其中，期望是对所有外生随机变量 X 和 $\{Z_i\}_{i\in[N]}$ 取的。这引出了贝叶斯纳什均衡（BNE）策略的概念。

定义 2（贝叶斯纳什均衡）：策略 $\boldsymbol{\gamma}^\star$ 是贝叶斯纳什均衡，当且仅当

$$\mathcal{J}_i(\gamma_i^\star, \boldsymbol{\gamma}_{-i}^\star) \geq \mathcal{J}_i(\gamma_i, \boldsymbol{\gamma}_{-i}^\star), \text{ for all } \gamma_i \in \Gamma, \;\; i \in [N],$$

其中，Γ 是所有可采纳协调策略/策略的空间。

D. 协调度量

由于智能体无法完美观测状态 X，因此通常无法实现完全协调。在确定性情形下，明确定义协调与一致的概念是一个适定问题。然而，在随机博弈情形下，定义协调效率度量有多种方式。这种度量的一个关键特征是能够反映：随着观测噪声增大，智能体围绕最优行动进行协调的程度会降低。引入齐次协调博弈框架后，我们便可以从数学上定义一种协调效率度量。

定义 3（协调效率）：设 $\boldsymbol{\gamma}$ 为齐次协调博弈（定义 1 中所定义）中的 N 个参与者的策略组合。我们将平均协调效率 $\varrho : \boldsymbol{\gamma} \mapsto [0, 1]$ 定义为

$$\varrho(\boldsymbol{\gamma}) \overset{\mathrm{def}}{=} \frac{1}{N} \sum_{i=1}^{N} \mathbb{P}\big(\gamma_i(Y_i) = a^\star(X)\big),$$

其中，$a^\star(x)$ 是定义 1 中定义的最优行动。

III. 重新审视全球博弈

齐次协调博弈的一个重要实例是一类二元行动全球博弈（即 $\mathcal{A} = \{0, 1\}$），其中每个代理人的效用为

$$u_i(a_i, \mathbf{a}_{-i}, x) = a_i \cdot \bigg(b\Big(\sum_{j\neq i} a_j\Big) - x\bigg), \quad (4)$$

其中 $b : [0, N-1] \to \mathbb{R}$ 是一个连续递增函数。函数 $b(\cdot)$ 称为收益函数。这种效用的一种应用是机器人团队中的分布式任务分配 [15], [16]，其中 x 表示任务的难度。如果足够多的其他智能体采取相同行动，智能体参与该任务便会获得收益。然而，如果智能体无法完全观测变量 x，那么智能体是否应该参与该任务就不明确了。

接下来的结果证明，具有效用结构 (4) 的全局博弈实际上是齐次协调博弈。

引理 1：对于任意递增连续函数 $b : [0, N-1] \to \mathbb{R}$，具有 $N \geq 2$ 个参与者、二元行动集 $\mathcal{A} = \{0, 1\}$ 和效用函数 (4) 的全局博弈都是齐次协调博弈。

证明：这类博弈的行动集和效用函数的齐次性可由定义直接得出。为证明性质（1），对于 $\mathbf{a} \in \mathcal{A}^N$ 和 $i \in [N]$，令 $p = \frac{\sum_{j\neq i} a_j}{N-1}$。则 $G(\mathbf{a}_{-i}) = (1-p, p)$。因此，令 $\hat{u}(a_i, \mathbf{g}, x) \overset{\mathrm{def}}{=} a_i(b((N-1)g_1) - x)$ 对所有 $a_i \in \mathcal{A}$ 均成立，$\mathbf{g} = (g_0, g_1) \in \Delta^2$，且 $x \in \mathbb{R}$，我们有

$$u_i(a_i, \mathbf{a}_{-i}, x) = a_i\Big(b\left((N-1)p\right) - x\Big) = \hat{u}(a_i, G(\mathbf{a}_{-i}), x). \quad (5)$$

为证明性质（2），固定 $x \in \mathbb{R}$，并令 $b(\cdot)$ 为一个递增的收益函数。于是，若 $x \leq b(0)$，则有

$$\hat{u}(1, \mathbf{g}, x) = b\Big((N-1)g_1\Big) - x \geq b(0) - x \geq \hat{u}_i(0, \mathbf{g}, x) = 0.$$

因此，对于任意 $x \leq b(0)$，(2) 在 $a^\star(x) = 1$ 和 $c^\star(x) = 0$ 时成立。类似地，可以证明，对于 $x \geq b(N-1)$，(2) 在 $a^\star(x) = 0$ 和 $c^\star(x) = 0$ 时成立。对于 $x \in (b(0)\ b(N-1))$，可以证明 $a^\star(x) = 0$ 和 $a^\star(x) = 1$ 都是可能的协调行动。为证明 $a^\star(x) = 1$，令 $c^\star(x) = \min\{q \in [0,1] \mid b(q(N-1)) \geq x\}$（注意，由于 $b(\cdot)$ 的连续性和紧性，该最小值存在）。那么，对于任意满足 $g_1 \geq c^\star(x) = q$ 的概率向量 $\mathbf{g} = (g_0, g_1) \in \Delta^2$，我们有

$$\begin{aligned}\hat{u}(1, \mathbf{g}, x) &= b\Big((N-1)g_1\Big) - x \\ &\geq b((N-1)q) - x \geq \hat{u}(0, \mathbf{g}, x) = 0.\end{aligned}$$

因此，条件 (2) 成立。类似地，可以证明，对于 $x \in (b(0)\ b(N-1))$，$a^\star(x) = 0$ 是一个协调行动，且 $c^\star(x) = \max\{q \in [0,1] \mid b((1-q)(N-1)) \geq x\}$。 ■

在本研究中，我们研究上述博弈所采用的所谓 最佳响应 策略。在我们的情形中，对于联合策略 $\boldsymbol{\gamma}$，代理人

i 的最佳响应策略是

$$\mathrm{BR}(y_i, \boldsymbol{\gamma}_{-i}) = \begin{cases} 1 & \text{if } \mathbb{E}[b\Big(\sum_{i\neq j}\gamma(Y_j)\Big) \mid Y_i = y_i] \\ & \qquad \geq \mathbb{E}[X \mid Y_i = y_i] \\ 0 & \text{otherwise} \end{cases}.$$

A. 阈值策略及其最佳响应

在许多不完全信息博弈中，包括全球博弈，我们关注阈值策略这一类策略。对于具有二元行动的全球博弈，阈值策略是指代理人将观察到的信号 $Y_i = y_i$ 与阈值 τ_i 进行比较，并决定是否采取风险行动（$a_i = 1$），即不采取该行动（$a_i = 0$）：

$$\gamma_i(y_i) = \begin{cases} 1, & y_i \leq \tau_i \\ 0, & \text{otherwise.} \end{cases}$$

利用下一个结果，我们将证明，对同质阈值策略的最佳响应也是阈值策略。

引理 *2*：如果函数 $b(\cdot)$ 非负且严格递增，并且所有其他智能体 $j \neq i$ 采用具有相同阈值 τ的阈值策略 γ，那么

$$\mathbb{E}\left[b\Big(\sum_{i\neq j}\gamma(Y_j)\Big) \;\middle|\; Y_i = y_i\right]$$

是关于 y_i的严格递减函数。

证明：令 B_{-i} 表示由随机变量 $\{Y_j\}_{j\neq i}$ 定义的函数 $B_{-i} \overset{\text{定义}}{=} b\Big(\sum_{j\neq i}\gamma(Y_j)\Big)$，其累积分布函数（CDF）为

$$F_{B_{-i}|Y_i=y_i}(\xi) \overset{\text{def}}{=} \mathbb{P}\Big(B_{-i} \leq \xi \;\Big|\; Y_i = y_i\Big).$$

由于 $b(\cdot) \geq 0$，B_{-i} 是一个非负随机变量。因此（参见 [17, 第 1.5 章，性质 E.6]），

$$\mathbb{E}[B_{-i} \mid Y_i = y_i] = \int_0^\infty \Big(1 - F_{B_{-i}|Y_i=y_i}(\xi)\Big)d\xi.$$

由于函数 b 严格递增，因此存在唯一的反函数 b^{-1}：$\mathbb{R}_+ \to \mathbb{R}_+$。此外，由于观测值 $\{Y_j\}_{j\in[N]}$ 在给定 X的条件下相互独立，因此，

$$F_{B_{-i}|Y_i=y_i}(\xi) = \int_{\mathbb{R}} \mathbb{P}\Big(\sum_{i\neq j}\gamma(Y_j) \leq b^{-1}(\xi) \;\Big|\; X = x\Big) \times f_{X|Y_i=y_i}(x)dx.$$

设 $A_j = \gamma(Y_j)$，其中 $j \neq i$。给定 $X = x$后，伯努利随机变量集合 $\{A_j\}_{j\neq i}$ 相互独立，且

$$\mathbb{P}\big(A_j = 1 \mid X = x\big) = \Phi\Big(\frac{\tau_j - x}{\sigma_Z}\Big),$$

其中 Φ (2 是标准高斯随机变量的累积分布函数。

[2]标准高斯随机变量的累积分布函数为

$$\Phi(x) \overset{\text{def}}{=} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{\xi^2}{2}\Big)d\xi.$$

在假设阈值策略组合均质的情况下，$\tau_j = \tau$ 对所有 $j \neq i$，$\{A_j\}_{j\neq i}$ 均服从相同分布，这意味着

$$\sum_{j\neq i} A_j \mid X = x \sim \mathcal{B}\left(N-1, \Phi\Big(\frac{\tau - x}{\sigma_Z}\Big)\right),$$

其中 $\mathcal{B}(k,\ p)$ 是参数为 $(k,\ p)$ 的二项分布。因此，

$$F_{B_{-i}|Y_i=y_i}(\xi) = \mathbb{E}_V\left[\sum_{\ell=0}^{\lfloor b^{-1}(\xi)\rfloor} \binom{N-1}{\ell} \Phi\Big(\frac{\tau - V - \alpha y_i}{\sigma_Z}\Big)^{\ell} \times \left(1 - \Phi\Big(\frac{\tau - V - \alpha y_i}{\sigma_Z}\Big)\right)^{N-1-\ell}\right],$$

其中，期望是相对于具有以下性质的随机变量 V 而言的

$$V \sim \mathcal{N}\Big(0, \frac{\sigma_X^2\sigma_Z^2}{\sigma_X^2 + \sigma_Z^2}\Big) \text{ and } \alpha \overset{\text{def}}{=} \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}. \tag{6}$$

令 $p(\tau,\ v,\ y_i) \overset{\text{定义为}}{=} \Phi\Big(\frac{\tau - v - \alpha y_i}{\sigma_Z}\Big)$。注意，$p(\cdot,\ \cdot,\ y_i)$ 关于 y_i 严格递减；又由于在给定点处计算的二项随机变量的累积分布函数是关于概率参数 p的严格递减函数，因此我们有

$$\frac{\partial}{\partial y_i}\Big(1 - F_{B_{-i}|Y_i=y_i}(\xi)\Big) < 0.$$

因此，

$$\frac{\partial}{\partial y_i}\int_0^\infty \Big(1 - F_{B_{-i}|Y_i=y_i}(\xi)\Big)d\xi < 0.$$

■

定理 *1*：若收益函数 $b: \mathbb{R}_+ \to \mathbb{R}_+$ 非负且严格递增，则对齐次阈值策略组合的最佳响应映射是一个阈值策略。

证明：令

$$f(y_i) \overset{\text{def}}{=} \mathbb{E}[b(\sum_{j\neq i}\gamma(Y_j)) \mid Y_i = y_i] - \mathbb{E}[X \mid Y_i = y_i].$$

引理 2 表明，$\mathbb{E}[b(\sum_{j\neq i}\gamma(Y_j)) \mid Y_i = y_i]$ 关于 y_i 单调递减，而

$$\mathbb{E}[X \mid Y_i = y_i] = \Big(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}\Big)y_i$$

是关于 y_i 的严格递增函数。因此，$f(y_i)$ 严格递减。此外，由于 b 是递增函数，$\mathbb{E}[b(\sum_{j\neq i}\gamma(Y_j)) \mid Y_i = y_i] \in [b(0)\ b(N-1)]$，因此，$\lim_{y_i\to-\infty} f(y_i) = \infty$ 且 $\lim_{y_i\to\infty} f(y_i) = -\infty$。因此，存在唯一交点 $\bar{\tau}$，使得当$f(y_i) > 0$时，并且当 $y_i < \bar{\tau}$时，$f(y_i) < 0$。 ■

B. 线性收益函数

定理 1 保证，对于广泛的一类全局博弈，对同质阈值的最佳响应是一种阈值策略。一旦找到新的阈值，其他智能体就通过采用相同的最佳响应阈值来模仿。我们递归地使用这一方案，该方案可能收敛到贝叶斯纳什均衡

策略。然而，对于任意收益函数 $b(\cdot)$，一般而言，这种方案是否收敛可能并不容易确定。此外，对于一般的收益函数，找到最优策略 $a^\star(x)$ 可能并不可行。不过，对于线性 收益函数这一类，这种刻画是可行的。

考虑以下由N, b_N^{lin} 索引的线性收益函数： $\mathbb{R} \to \mathbb{R}$ ，满足

$$b_N^{\text{lin}}(\xi) \stackrel{\text{def}}{=} \lambda \cdot \left(\frac{\xi}{N}\right), \tag{7}$$

其中 $\lambda > 0$。定义信念函数 π_{ij}： $\mathbb{R}^2 \to \mathbb{R}$ ，如下

$$\pi_{ij}(\tau_j, y_i) \stackrel{\text{def}}{=} \mathbb{P}\big(Y_j \le \tau \mid Y_i = y_i\big).$$

经过一些代数运算后，我们可以写成

$$\pi_{ij}(\tau_j, y_i) = \mathbb{E}\left[\Phi\left(\frac{\tau_j - V - \alpha y_i}{\sigma_Z}\right)\right], \tag{8}$$

其中 V 和 α 在式 (6) 中定义。

推论 *1*（定理 *1* 的推论）：假设收益函数为线性函数（由 (7) 给出），若每个代理人 $j \neq i$ 都使用阈值为τ_j 的阈值策略，则对于任意阈值策略组合，其最佳响应由 $\frac{\lambda}{N}\sum_{i\neq j}\pi_{ij}(\bar{y}_i) = \alpha\bar{y}_i$的唯一解$\bar{y}_i$ 给出。

根据推论 1，我们可以在阈值策略空间中定义一个 BR 映射，将一个 N 阈值向量映射为 N 阈值。令 $\mathcal{F}$： $\mathbb{R}^N \to \mathbb{R}^N$，其中

$$\mathcal{F}_i(\tau_i, \tau_{-i}) \stackrel{\text{def}}{=} \arg\min_{\xi\in\mathbb{R}} \left(\frac{\lambda}{N}\sum_{i\neq j}\pi_{ij}(\tau_j, \xi) - \alpha\xi\right)^2,$$

对所有 $i \in [N]$.

备注 *1*： 阈值策略中贝叶斯纳什均衡的存在性很容易证明，但其是否唯一取决于能否证明 $\mathcal{F}$ 满足压缩性质，这将留待未来研究。

C. 使用阈值 τ 的同质主体

当我们只关注齐性策略组合时，问题会更简单。在这种情况下，对于每个代理人都采用阈值 τ 的阈值策略组合，其最佳响应是以下方程的唯一解

$$\text{BR}(\tau) = \left\{\xi^\star : \lambda\frac{(N-1)}{N}\pi(\xi^\star; \tau) = \alpha\xi^\star\right\}, \tag{9}$$

here

$$\pi(\xi; \tau) \stackrel{\text{def}}{=} \mathbb{E}\left[\Phi\left(\frac{\tau - \tilde{\sigma}W - \alpha\xi}{\sigma_Z}\right)\right], \tag{10}$$

其中 $W \sim \mathcal{N}(0,\ 1)$ 且 $\tilde{\sigma}^2 = \alpha\sigma_Z^2$。

例 *1*：图 2 展示了式 (9) 中的最佳响应函数，以及噪声方差 σ_Z^2 不同时对应的纳什均衡（NE）阈值。这次数值实验得出两个结论：随着噪声方差增大，纳什均衡阈值也随之增大 $\tau^\star(\sigma_Z^2)$。较不明显的是，当噪声方差趋于零，即观测完美时，纳什均衡阈值的极限。在无噪声情形下，图 2

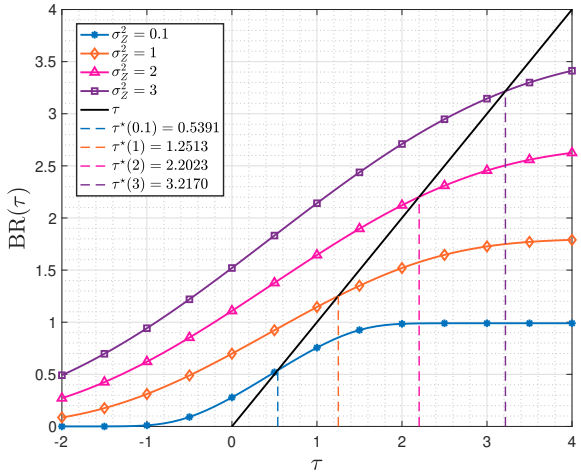


图 2. 对阈值为 τ 的齐性策略组合的最佳响应函数。其中 $N = 10$、 $\lambda = 1$ 和 $\sigma_X^2 = 1$。

显示， $\tau^\star = 1/2$，这意味着在本例中，$a^\star(x) = \mathbf{1}(x \le 0.5)$。

首先，我们讨论高斯累积分布函数的以下性质。由于篇幅限制，其证明略去。

引理 *3*： 设 $V \sim \mathcal{N}(0,\ 1)$。那么，对于任意 $c \in \mathbb{R}$ 和任意 $\epsilon \in \mathbb{R}$，有

$$|\mathbb{E}_V[\Phi(cV + \epsilon)] - \frac{1}{2}| \le |\epsilon|. \tag{11}$$

特别地，对于所有 $c \in \mathbb{R}$， $\Phi(cV) = \frac{1}{2}$。

利用这一结果，我们可以得到式 (10) 的不动点的以下重要估计值。

引理 *4*： 对于 σ_X^2， $\sigma_Z^2 > 0$，设 $\tau_N^\star = \tau_N^\star(\sigma_X^2,\ \sigma_Z^2)$ 为不动点方程的唯一解

$$\lambda\frac{N-1}{N}\pi(\tau_N^\star; \tau_N^\star) = \alpha\tau_N^\star. \tag{12}$$

那么，

$$\frac{1}{2} \le \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Z^2}\right)\frac{\tau_N^\star}{\lambda(N-1)/N} \le \left(\frac{1}{2} + \frac{\sigma_Z}{\sigma_X^2}\lambda(N-1)/N\right). \tag{13}$$

证明： 首先，注意到 $0 \le \Phi(\cdot) \le 1$ 和 (10) 可推出 $0 \le \pi(\tau_N^\star;\ \tau_N^\star) \le 1$。因此，(12) 的解满足：

$$0 \le \tau_N^\star \le \frac{\lambda(N-1)/N}{\alpha}. \tag{14}$$

使用 $\pi(\tau_N^\star;\ \tau_N^\star) = \mathbb{E}\left[\Phi\left(\frac{\tau_N^\star - \tilde{\sigma}W}{(1-\alpha)\ \sigma_Z}\right)\right]$，我们得到

$$\begin{aligned}
\frac{1}{2} &\stackrel{(a)}{=} \mathbb{E}\left[\Phi\left(\frac{-\tilde{\sigma}W}{\sigma_Z}\right)\right] \\
&\stackrel{(b)}{\le} \mathbb{E}\left[\Phi\left(\frac{(1-\alpha)\tau_N^\star - \tilde{\sigma}W}{\sigma_Z}\right)\right] \\
&\stackrel{(c)}{\le} \mathbb{E}\left[\Phi\left(\frac{(1-\alpha)\frac{\lambda(N-1)/N}{\alpha} - \tilde{\sigma}W}{\sigma_Z}\right)\right]
\end{aligned}$$

$$\overset{(d)}{\leq} \frac{1}{2} + \frac{(1-\alpha)}{\sigma_Z}\lambda(N-1)/N.$$

这里，(a)由 $\epsilon = 0$的式(11)得出，(b)和(c)由式(14)、事实 $0 < \alpha < 1$、期望的单调性以及 $\Phi(\cdot)$的单调性得出，(d)由式(11)得出。利用上述不等式以及 $\tau_N^\star$满足(12)这一事实，我们有

$$\frac{1}{2} \leq \alpha\frac{\tau_N^\star}{\lambda(N-1)/N} \leq \frac{1}{2} + \frac{(1-\alpha)}{\sigma_Z}\lambda(N-1)/N.$$

最后，注意到 $\alpha = \frac{\sigma_X^2}{\sigma_X^2+\sigma_Z^2}$，证明得证。 ■

以下结果可由引理 4 立即推出。

定理 *2*（弥散高斯先验）：考虑一个收益函数如式（7）所示的全局博弈。那么，对于所有 $\sigma_Z^2 > 0$,

$$\lim_{\sigma_X^2\to\infty} \tau_N^\star = \frac{\lambda}{2}\cdot\Big(1-\frac{1}{N}\Big).$$

注释*2*：在关于全球博弈的经济学文献中，[6]–[8],通常假设 X 服从弥散先验分布。然而，从工程学视角来看，假设 X 上存在弥散高斯分布，且满足 $\sigma_X^2 \to \infty$，实际上相当于拥有并行的高斯通信信道，其容量无限大[3]。因此，从信息论视角来看，信道噪声的影响变得可以忽略不计，从而能够在均方误差意义下，根据 Y_i 对输入 X 做出完美估计 [18]。

注释*3*：我们的结果的一个特例是，当 $N \to \infty$ 且 $\lambda = 1$时，我们得到 $\tau_\infty^\star(1) \to 1/2$，该结果由 Morris and Shin [6] 建立。

定理 2 的重要意义在于，它导出了一个贝叶斯纳什均衡阈值策略，对应于通过完美信道观测信号的情形。值得注意的是，对于线性收益函数，这种策略具有闭式解。我们将其称为预言机策略，定义如下：

$$\gamma^\star_{\text{oracle}}(x) \overset{\text{def}}{=} \mathbf{1}\big(x \leq \tau_N^\star(\sigma_Z^2)\big). \tag{15}$$

D. 确定性等价策略

对 $\sigma_X^2 \to \infty$最优策略的了解促使我们定义确定性等价（CE）策略[19]：当 $\sigma_X^2 < \infty$ 时，智能体首先使用式 (3) 中的最小均方误差估计器估计基本面，然后使用式 (15) 中的预言机策略：

$$\gamma_{\text{ce}}(y) \overset{\text{def}}{=} \mathbf{1}\big(\hat{x}_{\text{mmse}}(y) \leq \tau_N^\star(\sigma_Z^2)\big).$$

因此，

$$\gamma_{\text{ce}}(y) = \mathbf{1}\left(y \leq \Big(1+\frac{\sigma_Z^2}{\sigma_X^2}\Big)\cdot\frac{\lambda}{2}\cdot\Big(1-\frac{1}{N}\Big)\right).$$

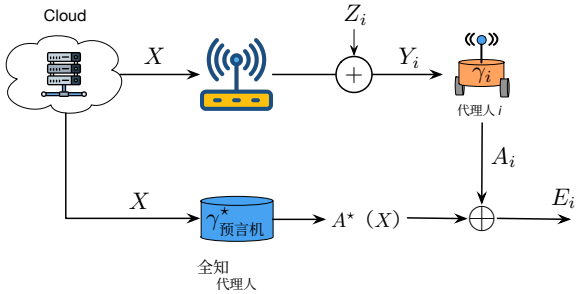


图3. 示意图展示了如何计算一般主体与可获取完美信息、参与随机协调博弈的全知智能体之间的协调错误事件。

IV. 协调的基本面限制

现在，我们得到一个适用于任何策略的效率通用上界，不论其结构如何。我们的结果基于*Fano*不等式 [20]。Fano不等式给出了基于侧信息对离散随机变量进行估计时，估计值发生错误的概率上界。

定理*3*（协调效率上界）：对于具有线性收益函数的全局博弈，任意齐性策略组合的协调效率均满足以下界

$$\varrho \leq 1 - h^{-1}\Big(H\big(A^\star(X) \mid Y_i\big)\Big),$$

其中 $H(\cdot \mid \cdot)$是条件熵函数[4]，，而$h^{-1}(\cdot)$是区间 $[0, 1/2]$ 上二元熵函数的反函数。

Proof: Let $A^\star(X) \in \{0,1\}$ such that

$$A_i^\star(X) = \mathbf{1}\big(X \leq \tau_N^\star(\sigma_Z^2)\big)$$

并令$\hat{}$ $A_i(Y) \in \{0,1\}$ 表示基于 Y_i对 $A^\star(X)$的任意估计值。然后注意，满足以下马尔可夫关系

$$A^\star(X) \leftrightarrow X \leftrightarrow Y \leftrightarrow \hat{A}_i(Y_i).$$

考虑图3中的框图，定义错误随机变量

$$E_i \overset{\text{def}}{=} \mathbf{1}\big(\hat{A}_i(Y) \neq A^\star(X)\big),$$

并注意，估计 $A^\star(X)$时出错的概率至多为 $1/2$。Fano不等式 [20] 给出了利用预言机策略，根据第 i个代理人可用的信号 Y_i 计算出的最优决策的条件熵上界

$$H\big(A^\star(X) \mid Y_i\big) \leq h(E_i) + \mathbb{P}(E_i = 1)\log_2(|\mathcal{A}| - 1),$$

[3]高斯信道的香农容量为

$$C = \frac{1}{2}\log_2\Big(1+\frac{\sigma_X^2}{\sigma_Z^2}\Big) \text{ bits/channel use.}$$

[4]随机变量 $X \sim f(x)$ 的熵定义为

$$H(X) \overset{\text{def}}{=} -\mathbb{E}\Big[\log_2\big(f(X)\big)\Big].$$

其中，$|\mathcal{A}|$ 是决策变量 A_i的基数。由于我们的决策变量是二元变量，因此有

$$H\big(A^\star(X) \mid Y_i\big) \le h(E_i). \tag{16}$$

假设我们可以计算式（16）的左侧（LHS），则通过在区间 $[0, 0.5]$ 内求二元熵函数[5] 的反函数，可以得到 $\mathbb{P}(E_i = 1)$ 的一个上界。最后，注意到对于齐次策略剖面，协同效率为

$$\varrho = 1 - \mathbb{P}(E_i = 1) \le 1 - h^{-1}\Big(H\big(A^\star(X) \mid Y_i\big)\Big).$$

■

A. 计算协调效率的上界

利用熵函数的性质，我们得到:

$$H\big(A^\star(X) \mid Y_i\big) = H\big(A^\star(X)\big) - H(Y_i) + H\big(Y_i \mid A^\star(X)\big).$$

接下来计算这三个项：第一项是预言机计算得到的最优决策变量的熵:

$$h\big(A^\star(X)\big) = h\left(\mathbb{P}\left(X \le \frac{\lambda}{2}\Big(1 - \frac{1}{N}\Big)\right)\right),$$

其中，$h(\cdot)$ 表示二元熵函数。

第二项是信号 Y_i 的微分熵，其中信号是方差为 $\sigma_X^2 + \sigma_Z^2$ 的高斯随机变量。因此，

$$H(Y_i) = \frac{1}{2}\log_2\big(2\pi e(\sigma_X^2 + \sigma_Z^2)\big).$$

第三项更具挑战性，必须通过数值方法计算。

$$H(Y_i \mid A^\star(X) = 1) = H\left(Y_i \mid X \le \frac{\lambda}{2}\Big(1 - \frac{1}{N}\Big)\right).$$

为了计算这个熵，我们必须使用条件概率密度函数

$$f_{Y_i \mid X \le \frac{\lambda}{2}(1-\frac{1}{N})}(y_i) = \frac{\int_{-\infty}^{\frac{\lambda}{2}(1-\frac{1}{N})} f_Z(y_i - x) f_X(x) dx}{\int_{-\infty}^{\frac{\lambda}{2}(1-\frac{1}{N})} f_X(x) dx}.$$

同样，

$$H(Y_i \mid A^\star(X) = 0) = H\left(Y_i \mid X > \frac{\lambda}{2}\Big(1 - \frac{1}{N}\Big)\right).$$

为了计算这个熵，我们必须使用条件概率密度函数

$$f_{Y_i \mid X > \frac{\lambda}{2}(1-\frac{1}{N})}(y_i) = \frac{\int_{\frac{\lambda}{2}(1-\frac{1}{N})}^{\infty} f_Z(y - x) f_X(x) dx}{\int_{\frac{\lambda}{2}(1-\frac{1}{N})}^{\infty} f_X(x) dx}.$$

最后，我们可以计算:

$$\begin{aligned} H(Y_i \mid A^\star(X)) = {} & H(Y_i \mid A^\star(X) = 0)\mathbb{P}\big(A^\star(X) = 0\big) \\ & + H(Y_i \mid A^\star(X) = 1)\mathbb{P}\big(A^\star(X) = 1\big), \end{aligned}$$

[5]二元熵函数定义为

$$h(p) \stackrel{\text{def}}{=} -p\log_2 p - (1-p)\log_2(1-p).$$

其中

$$\begin{aligned} H(Y_i \mid A^\star(X) = 0) = {} & -\int_{\mathbb{R}} f_{Y_i \mid X > \frac{\lambda}{2}(1-\frac{1}{N})}(y_i) \\ & \times \log_2\Big(f_{Y_i \mid X > \frac{\lambda}{2}\left(1-\frac{1}{N}\right)}(y_i)\Big) dy_i \end{aligned}$$

and

$$\begin{aligned} H(Y_i \mid A^\star(X) = 1) = {} & -\int_{\mathbb{R}} f_{Y_i \mid X \le \frac{\lambda}{2}(1-\frac{1}{N})}(y_i) \\ & \times \log_2\Big(f_{Y_i \mid X \le \frac{\lambda}{2}(1-\frac{1}{N})}(y_i)\Big) dy_i. \end{aligned}$$

最后，可以通过数值方法高效地计算二元熵函数的反函数。

V. 数值结果

迄今为止，我们给出的刻画假设智能体根据某种策略选择行动，该策略理想情况下会跟踪全知智能体的行为，而全知智能体能够获得有关状态的完美信息。由于智能体接收到有关状态的含噪信号，因此它们无法通过以 $\tau^\star$ 为索引的阈值策略与全知智能体实现完美协同。

假设智能体采用以 τ 为索引的一般同质阈值策略，则未能协调的概率 $X = x$ 为:

$$\begin{aligned} \mathbb{P}(E_i = 1 \mid X = x) = {} & \left(1 - \Phi\Big(\frac{\tau - x}{\sigma_Z}\Big)\right)\mathbf{1}(x \le \tau^\star) \\ & + \Phi\Big(\frac{\tau - x}{\sigma_Z}\Big)\mathbf{1}(x > \tau^\star). \end{aligned} \tag{17}$$

因此，

$$\varrho(\boldsymbol{\gamma}) = 1 - \int_{\mathbb{R}} \mathbb{P}(E_i = 1 \mid X = x) f(x) dx. \tag{18}$$

假设存在一个具有线性收益函数且包含 $N \to \infty$ 个智能体的全局博弈。全知智能体采用的最优阈值为 $\tau^\star = \lambda/2$。从智能体的角度来看，我们考虑两种策略：1. 利用先验信息 σ_X^2 和 σ_Z^2 以及参数 λ，计算该全局博弈的纳什均衡阈值；2. 使用最小均方误差估计器估计状态变量 X，并采用确定性等价策略。

图4（左）展示了系统 $\sigma_X^2 = 1$ and$\lambda = 1$ 中这两类协调策略对应的阈值随噪声方差 σ_Z^2变化的情况。我们可以清楚地看到这两种策略有多么不同。此外，求解第一种策略中的 NE 计算成本也更高，而在这种情况下，CE 策略可以通过闭式解得到。图4（中）展示了这两种策略的期望效用。采用 NE 的代理人与采用 CE 的代理人，其效用之间存在显著差距。这一点也很明显，因为在随机控制和优化中，CE 通常是一种次优策略。更令人惊讶的是，CE 带来了更高的协调效率，如图4（右）所示。

最后，图4（右）还展示了任意齐性策略组合（而不仅仅是阈值策略）的协调效率信息论上界。该图的意义在于，它表明某些协调

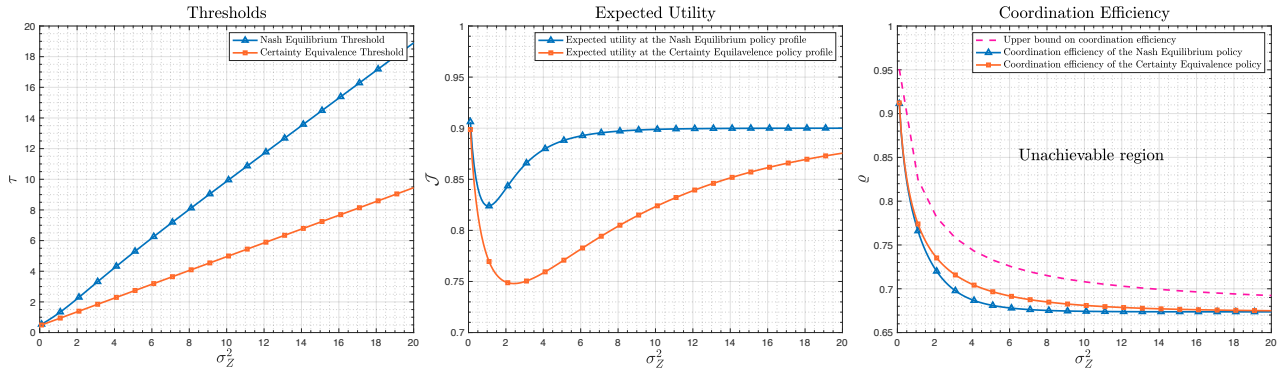


图4. 对于具有线性收益函数的全局博弈， $\lambda = 1$ 在方差为 $\sigma_X^2 = 1$的非弥散先验以及可数无穷多个智能体条件下的数值结果。纳什均衡和协同效率阈值随噪声方差变化的关系（左）；期望效用随噪声方差变化的关系（中）；协调效率及其通用下界（右）。

效率在实际应用中，对于网关与机器人智能体之间通信信道中的给定噪声水平，无法通过任何策略实现。因此，在适当规划由具有策略性的自利智能体执行集体任务的分布式实现时，系统设计者需要以一定的信噪比进行通信；该信噪比并非由接收器处的误码率决定，而是取决于其希望实现的集体协调水平。

VI. 结论与未来工作

我们定义了齐次协调博弈这一类别，其中包含广受欢迎的全球博弈。随后，我们提出了一种贝叶斯协调度量，该度量基于智能体采取“正确”行动的概率，即它们将自己的决策与能够获取系统完美状态的全知智能体的决策保持一致的概率。我们表明，可以利用信息论不等式对这种协调效率度量进行界定，从而确定某些协调水平无法实现的情形。据我们所知，这是首次将此类方法用于全球博弈。

未来关于这一主题的研究将包括：在智能体拥有本地数据的情况下，为阈值策略设计新的学习算法；考虑智能体效益函数上部分连接的影响图；以及刻画更优的协调效率上界，并将策略结构（例如阈值）纳入考量。

参考文献

[1] L. Arditti、G. Como、F. Fagnani 和 M. Vanelli，“混合网络协调/反协调博弈中的均衡与学习动态”，载于*2021*年 *IEEE* 第*60*届决策与控制会议（*CDC*），2021年，第4982–4987页。[2] P. Ramazi 和 M. H. Roohi，“协调者与反协调者异质群体中的振荡特征”，载于*2022*年 *IEEE* 第*61*届决策与控制会议（*CDC*），2022年，第4615–4620页。[3] K. Paarporn、B. Canty、P. N. Brown、M. Alizadeh 和 J. R. Marden，“复杂且掌握信息的对抗行为对图协调博弈的影响”，*IEEE* 网络系统控制汇刊，第8卷，第1期，第200–211页，2021年。

[4] K. Paarporn、M. Alizadeh 和 J. R. Marden，“图协调博弈中的风险-安全权衡”，*IEEE*自动控制汇刊，第66卷，第5期，第1973–1985页，2021年。[5] S. Das 和 C. Eksin，“网络博弈中反协调最大化的近似次模性”，载于*2022*年 *IEEE* 第*61*届决策与控制会议（*CDC*），2022年，第3151–3157页。[6] S. Morris 和 H. S. Shin，《全局博弈：理论与应用》，计量经济学会专著丛书。剑桥大学出版社，2003年，第1卷，第56–114页。[7] H. Carlsson 和 E. Van Damme，“全球博弈与均衡选择”，《计量经济学》：计量经济学会期刊，第989–1018页，1993年。[8] M. A. Dahleh、A. Tahbaz-Salehi、J. N. Tsitsiklis 和 S. I. Zoumpoulis，“本地信息下的协调”，《运筹学》，第64卷，第3期，第622–637页，2016年。[9] B. Touri 和 J. Shamma，“噪声环境下的信息共享与全局博弈”，载于第*53*届*IEEE*决策与控制会议。IEEE，2014年，第4473–4478页。[10] H. Mahdavifar、A. Beirami、B. Touri 和 J. S. Shamma，“噪声环境下的信息共享与全局博弈”，*IEEE* 网络信号与信息处理汇刊，第4卷，2017年，第509–497页。[11] C. M. Leister、Y. Zenou 和 J. Zhou，“社会连通性与局部传染”，《经济研究评论》，第89卷，第1期，第372–410页，2022年。[12] M. M. Vasconcelos，“基于泊松观测的仿生多智能体协调博弈”，*IFAC-PapersOnLine*，第55卷，第13期，第180–185页，2022年。[13] Y. Wei 和 M. M. Vasconcelos，“在计算与通信能力有限的机器人正则网络上进行策略性多任务协调”，*arXiv*预印本 *arXiv:2212.10968*，2022年。[14] E. J. Hoffmann 和 T. Sabarwal，“具有策略互补与替代关系的全局博弈”，《博弈与经济行为》，第118卷，第72–93页，2019年。[15] A. Kanakia、B. Touri 和 N. Correll，“将信息有限条件下的多机器人任务分配建模为全局博弈”，《群体智能》，第10卷，第2期，第147–160页，2016年。[16] S. Berman、A. Halasz、M. A. Hsieh 和 V. Kumar，“机器人群任务分配的优化随机策略”，*IEEE*机器人学汇刊，第25卷，第4期，第927–937页，2009年。[17] B. Hajek，《工程师随机过程》。剑桥大学出版社，2015年。[18] D. Guo、S. Shamai 和 S. Verdu，“高斯信道中的互信息与最小均方误差”，*IEEE* 信息论汇刊，第51卷，第4期，第1261–1282页，2005年。[19] D. Bertsekas，《动态规划与最优控制：第一卷》。Athena Scientific，2012年，第1卷。[20] T. M. Cover 和 J. A. Thomas，《信息论基础》，第2版，Wiley电信与信号处理丛书。Wiley，2006年。