EURECOM
Department of Mobile Communications
2229, route des Crêtes
B.P. 193
06904 Sophia-Antipolis
FRANCE

Research Report RR-12-264

# Confidence intervals for Shapley value in Markovian dynamic games

January 5$^{th}$, 2012

Prof. Konstantin AVRACHENKOV
Prof. Laura COTTATELLUCCI
Lorenzo MAGGI

Tel : (+33) 4 93 00 81 00
Fax : (+33) 4 93 00 82 00
Email : k.avrachenkov@sophia.inria.fr, {laura.cottatellucci, lorenzo.maggi}@eurecom.fr,

[1]EURECOM's research is partially supported by its industrial members: BMW Group, Cisco, Monaco Telecom, Orange, SAP, SFR, Sharp, STEricsson, Swisscom, Symantec, Thales.

# Confidence intervals for Shapley value in Markovian dynamic games

Konstantin Avrachenkov, Laura Cottatellucci, Lorenzo Maggi

**Abstract**

We consider a dynamic multiagent system in which several states succeed each other, following a Markov chain process. In each state, a different single stage game among the agents, or players, is played, and cooperation among subsets of players can arise in order to achieve a common goal. We assume that each coalition can ensure a certain value for itself. The Shapley value is a well known method to share the value of the grand coalition, formed by all the players, among the players themselves. It reflects the effective incremental asset brought by each player to the community. Unfortunately, the exact computation of the Shapley value for each player requires an exponential complexity in the number of players. Moreover, we prove that an exponential number of queries is necessary for any deterministic algorithm even to approximate the Shapley value in a Markovian dynamic game with polynomial accuracy. Motivated by these reasons, we propose three different methods to compute a confidence interval for the Shapley value in the Markovian game. Our approaches require a polynomial number of queries to achieve a polynomial accuracy. We compare our confidence intervals in terms of their tightness and we provide a straightforward sampling strategy to optimize the tightness of one of them.

**Index Terms**

Economic paradigms: Game theory (cooperative and non-cooperative), Agent Cooperation: Implicit Cooperation

# Contents


**1 Introduction** **1**

**2 Model and Background results** **2**

**3 Motivations** **4**

**4 Complexity of deterministic algorithms** **4**

**5 Static approach** **6**

**6 Dynamic approach** **8**
- 6.1 First dynamic method . . . . . . . . . . . . . . . . . . . . . . . . 8
  - 6.1.1 Optimum sampling strategy . . . . . . . . . . . . . . . . . . 9
- 6.2 Second dynamic method . . . . . . . . . . . . . . . . . . . . . . . 11
  - 6.2.1 Confidence interval for single stage games: example for simple games . . . . . . . . . . . . . . . . . . . . . . . . 13

**7 Comparison among the proposed approaches** **14**

**8 Complexity of confidence intervals** **16**

**9 Conclusions** **17**

**10 Acknowledgments** **17**

# 1 Introduction

Cooperative game theory is perhaps the most powerful tool to analyze, predict and, especially, influence the interactions among several agents - or players - capable to form subcoalitions in order to pursue a common interest. It is crucial to have at our disposal an allocation rule that shares the payoff of the grand coalition, i.e. the coalition made up of all the players, among all the players themselves.

Introduced by Lloyd S. Shapley in its seminal paper [17], the Shapley value (Sv) is one of the most well known and useful allocation rules among the participants of a transferable utility cooperative game. The Sv always exists under a reasonable superadditive assumption and it is the only allocation fulfilling three reasonable conditions of symmetry, additiveness and dummy player compensation (see [17] for details). The significance of the Sv is evidenced by the broadness of its applications, spanning from pure economics [3] to internet economics [6, 14, 19], to politics [4], and to wireless communications [13]. Moreover, the concept of Sv was successfully applied to the case of weighted voting games, and referred to as Shapley-Shubik power index [18]. Such games typically imply that several agents possess resources and they need to collect a certain amount of resources to accomplish a task.

The computation of the Sv requires an exponential complexity in the number of players $P$, under oracle access to the characteristic function. Hence, when $P$ grows large, it becomes more and more crucial to find a suitable way to approximate it with a manageable number of queries. This problem was addressed for static games in [7], where the authors provided a confidence interval for two power indeces, i.e. the Banzhaf index and the Shapley-Shubik index.

In this work, we consider that the game is not played one-shot but over an infinite horizon: there exist several single stage games that come one after the other over time, following a discrete time homogeneous Markov chain process. We take into account two criteria to sum over time the allocations earned in each single stage game, specifically the average and the discount one. Our model is equivalent to the one in [16], except that we consider the utility to be transferable. We study in Section 4 the tradeoff complexity/accuracy of deterministic algorithms to compute the Sv in the Markovian game (SvM). We propose three different approaches to compute a confidence interval for SvM in Sections 5, 6.1, and 6.2. We propose in Section 6.1.1 a straightforward way to optimize the tightness of the second interval. We compare in Section 7 our three approaches in terms of tightness of the confidence interval. Finally, in Section 8 we provide a tradeoff complexity/accuracy for our randomized algorithm.

All the results found for SvM are also valid for the Shapley-Shubik index in Markovian weighted voting games. The extension of our results in simple games to the Banzhaf index [8] is straightforward.

## 2 Model and Background results

In this paper we consider Markovian cooperative games. We have a set of states $S = \{s_1, \dots, s_N\}$; in each state $s \in S$ the game $\Psi_s \equiv (\mathcal{P}, v)$ is played among $P$ players. Let $\mathcal{P} = \{1, \dots, P\}$ be the grand coalition of players. In state $s$ each coalition $\Lambda \subseteq \mathcal{P}$ can ensure for itself the value $v_s(\Lambda)$, that, under a TU (transferable utility) assumption, can be shared in any manner among the players. Let $\mathcal{V}_s(\Lambda)$ be the half-space of all feasible allocations for coalition $\Lambda$ in the TU game $\Psi_s$, i.e. the set of real $|\Lambda|$-tuple $\mathbf{a} \in \mathbb{R}^{|\Lambda|}$ such that $\sum_j \mathbf{a}_j \leq v_s(\Lambda)$. We suppose that the coalition values are superadditive, i.e.

$$v_s(\Lambda_1 \cup \Lambda_2) \geq v_s(\Lambda_1) + v_s(\Lambda_2), \qquad \forall\, \Lambda_1 \cap \Lambda_2 = \emptyset.$$

The succession of the states is a discrete time homogeneous Markov chain, whose transition probability matrix is $\mathbf{P}$. Let $\mathbf{x}(\Psi_s) \in \mathbb{R}^P$ be a payoff allocation among the players in the single stage game $\Psi_s$. Under the $\beta$-discounted criterion, where $\beta \in [0; 1)$, the discounted allocation in the Markovian dynamic game $\Gamma_{s_k}$, starting from state $s_k$, can be expressed as

$$\sum_{t=0}^{\infty} \beta^t \sum_{i=1}^{N} p_t(s_i|s_k)\, \mathbf{x}(\Psi_{s_i}) = \sum_{i=1}^{N} \boldsymbol{\nu}_i(s_k)\, \mathbf{x}(\Psi_{s_i})$$

where $p_t(s_i|s_k)$ is the probability for the process to be after $t$ steps in state $s_i$ when the initial state is $s_k$, and $\boldsymbol{\nu}(s_k)$ is the $k$-th row of the non negative matrix $(\mathbf{I} - \beta\mathbf{P})^{-1}$. Under the average criterion, if the transition probability matrix $\mathbf{P}$ is irreducible, then the allocation in the long run game $\Gamma_{s_k}$ can be written as

$$\lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \sum_{i=1}^{N} p_t(s_i|s_k)\, \mathbf{x}(\Psi_{s_i}) = \sum_{i=1}^{N} \boldsymbol{\pi}_i\, \mathbf{x}(\Psi_{s_i})$$

where $\boldsymbol{\pi}$ is the stationary distribution of the matrix $\mathbf{P}$. All the results in this paper are valid both for the discounted criterion and, provided that $\mathbf{P}$ is irreducible, for the average one.

We define $\mathcal{V}(\Lambda, \Gamma_s)$ as the set of feasible allocations in the long run game $\Gamma_s$ for coalition $\Lambda$, i.e. the Minkowski sum:

$$\mathcal{V}(\Lambda, \Gamma_s) \equiv \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\, \mathcal{V}_{s_i}(\Lambda).$$

where $\boldsymbol{\sigma}_i(s) \equiv \boldsymbol{\nu}_i(s)$ if the discounted criterion is adopted, and $\boldsymbol{\sigma}_i(s) \equiv \boldsymbol{\pi}_i$ in case of average criterion.

**Proposition 1** ( [5]). *$\mathcal{V}(\Lambda, \Gamma_s)$ is equivalent to the set $\mathcal{A}$ of real $\mathbb{R}^{|\Lambda|}$-tuples $\mathbf{a}$ such that $\sum_{i=1}^{|\Lambda|} \mathbf{a}_i \leq v(\Lambda, \Gamma_s)$, where $v(\Lambda, \Gamma_s) = \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\, v_{s_i}(\Lambda)$, for all $s \in S$, $\Lambda \subseteq \mathcal{P}$.*

Thanks to Proposition 1, we are legitimated to define $v(\Lambda, \Gamma_s)$ as the value of coalition $\Lambda \subseteq \mathcal{P}$ in the long run game $\Gamma_s$.

Let us now define the Sv [17].

**Definition 1.** *Let $\Delta = (P, v)$ be a TU cooperative game. The* Shapley value $\mathbf{Sv}(\Delta)$ *is a real $P$-tuple whose $j$-th component is the payoff allocation to player $j$:*

$$\mathbf{Sv}_j(\Delta) = \sum_{\Lambda \subseteq \mathcal{P}/\{j\}} \frac{|\Lambda|!(P - |\Lambda| - 1)!}{P!} \big[v(\Lambda \cup \{j\}, \Delta) - v(\Lambda, \Delta)\big].$$

In this paper we are mainly interested in providing confidence intervals for the SvM $\mathbf{Sv}(\Gamma_s)$, that can be expressed, thanks to Proposition 1 and to the standard linearity property of Sv, as

$$\mathbf{Sv}_j(\Gamma_s) = \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\, \mathbf{Sv}_j(\Psi_{s_i}), \quad \forall\, s \in S,\ 1 \leq j \leq P. \tag{1}$$

In the next sections we will exploit Hoeffding's inequality [12] to derive some confidence intervals for the Shapley value of Markovian games.

**Theorem 1** (Hoeffding's inequality)**.** *Let $A_1, \ldots, A_n$ be $n$ independent random variables, where $A_i \in [a_i, b_i]$. Then, for all $\delta \in (0;1)$, there exists $\epsilon(n, \delta) > 0$ such that*

$$\Pr\left(\sum_{i=1}^{n} A_i - E\Big[\sum_{i=1}^{n} A_i\Big] \geq n\epsilon\right) \leq 2\exp\left(-\frac{2n^2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

In this paper we will derive some results especially for simple and weighted voting games.

**Definition 2.** *A* simple Markovian game *is a Markovian cooperative game in which the coalition values associated to each single stage TU cooperative game are binary, i.e. can assume only the values 0 and 1.*

We say that player $i$ is critical for coalition $\Lambda \subseteq \mathcal{P} \backslash \{i\}$ in state $s$ if $v_s(\Lambda \cup \{i\}) - v_s(\Lambda) = 1$. Technically, weighted voting games are not TU cooperative game, since the value of a coalition does not represent a payoff to be shared among the players, but rather it indicates the effectiveness of a coalition at completing a specific task.

**Definition 3.** *A* weighted voting Markovian game *is a Markovian game in which each single stage game $\Psi_s$ is associated to a 3-tuple $(P, \mathcal{T}_s, v)$ in which $1, \ldots, P$ are the players, $\mathcal{T}_s$ is a task, and $v$ are binary coalition values, such that $v_s(\Lambda) = 1$ iff the coalition $\Lambda \subseteq \mathcal{P}$ can complete the task $\mathcal{T}_s$ in state $s \in S$.*

Even in the case of weighted voting games, we will still assume that the coalition values in the long run game $\Gamma$ possess the linearity property of Proposition 1. In other words, $v(\Lambda, \Gamma_s)$ can be interpreted as the expected effectiveness in the long run game $\Gamma_s$ of coalition $\Lambda$ at completing the different tasks associated to each single stage game.
The concept of Sv applied to weighted voting games is called Shapley-Shubik index, and its formulation is the same as in Definition 1.

# 3 Motivations

Games on Markov chains arise especially in telecommunications. In the literature, the channel model under which the channel coefficients follow a discrete Markov chain over time has been extensively studied (e.g. see [10, 11]). Moreover, it is well known that the static gaussian multiple access channel can be seen as a concave game, hence its Sv lies in the capacity region and is the centroid of its Core (e.g. see [15]). Hence, by considering the channel coefficients as constant throughout the whole duration of a codeword, the multiple access channel with Markovian channel coefficients is an example of our model.

For equation (1), the SvM $\mathbf{Sv}(\Gamma_s)$ can be distributed in the course of the game by assigning the single stage Shapley value $\mathbf{Sv}(\Psi_{s'})$ to the agents in each state $s' \in S$. In such a case, $\mathbf{Sv}(\Gamma_s)$ is the long run expected payoff allocation in the long run game starting from state $s$. Moreover, such allocation procedure is time consistent [5], i.e. the expected long run allocation is still $\mathbf{Sv}(\Gamma_{S_t})$ for any subgame starting from state $S_t$ at a later time step $t$. Therefore, it is useful to provide a confidence interval for both $\mathbf{Sv}(\Psi_s)$ and $\mathbf{Sv}(\Gamma_s)$. The former issue is already addressed in [7], while we study the latter. Also, Section 6.2 provides a bridge between the two.

# 4 Complexity of deterministic algorithms

Since the exact computation of the Sv - or, equivalently, of the Shapley-Shubik index - involves the calculation of the incremental asset brought by a player to each coalition, then its complexity is proportional to the number of such coalitions, i.e. $2^{P-1}$. We mean by "game instance" a particular coalition value distribution.

**Definition 4.** *Let us assume that the Sv for player $j$ in the game $\Delta$ is $\mathbf{Sv}_j(\Delta) = a$. We say that a deterministic algorithm has an* accuracy *of at least $d > 0$ whenever, for all $d' > d$, there exist no game instances for which the algorithm answers $\mathbf{Sv}_j(\Delta) \in a \pm d'$.*

A query consists in evaluating the marginal contribution of player $j$ with respect to a coalition. We will first show that an exponential number of queries is necessary in order to achieve a polynomial accuracy for any deterministic algorithm aiming to approximate the Shapley-Shubik index in the static case. This is an extension of Theorem 3 in [7] to the Shapley-Shubik index.

**Theorem 2.** *Any deterministic algorithm computing the Shapley value in the simple single stage game $\Psi_s$ requires $\Omega(2^P/\sqrt{P})$ samples to achieve an accuracy of at least $1/(2P)$, for all $s \in S$.*

*Proof.* We will prove that there exists a class $\mathcal{F}$ of game instances for which any deterministic algorithm computing $\mathbf{Sv}_j(\Psi_s)$ with accuracy of at least $1/(2P)$ must utilize $\Omega(2^P/\sqrt{P})$ queries. Similarly to [7], let us construct $\mathcal{F}$ when $P$ is odd. Let $\Lambda \subseteq \mathcal{P}\backslash\{j\}$. There exists a set $D_o$ of $\binom{P-1}{[P-1]/2}/2$ coalitions of cardinality $[P-1]/2$ such that player $j$ is critical only for $D_o$. In particular, for $|\Lambda| \leq [P-1]/2$, $v_s(\Lambda) = 0$; if $|\Lambda| = [P-1]/2$, then, if $\Lambda \in D_o$, $v_s(\Lambda \cup \{j\}) = 1$, otherwise $v_s(\Lambda \cup \{j\}) = 0$. The values of the remaining coalitions are 1 if and only if they contain a winning coalition among the ones constructed so far. The Sv for player $j$ is thus:

$$\mathbf{Sv}_j(\Psi_s) = \frac{([P-1]/2)!\,([P-1]/2)!}{2(P)!}\binom{P-1}{[P-1]/2} = \frac{1}{2P}$$

Hence, for any deterministic algorithm $\mathcal{A}_o$ employing a number of queries smaller than $\mu_o(P)$, where

$$\mu_o(P) = \frac{1}{2}\binom{P-1}{[P-1]/2},$$

there always exists an instance belonging to $\mathcal{F}$ for which $\mathcal{A}_o$ would answer $\mathbf{Sv}_j(\Psi_s) = 0$. By Stirling's approximation, we can say that $\mu_o(P) \in \Omega(2^P/\sqrt{P})$. Let us now construct the class $\mathcal{F}$ of instances when $P$ is even and $P > 2$. Let $D_e$ be a set of $\binom{P-2}{[P-2]/2}$ coalitions of cardinality $[P-2]/2$, belonging to $\mathcal{C}\backslash\{j\}$, such that player $j$ is critical only for $D_e$. Then,

$$\mathbf{Sv}_j(\Psi_s) = \frac{(P/2-1)!\,(P/2)!}{(P)!}\binom{P-2}{[P-2]/2} = \frac{1}{2[P-1]} > \frac{1}{2P}.$$

Similarly to before, for any deterministic algorithm $\mathcal{A}_e$ using a number of queries smaller than

$$\mu_e(P) = \binom{P-1}{[P-2]/2} - \binom{P-2}{[P-2]/2} = \frac{P-2}{P}\binom{P-2}{[P-2]/2}$$

there always exists an instance belonging to $\mathcal{F}$ for which $\mathcal{A}_e$ would answer $\mathbf{Sv}_j(\Psi_s) = 0$. By Stirling approximation, we can say that $\mu_e(P) \in \Omega(2^P/\sqrt{P})$. Hence, a number of samples $\mu \in \Omega(2^P/\sqrt{P})$ is needed to achieve an accuracy of at least $1/(2P)$. Hence, the thesis is proved. □

We are ready to derive the complexity of a deterministic algorithm computing SvM in a simple game, as a function of the number of players $P$.

**Corollary 1.** *There exists $c > 0$ such that any deterministic algorithm computing the Shapley value in the simple Markovian game $\Gamma_s$ requires $\Omega(2^P/\sqrt{P})$ samples to achieve an accuracy of at least $c/P$, for all $s \in S$.*

*Proof.* Any deterministic algorithm will employ a certain number of queries in each state $s$ in order to compute $\mathbf{Sv}_j(\Gamma_s) = \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\mathbf{Sv}_j(\Psi_{s_i})$. Let $I_0$ be the instance for which player $j$ is a dummy player in all the single stage games $\{\Psi_s\}_{s\in S}$, i.e. $\mathbf{Sv}_j(\Psi_s) = 0$ for all $s \in S$. Let $I_1$ be the instance such that $\mathbf{Sv}_j(\Psi_s) = 0$ for all $s$ except for $s_k$, for which $\sigma(s_k) \neq 0$, and such that the game $\Psi_{s_k}$ belongs to the class $\mathcal{F}$ of instances described in the proof of Theorem 2. Therefore,

$$\mathbf{Sv}_j(\Gamma_s) = \frac{\boldsymbol{\sigma}_k(s)}{2P}$$

in the case that $P$ is odd and

$$\mathbf{Sv}_j(\Gamma_s) = \frac{\boldsymbol{\sigma}_k(s)}{2[P-1]}$$

if $P$ is even. Hence, any deterministic algorithm needs $\Omega(2^P/\sqrt{P})$ queries in state $s_k$ to achieve an accuracy better than $\boldsymbol{\sigma}_k(s)/(2P)$. Set $c = \boldsymbol{\sigma}_k(s)/2$. Hence, the thesis is proved. □

The results of this section clearly discourage from computing exactly or even approximating SvM with a deterministic algorithm when the number of players $P$ is high. Motivated by this, in the next sections we will propose three methods to construct confidence intervals for Sv, whose complexity does not even depend on $P$.

# 5 Static approach

In this section we will propose our first approach to compute a confidence interval for SvM. We assume to have at our disposal beforehand the value of all coalitions in each single stage games.

**Assumption 1.** *The estimator agent has access to all the coalition values in each state:*

$$\{v_s(\Lambda),\ \forall\, \Lambda \subseteq \mathcal{P},\ s \in S\}$$

*at the same time.*

Let us first find a formulation of Sv which is suitable for our purpose. Let $X$ be the set of all the permutations of $\{1, \ldots, P\}$. Let $\mathcal{C}_\chi(j)$ be the coalition of all the players whose index precedes $j$ in the permutation $\chi \in X$, i.e.

$$\mathcal{C}_\chi(j) \equiv \{i : \chi(i) < \chi(j)\}. \tag{2}$$

We can write the Sv of the long run game $\Gamma_s$, either for the discount or for the average criterion, as

$$\begin{aligned}
\mathbf{Sv}_j(\Gamma_s) &= \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\,\mathbf{Sv}_j(\Psi_{s_i}) \\
&= \frac{1}{P!}\sum_{\chi\in X}\sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\big[v_{s_i}(\mathcal{C}_\chi(j)\cup\{j\}) - v_{s_i}(\mathcal{C}_\chi(j))\big] \\
&= E_\chi\Big[\sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\big[v_{s_i}(\mathcal{C}_\chi(j)\cup\{j\}) - v_{s_i}(\mathcal{C}_\chi(j))\big]\Big],
\end{aligned}$$

where $E_\chi$ is the expectation over all the permutations $\chi \in X$, each having the same probability $1/P!$.

We now propose the first algorithm to compute a confidence interval for $\mathbf{Sv}_j(\Gamma_s)$, for each player $j$ and initial state $s$. For each query $t = 1, \ldots, n$, let us select independently a random permutation $\chi_t$ of $\{1, \ldots, P\}$. Let us define $Z$ as the random (over $X$) variable

$$\begin{aligned}
Z &\equiv \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\big[v_{s_i}(\mathcal{C}_\chi(j)\cup\{j\}) - v_{s_i}(\mathcal{C}_\chi(j))\big] \\
&= v(\mathcal{C}_\chi(j)\cup\{j\},\Gamma_s) - v(\mathcal{C}_\chi(j),\Gamma_s)
\end{aligned} \tag{3}$$

and let $Z_t$ be the $t$-th realization of $Z$, over the permutation $\chi_t$. Thanks to Hoeffding's inequality, we can write that, for all $\delta \in (0; 1)$,

$$\Pr\left(\left|\frac{1}{n}\sum_{t=1}^{n} Z_t - \mathbf{Sv}_j(\Gamma_s)\right| \geq \epsilon\right) \leq\ 2\exp\left(-\frac{2n\epsilon^2(n,\delta)}{[\overline{y}-\underline{y}]^2}\right)$$

where

$$\begin{aligned}
\overline{y} &= \max_{\mathcal{C}\subseteq\mathcal{P}} \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\big[v_{s_i}(\mathcal{C}\cup\{j\}) - v_{s_i}(\mathcal{C})\big] \\
\underline{y} &= \min_{\mathcal{C}\subseteq\mathcal{P}} \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\big[v_{s_i}(\mathcal{C}\cup\{j\}) - v_{s_i}(\mathcal{C})\big]
\end{aligned}$$

Hence, we can propose our first confidence interval.

**Confidence interval 1 (SCI).** *Let $1 \leq j \leq P$, $s \in S$. With probability at least $1-\delta$, $\mathbf{Sv}_j(\Gamma_s)$ belongs to the confidence interval*

$$\left[\frac{1}{n}\sum_{t=1}^{n} Z_t - \epsilon(n,\delta)\ ;\ \frac{1}{n}\sum_{t=1}^{n} Z_t + \epsilon(n,\delta)\right],$$

*where*

$$\epsilon(n,\delta) = \sqrt{\frac{[\overline{y}-\underline{y}]^2\log(2/\delta)}{2n}} \tag{4}$$

*In the case of simple games,*

$$\Big[\overline{y}-\underline{y}\Big]^2 \leq \Big[\sum_{i=1}^{N}\boldsymbol{\sigma}_i(s)\Big]^2,$$

*hence (4) becomes*

$$\epsilon(n,\delta) = \sqrt{\frac{\Big[\sum_{i=1}^{N}\boldsymbol{\sigma}_i(s)\Big]^2\log(2/\delta)}{2n}}\,. \tag{5}$$

# 6 Dynamic approach

Not surprisingly, the confidence interval SCI is analogous to the one found in [7] for single shot games. Indeed, the intrinsic notion of dynamicity of the game is surpassed by Assumption 1, for which the estimator has global knowledge of all the coalition values.

In the following, we will we propose two methods to compute a confidence interval for SvM, for which Assumption 1 on global knowledge of coalition values is no longer necessary. Their conception naturally arises from the assumption that the estimator learns the coalition values in each single stage game while the Markov chain process unfolds.

**Assumption 2.** *The state in which the estimator agent finds itself at each time step follows the same Markov chain process of the Markovian game itself. Hence, at step $t$, the estimator agent has access only to the coalition values associated to the game $\Psi_{S_t}$ of the current state $S_t$.*

Of course, the following approaches can also be employed under Assumption 1, and we will show that, in this case, they outperform the confidence interval SCI in terms of tightness.

## 6.1 First dynamic method

Let $\chi \in X$ be, as before, a uniformly random permutation of $\{1,\ldots,P\}$. Let us define $Y^{(i)}$ as the random variable associated to state $s_i$:

$$Y^{(i)} \equiv \; v_{s_i}(\mathcal{C}_\chi(j)\cup\{j\}) - v_{s_i}(\mathcal{C}_\chi(j)) \tag{6}$$

Let us assume that $Y^{(i)}$ has been sampled $n_i$ times in state $s_i$, and $\sum_{i=1}^{N} n_i = n$. We can still exploit Hoeffding's inequality to say that, for all $\delta \in (0;1)$,

$$\Pr\left(\left|\sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i(s)}{n_i} \sum_{t=1}^{n_i} Y_t^{(i)} - \mathbf{Sv}_j(\Gamma_s)\right| \geq n\epsilon'(n,\delta)\right) \leq$$
$$2\exp\left(-\frac{2[n\epsilon'(n,\delta)]^2}{\sum_{i=1}^{N} \boldsymbol{\sigma}_i^2(s)[\overline{x}(i) - \underline{x}(i)]^2/n_i}\right)$$

where, for all $i = 1, \dots, N$,

$$\overline{x}(i) = \max_{\mathcal{C} \subseteq \mathcal{P}} v_{s_i}(\mathcal{C} \cup \{j\}) - v_{s_i}(\mathcal{C})$$
$$\underline{x}(i) = \min_{\mathcal{C} \subseteq \mathcal{P}} v_{s_i}(\mathcal{C} \cup \{j\}) - v_{s_i}(\mathcal{C})$$

Set $\widetilde{\epsilon}(n,\delta) = n\,\epsilon'(n,\delta)$. We are now ready to propose our second confidence interval for $\mathbf{Sv}_j(\Gamma_s)$.

**Confidence interval 2 (DCI1).** *Let $1 \leq j \leq P$, $s \in S$. With a probability of confidence of at least $1-\delta$, $\mathbf{Sv}_j(\Gamma_s)$ belongs to the confidence interval*

$$\left[\sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i(s)}{n_i} \sum_{t=1}^{n_i} Y_t^{(i)} - \widetilde{\epsilon}(n,\delta)\,;\, \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i(s)}{n_i} \sum_{t=1}^{n_i} Y_t^{(i)} + \widetilde{\epsilon}(n,\delta)\right],$$

*where*

$$\widetilde{\epsilon}(n,\delta) = \sqrt{\frac{\log(2/\delta)}{2} \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i^2(s)}{n_i} [\overline{x}(i) - \underline{x}(i)]^2} \tag{7}$$

*In the case of simple games, $\overline{x}(i) = 1$ and $\underline{x}(i) = 0$ for all $i = 1, \dots, N$, hence (7) becomes*

$$\widetilde{\epsilon}(n,\delta) = \sqrt{\frac{\log(2/\delta)}{2} \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i^2(s)}{n_i}}\,. \tag{8}$$

### 6.1.1 Optimum sampling strategy

It is interesting to investigate the optimum number of times the variable $Y^{(i)}$ should be sampled in each state $s_i$, in order to minimize the length of the confidence interval DCI1, keeping the confidence probability fixed. We notice that, by fixing $1-\delta$, we can find the optimal values for $n_1, \dots, n_N$ by setting up the following minimization problem over integers:

$$\begin{cases} \min\limits_{n_1,\dots,n_N} \sum_{i=1}^{N} \boldsymbol{\sigma}_i^2(s)[\overline{x}^2(i) - \underline{x}^2(i)]/n_i \\ \sum_{i=1}^{N} n_i = n, \qquad n_i \in \mathbb{N} \end{cases} \tag{9}$$

If the static Assumption 1 holds, then the computation of the optimum values $n_1^*, \dots, n_N^*$ in (9) is sufficient to optimize the tightness of the confidence interval DCI1, since the sampling is done offline. Otherwise, if Assumption 2 holds, the estimator does not know in advance the succession of states hit by the process, hence it is crucial to plan a sampling strategy of the variable $Y^{(i)}$ along the Markov chain. Of course, a possible strategy would be, when $n$ is fixed, to sample $n_i^*$ times the variable $Y^{(i)}$ only the first time the state $s_i$ is hit, until all the states are hit. Nevertheless, this approach is clearly not efficient, since in several time steps the estimator is forced to remain idle.
Hence, we now assess the performance of an efficient and straightforward sampling strategy, consisting in sampling $Y^{(i)}$, *each* time the state $s_i$ is hit, an equal number of times over all $i = 1, \dots, N$. Let us first show a useful classical result for Markov chains (e.g., see [2]). Let $\eta$ be the number of steps performed by the Markov chain $\{S_t,\ t \in [0; \eta - 1]\}$. Let $\eta_i$ be the number of visits to state $s_i$, i.e.

$$\eta_i = \sum_{t=0}^{\eta-1} \mathbb{I}(S_t = s_i).$$

**Theorem 3.** *Let $\{S_t,\ t \geq 1\}$ be an ergodic Markov chain. Let $\hat{\boldsymbol{\pi}}_i^{(\eta)} \equiv \eta_i/\eta$. Then, for any initial distribution and for all $i = 1, \dots, N$,*

$$\hat{\boldsymbol{\pi}}_i^{(\eta)} \stackrel{\eta\uparrow\infty}{\longrightarrow} \boldsymbol{\pi}_i \qquad \text{with probability } 1,$$

*where $\boldsymbol{\pi}$ is the stationary distribution of the Markov chain.*

We will now show under which conditions the straightforward and efficient sampling strategy described above is also optimal, under the average criterion.

**Theorem 4.** *Suppose that Assumption 2 holds and that the Markov chain of the Markovian game is ergodic. Fix the confidence probability $1 - \delta$. Under the average criterion and in the case of simple Markovian game, if in each state $s_i$ the estimator agent samples the random variable $Y^{(i)}$, defined in (6), an equal number of times, then*

$$\sqrt{n}\,\widetilde{\epsilon}(n, \delta) \stackrel{n\uparrow\infty}{\longrightarrow} \inf_{n \in \mathbb{N}} \min_{\substack{n_1, \dots, n_N: \\ \sum_i n_i = n}} \sqrt{n}\,\widetilde{\epsilon}(n, \delta) = \sqrt{\frac{\log(2/\delta)}{2}},$$

*where the equality occurs with probability 1.*

*Proof.* Let us consider the following constrained minimization problem over the reals:

$$\begin{cases} \min\limits_{\omega_1, \dots, \omega_N} \sum_{i=1}^N \boldsymbol{\sigma}_i^2(s)/\omega_i \\ \sum_{i=1}^N \omega_i = n, \qquad \omega_i \in \mathbb{R} \end{cases} \tag{10}$$

By using, e.g., the Lagrangian multiplier technique, it is easy to see that the optimum value for $\omega_i$ is

$$\omega_i^* = \frac{\boldsymbol{\sigma}_i(s)\, n}{\sum_{k=1}^{N} \boldsymbol{\sigma}_k(s)} \tag{11}$$

and that the minimum value of the objective function is

$$\xi^* = \frac{\left[\sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\right]^2}{n} \tag{12}$$

The value $\xi^*$ clearly represents a lower bound for the optimization problem over the integers in the case of simple games. Since we deal with the average criterion, let $\boldsymbol{\sigma}_i(s) \equiv \boldsymbol{\pi}_i$. We can find now a lower bound for $\sqrt{n}\,\widetilde{\epsilon}(n,\delta)$ over $n$ that does not depend on the number of samples $n$:

$$\begin{aligned}
&\inf_{n\in\mathbb{N}} \min_{\substack{n_1,\dots,n_N:\\ \sum_i n_i = n}} \sqrt{n}\,\widetilde{\epsilon}(n,\delta) = \\
&= \inf_{n\in\mathbb{N}} \min_{\substack{n_1,\dots,n_N\in\mathbb{N}:\\ \sum_i n_i = n}} \sqrt{\frac{n\,\log(2/\delta)}{2}\sum_{i=1}^{N}\frac{\boldsymbol{\pi}_i^2}{n_i}} && (13)\\
&= \inf_{\substack{q_1,\dots,q_N\in\mathbb{Q}^+:\\ \sum_i q_i = 1}} \sqrt{\frac{\log(2/\delta)}{2}\sum_{i=1}^{N}\frac{\boldsymbol{\pi}_i^2}{q_i}}\\
&= \min_{\substack{x_1,\dots,x_N\in\mathbb{R}^+:\\ \sum_i x_i = 1}} \sqrt{\frac{\log(2/\delta)}{2}\sum_{i=1}^{N}\frac{\boldsymbol{\pi}_i^2}{x_i}} && (14)\\
&= \sqrt{\frac{\log(2/\delta)}{2}}
\end{aligned}$$

and the optimum value of $x_i$ in (14) is

$$x_i^* = \frac{\boldsymbol{\pi}_i}{\sum_{k=1}^{N}\boldsymbol{\pi}_k} = \boldsymbol{\pi}_i\,.$$

For Theorem 3,

$$n_i/n \overset{n\uparrow\infty}{\longrightarrow} \boldsymbol{\pi}_i \qquad \text{with probability 1.}$$

Hence, $n_i/n$ converges with probability 1 to the optimum value $x_i^*$ and, by continuity, the thesis is proved. □

## 6.2 Second dynamic method

Since Hoeffding's inequality has a very general applicability and does not refer to any particular probability distribution of the random variables at issue, it is natural to look for confidence intervals suited to particular instances of games. Suppose now that we can compute the confidence intervals $[l_i; r_i]$ for the Svs of the single stage games $\mathbf{Sv}_j(\Psi_{s_i})$. In this section we will show a third confidence interval for the Sv of the long run game $\Gamma$ which is tighter *i*) the higher the confidence probability $1-\delta$ is and *ii*) the tighter the confidence intervals $[l_i; r_i]$ are. As an example, in section 6.2.1 we will show a tight confidence interval for simple Markovian games.

The confidence interval proposed in this section is based on the following Lemma.

**Lemma 1.** *Let* $A_1, \ldots, A_k$ *be* $k$ *random variables such that* $\Pr(A_i \in [l_i; r_i]) \geq 1-\delta_i$. *Let* $c_i \geq 0$, *for* $i = 1, \ldots, k$. *Then,*

$$\Pr\left(\sum_{i=1}^{k} c_i A_i \in \left[\sum_{i=1}^{k} c_i l_i \; ; \; \sum_{i=1}^{k} c_i r_i\right]\right) \geq \prod_{i=1}^{k}(1-\delta_i)$$

*Proof.* We will provide the proof for continuous random variables; the proof for the discrete case is totally similar. By induction, it is sufficient to prove that, if $\Pr(A_1 \in [l_1; r_1]) \geq 1-\delta_1$ and $\Pr(A_2 \in [l_2; r_2]) \geq 1-\delta_2$, then

$$\Pr\left(A_1 + A_2 \in [l_1 + l_2 \; ; \; r_1; r_2]\right) \geq \; (1-\delta_1)(1-\delta_2)$$

Let $f_A$ be the probability density function of the random variable $A$. Let $\overline{f}_{A_i}(x) = f_{A_i}(x)\mathbb{1}(x \in [l_i; r_i])$, $i = 1, 2$. Then,

$$\begin{aligned}
\Pr\big(A_1 + A_2 \in [l_1 + l_2; r_1; r_2]\big) = & \int_{l_1+l_2}^{r_1+r_2} f_{A_1+A_2}(x)dx \\
& = \int_{l_1+l_2}^{r_1+r_2} \int_{\mathbb{R}} f_{A_1}(x-\tau)\, f_{A_2}(\tau)\, d\tau\, dx \\
& \geq \int_{l_1+l_2}^{r_1+r_2} \int_{\mathbb{R}} \overline{f}_{A_1}(x-\tau)\, \overline{f}_{A_2}(\tau)\, d\tau\, dx \\
& = \int_{\mathbb{R}} \int_{\mathbb{R}} \overline{f}_{A_1}(x-\tau)\, \overline{f}_{A_2}(\tau)\, d\tau\, dx \\
& = \int_{\mathbb{R}} \overline{f}_{A_1}(x)\, dx \int_{\mathbb{R}} \overline{f}_{A_2}(x)\, dx \\
& = \Pr(A_1 \in [l_1; r_1]) \; \Pr(A_2 \in [l_2; r_2]) \\
& \geq (1-\delta_1)(1-\delta_2)
\end{aligned}$$

□

It is straightforward to see that the lower bound on the confidence probability $\prod_{i=1}^{k}(1-\delta_i)$ in Lemma 1 is tighter the smaller the single confidence levels $\delta_1, \ldots, \delta_k$ are. Lemma 1 suggests a new confidence interval for SvM.

We are now ready to show our second dynamic approach. Let the random variable $Y^{(i)}$ be defined as in (6). Also this method implies that $Y^{(i)}$ is sampled $n_i$ times in state $s_i$, for all $i = 1, \ldots, N$.

**Confidence interval 3 (DCI2).** *Set the confidence levels* $\delta_i \in (0;1)$*, for all* $i = 1,\dots,N$*. Let*

$$\left[ l^{(s_i)}\left(n_i, \sum_{t=1}^{n} Y_t^{(s_i)}, \delta_i\right) \; ; \; r^{(s_i)}\left(n_i, \sum_{t=1}^{n} Y_t^{(s_i)}, \delta_i\right)\right]$$

*be the confidence interval for* $\mathbf{Sv}(\Psi_{s_i})$*, for all* $i = 1,\dots,N$*. Let* $1 \le j \le P$*,* $s \in S$*. With a confidence probability* $\prod_{i=1}^{N}(1-\delta_i)$*,* $\mathbf{Sv}_j(\Gamma_s)$ *belongs to the confidence interval*

$$\left[ \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\, l^{(s_i)}\left(n_i, \sum_{t=1}^{n} Y_t^{(s_i)}, \delta_i\right) \; ; \; \sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\, r^{(s_i)}\left(n_i, \sum_{t=1}^{n} Y_t^{(s_i)}, \delta_i\right)\right] \tag{15}$$

It is interesting to note that, in the case of the confidence interval DCI1, we could optimize its tightness by modifying the number of samples $n_1,\dots,n_N$. Here, in addition, we can optimize DCI2 also over the set of confidence levels $\delta_1,\dots,\delta_N$, under the nonlinear constraint:

$$\prod_{i=1}^{N}(1-\delta_i) = 1-\delta\,.$$

#### 6.2.1 Confidence interval for single stage games: example for simple games

In this section we will show an example in which the employment of the confidence interval DCI2 is well justified. In [7], the authors derived a confidence interval for the Sv of a single stage game, based on Hoeffding's inequality. In the case of simple games, a tighter confidence interval can be obtained. Let $\chi \in X$ be a permutation of $\{1,\dots,P\}$. Let us assume that all $\chi \in X$ have the same probability $1/P!$. Let us define the Bernoulli variable $B(s)$ as

$$B(s) = v_s\big(\mathcal{C}_\chi(j) \cup \{j\}\big) - v_s\big(\mathcal{C}_\chi(j)\big),$$

where $\mathcal{C}_\chi(j)$ is defined as in (2). As pointed out in [7], we can interpret the Sv $\mathbf{Sv}_j(\Psi_s)$ as

$$\mathbf{Sv}_j(\Psi_s) = \Pr\left(B(s) = 1\right).$$

Let $B_1,\dots,B_n$ be $n$ independent realization of $B(s)$. It is evident that $\sum_{t=1}^{n} B_t \sim \mathcal{B}(n, \mathbf{Sv}_j(\Psi_s))$, where $\mathcal{B}(a,b)$ is the binomial distribution with parameters $a, b$. Hence, computing a confidence interval for $\mathbf{Sv}_j(\Psi_s)$ boils down to the computation of confidence intervals of the the probability of success of the Bernoulli variable $B(s)$ given the proportion of successes $\sum_{t=1}^{n} B_t/n$, which is a well know problem in literature. Of course, one might still utilize Hoeffding's inequality to do that, but over the last decades some more efficient methods have been proposed, like Wilson's score interval [21], the Wald interval [20], the adjusted Wald interval [1], and the "exact" Clopper-Pearson interval [9].

# 7 Comparison among the proposed approaches

If the dynamic Assumption 2 fails to hold, but instead we consider the static Assumption 1, then we are allowed to use any of the three methods presented in this paper, SCI, DCI1, and DCI2, to compute a confidence interval for the Sv. In fact, DCI1 and DCI2 involve independent queries over the different states, and this can be also done under Assumptions 1. Hence, in this case, we are allowed to compare the tightness of the two confidence intervals SCI and DCI1.

**Lemma 2.** *In the case of simple Markovian games, for any integer $n$ and for any confidence probability $1-\delta$,*

$$\epsilon(n,\delta) \leq \widetilde{\epsilon}(n,\delta).$$

*Proof.* In the case of simple Markovian games, the optimization problem (9) turns into

$$\begin{cases} \min\limits_{n_1,\ldots,n_N} \sum_{i=1}^{N} \boldsymbol{\sigma}_i^2(s)/n_i \\ \sum_{i=1}^{N} n_i = n, \qquad n_i \in \mathbb{N} \end{cases} \tag{16}$$

Let us consider the constrained minimization problem over the reals in (10). Since evidently $\xi^*$, defined in (12), is not greater than the mimimum value of the objective function in (16), then the thesis is proved by straighforward inspection over the expressions (5) and (8). □

The reader should not be misled by the result in Lemma 2. In fact, $n$ being equal in the two cases, the number of queries needed for confidence interval SCI is $N$ times bigger than for DCI1, since each sampling of the variable $Z$, defined in (3), requires the query of $N$ incremental coalition values, one per each state. The complexity in the two cases would be the same only if the estimator agent had access to the coalition values of the long run game $\{v(\Lambda,\Gamma_s)\}_{s,\Lambda}$.
Hence, it is reasonable to compare the length of the confidence interval for the static case $2\,\epsilon(n,\delta)$ with the one for the dynamic case, $2\,\widetilde{\epsilon}(Nn,\delta)$, calculated with $N$ times many queries. It is interesting to notice that, in this case, the relation between the tightness of SCI and DCI can be reversed.

**Lemma 3.** *In the case of simple Markovian games, for all the integer $n$,*

$$\min_{\substack{n'_1,\ldots,n'_N: \\ \sum_i n'_i = Nn}} \widetilde{\epsilon}(Nn,\delta) \leq \epsilon(n,\delta)$$

*Proof.* We can write

$$\begin{aligned} \min_{\substack{n'_1,\ldots,n'_N: \\ \sum_i n'_i = Nn}} \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i^2(s)}{n'_i} \leq \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i^2(s)}{\sum_{j=1}^{N} n'_j/N} = \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i^2(s)}{n} \\ \leq \frac{\left[\sum_{i=1}^{N} \boldsymbol{\sigma}_i(s)\right]^2}{n} \end{aligned}$$

where the last inequality holds since $\boldsymbol{\sigma}_i(s) \geq 0$. Hence, by inspection over the expressions (5) and (8), the thesis is proved. $\square$

Lemmas 2 and 3 clarify the relation between the confidence intervals SCI and DCI1, under the condition of simple Markovian games. Indeed, the dynamic approach is better than the static one in terms of accuracy of the confidence interval, when the number of queries is equal for the two methods. Moreover, simulations showed that, when the number of samples $n$ and the confidence level $\delta$ are equal for the two methods, then the *effective* confidence probability for SCI is generally higher than for DCI1, i.e. the lower bound $1-\delta$ is less tight. We explain this by reminding that the centres of the confidence intervals SCI and DC1, respectively

$$\frac{1}{n}\sum_{t=1}^{n} Z_t \quad ; \quad \sum_{i=1}^{N} \frac{\boldsymbol{\sigma}_i(s)}{n_i} \sum_{t=1}^{n_i} X_t^{(i)}$$

are already two estimators for $\mathbf{Sv}(\Gamma_s)$, and the former possesses a smaller variance than the second one.

About the performances of confidence interval DCI2, the simulations confirmed our intuitions. We utilized both the Hoeffding inequality and the Clopper-Pearson interval to compute a confidence interval for the Sv of simple single stage games, and we saw that the confidence interval is more and more tight when the confidence probability approaches 1 and the Clopper-Pearson method gives generates intervals.

In Table 7 we show, for each value of confidence probability $1-\delta$, the percentage $a_{3\succ 2}$ of cases in which the confidence iterval DCI2 is narrower than confidence interval DCI1. We see that, for $1-\delta < 0.8$, the two confidence interval have a comparable length. For $1-\delta \geq 0.8$, the confidence iterval DCI2 is to be strongly preferred to DCI1, in terms of accuracy, under these settings.

| $1-\delta$ | $a_{3\succ 2}$ (%) |
|---|---|
| .97 | 100 |
| .95 | 99.9 |
| .9 | 87.5 |
| .8 | 57.7 |

# 8 Complexity of confidence intervals

In Section 4 we called for a randomized algorithm who could reach a polynomial accuracy in the number of players $P$ without the need of an exponential number of queries. In this case, by "accuracy" we mean the length of confidence interval, keeping fixed the confidence probability. In the previous sections we derived three confidence intervals, SCI, DCI1, and DCI2, whose expression do not depend on the number of players $P$.

**Proposition 2.** *Fix the confidence level $\delta$ and the length of confidence interval $2\,\epsilon$. Then $n$ queries are required for the confidence interval SCI, where*

$$n = \frac{\left[\overline{y} - \underline{y}\right]^2 \log(2/\delta)}{2\,\epsilon^2}.$$

*Proof.* The proof follows straightforward from the expression of confidence interval SCI. □

**Proposition 3.** *Fix the confidence level $\delta$ and the length of confidence interval $2\,\widetilde{\epsilon}$. Then, there exists a set of $n_1, \ldots, n_N$, $\sum_i n_i = n$, such that $n$ queries are required for the confidence interval DCI1, where*

$$n \leq \frac{N\left[\overline{y} - \underline{y}\right]^2 \log(2/\delta)}{2\,\widetilde{\epsilon}^2}$$

*Proof.* The proof follows straightforward from Lemma 3. □

From Propositions 2 and 3 the following result follows.

**Theorem 5.** *The number of required queries to achieve an accuracy of $1/p(P)$, where $p(P)$ is a polynomial of $P$, is $O(p^2(P))$, for both the confidence interval SCI and DCI1.*

Since we do not provide an explicit expression for the confidence interval DCI2, then we can not provide an analogous result for DCI2. Nevertheless, we notice that its expression does not depend on the number of players $P$, and that if the Hoeffding's inequality is used to compute the confidence interval for the Sv in the single stage games, then a similar result of Theorem 5 can be derived.

Corollary 1 and Theorem 5 confirm that our randomized approaches are better than any deterministic approach for a number of players sufficiently high.

*Remark*: All the results in this paper for simple Markovian games are also valid for the Shapley-Shubik index in the case of weighted voting Markovian games. In fact, for our purposes, the Shapley-Shubik index and the Sv are totally equivalent, since they both possess the linearity property (1).

# 9 Conclusions

We proved in Section 4 that an exponential number of queries is necessary for any deterministic algorithm even to approximate SvM with polynomial accuracy. Hence, we focused on randomized algorithms and we proposed three different methods to compute a confidence interval for SvM. The first, described in Section 5 and called SCI, is a static one, since it assumes that we have at our disposal the values in each state at the same time. The remaining two methods, DCI1 in Sections

6.1 and DCI2 in Section 6.2, are also valid if we assume that the estimator learns the coalition values in each single stage game while the Markov chain process unfolds. We focused then on simple Markovian games. We proposed in Section 6.1.1 a straightforward way to optimize the tightness of DCI1. We compare in Section 7 our three approaches in terms of tightness of the confidence interval. The simulations confirmed that DCI2 is better than the first two when both the confidence probability is close to 1 and a tight confidence interval for the Sv of the single stage games is available. We prove that DCI1 is tighter than SCI, with an equal number of queries. Hence, the dynamic approach is better than the static one, although it relies on milder assumptions. This is essentially because it allows to tune the number of samples according to the weight of the state. Finally, in Section 8 we show that a polynomial number of queries is sufficient to achieve a polynomial accuracy for our algorithms. Hence, in order to compute SvM, our randomized approaches are better than any deterministic approach for a number of players sufficiently high.

## 10 Acknowledgments

This research was supported by "Agence Nationale de la Recherche", with reference ANR-09-VERS-001, and by the European research project SAPHYRE, which is partly funded by the European Union under its FP7 ICT Objective 1.1 - The Network of the Future.

## References

[1] A. Agresti and B. Coull. Approximate is better than "exact" for interval estimation of binomial proportions. *American Statistician*, pages 119–126, 1998.

[2] K. B. Athreya and C. D. Fuh. Bootstrapping Markov chains: countable case. *Journal of Statistical Planning and Inference*, 33:311–331, 1992.

[3] R. J. Aumann. Economic Applications of the Shapley Value. In S. S. J.-F. Mertens, editor, *Game-Theoretic Methods in General Equilibrium Analysis*, pages 121–133. Kluwer Academic Publisher, 1994.

[4] R. J. Aumann and S. Hart. *Handbook of Game Theory with Economic Applications*, volume 2. Elsevier, 1994.

[5] K. Avrachenkov, L. Cottatellucci, and L. Maggi. Cooperative Markov Decision Processes: Time Consistency, Greedy Players Satisfaction, and Cooperation Maintenance. Technical Report RR-11-248, EURECOM, 2010. submitted.

[6] K. Avrachenkov, J. Elias, F. Martignon, G. Neglia, and L. Petrosyan. A Nash bargaining solution for Cooperative Network Formation Games. *NETWORKING 2011*, pages 307–318, 2011.

[7] Y. Bachrach, E. Markakis, E. Resnick, A. D. Procaccia, J. S. Rosenschein, and A. Saberi. Approximating power indices: theoretical and empirical analysis. *Autonomous Agents and Multi-Agent Systems*, 20:105–122, March 2010.

[8] J. F. Banzhaf III. Weighted voting doesn't work: A mathematical analysis. *Rutgers L. Rev.*, 19:317–343, 1964.

[9] C. J. Clopper and E. S. Pearson. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, 26(4):404–413, 1934.

[10] A. J. Goldsmith and P. P. Varaiya. Capacity, mutual information, and coding for finite-state Markov channels. *Information Theory, IEEE Transactions on*, 42(3):868–886, 1996.

[11] S. Hanly and D. Tse. Multiaccess fading channels. II. delay-limited capacities. *Information Theory, IEEE Transactions on*, 44(7):2816–2831, 1998.

[12] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.

[13] F. Javadi, M. R. Kibria, and A. Jamalipour. Bilateral Shapley Value Based Cooperative Gateway Selection in Congested Wireless Mesh Networks. In *IEEE GLOBECOM 2008*, pages 1–5. IEEE, 2008.

[14] R. T. B. Ma, D. Chiu, J. Lui, V. Misra, and D. Rubenstein. Internet Economics: The use of Shapley value for ISP settlement. In *Proceedings of CoNEXT 2007*, Columbia University, New York, December 2007.

[15] M. Madiman. Playing games: A fresh look at rate and capacity regions. In *Information Theory, 2008. ISIT 2008. IEEE International Symposium on*, pages 2528 –2532, july 2008.

[16] A. Predtetchinski. The strong sequential core for stationary cooperative games. *Games and economic behavior*, 61(1):50–66, 2007.

[17] L. S. Shapley. A value for n-person games. *Contributions to the Theory of Games*, 2:31–40, 1953.

[18] L. S. Shapley and M. Shubik. A method for evaluating the distribution of power in a committee system. *The American Political Science Review*, 48(3):787–792, 1954.

[19] R. Stanojevic, N. Laoutaris, and P. Rodriguez. On economic heavy hitters: Shapley value analysis of 95th-percentile pricing. In *Proceedings of the 10th annual conference on Internet measurement*, pages 75–80. ACM, 2010.

[20] A. Wald and J. Wolfowitz. Confidence limits for continuous distribution functions. *Annals of Mathematical Statistics*, 10:105–118, 1939.

[21] E. B. Wilson. Probable inference, the law of succession, and statistical inference. *JASA*, pages 209–212, 1927.