NUS National University of Singapore
School of Computing

On the Convergence of the Shapley Value in Parametric Bayesian Learning Games

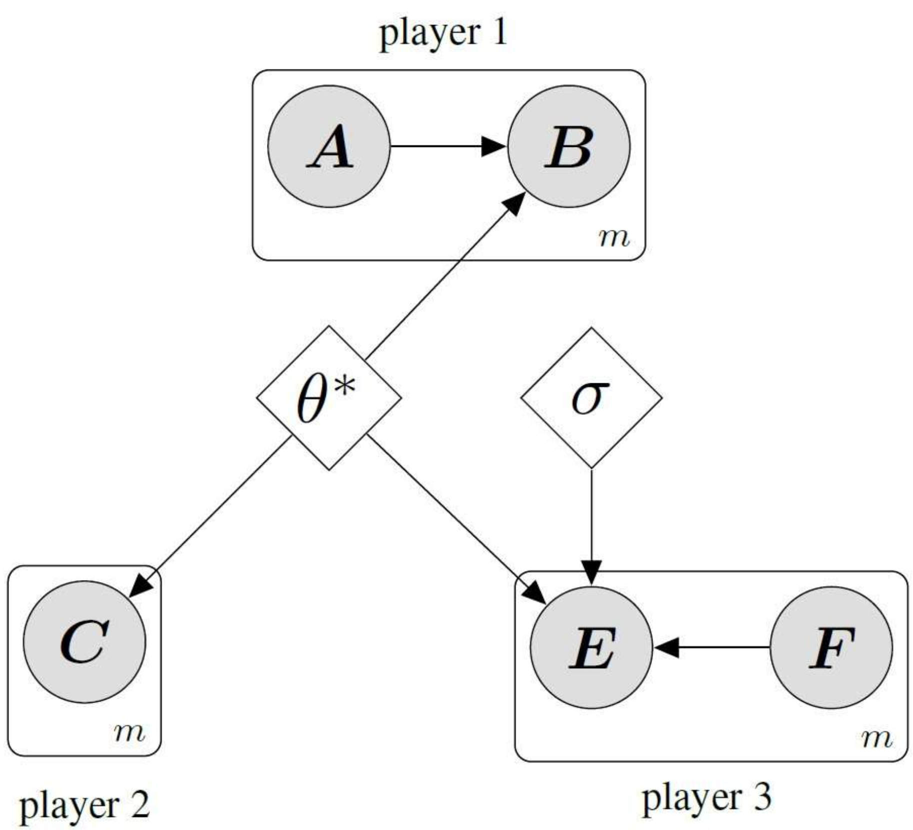
Lucas Agussurja Xinyi Xu Bryan Kian Hsiang Low

Introduction

National University of Singapore
School of Computing

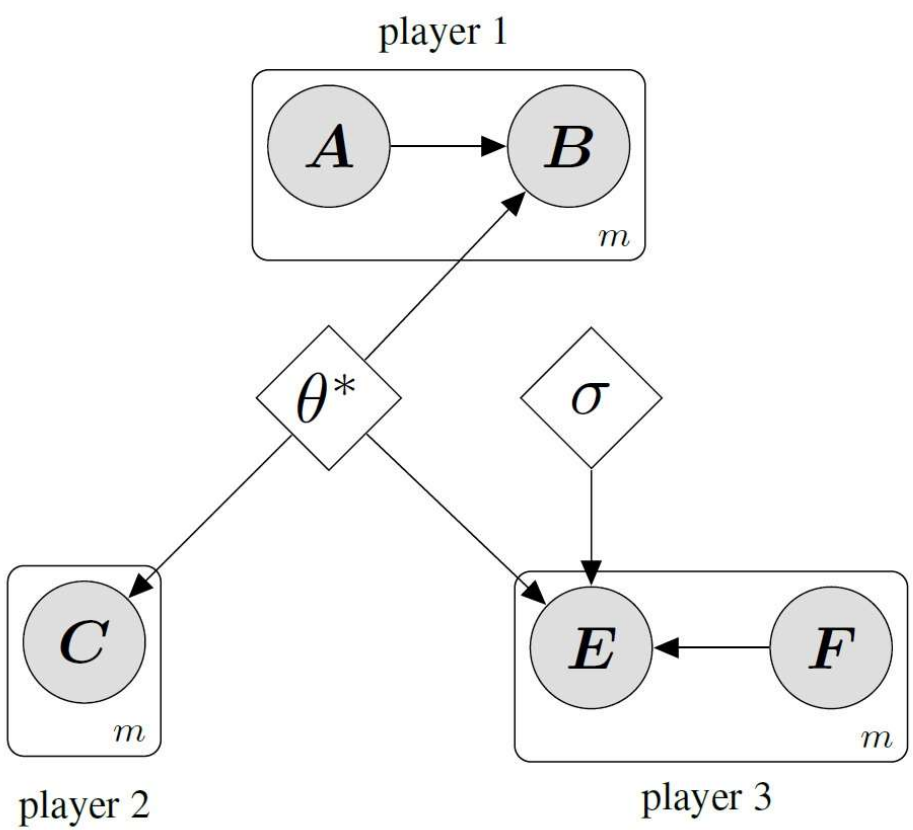
- Measuring datasets **contributions** in a **joint Bayesian inference**
- **Cooperative game theory** perspective
- Characteristic function: **Posterior-prior information gain**
- **Convergence** properties of the **Shapley value**

Collaborative Setting: Example



Collaborative Setting: Inference

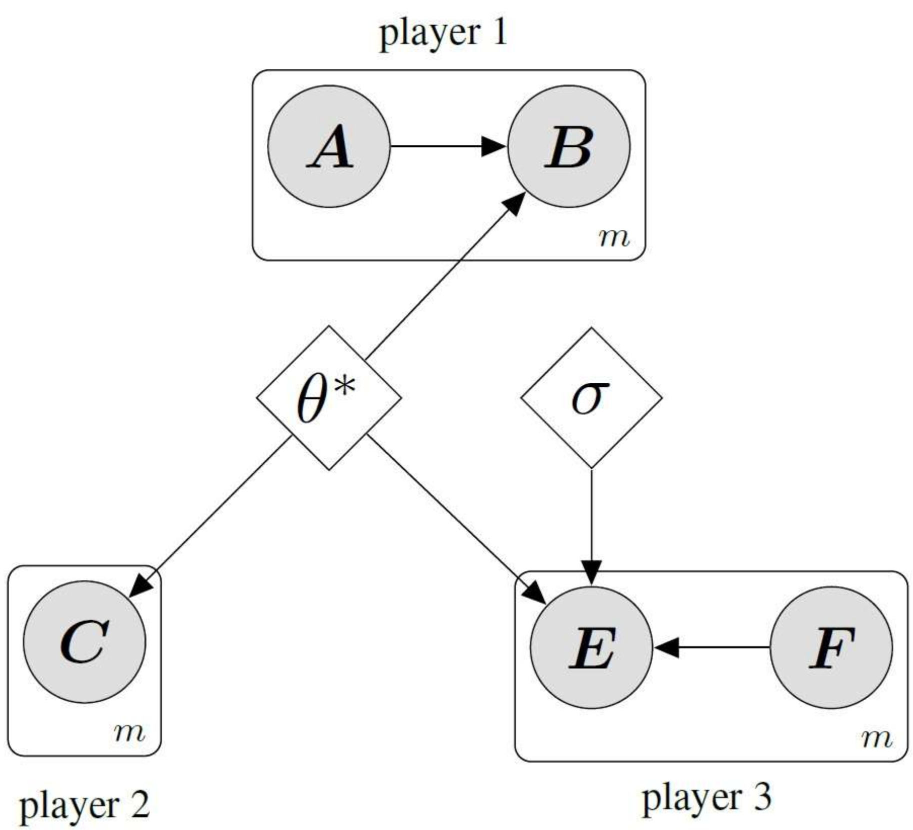
NUS National University of Singapore
School of Computing



Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\ \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\ \mathrm{d}\Pi(\theta')}$$

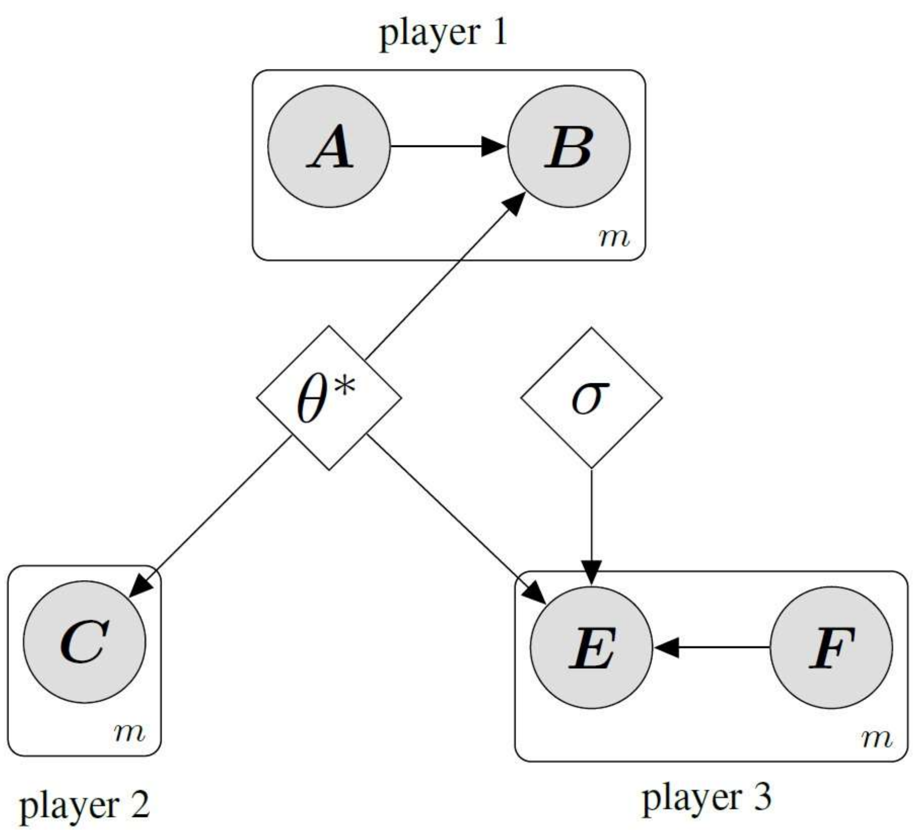
Collaborative Setting: Inference



Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\, \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\, \mathrm{d}\Pi(\theta')}$$

Collaborative Setting: Inference

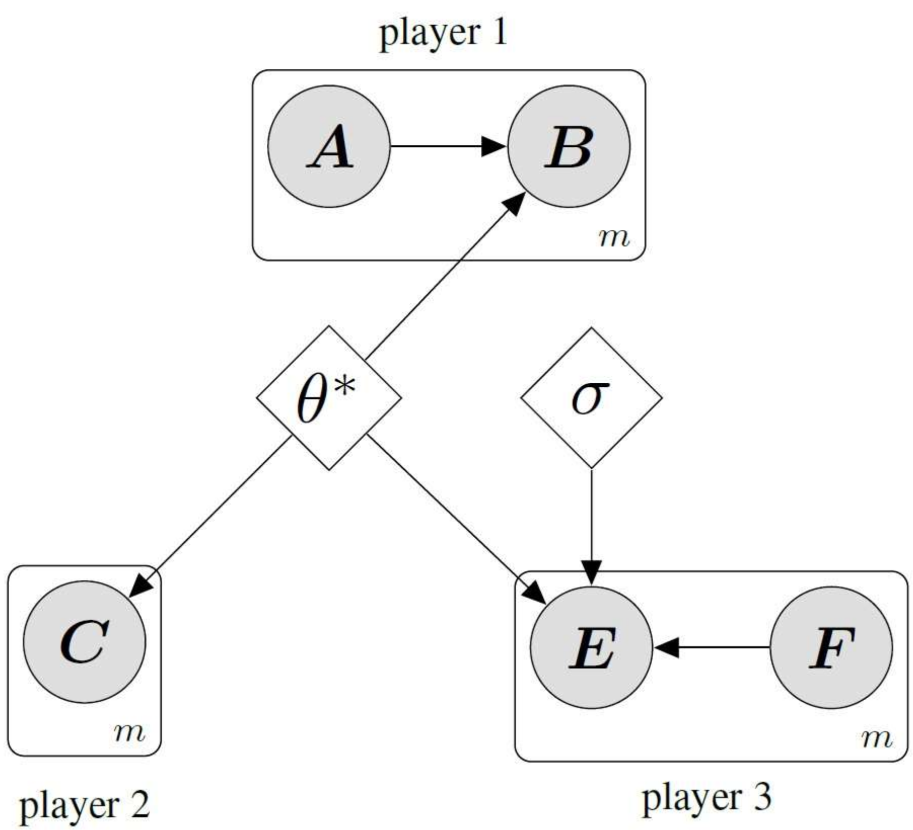


Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\, \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\, \mathrm{d}\Pi(\theta')}$$

Collaborative Setting: Assumptions

NUS National University of Singapore
School of Computing



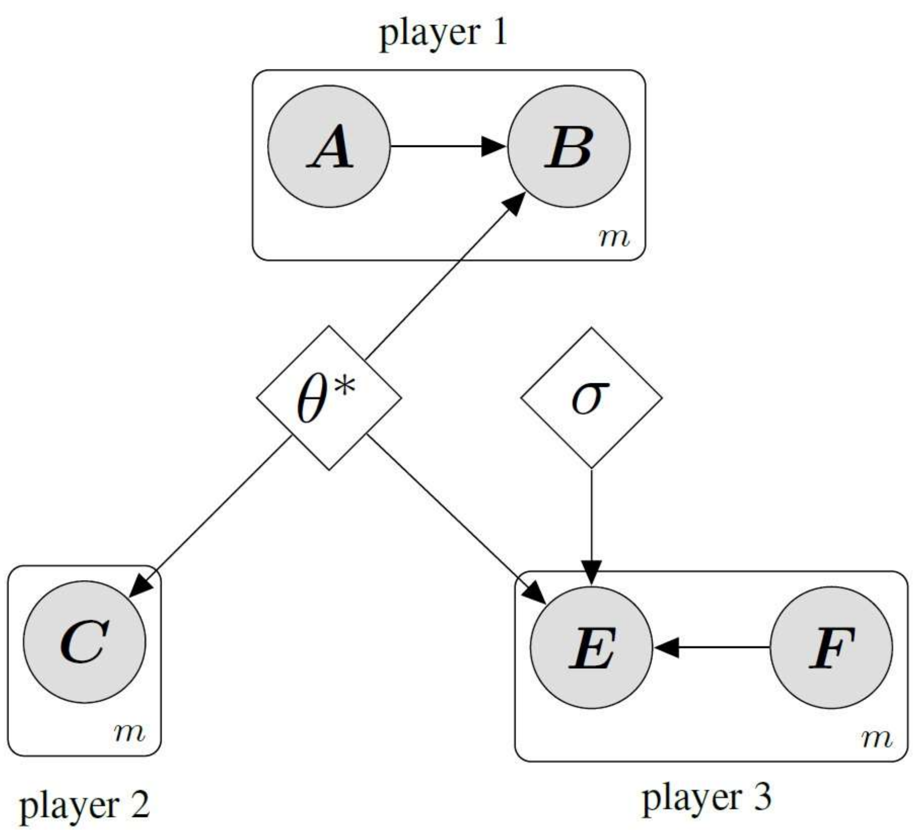
Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\, \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\, \mathrm{d}\Pi(\theta')}$$

Assumption 1: Conditional Independence

$$\mathcal{L}_S(\theta; \mathbf{X}_S^m) = \prod_{i \in S} \mathcal{L}_i(\theta; \mathbf{X}_i^m)$$

Collaborative Setting: Assumptions



Joint Inference:

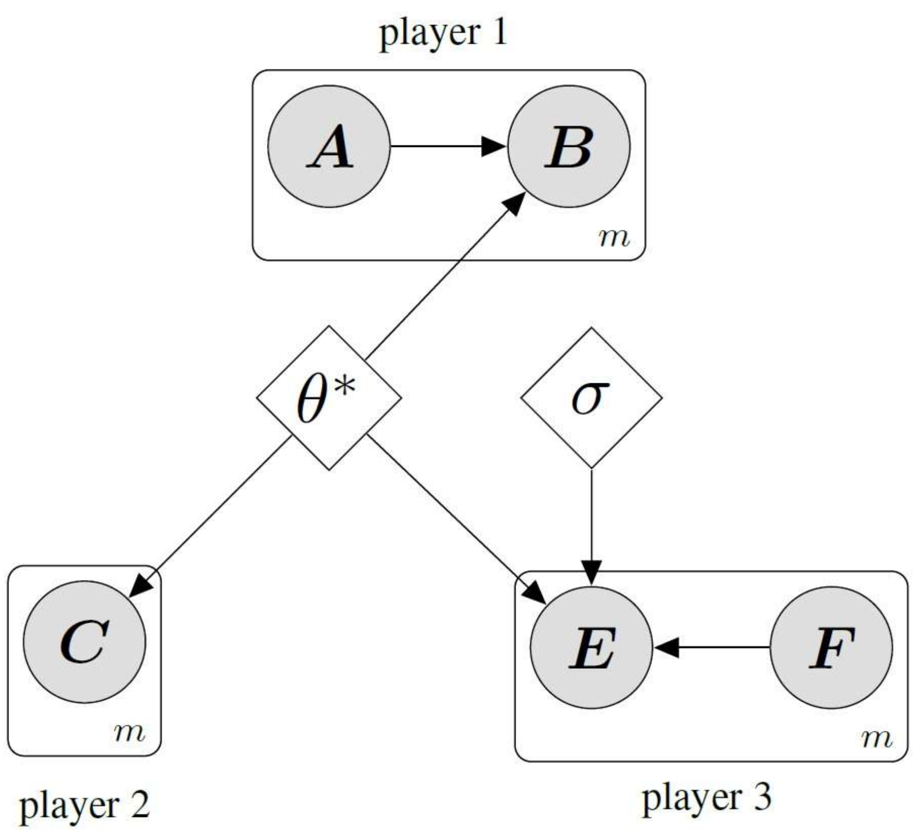
$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\,\pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\,\mathrm{d}\Pi(\theta')}$$

Assumption 2: Twice Differentiability

$$\mathcal{L}_S(\theta; \mathbf{X}_S^m) = \prod_{i \in S} \mathcal{L}_i(\theta; \mathbf{X}_i^m)$$

Collaborative Setting: Assumptions

NUS National University of Singapore
School of Computing



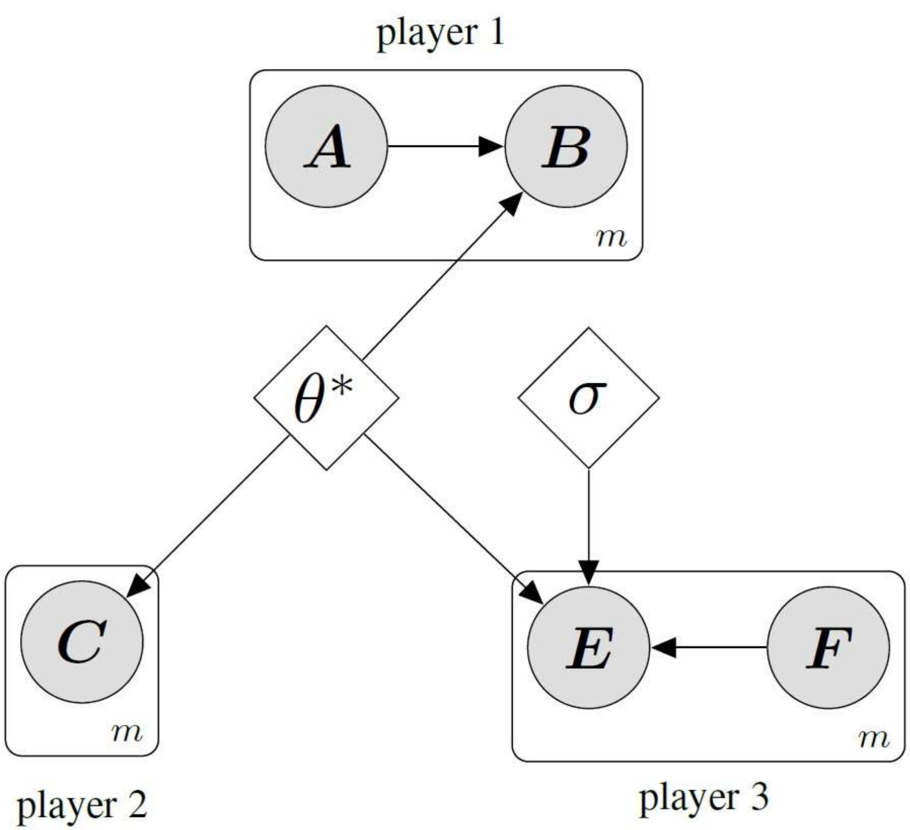
Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\, \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\, \mathrm{d}\Pi(\theta')}$$

Assumption 3:
Distinguisability

$$\mathcal{L}_S(\theta; \mathbf{X}_S^m) = \prod_{i \in S} \mathcal{L}_i(\theta; \mathbf{X}_i^m)$$

Collaborative Setting: Assumptions



Joint Inference:

$$\boldsymbol{p}_S^m(\theta) = \frac{\mathcal{L}_S(\theta; \mathbf{X}_S^m)\, \pi(\theta)}{\int_{\theta' \in \Theta} \mathcal{L}_S(\theta'; \mathbf{X}_S^m)\, \mathrm{d}\Pi(\theta')}$$

Assumption 4: The Prior supports the true parameter

Cooperative Game Formulation

- **Characteristic function**:

$$\boldsymbol{\mathcal{V}}^m(S) := \mathrm{KL}(\boldsymbol{P}_S^m \| \Pi) := \int_{\Theta} \log\left(\mathrm{d}\boldsymbol{P}_S^m / \mathrm{d}\Pi\right) \mathrm{d}\boldsymbol{P}_S^m$$

- **Shapley value**:

$$\phi(i; \boldsymbol{\mathcal{V}}^m) := \sum_{S \subseteq N \setminus \{i\}} w_S \left[\boldsymbol{\mathcal{V}}^m(S \cup \{i\}) - \boldsymbol{\mathcal{V}}^m(S)\right]$$

Main Theoretical Result

NUS National University of Singapore
School of Computing

- Define the **limiting game**:

$$V(S) := 0.5 \log |\mathcal{I}_S| \quad \text{if } S \text{ is nonempty,}$$
$$V(S) := 0 \quad \text{otherwise}$$

- Under **Assumptions 1-4**
- For a **uniformly* distributed prior**
- The following holds:

$$\phi(i; \boldsymbol{\mathcal{V}}^m) - \phi(j; \boldsymbol{\mathcal{V}}^m) \xrightarrow{\text{p}} \phi(i; V) - \phi(j; V)$$

for any two players i and j.

High Level Proof Ideas

- Assumptions 1-4 allow for **closed form approximation** of joint posteriors
- This is given by **Bernstein-von Mises theorem**
- Approximate KL divergence using the **normal posterior**
- Characteristic function **decomposes**
- By **linearity** of Shapley value, when taking differences:
  - Constant terms cancel out
  - Left with term that depends only on Fisher information

Collaborative Online Framework

National University of Singapore
School of Computing

- **Bundling of data points:**

$$\mathrm{Var}\left[\frac{\partial}{\partial\theta}\log\prod_{j=1}^{r}\mathcal{L}_1(\theta^*;\mathbf{X}_{1j})\right]=\sum_{j=1}^{r}\mathrm{Var}\left[\frac{\partial}{\partial\theta}\log\mathcal{L}_1(\theta^*;\mathbf{X}_{1j})\right]=r\mathcal{I}_1$$

- **Setting:**

$$r_1 : r_2 = |\mathcal{I}_2|^{1/k} : |\mathcal{I}_1|^{1/k}$$

- **Results in:**

$$|r_1\mathcal{I}_1| = r_1^k|\mathcal{I}_1| = \left(\sqrt[k]{\frac{|\mathcal{I}_2|}{|\mathcal{I}_1|}}\,r_2\right)^k \cdot |\mathcal{I}_1| = r_2^k|\mathcal{I}_2| = |r_2\mathcal{I}_2|$$

Collaborative Online Framework

NUS National University of Singapore
School of Computing

In each iteration:

1. Use current available data points to compute an estimate $\bar{\theta}$ of the true parameter.

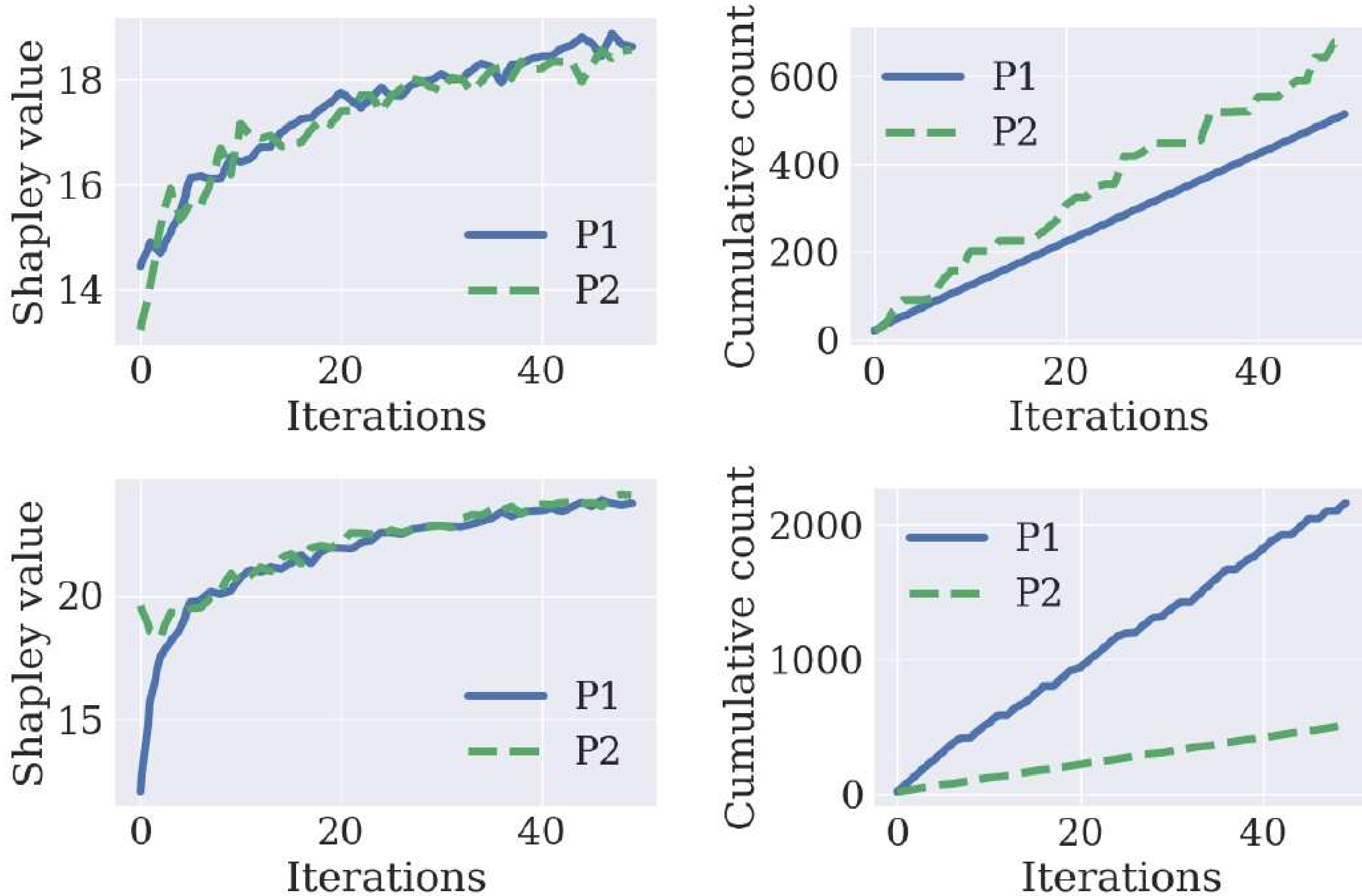
2. Let m_i denote the number of player i's data points received so far. For $i = 1, 2$, use sample approximation to estimate the Fisher information at $\bar{\theta}$:

$$\hat{\mathcal{I}}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \left(\frac{\partial}{\partial \theta} \log \mathcal{L}_i(\bar{\theta}; x_{ij}) \right) \left(\frac{\partial}{\partial \theta} \log \mathcal{L}_i(\bar{\theta}; x_{ij}) \right)^\top .$$

3. Collect r_1 and r_2 data points from the respective players 1 and 2 s.t. the proportion $m_1 + r_1 : m_2 + r_2$ of their cumulative data points is equal to $|\hat{\mathcal{I}}_2|^{1/k} : |\hat{\mathcal{I}}_1|^{1/k}$.

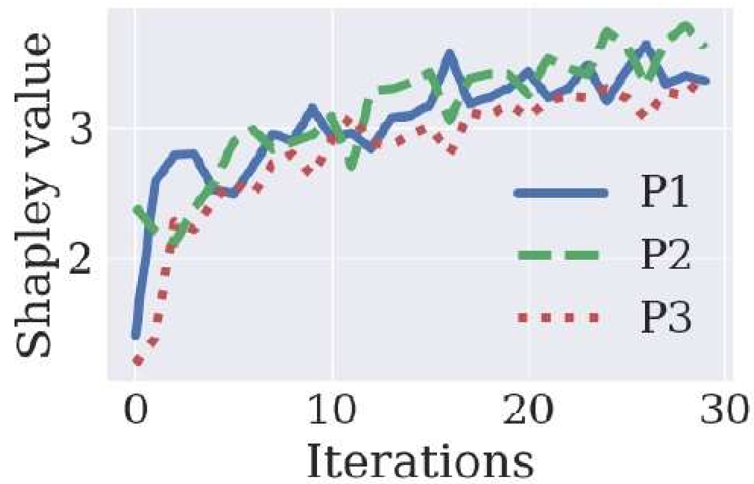
Experimental Results

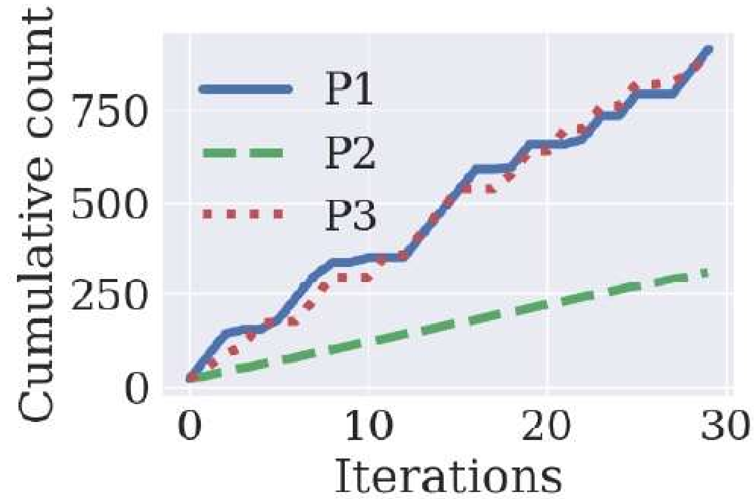
NUS National University of Singapore
School of Computing

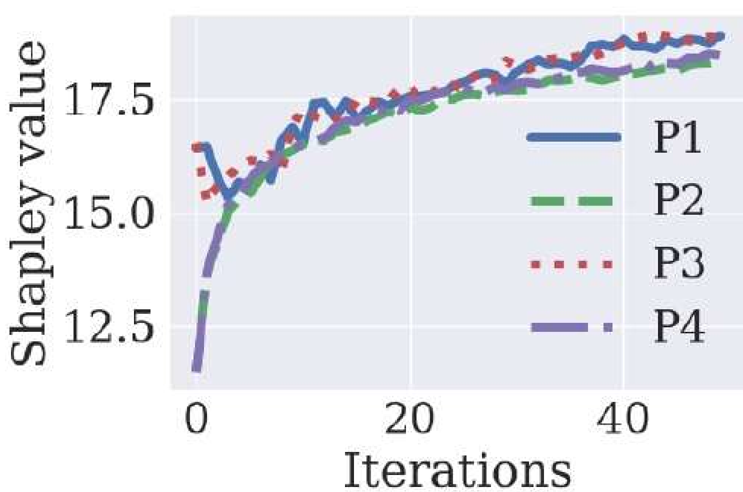


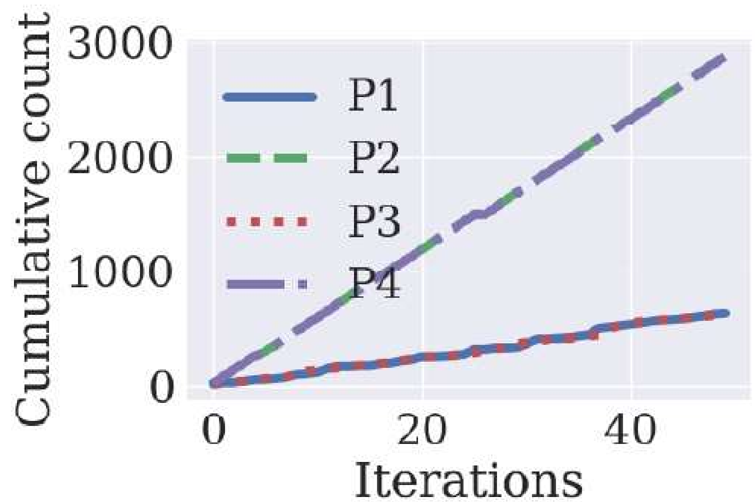
Experimental Results

School of Computing









National University of Singapore
School of Computing

On the Convergence of the Shapley Value in Parametric Bayesian Learning Games

Lucas Agussurja Xinyi Xu Bryan Kian Hsiang Low