K-SNCC: Group Deviations in Subsidized Non-Cooperative Computing

Andrey Klinger
Computer Science Department
Technion
andreyklinger@gmail.com

Moshe Tennenholtz
Faculty of Industrial Eng. & Mgmt, Technion
&
Microsoft Israel R&D Center
moshet@microsoft.com

ABSTRACT

A function is subsidized non-cooperative computable [SNCC] if honest agents can compute it by reporting truthfully their private inputs, while unilateral deviations by the players are not beneficial: if a deviation from truth revelation can mislead other agents, this deviation will decrease the deviator's chances of correct computation, or, it will not affect these chances but the expected payment to the deviator will decrease; in addition, deviations can not increase the expected monetary payments to a deviator without decreasing his chances of correct computation. This paper extends the study of SNCC functions to the context of group deviations. A function is K-SNCC if deviations by a group of at most K agents are not beneficial. We provide a full characterization of the K-SNCC functions, both for the independent values and the correlated values settings.

1. INTRODUCTION

There has been much interest in the recent years in the idea of joint activity in computational settings. This topic has been facilitated by the emergence of social communities, in which joint activity by self-motivated agents is coordinated through a social platform. If the aim of participants in that setting is to aggregate privately available information into some meaningful output, captured as a function of that information, we arrive at the *non-cooperative computing* [NCC] setting [**?**]. NCC deals with the desire to compute a function defined on agents' private inputs where the agents might have incentives not to report truthfully. This can be viewed as a task of *informational mechanism design.* While in a classical mechanism design context (see [3] Chapter 23) the essence of the problem is the lack of information about the agents' preferences, in NCC the agents' preferences are known but other information they possess which is needed for the joint activity is private. NCC introduces a game-theoretic version of the problem of multi-party computation.[1]

[1]Indeed, the work in [2, **?**] deals with NCC when there is no center in the system, bridging the gap to the classical assumptions in the cryptographic and distributed computing literature

Cite as: Title, Author(s), *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX-XXX.

In order to see the basic idea behind NCC consider for example the situation where each agent's secret is a bit, and the function to be computed is the parity function. If all agents report their bits honestly then the parity can be easily computed. However, if an agent reports 1 (resp. 0) instead of 0 (resp. 1), while all other agents report honestly, then this agent will be able to recover the true result by reversing the reported outcome, while misleading the other agents. Hence, the parity function is **not** non-cooperatively computable. On the other hand, if the function is the majority function, then false report might make the deviator unclear about the true result, given that the result of the majority function is computed and reported to the participants using a trusted center based on the information provided by them; this makes this function non-cooperatively computable.

The early results on NCC provided complete characterization of the functions which are non-cooperatively computable. Additional work has been carried out on extending this setting [4], as well as on considering the agents' costs, which lead to other forms of deviations [**?**]. Most recently, the challenging topic of deviations by coalitions in the NCC setting has been considered [**?**]. While NCC is associated with honest computation being in equilibrium, group deviations in that context can be associated with the concept of strong equilibrium as introduced by Aumann [1]; therefore, a function is strong-NCC if no coalition can mislead in some cases at least one member which is not part of the coalition, without taking the risk this would cause at least one member of the coalition not to know the function value. More generally, in [**?**] the authors provide a characterization of the K-NCC functions, in which deviations of coalitions of size at most K are considered. The case of NCC is then associated with 1-NCC functions.

The NCC setting ignores the idea that the center (e.g. a social platform moderator, or any other organizer) can attempt and lead agents to desired behavior by introducing (even small) monetary incentives. For example, if the participants aim at computing the AND function, an agent holding the value *false* may be tempted to report *true*, potentially misleading the other agents, without taking any risk. If the center will offer some monetary payment to agents who report *false* then this problem may be overcome: an agent who has the value *false* will have an incentive to report truthfully, while an agent who has the value *true* will have to risk knowing the true value by reporting another value. This issue is captured in [**?**] using the notion of *subsidized non-cooperative computing [SNCC].* In SNCC, an agent's first desire is the correct computation of the function. However, if the function is computed correctly, the agent will prefer to maximize his payment; in particular, according to SNCC the agent will prefer a positive payment to the opportunity of misleading other agents. Only when deviation will not decrease chances for computing the correct answer, nor decrease expected payments, the agent may be tempted to mislead others. The SNCC concept is very realistic. The main concern is that in general one may expect large payments to become more important than correctness; however, since in all results so far on SNCC, payments could have been made as small as one wishes, this concern did not have a bite.

In this paper we take the challenge of considering group deviations in the SNCC setting. Namely, we wish to initiate the study of K-SNCC functions. Roughly speaking, a function is K-SNCC, if it is stable against deviations by coalitions of size up to K, where each agent utility is modelled as in the SNCC setting. The objective is to provide a full characterization of the functions that are K-SNCC; namely, we wish to find a set of necessary and sufficient conditions on a function to be K-SNCC. Needless to say, such characterization should refer only to the function in question, and should not refer to payment policies.

As it turns out, the above characterization is non-trivial. In order to provide such a characterization we consider a Bayesian model, where the agents' values (inputs) are drawn from a commonly known distribution. We first consider the case of independent values, for which we provide full characterization of the K-SNCC functions. Then we deal with the general case of correlated values. We provide a full characterization for this general situation, under the assumption that all agents are relevant to the computation.

In the next section we present the basic concepts and definitions. In Section 3.1 we provide several results, leading to complete characterization of the K-SNCC functions in the independent values case. In Section 3.2 we deal with the correlated values case. This requires additional definitions and several more technical results.

2. DEFINITIONS

Given a set of agents $N = \{1, 2, \ldots, n\}$, and a special agent termed "the center", we assume that there exists a private secure communication channel between every agent $i \in N$ and the center. The type v_i of agent i is selected from some domain B_i. We concentrate on a Boolean domain, where $B_i = B = \{0, 1\}$. The types of the agents are drawn from a joint probability distribution p over B^n.

Given a function $w : B^n \to B$, we consider the following protocol:

1. For any instantiated type vector $v \in B^n$, each agent i declares his type $\hat{v_i}$ to the center (truthfully or not; $\hat{v_i} = v_i$ may or may not hold).
2. The center computes the value $w(\hat{v}) = w(\hat{v_1}, \ldots, \hat{v_n})$ and announces it to all agents.
3. The center pays to each agent i an amount $m_i(\hat{v})$
4. Each agent i determines $w(v)$ based on $w(\hat{v})$, $m_i(\hat{v})$ and v_i (his true input).

The protocol above defines a strategy space for each agent. A strategy for agent i is a pair of functions (f_i, g_i). $f_i : B \to B$, the *declaration function*, determines the input declared to the center based on the agent's true input. The *truthful* declaration function is the identity function $f^t(v) = v$. $g_i : B \times \Re^+ \times B \to B$, the *interpretation function*, is used by the agent to decide on the value of the function based on the announcement by the center, his payment and his true input. The *trusting* interpretation function is the projection function $g^t(r, \tau, v) = r$ in which the agent simply accepts the value announced by the center. We will refer to the strategy (f^t, g^t) as the *straightforward* strategy.

Note that the strategy profile consisting only of straightforward strategies results in each agent computing w correctly for all input vectors. We are looking for functions for which such a strategy profile forms an equilibrium, and more generally a (k-)strong equilibrium which is stable against deviations of coalitions (of size at most k).

The payment functions of the center are probabilistic, meaning that $m_i(\hat{v})$ is not the actual payment agent i gets, but merely an expected value of the payment. The exact payments can be 0 and 1 (or any other constants) with corresponding probabilities. This is done so the payment doesn't automatically reveal to the agent the value of $m_i(\hat{v})$. This allows our proofs to ignore the additional information provided by the payment functions.

We will use the following notations:

Definition 1. For a set of agents $C = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$, B_C is defined as $\prod_{j \in C} B_j$ and B_{-C} is defined as $\prod_{j \notin C} B_j$. In the same way, $v_C \in B_C$ is a tuple of types of agents participating in the set C, and $v_{-C} \in B_{-C}$ is a tuple of types of agents not participating in the set C.

Definition 2. The declared types of the agents are denoted by $\hat{v}$, i.e. $\hat{v} = (\hat{v_1}, \ldots, \hat{v_n}) = (f_1(v_1), \ldots, f_n(v_n))$

Definition 3. An agent i is termed *independent* if $p(v_i | x_{-\{i\}}) = p(v_i | y_{-\{i\}})$ for every $v_i \in B_i$, $x_{-\{i\}}, y_{-\{i\}} \in B_{-\{i\}}$.

A probability distribution over B_N is termed *independent* if $\forall i$, $1 \le i \le n$, $p(v_i | x_{-\{i\}}) = p(v_i | y_{-\{i\}})$ for every $v_i \in B_i$, $x_{-\{i\}}, y_{-\{i\}} \in B_{-\{i\}}$, i.e. if all the agents are independent.

Definition 4. Tuples of types for agents belonging to different sets will be denoted by a comma separated list. For example, for $A = \{i_{a_1}, \ldots i_{a_k}\}$, $B = \{i_{b_1}, \ldots i_{b_l}\}$, $C = N \setminus (A \cup B)$, the notation for the types of the agents would be (v_A, v_B, v_C) or $(v_A, v_B, v_{-(A \cup B)})$

The notation for $(1 - v_{a_1}, \ldots, 1 - v_{a_k}, v_{b_1}, \ldots, v_{b_l}, v_C)$ would be $(\overline{v}_{a_1}, \ldots, \overline{v}_{a_k}, v_{b_1}, \ldots, v_{b_l}, v_C)$ or $(\overline{v_A}, v_B, v_C)$.

Definition 5. An agent i is termed *balanced* if $p(v_i) = p(1 - v_i) = 1/2$ for every $v_i \in B_i$

Definition 6. A function w is called dominated if the following holds: $\exists i \in \{1, \ldots, n\}, v_i \in B$, such that $\forall x_{-\{i\}} \in B_{-\{i\}}, y_{-\{i\}} \in B_{-\{i\}}$, $w(v_i, x_{-\{i\}}) = w(v_i, y_{-\{i\}})$, and there exists some $z_{-\{i\}} \in B_{-\{i\}}$, for which $w(1 - v_i, z_{-\{i\}}) = 1 - w(v_i, z_{-\{i\}})$.

A function w is dominated by agent i, if when agent i has the type v_i then he already knows the value of w, but w is not a constant function.

Definition 7. A function w is termed k-reversible if the following holds: $\exists C = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ such that $\forall v_{-C} \in B_{-C}, \forall j, 1 \le j \le k, \forall v_{i_j} \in B_{i_j}$,

$$w(v_{i_1}, \ldots, v_{i_k}, v_{-C}) = 1 - w(1 - v_{i_1}, \ldots, 1 - v_{i_k}, v_{-C})$$

We will say that w is k-reversible *by* such C. When a function w is k-reversible by a coalition C, it means that when all the agents in C reverse their declarations, the value of w reverses. Note that in this definition, the set C contains exactly k agents.

Definition 8. A function w is termed k-irreversible if the following holds: $\exists C = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ such that $\forall v_{-C} \in B_{-C}, \forall j, 1 \leq j \leq k, \forall v_{i_j} \in B_{i_j}$,

$$w(v_{i_1}, \ldots, v_{i_k}, v_{-C}) = w(1 - v_{i_1}, \ldots, 1 - v_{i_k}, v_{-C})$$

We will say that w is k-irreversible *by* such C. When a function w is k-irreversible by a coalition C, it means that when all the agents in C reverse their declarations, the value of w doesn't change. Note that reversibility and irreversibility are not complementary properties: for a coalition C of agents that reverse their declaration, the function w can sometime keep its value and sometime reverse it depending on the types of agents not in C.

Definition 9. An agent i is irrelevant to the function w if $\forall v_{-\{i\}} \in B_{-\{i\}}, w(0, v_{-\{i\}}) = w(1, v_{-\{i\}})$

We will call such agent an *irrelevant* agent. All the agents who are not irrelevant will be called *relevant* agents. The set of irrelevant agents will be denoted by I.

Definition 10. Let w be an n-ary Boolean function. Then $\widetilde{w_r^k}$ is the set of all possible k-reversible coalitions, i.e $\widetilde{w_r^k} = \{C | w(v_C, v_{-C}) = 1 - w(\overline{v}_C, v_{-C}), |C| = k, \forall v_C \in B_C, v_{-C} \in B_{-C}\}$

Let w be an n-ary Boolean function. Then $\widetilde{w_i^k}$ is the set of all possible k-irreversible coalitions, i.e $\widetilde{w_i^k} = \{C | w(v_C, v_{-C}) = w(\overline{v}_C, v_{-C}), |C| = k, \forall v_C \in B_C, v_{-C} \in B_{-C}\}$

We will also use the following notations:
$\widetilde{w^k} = \widetilde{w_i^k} \bigcup \widetilde{w_r^k}$
$\mathring{w}_i^k = \bigcup_{1 \leq j \leq k} \widetilde{w_i^j}$
$\mathring{w}_r^k = \bigcup_{1 \leq j \leq k} \widetilde{w_r^j}$
$\mathring{w}^k = \mathring{w}_i^k \bigcup \mathring{w}_r^k$

K-SNCC

Let C be a deviating coalition, whose members $i_j \in C$ are playing the tuples of strategies (f_{i_j}, g_{i_j}), and the rest of the agents are playing the straightforward strategy. The expected payoff of agent i when coalition C is declaring f_C is: $Em_i^{(C,f_C)} = \sum_{v_C \in B_C, v_{-C} \in B_{-C}} p(v_C, v_{-C}) m_i(f_C(v_C), v_{-C})$. With those, K-SNCC is defined as follows.

Definition 11. Let N, p and w be as above. Let $\epsilon > 0$. Then w is ϵ-K-SNCC, (i.e. non-cooperatively computable with subsidy ϵ while stable against deviations by coalitions of size no more than K) if there exist payment functions m_i for each $i \in N$, such that the following conditions hold:

- For every $i \in N$, $v_N \in B_N$; $0 \leq m_i(v_N) \leq \epsilon$
- For every coalition of deviating agents C, $|C| \leq K$ there exists an agent $i \in C$, such that one of the following is true:
  - $p(g_i(w(\hat{v}), m_i(\hat{v}), v_i) \neq w(v)) > 0$
  - $p(g_i(w(\hat{v}), m_i(\hat{v}), v_i) \neq w(v)) = 0$ and $Em_i^{(C,f_C)} < Em_i^{(C,f^t)}$
  - $p(g_i(w(\hat{v}), m_i(\hat{v}), v_i) \neq w(v)) = 0$ and $Em_i^{(C,f_C)} = Em_i^{(C,f^t)}$ and $p(w(v) \neq w(\hat{v})) = 0$

In other words, any deviation would either increase the possibility of the deviating agent's mistake, or it won't change the mistake probability but would decrease the expected amount of that agent's payoff, or both correctness and payoff would stay the same, but no other agent is misled. A function will be termed K-SNCC if it is ϵ-K-SNCC for any $\epsilon > 0$.

All the payments proposed in this paper can be scaled to be arbitrary small. Therefore, for the ease of exposition the requirement that the agents are paid no more than ϵ will be discarded throughout this paper. If one wishes to build a mechanism with payment functions, all the functions proposed can be downscaled to include only payments of ϵ at most.

3. SNCC CHARACTERIZATION

3.1 The independent values setting

In this subsection we provide a full characterization of the K-SNCC functions when all agents' values are independent.

Theorems 1 and 2 below, provide conditions that should be satisfied by any beneficial group deviation.

THEOREM 1. *In every worthwhile minimal deviating coalition C of size $K \geq 2$ all relevant agents employ a reverse declaration strategy.*

PROOF. A deviation can be worthwhile by either of the utility components. Since in the straightforward strategy all the agents are always correct, the deviators can not benefit on the correctness component. Suppose the coalition C has a deviation strategy that has a higher expected payment or is misleading non-deviating agents while still being always correct. Let us analyze the possible declaration strategies of a relevant agent $i \in C$. If he is declaring truthfully then by deviating from trusting interpretation he doesn't affect the utilities of others. Thus he could be ejected from the coalition C in contradiction to it being minimal.

If i uses a constant declaration function then other coalition members would sometimes be wrong: Suppose wlog that i always declares $f_i(v) = 0$. There exists $v_{-\{i\}} \in B_{-\{i\}}$ such that $r_0 = w(0, v_{-\{i\}}) \neq w(1, v_{-\{i\}}) = r_1$, because otherwise i would be irrelevant. No interpretation function of an agent $j \in C \setminus \{i\}$ would make him always correct for the types $v_{-\{i\}}$, since the center's announcement will always be r_0 and the real value of w can be both r_0 and r_1. Thus constant declaration functions are not feasible for a relevant agent in a worthwhile coalition C, $|C| \geq 2$.

The only possible declaration function left for agent i is always to declare the opposite of his real type: $\forall i \in C, v_i \in B_i$ it must be that $f_i(v_i) = \overline{v_i}$. □

THEOREM 2. *If a payoff of each agent is dependent only on his own declaration then a worthwhile minimal deviating coalition C of size $K \geq 2$ satisfies $C \in \widetilde{w^K}$, i.e. w is either K-reversible or K-irreversible.*

PROOF. According to Theorem 1, the declaration strategies of all the relevant agents of the deviating coalition C

are always to declare the opposite type. Suppose $C \notin \widetilde{w^K}$. Then since w is not K-reversible by C,

$$\exists v_C \in B_C, v_{-C} \in B_{-C}\ ,\ w(v_C, v_{-C}) = w(\overline{v}_C, v_{-C}) = r \quad (1)$$

And since w is not K-irreversible by C,

$$\exists v'_C \in B_C, v'_{-C} \in B_{-C}\ ,\ r = w(v'_C, v'_{-C}) \neq w(\overline{v'}_C, v'_{-C}) = \overline{r} \quad (2)$$

Consider a possible interpretation function g_i of an agent $i \in C$, $g_i(r, \tau, v)$ for the specific value of r given in the equations above. Equation 1 states that when the coalition members have types v_C and they declare $f_i(v_i) = \overline{v}_i$, and the center announces the correct value of r, the interpretation function $g_i(r, \tau, v_i)$ must be r (trusting) for the specific values of r and v_i. Although only the relevant agents must declare reverse types, irrelevant agent don't affect the value of w and thus the equation holds. Equation 1 also states that when the coalition members have types $\overline{v}_C$ and they declare $f_i(\overline{v}_i) = v_i$, and the center announces the correct value of r, the interpretation function $g_i(r, \tau, \overline{v}_i)$ must be r (trusting) for the specific values of r and $\overline{v}_i$. Thus agent i must trust the center's announcement r regardless of his own type.

On the other hand, equation 2 states that for types $\overline{v'}_C$ the interpretation must be $\overline{r}$ because the declarations would be v'_C and the center would announce r while the true value is $\overline{r}$. The agent can not distinguish between these cases by a value of τ (the payoff) because it is dependent only on his own declaration. From this contradiction we can conclude that under no interpretation strategy members of the deviating coalition C would be always correct. Hence, since correctness is the main utility component, $C \notin \widetilde{w^K}$ is a contradiction to C being a worthwhile minimal deviation. □

Using Theorems 1 and 2, we provide the desired characterization. This is done in Theorems 3–5. Theorem 3 recalls the result from [?] about 1-SNCC functions. Theorems 4–5 provide the extension to the general case of group deviations, dealing with general K-SNCC functions.

THEOREM 3. *If a function w is 1-reversible by an independent agent i then w is not 1-SNCC.*

PROOF. The proof was given in [?] and it is repeated here for completeness.

Assume that w is 1-reversible by agent i. Then $w(v_i, y_{-i}) = 1 - w(\overline{v_i}, y_{-i})$ for every $y_{-i} \in B_{-i}$. Assume that all the other agents employ the straightforward strategy (f^t, g^t). Let μ_0 and μ_1 be the expected payoffs of agent i when truthfully declaring his type as 0 and 1 correspondingly. Since the distribution of his type is independent of others, the payoffs are not dependent on i's real type. If $\mu_0 > \mu_1$ then for i it is better always to declare 0 and have $g_i(r, \tau, 1) = 1 - r, g_i(r, \tau, 0) = r$ to always have the correct value and better payoff than under the truthful declaration. In the same way, in case $\mu_1 > \mu_0$ then for i it is better always to declare 1 and have $g_i(r, \tau, 0) = 1 - r, g_i(r, \tau, 1) = r$ to always have the correct value and better payoff than under the truthful declaration. If $\mu_1 = \mu_0$ then i can implement the same strategy as in the $\mu_1 > \mu_0$ case. The payoff will be the same, but he will mislead other agents. □

THEOREM 4. *In the independent values setting, for $1 \leq K \leq n$, if a function w is dominated and not 1-reversible then it is K-SNCC.*

PROOF. Let D be the set of agents that w is dominated by. For each $i \in D$ let $v_i \in B_i$ be the type for which agent i knows the output, i.e. $w(v_i, x_{-\{i\}}) = w(v_i, y_{-\{i\}})$ for every $x_{-\{i\}} \in B_{-\{i\}}, y_{-\{i\}} \in B_{-\{i\}}$.

Set the payment functions $m_i(v_i) = \epsilon$, $m_i(\overline{v_i}) = 0$ for every $i \in D$ like in [?]. For non-dominating agents set $m_{\overline{D}} = 0$. Using such payments, [?] gives a proof for $K = 1$.

By definition of dominated functions there $\exists d \in B, \forall i \in D \exists v_i \in B_i$ such that $w(v_i, v_{-\{i\}}) = d$ for any $v_{-\{i\}} \in B_{-\{i\}}$. Since w is not constant (constant functions are not dominated), $\exists v'_{-D} \in B_{-D}$ such that $w(\overline{v}_D, v'_{-D}) = \overline{d}$.

For $K \geq 2$, by Theorem 1 the only possible declaration strategy for the relevant agents of the minimal deviating coalition is to declare the reversed types. Notice that all the agents in D are relevant because otherwise they wouldn't dominate w.

But when an agent's type is $\overline{v}_i$ and he declares v_i, the center will announce d while the real value of w might be either d or $\overline{d}$, thus the agent might be mistaken. Therefore coalitions of at least two agents are also not worthwhile. □

THEOREM 5. *In the independent values setting, for $K \geq 2$ a function w is K-SNCC iff it is (K-1)-SNCC and $\forall C \in \widetilde{w_r^K}$, $\exists i \in C$, such that i is not* balanced.

PROOF. $\Longrightarrow$ From the definition of K-SNCC, it follows that K-SNCC$\subseteq$ $(K-1)$-SNCC. So a function that is not $(K-1)$-SNCC can't be K-SNCC.

Assume that there exists $C \in \widetilde{w_r^K}$ such that $\forall i \in C$, agent i is balanced. Then consider the following deviation from straightforward strategies by a coalition C: $f_C(v) = 1 - v$; $g_C(r, \tau, v) = 1 - r$. Since $C \in \widetilde{w_r^K}$, $w(f_C(v_C), f^t(v_{-C})) = 1 - w(v_C, v_{-C})$ and thus all the agents in C are always right and all the agents not in C are always wrong. Let us look at some agent $i \in C$. Let $u_0 = \sum_{v_{-\{i\}} \in B_{-\{i\}}} p(v_{-\{i\}}) m_i(0, v_{-\{i\}})$ be the expected payment of the agent i if he declares $f_i = 0$. Since we are in the independent values setting, u_0 is not dependent on the real type of agent i. In the same way let u_1 be the expected payment of the agent i if he declares $f_i = 1$. Since i is balanced, $p(v_i = 0) = p(v_i = 1) = 1/2$. Thus the expected payment of the agent i, playing the above strategy is $\frac{u_0 + u_1}{2}$, just as if i played the straightforward strategy. Since i is always correct, and has the same expected payment and the agents not in C are misled, i benefits from the deviation. This analysis is correct for every agent $i \in C$; they all benefit from the deviation, and thus w is not K-SNCC.

$\Longleftarrow$ The opposite direction states that a function w that is (K-1)-SNCC and $\forall C \in \widetilde{w_r^K}$ and $\exists i \in C$, such that i is not balanced, then w is K-SNCC. The function w is not 1-reversible, otherwise by Theorem 3 it won't be 1-SNCC and thus not (K-1)-SNCC. If w is dominated then by Theorem 4 it is K-SNCC. So it is enough to analyze non-dominated, non 1-reversible functions.

Consider the following payment functions of the agents: For each relevant agent i, $\forall v_i \in B_i, v_{-\{i\}} \in B_{-\{i\}}$ $m_i(v_i, v_{-\{i\}}) = p(v_i)$, i.e. each agent is paid by the probability of his declaration. Irrelevant agents are not paid. Since we are in the independent values setting, $p(v_i)$ is not dependent on $v_{-\{i\}}$ and thus an agent's payment is only affected by his own declaration. An irrelevant agent can not participate in a minimal deviating coalition, one that will make (K-1)-SNCC function not K-SNCC, because he doesn't contribute to misleading agents and doesn't contribute to payoffs of others.

Let us analyze possible deviating coalitions that could make w not K-SNCC. Let C be such a coalition, $|C| = K$. By Theorem 2 $C \in \widetilde{w^K}$.

In the case where $C \in \widetilde{w_r^K}$, then, by the condition of the theorem there exists an agent $i \in C$ that is not balanced. Wlog let $p_1 = p(v_i = 1) > p(v_i = 0) = p_0$. If i would play the straightforward strategy then his expected payment would be $(p_0^2 + p_1^2)$. So in order not to reduce the expected payment, agent i must declare $f_i = 1$ with probability of at least p_1. By Theorem 1, for $|C| \geq 2$ the only worthwhile declaration function is for relevant agents to declare a reversed type. But this would declare $f_i = 1$ with probability $p_0 < p_1$. Thus such coalition C is not worthwhile. For $|C| = 1$, the single deviating agent could also have a constant declaration function, but since w is not dominated and not 1-reversible, such constant declaration would cause the agent to be mistaken on some agent types. Therefore, with the given payment functions a deviation from the straightforward strategies is not worthwhile.

The other case is where $C \in \widetilde{w_i^K}$. Then w is K-irreversible by C, and by deviating from the straightforward strategies the coalition members could only hope for a better payment and not better exclusivity, since they don't mislead any non-deviating agents.

The truthful declaration would yield in a payment of $p(v_i = 0)^2 + p(v_i = 1)^2$. For $|C| \geq 2$ the coalition members must reverse their declaration and thus for the same reason as in the first case the payment won't improve. For $|C| = 1$ that would make the single agent of the deviating coalition C irrelevant, and thus he is not paid and doesn't benefit from the deviation.

By those two cases we have shown that under the conditions of the theorem no worthwhile deviating coalition of size up to K exists, and thus w is K-SNCC. ☐

To summarize, in the independent values setting, Theorems 3, 4 and 5 imply the following:

Corollary 1. *In the independent values setting, a function w is K-SNCC for $1 \leq K \leq n$ iff it is not 1-reversible and $\forall C \in \mathring{w}_r^K, \exists i \in C$, such that i is not balanced.*

If a function w is 1-reversible then it is not 1-SNCC and thus not K-SNCC by Theorem 3. Non 1-reversible dominated functions are K-SNCC by theorem 4. Non 1-reversible non-dominated functions are 1-SNCC by [?], and Theorem 5 ties between K-SNCC and $\forall C \in \mathring{w}_r^K, \exists i \in C$, such that i is not balanced.

3.2 The correlated values setting

In the correlated values setting we refer to the general case where we are not restricted by the assumption that all agents' types are independent. As illustrated in [?] a careful designer may exploit dependencies between agents' types in order to make a function non-cooperatively computable. This makes a full characterization of the K-SNCC functions a highly desired but non-trivial challenge.

Our characterization makes use of the following definitions:

Definition 12. For a function w, its correlation graph $G = (V, E)$ is defined as follows: Each relevant agent is a vertex, thus $V = N \setminus I$. For each pair of vertices $v_1, v_2 \in V$, there exists an undirected edge between them iff the agents' v_1 and v_2 types distributions are dependent, i.e. $p(v_1 = x|v_2 = y) \neq p(v_1 = x|v_2 = z)$ for some $x \in B_{v_1}, y \neq z \in B_{v_2}$. Note that in the Boolean domain if the inequality holds for some x and y then it holds for any x and y.

Following this definition we can now define a *connected component M* of an agent i.

Definition 13. For a function w, and its correlation graph G, a connected component $M(i)$ of a relevant agent i is the set of all the relevant agents for which there is an undirected path in G from them to i.

Given the above terminology, we provide a sufficient condition for preventing an independent agent from participating in any beneficial deviation.

Theorem 6. *Let G be the correlation graph of a function w. Then if an agent i's connected component in G is of size 1, (i.e $M(i) = \{i\}$, meaning that i is independent) and i is not 1-reversible and not balanced, then i cannot participate in a worthwhile deviating coalition of any size.*

Proof. If w is dominated by i, then set $m_i = \epsilon > 0$ for the type for which i knows w's value, and $m_i = 0$ for the other type, like in theorem 4. By its proof it wouldn't be worthwhile for i to deviate (even together with other agents).

If w is not dominated and not 1-reversible by i, then i can't deviate unilaterally, and since i is also not balanced, then by setting $m_i(v_i) = p(v_i)$ for every $v_i \in B_i$, by the proof of theorem 5 it is not worthwhile for i to participate in any deviating coalition of a size at least 2.

Thus for such i it is not worthwhile to deviate from the straightforward strategy. ☐

The previous theorem dealt with independent agents. We now consider agents which are not independent, but are unbalanced:

Theorem 7. *Let G be the correlation graph of a function w. Then if some connected component M in G is of size at least 2, and contains some unbalanced agent i, then agent i cannot participate in any worthwhile deviating coalition of any size.*

Proof. Let $j \neq i \in M$ be some agent that his type's probability distribution is dependent on i's type probability distribution. Let $y_j \in B_j$ be the value such that:

$$p(v_j = y_j|v_i = 0) + p(v_j = y_j|v_i = 1) \geq \\ p(v_j = \overline{y_j}|v_i = 0) + p(v_j = \overline{y_j}|v_i = 1)$$

Since the sum of the two sides equals 2, the left side is at least 1. Then define $\delta_i \geq 0$ as the difference by which the larger side exceeds 1:

$$\delta_i = p(v_j = y_j|v_i = 0) + p(v_j = y_j|v_i = 1) - 1$$

If $\delta_i = 0$ then $y_j \in B_j$ can take *any* value. So we will decide on it later.

Since i's and j's types probability distributions are dependent: $p(v_j = y_j|v_i = 0) \neq p(v_j = y_j|v_i = 1)$. Let $p_{i,l} = \min(p(v_j = y_j|v_i = 0), p(v_j = y_j|v_i = 1))$ be the smaller probability of the two, and $p_{i,h} = \max(p(v_j = y_j|v_i = 0), p(v_j = y_j|v_i = 1))$ be the larger probability of the two.

Let $v_{i,h}$ and $v_{i,l}$ be the types of agent i corresponding to $p_{i,h}$ and $p_{i,l}$ respectively. Since we are discussing a correlation of two Boolean random variables, if y_j and $v_{i,h}$ have positive correlation, $\overline{y_j}$ and $v_{i,l}$ also have positive correlation. If $\delta_i = 0$ then we were free to choose any y_j, which imposes $v_{i,h}$ that it is positively correlated to. In that case we will choose y_j such that $p(v_{i,h}) > p(v_{i,l})$. We can do that since i is unbalanced and thus the two probabilities are unequal.

Now we can define the payoff function for agent i. If he declares $v_{i,l}$, he gets a constant value regardless of what other agents declare: $m_i(v_{i,l}, v_{-\{i,j\}}, v_j) = p_{i,h} - \epsilon_{i1}$ for every $v_j \in B_j, v_{-\{i,j\}} \in B_{-\{i,j\}}$. If i declares $v_{i,h}$, he gets a lottery, with expected value $p_{i,h}$ which is greater than $p_{i,h} - \epsilon_{i1}$ under truthful declaration (if agent j also declares truthfully): $m_i(v_{i,h}, v_{-\{i,j\}}, y_j) = 1$. And $m_i(v_{i,h}, v_{-\{i,j\}}, z_j) = 0$ for every $z_j \neq y_j$. Recall that y_j is fixed here. Under truthful declaration of both agents, the payoff of agent i declaring $v_{i,l}$ is $p_{i,h} - \epsilon_{i1}$, and the expected payoff of agent i declaring $v_{i,h}$ is $p_{i,h}$. All payments can be multiplied by $\epsilon_{i2} \to 0$ to make them arbitrary small.

We take $0 < \epsilon_{i1} < p_{i,h} - p_{i,l}$ so the expected payment for truthful declaration of $v_{i,h}$ is greater than for a deviation to $v_{i,l}$. If the real type of agent i is $v_{i,l}$ then it is also not worthwhile to deviate, because instead of a sure payoff of $p_{i,h} - \epsilon_{i1} > p_{i,l}$ he will get a lottery with an expected payoff of $p_{i,l}$. Thus if j is declaring truthfully, it is not worthwhile for i to deviate.

If j also participates in the deviating coalition then by Theorem 1 all relevant coalition member declare the opposite of their true type. Let us analyze the expected payoff of agent i under those declarations. If i and j played the straightforward strategy, i's expected payoff would be:

$$p(v_{i,l})(p_{i,h} - \epsilon_{i1}) + p(v_{i,h})p(y_j|v_{i,h}) = \\ p(v_{i,l})p_{i,h} + p(v_{i,h})p_{i,h} - \epsilon_{i1}p(v_{i,l})$$

If both i and j declared the opposite of their true type, the expected payoff of the agent i would be:

$$p(v_{i,h})(p_{i,h} - \epsilon_{i1}) + p(v_{i,l})p(\overline{y_j}|v_{i,l}) = \\ p(v_{i,h})p_{i,h} + p(v_{i,l})(1 - p_{i,l}) - \epsilon_{i1}p(v_{i,h})$$

Thus the change in payoff is $p(v_{i,l})(1-p_{i,l}-p_{i,h})+\epsilon_{i1}(p(v_{i,l})-p(v_{i,h})) = p(v_{i,l})(-\delta_i) + \epsilon_{i1}(p(v_{i,l}) - p(v_{i,h}))$.

If $\delta_i = 0$, then y_j was chosen in a way such that $p(v_{i,h}) > p(v_{i,l})$ and thus the change is negative. If $\delta_i > 0$ then ϵ_{i1} can be taken small enough: $\epsilon_{i1} < \frac{\delta_i p(v_{i,l})}{|p(v_{i,l}) - p(v_{i,h})|}$ so the change is again negative.

Thus if j also deviates, i's expected payoff get smaller. Therefore regardless of whether j deviates or not, it is not worthwhile for i to deviate. □

Theorem 7 showed a way to make deviations not beneficial for an agent. Theorem 8 shows that this enables to prevent deviations by all agents in that agent's connected component.

THEOREM 8. *Let i be an agent who plays the straightforward strategy. Then, deviation from the straightforward strategy is not beneficial for all the agents in i's connected component $M(i)$.*

PROOF. The proof is by construction. Let $M_0 = \{i\}$. All the agents in M_0 are playing the straightforward strategy. Let $M_1 \subseteq M(i) \setminus M_0$ be the agents that are dependent on agents in M_0 (on agent i in this case). The probability distribution of an agent $j_0 \in M_0$'s types is dependent on the type of $i_1 \in M_1$. The following is similar to proof of Theorem 7, where it was proven that it is not worthwhile for unbalanced i to deviate unilaterally.

Since i_1's and j_0's types probability distributions are dependent: $p(v_{j_0} = y_{j_0}|v_{i_1} = 0) \neq p(v_{j_0} = y_{j_0}|v_{i_1} = 1)$ for some y_{j_0}. Let $p_{i_1,l} = \min(p(v_{j_0} = y_{j_0}|v_{i_1} = 0), p(v_{j_0} = y_{j_0}|v_{i_1} = 1))$ be the smaller probability of the two, and $p_{i_1,h} = \max(p(v_{j_0} = y_{j_0}|v_{i_1} = 0), p(v_{j_0} = y_{j_0}|v_{i_1} = 1))$ be the larger probability of the two. Let $v_{i_1,h}$ and $v_{i_1,l}$ be the types of agent i_1 corresponding to $p_{i_1,h}$ and $p_{i_1,l}$ respectively.

If i_1 declares $v_{i_1,l}$, he gets a constant value: $m_{i_1}(v_{i_1,l}, v_{-\{i_1,j_0\}}, v_{j_0}) = p_{i_1,h} - \epsilon_{i_1 1}$ for every $v_{j_0} \in B_{j_0}$, $v_{-\{i_1,j_0\}} \in B_{-\{i_1,j_0\}}$. If i_1 declares $v_{i_1,h}$, he gets a lottery, whose expected value $p_{i_1,h}$ is greater than $p_{i_1,h} - \epsilon_{i_1 1}$ only under truthful declaration (if the agent j_0 also declares truthfully): $m_{i_1}(v_{i_1,h}, v_{-\{i_1,j_0\}}, y_{j_0}) = 1$. And $m_{i_1}(v_{i_1,h}, v_{-\{i_1,j_0\}}, z_{j_0}) = 0$ for every $z_{j_0} \neq y_{j_0}$. Recall that y_{j_0} is fixed here. Under truthful declaration of both agents, the payoff of agent i_1 declaring $v_{i_1,l}$ is $p_{i_1,h} - \epsilon_{i_1 1}$, and the expected payoff of agent i_1 declaring $v_{i_1,h}$ is $p_{i_1,h}$. All payments can be multiplied by $\epsilon_{i_1 2} \to 0$ to make them arbitrary small.

We take $0 < \epsilon_{i_1 1} < p_{i_1,h} - p_{i_1,l}$ so the expected payment for truthful declaration of $v_{i_1,h}$ is greater than for a deviation to $v_{i_1,l}$. If the real type of agent i_1 is $v_{i_1,l}$ then it is also not worthwhile to deviate, because instead of a sure payoff of $p_{i_1,h} - \epsilon_{i_1 1} > p_{i_1,l}$ he will get a lottery with an expected payoff of $p_{i_1,l}$. Thus if j_0 is declaring truthfully (and he is, because $j_0 \in M_0$), it is not worthwhile for i_1 to deviate.

Therefore all the agents in M_1 will be playing the straightforward strategy. The proof for the rest of the agents in $M(i)$ is similar. Let $M_k \subseteq M \setminus \bigcup_{0 \leq l < k} M_l$ be the agents, whose payment we didn't set yet, who are dependent on M_{k-1}. For each agent $i_k \in M_k$ his payment will be dependent upon the declaration of $j_{k-1} \in M_{k-1}$ that i_k's type is dependent upon. Since all the agents in M_{k-1} are playing the straightforward strategy, all the agents in M_k must play the straightforward strategy. Thus all the agents in $M(i)$ are playing the straightforward strategy and no one deviates. □

In order to provide the desired characterization we are left with the need to deal with components of the correlation graph in which all agents are balanced. This case will be handled in Theorems 9 and Theorem 10 below, and will make use of the following definitions.

Definition 14. Let p_1, p_2 be the probability distributions of t Boolean random variables v_i, $1 \leq i \leq t$. $p_1, p_2 \in \Delta(B^t)$. Those probability distributions are said to be *co-permutations* of each other, if there exists a permutation $\pi : B^t \to B^t$ such that $p_1(v) = p_2(\pi(v))$ for every $v \in B^t$.

Those probability distributions are said to be *reversible co-permutations* of each other if the permutation π always reverses some inputs and never changes the others, i.e.$\exists T \subseteq \{1, \ldots, t\}$ such that $\pi(v_T, v_{-T}) = (\overline{v}_T, v_{-T})$ for every v_T and v_{-T}.

Since all agents in a deviating coalition C of size at least 2 always declare the opposite of their types (and agents not in C declare their true type) we will say that the coalition C imposes a reversible permutation $\pi_C(v_C, v_{-C}) = (\overline{v}_C, v_{-C})$.

Theorem 9 below shows that if for a balanced agent i, there exists a set of relevant agents Q, $i \notin Q$, such that the

Table 1: Conditional distribution table

	y_1	$\dots$	y_q
$v_i = 0$	$p(v_Q = y_1 \mid v_i = 0)$	$\dots$	$p(v_Q = y_q \mid v_i = 0)$
$v_i = 1$	$p(v_Q = y_1 \mid v_i = 1)$	$\dots$	$p(v_Q = y_q \mid v_i = 1)$

Table 2: Joint distribution table

	y_1	$\dots$	y_q
$v_i = 0$	$p(v_Q = y_1, v_i = 0)$	$\dots$	$p(v_Q = y_q, v_i = 0)$
$v_i = 1$	$p(v_Q = y_1, v_i = 1)$	$\dots$	$p(v_Q = y_q, v_i = 1)$

probability distributions over these agents' types induced by the two possible types of i are not reversible co-permutations of each other, then i can not take part in any beneficial deviation.

THEOREM 9. *Let i be a balanced agent. Let Q be some set of relevant agents, $i \notin Q$. Then if the probability distributions of $y_Q \in B_Q$, $p(y_Q|v_i = 0)$ and $p(y_Q|v_i = 1)$ are not reversible co-permutations of each other, then agent i cannot participate in any deviating coalition.*

PROOF. Let us examine the probability distribution of types of agents in Q conditioned on a specific type of agent i. They can be shown as a table with two rows, where each row is a conditioning on another i's type. For $q = 2^{|Q|}$ it is shown in Table 1 above.

Since agent i is balanced, i.e. $p(v_i = 0) = p(v_i = 1) = 1/2$, we can instead talk about the joint distribution which is shown in Table 2. The entries in Table 2 are entries of Table 1 multiplied by 0.5.

Now we will define the payment function for agent i that will prevent him from participating in a deviating coalition. We will identify some joint types of agents in Q as *payable lotteries*. The set P will denote those payable lottery types. For one out of the two possible types of agent i the payment will be as following:

$$m_i(v_i, v_Q, v_{-(Q\cup\{i\})}) = \begin{cases} 1 & \text{if } (v_Q, v_i) \in P \\ 0 & \text{otherwise} \end{cases}$$

For the other type of agent i the payoff will be constant (calculated later).

The algorithm for determining the types in P is as follows:

1. Set $P_0 = \emptyset, P_1 = \emptyset$,
   $C = 2^{Q\cup\{i\}} = \{(y_1, 0), \dots, (y_q, 0), (y_1, 1), \dots (y_q, 1)\}$.
2. Look for the tuples of types in C whose probability is maximal.
   $max = \max_{v_Q \in B_Q, v_i \in B_i}(p(v_Q, v_i))$.
3. For each $(y_m, h) \in C$ such that $p(v_Q = y_m, v_i = h) = max$ do:
   (a) Add (y_m, h) to P_h
   (b) Remove (y_m, h) from C.
4. If the sizes of P_0 and P_1 are not equal, then set P to be the larger one and terminate.
5. If there exists no *reversible permutation* π such that for every $(y_m, 0) \in P_0$, $(\pi(y_m), 1) \in P_1$ then set $P = P_0$ and terminate.
6. If C is empty then terminate, otherwise goto step 2.

At each iteration the algorithm removes elements from C, so it will eventually terminate. If the algorithm terminates at step 6, then each time step 3 was executed, an equal number of elements was added to P_0 and P_1, until P_0 contained the whole first line of the table, and P_1 the whole second line. Since we haven't terminated at step 5, $p(y_Q|v_i = 0)$ and $p(y_Q|v_i = 0)$ are reversible co-permutations contrary to conditions of the theorem.

Let h be the index of the payable lottery set chosen, i.e. $P = P_h$. Let p_{Ph} be the sum of the probabilities of all the elements in P given the appropriate value of v_i, i.e. $p_{Ph} = \sum_{(y_m,h)\in P_h} p(v_Q = y_m, v_i = h|v_i = h)$. Let p_{Pl} be the sum of probabilities of all the elements in P given the opposite value of v_i, i.e. $p_{Pl} = \sum_{(y_m,h)\in P} p(v_Q = y_m, v_i = \overline{h}|v_i = \overline{h})$. Set:

$$m_i(v_i, v_Q, v_{-Q\cup\{i\}}) = \begin{cases} 1 & \text{if } v_i = h \text{ and } (v_Q, v_i) \in P \\ 0 & \text{if } v_i = h \text{ and } (v_Q, v_i) \notin P \\ p_{Ph} - \epsilon_{i1} & \text{if } v_i \neq h, \text{ for some } \epsilon_{i1} > 0 \end{cases}$$

If agent i plays the straightforward strategy, then when his type is h he gets a lottery with an expected value of p_{Ph}, and when his type is $\overline{h}$ he gets a constant value of $p_{Ph} - \epsilon_{i1}$. If agent i would unilaterally deviate from h to $\overline{h}$, his expected payoff would decrease from p_{Ph} to $p_{Ph} - \epsilon_{i1}$. If agent i would unilaterally deviate from $\overline{h}$ to h, his expected payoff would change from $p_{Ph} - \epsilon_{i1}$ to p_{Pl}. Since the algorithm picked the maximal values for P, and the algorithm didn't terminate at step 6, it must be that $p_{Ph} > p_{Pl}$ and thus ϵ_{i1} could be picked such that $0 < \epsilon_{i1} < p_{Ph} - p_{Pl}$, making unilateral deviation not worthwhile.

Let us now consider deviations by coalitions. If agent i participates in a coalition of size at least 2, then by theorem 1 all the relevant deviating agents declare the opposite of their true types. Thus the probability distribution of the *declarations* is a reversible permutation of the probability distribution of the real types. Let us call that reversible permutation π and let us calculate the expected payoff of such deviation. In the case where the real type of agent i is h, the payoff will decrease to $p_{Ph} - \epsilon_{i1}$. In the case where the real type of agent i is $\overline{h}$, the payoff will change to: $p_{\text{dev}} = \sum_{(y_m,h)\in P} p(v_Q = \pi(y_m), v_i = \overline{h}|v_i = \overline{h})$.

If the algorithm terminated at step 4, then on the last iteration more elements were added to $P = P_h$ than to $P_{\overline{h}}$. Since the probabilities of the elements added at the same iteration of step 3 are equal, $p_{\text{dev}} < p_{Ph}$ and thus by picking ϵ_{i1} small enough, i.e. $0 < \epsilon_{i1} < p_{Ph} - p_{\text{dev}}$, the deviation won't be worthwhile.

If the algorithm terminated at step 5, then it could be that $p_{\text{dev}} = p_{Ph}$, but that would mean that the lines of the distributions table are reversible permutations of each other (by the permutation π used in the calculation of p_{dev}). This would be a contradiction to the termination condition of step 5.

Therefore under the conditions stated in the theorem, it is not worthwhile for agent i to deviate either unilaterally or as a part of a deviating coalition. □

The above theorem shows a way to prevent a particular agent from deviation; By applying Theorem 8 we get that it is not worthwhile for the rest of the agents in that agent's connected component to deviate.

Theorem 10 complements Theorem 9, by showing a condition for the existence of beneficial deviation by a group of balanced agents, using the notion of reversible co-permutations as discussed above.

THEOREM 10. *Let $C \subset N$, $|C| \geq 2$ be a reversing deviating coalition consisting of relevant agents, such that all the agents in C are never mistaken on w's value. Let G be the correlation graph of a function w. If for every agent $i \in C$ the connected component $M(i)$ of agent i contains only balanced agents, and for a reversible permutation $\pi_C(v_C, v_{-C}) = (\overline{v}_C, v_{-C})$ it holds that $p(v_{-i}|v_i = 0) = p(\pi_C(v_{-i})|v_i = 1)$ (i.e. $p(v_{-i}|v_i = 0)$ and $p(v_{-i}|v_i = 1)$ are reversible co-permutations) for every $v_{-i} \in B_{-i}$, then C is a worthwhile deviation in the SNCC setting.*

PROOF. Let us consider an agent $i \in C$. The payment function m_i is dependent on the declarations of the agents. For each declared tuple of types (v_i, v_{-i}) there is a payoff $u_{(v_i,v_{-i})}$ that the agent i would get with a probability $p(v_i)p(v_{-i}|v_i)$ if he played the straightforward strategy. However when participating in a deviating coalition C, the probability to declare (v_i, v_{-i}) is $p(\overline{v}_i)p(\pi_C(v_{-i})|\overline{v}_i) = p(v_i)p(v_{-i}|v_i)$. The above equation holds if agent i always declares the opposite of his type, which is his only option if $|C| \geq 2$. Thus the expected payoff stays the same. Therefore the deviation is worthwhile for agent i, since he is paid the same, but he misleads agents not in C. This analysis is correct for every $i \in C$, thus the coalition C is a worthwhile deviation. □

Full characterization

We can now state the full characterization of when a Boolean function w is K-SNCC.

The function w is K-SNCC if for every possible coalition $C \in \mathring{w}_r^K$ there exists an agent $i \in C$ such that one of the following holds:

- The connected component of agent i is of size 1 (i is independent), i is not 1-reversible and it is either unbalanced (Theorem 6) or is a single agent in the coalition ([?]).
- The connected component of agent i is at least of size 2 and contains an unbalanced agent. (Theorem 7)
- The connected component of agent i is at least of size 2, all the agents in the component are balanced and there exists an agent j in i's component such that the probability distributions $p(v_{-j}|v_j = 0)$ and $p(v_{-j}|v_j = 1)$ are not reversible co-permutations. (Theorem 9)

In the cases not covered above, w is not K-SNCC, i.e. there exists a deviating coalition C, $|C| \leq K$ such that for every agent i either of the following holds:

- The connected component of agent i is of size 1 and i is 1-reversible. (Theorem 3)
- The connected component of agent i is of size 1 and i is balanced and i is not the only agent in the coalition C. (Theorem 5)
- Both of the following are true:
  - The connected component $M(i)$ of agent i is of size at least 2 and it contains only balanced agents.
  - Under the reversible permutation π_C imposed by the deviating coalition C, the probability distributions $p(v_{-\{i\}}|v_i = 0)$ and $p(v_{-\{i\}}|v_i = 1)$ are reversible co-permutations. (Theorem 10)

3.3 Irrelevant agents

Notice that our analysis considered only relevant agents. The irrelevant agents can be thought of spectators that don't influence the value of the function, but still wish to know its value. Although it is natural to simply ignore irrelevant agents, formally speaking our characterization should be revisited if we don't simply ignore them. More specifically, if the irrelevant agents' types probability distributions are independent of the types of the relevant agents, then they possess no information altogether and can be completely disregarded; however, if their types are dependent on the types of relevant agents, then the characterization above becomes incomplete, since the payment function may consider the declarations of the irrelevant agents.

4. CONCLUSIONS

Non-cooperative computing captures a basic problem in information aggregation within communities. SNCC is a natural framework for the study on how a moderator of a social platform can lead agents to desired joint activity. This paper extends the study of SNCC to deal with deviations by coalitions, and provide full characterization of the functions which are K-SNCC for both the independent and the correlated values settings. We believe that the understanding of the structure of the K-SNCC functions is an essential step in building a rigorous theory of computation in communities. Future work may integrate into this setting other ingredients, such as the cost of accessing information by individual agents, or the lack of a single moderator for the social platform. This will allow further understanding of the possibilities and challenges for joint activity within communities.

5. REFERENCES

[1] R. Aumann. Acceptable points in general cooperative n-person games. In A. Tucker and R. Luce, editors, *Contribution to the Theory of Games, Vol. IV, Annals of Mathematics Studies, 40*, pages 287–324. 1959.

[2] J. Halpern and V. Teague. Rational secret sharing and multiparty computation. In *Proc. of STOC 2004*, 2004.

[3] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory.* Oxford University Press, 1995.

[4] R. McGrew, R. Porter, and Y. Shoham. Towards a general theory of non-cooperative computation. In *Proc. of the 9th conference of theoretical aspects of rationality and knowlede (TARK 2003)*, pages 59–71, 2003.

Appendix

The following theorem explains why Theorem 9 and Theorem 10 are complementary. If the conditions of Theorem 10 hold, then using Theorem 11 the set of all relevant agents can be reduced to any subset Q that still has reversible co-permutations and thus the conditions of Theorem 9 do not hold.

THEOREM 11. *If i is a balanced relevant agent, and R is the set of all the other relevant agents, and the probability distributions $v_R|v_i = 0$ and $v_R|v_i = 1$ are reversible co-permutations of each other, then for $\emptyset \neq Q \subset R$ the probability distributions $v_Q|v_i = 0$ and $v_Q|v_i = 1$ are reversible co-permutations of each other.*

PROOF. Unless $|R| = 1$, let us pick some agent $j \in R \backslash Q$ and define $R_1 = R \backslash \{j\}$. We will show that $v_{R_1}|v_i = 0$ and $v_{R_1}|v_i = 1$ are reversible permutations of each other.

Let π be the reversible permutation such that $p(v_R|v_i = 0) = p(\pi(v_R)|v_i = 1)$ for every $v_R \in B_R$. Since π is a reversible permutation there exist some set of agents C such that $\pi(v_C, v_{R\backslash C}) = (\overline{v}_C, v_{R\backslash C})$. In case $j \in C$ we will define $\pi_1(v_{R_1}) = (\overline{v}_{C\backslash\{j\}}, v_{R\backslash C}) = (\overline{v}_{C\backslash\{j\}}, v_{R_1\backslash C})$. If $j \notin C$ we will define $\pi_1(v_{R_1}) = (\overline{v}_C, v_{R\backslash(C\cup\{j\})}) = (\overline{v}_C, v_{R_1\backslash C})$.

$p(\pi_1(v_{R_1})|v_i = 1) =$
$= p(\pi_1(v_{R_1}), v_j = 0|v_i = 1) + p(\pi_1(v_{R_1}), v_j = 1|v_i = 1) =$
$= p(\pi(v_{R_1}, v_j = 0)|v_i = 1) + p(\pi(v_{R_1}, v_j = 1)|v_i = 1) =$
$= p(v_{R_1}, v_j = 0|v_i = 0) + p(v_{R_1}, v_j = 1|v_i = 0) = p(v_{R_1}|v_i = 0)$

In the same way $p(\pi_1(v_{R_1})|v_i = 0) = p(v_{R_1}|v_i = 1)$. Thus $v_{R_1}|v_i = 0$ and $v_{R_1}|v_i = 1$ are reversible permutations of each other.

Just like we removed j from R, we can continue removing agents until we are left with Q. $\square$

If we wish to considerer smaller tables when using Theorem 9, we can make use of the following theorem:

THEOREM 12. *Let i be a balanced agent, $Q_N = N\backslash\{i\}$ be the set of rest of the agents, $Q_M = M\backslash\{i\}$ be the agents in i's connected component except for i. If the probability distributions $v_{Q_M}|v_i = 0$ and $v_{Q_M}|v_i = 1$ are reversible co-permutations of each other, then the probability distributions $v_{Q_N}|v_i = 0$ and $v_{Q_N}|v_i = 1$ are reversible co-permutations of each other.*

PROOF. Let π_{Q_M} be that reversible permutation. Consider the permutation π_{Q_N} that reverses the same inputs as π_{Q_M} (rest of the inputs stay as they are). Then for every input type $(v_i, v_{Q_M}, v_{Q_N\backslash Q_M})$ the probability $p(v_i, v_{Q_M}, v_{Q_N\backslash Q_M}) =$
$p(v_{Q_N\backslash Q_M}|v_i, v_{Q_M})p(v_{Q_M}|v_i)p(v_i) =$
$p(v_{Q_N\backslash Q_M}|\overline{v_i}, \pi_{Q_M}(v_{Q_M}))p(\pi_{Q_M}(v_{Q_M})|\overline{v_i})p(\overline{v_i}) =$
$(\overline{v_i}, \pi_{Q_M}(v_{Q_M}), v_{Q_N\backslash Q_M}) = (\overline{v_i}, \pi_{Q_N}(v_{Q_N}))$.

The first parentheses:
$p(v_{Q_N\backslash Q_M}|v_i, v_{Q_M}) = p(v_{Q_N\backslash Q_M}|\overline{v_i}, \pi_{Q_M}(v_{Q_M}))$, are equal because $Q_N\backslash Q_M$ are not in the connected component M and thus are independent with inputs in M. The second parentheses: $p(v_{Q_M}|v_i) = p(\pi_{Q_M}(v_{Q_M})|\overline{v_i})$, are equal because π_{Q_M} is a reversible permutation. The last parentheses: $p(v_i) = p(\overline{v_i})$, are equal because i is a balanced agent. The multiplication of those 3 parentheses is the probability of the input $(\overline{v_i}, \pi_{Q_M}(v_{Q_M}), v_{Q_N\backslash Q_M})$ which is $(\overline{v_i}, \pi_{Q_N}(v_{Q_N}))$. Thus π_{Q_N} is a reversible permutation. $\square$