

MPRA

Munich Personal RePEc Archive

A minority-proof cheap-talk protocol

Heller, Yuval

Tel-Aviv University

19 August 2005

Online at <https://mpra.ub.uni-muenchen.de/7716/>

MPRA Paper No. 7716, posted 12 Mar 2008 16:21 UTC

A Minority-Proof Cheap-Talk Protocol

Yuval Heller¹

Revision 5 - February 26, 2008.

First version received in August 19, 2005.

Corresponding author: Yuval Heller

School of Mathematical Sciences,

Tel-Aviv University,

Tel-Aviv 69978, Israel.

Email: helleryu@post.tau.ac.il

Phone: 972-3-6405386, 972-52-5282182

Fax: 972-3-640-9357

¹ The School of Mathematical Sciences, Tel Aviv University. This paper is based on a Master thesis the author done under the supervision of Prof. Ehud Lehrer. I would like to thank Prof. Lehrer for his careful supervision and for the continuous help he offered. My deep gratitude is also given to Prof. Eilon Solan for many discussions and useful ideas concerning the subject, and for the associate editor and the anonymous referees for many useful comments during the process of writing this paper.

Abstract

This paper analyzes the implementation of correlated equilibria that are immune to joint deviations of coalitions by cheap-talk protocols. We construct a universal cheap-talk protocol (a polite protocol that uses only 2-player private channels) that is resistant to deviations of fewer than half the players, and using it, we show that a large set of correlated equilibria can be implemented as Nash equilibria in the extended game with cheap-talk. Furthermore, we demonstrate that in general there is no cheap-talk protocol that is resistant for deviations of half the players.

JEL classification: C72

Keywords: non-cooperative games, cheap-talk, correlated equilibrium, strong equilibrium, coalition-proof equilibrium, fault-tolerant distributed computation.

1. Introduction

The self-enforcing outcomes of a "silent" game (with no communication) are the set of Nash equilibria (Nash, 1951). Pre-play mediated communication allows players to implement the larger set of correlated equilibria (Aumann, 1974, 1987). In many environments, it is hard to find a fair and discrete mediator, and it is natural to ask which agreements can be implemented using only cheap-talk: pre-play, unmediated, non-binding, non-verifiable communication among the players (See Crawford & Sobel, 1982; Forges, 1990; Barany, 1992; Aumann & Hart, 2003; Gerardi, 2004, and the references within). In this paper we are going to extend two existing results:

- Ben-Porath (2003): The existence of a cheap-talk protocol that allows three or more players to implement any Nash dominating correlated equilibrium as a Nash equilibrium of the extended game.
- Ben-Porath (1998) and Krishna (2004): In general such a protocol does not exist for two players.

In "silent" games, coalitional deviations might be considered improbable, due to not having a device to coordinate the joint deviation. However, cheap-talk is a natural coordination device. Resistance against coalitional deviations was discussed in Aumann (1959) and Bernheim et al. (1987), where strong Nash equilibrium was defined as a Nash equilibrium, which is resistant against a deviation of any coalition (i.e. there is no coalitional deviation that is profitable for all the deviators). Moreno & Wooders (1996) give the correlated counterpart definition: a *strong correlated equilibrium*.²

In some real-world environments, while it is easy for small coalitions to deviate from an equilibrium, it is much harder for large coalitions to secretly coordinate a deviation while hiding it from the non-deviating players.³ One example for such environment is the field of foreign affairs: there are no known examples of secret joint deviations of large coalitions (a few dozens countries), but there are secret joint deviations of small coalitions (a few countries), as in the following examples:

- The secret additional protocol of Molotov-Ribbentrop Pact (1939) in which two countries have secretly divided between them six neighboring countries.
- The surprising joint attack of Egypt and Syria against Israel in 1973 October war.

We therefore introduce two new concepts:⁴ *k-strong correlated equilibrium* and *k-strong Nash equilibrium*, which require resistance against coalitional deviations of up to k players. We are

² Alternative definitions can be found in Milgrom & Roberts (1996), Ray (1996, 1998), Einy & Peleg (1995) and Bloch & Dutta (2007).

³ While it is true that a deviation of the grand coalition (all n players) may sometimes be easily coordinated, we would like to note that such a deviation does not threaten any player (because everyone earns from it).

⁴ Those concepts somewhat resemble Eliaz (1999) concept of k -Fault-Tolerant Nash Equilibrium.

interested in the question of existence of cheap-talk protocols that can implement a k -strong correlated equilibrium, as a k -strong Nash equilibrium in an extended game with cheap-talk. We would like to note that existing cheap-talk protocols (like Ben-Porath, 2003) can make such an implementation only for the case $k=1$. We say that a strategy profile q is k -strong *punishable* if there exists another strategy profile \tilde{q} that is dominated by q even when the players of a coalition with up to k players do a joint scheme (as defined formally in definition 2.5). Our main result is the existence of a general "minority-proof" cheap-talk protocol (a finite polite protocol that uses only 2-player private channels) that for every $k < n/2$, can implement any k -strong correlated equilibrium (with rational parameters) which is k -strong punishable as a k -strong Nash equilibrium in the extended cheap-talk game. Furthermore, we prove that in general such a protocol does not exist when $k \geq n/2$. Our protocol is based on the k -private protocol presented in Ben-Or et al. (1988), which deals with fault-tolerant distributed computation (as described in Section 3).

The paper is organized as follows: Section 2 presents the model and formal definitions. Section 3 presents the main result. Section 4 shows that in general there are no similar protocols when $k \geq n/2$. Section 5 gives an example for the applicative use of our protocol. We conclude in Section 6.

2. Model and Definitions

A finite game in strategic form G is defined as $G = (N, (A^i)_{i \in N}, (u^i)_{i \in N})$, where $N = \{1, \dots, n\}$ is a non-empty finite set of players, and for each $i \in N$, A^i is player i 's finite (and non-empty) set of pure actions, and u^i is player i 's payoff function, a real-valued function on $A = \prod_{i \in N} A^i$. The multi-linear

extension of u^i to $\Delta(A)$ is still denoted by u^i . A member of $\Delta(A)$ is called a (correlated) *strategy profile*. A coalition S is a non-empty member of 2^N . Given a coalition $S \subseteq N$, let $A^S = \prod_{i \in S} A^i$, and let

$-S = \{i \in N \mid i \notin S\}$ denote the complementary coalition. A member of $\Delta(A^S)$ is called an S -*strategy profile*. Given $q \in \Delta(A)$ and $a^S \in A^S$, we define $q_{|S} \subseteq \Delta(A^S)$ to be $q_{|S}(a^S) = \sum_{a^{-S} \in A^{-S}} q(a^S, a^{-S})$, and

for simplicity we omit the subscript: $q(a^S) = q_{|S}(a^S)$. Given $a^S, s.t. q(a^S) > 0$, we define

$q(a^{-S} | a^S) = \frac{q(a^S, a^{-S})}{q(a^S)}$. We say that $q \in \Delta(A)$ is an *uncorrelated strategy profile* if for every

$a = (a^1, \dots, a^n) \in A$, $q(a) = q(a^1) * \dots * q(a^n)$. Similarly, given a coalition $S \subset N$, we say that $q^S \in \Delta(A^S)$ is an *uncorrelated S -strategy profile* if for every $a^S \in A^S$, $q^S(a^S) = \prod_{i \in S} q^S(a^i)$. Let

$IA \subset \Delta(A)$ be the set of uncorrelated strategy profiles, and let $IA^S \subset \Delta(A^S)$ be the set of uncorrelated S -strategy profiles. Given $q \in IA$, we write $q = (q^S, q^{-S})$ where $q^S \in IA^S$ and $q^{-S} \in IA^{-S}$.

Definition 2.1: An uncorrelated strategy profile $q \in IA$ is a k -strong Nash equilibrium if and only if for all coalitions $S \subseteq N$ satisfying $|S| \leq k$, and for every uncorrelated S -strategy profile $p^S \in IA^S$, there exists a player $i \in S$ such that $u^i(q) \geq u^i(p^S, q^{-S})$. Observe that a 1-strong Nash equilibrium is a Nash equilibrium, an n -strong Nash equilibrium is a strong Nash equilibrium (Bernheim et al., 1987), and that the set of $(k+1)$ -strong Nash equilibria is a subset of the set of k -strong Nash equilibria.

Definition 2.2: Given a coalition $S \subseteq N$, we define an S -deviating scheme as a function $d^S : A^S \rightarrow \Delta(A^S)$. Given a strategy profile q we say that $p \in \Delta(A)$ is an S -deviation from the strategy profile q , if there exists a deviating scheme d^S , such that for all $a \in A$, we have $p(a) = \sum_{b^S \in A^S} q(b^S, a^{-S}) d^S(a^S | b^S)$. Let $D(q, S) \subseteq \Delta(A)$ denote the set of all S -deviations from q .

Thus, a correlated strategy profile p is an S -deviation from another strategy profile q if the members of S , using a plan to correlate their play (which may depend on S -part of the recommendations $a^S \in A^S$, chosen by a correlation device according to q), can induce the correlated strategy profile p when each member of the complementary coalition obeys his recommendation.

Definition 2.3: $q \in \Delta(A)$ is a k -strong correlated equilibrium if and only if for every coalition $S \subseteq N$ satisfying $|S| \leq k$, and for every S -deviation $p \in D(q, S)$, there is a player $i \in S$, s.t. $u^i(q) \geq u^i(p)$.

A k -strong correlated equilibrium is a correlated strategy profile, from which no coalition (with up to k players) has a joint deviation, which makes every member of the coalition better off. Observe that a 1-strong correlated equilibrium is a correlated equilibrium, the set of $(k+1)$ -strong correlated equilibria is a subset of the set of k -strong correlated equilibria, and that an n -strong correlated equilibrium is a strong correlated equilibrium as defined in Moreno & Wooders (1996). Other definitions of a strong correlated equilibrium in the literature differ in their assumptions about the gaming framework: whether the players of a deviating coalition can transmit private information and construct a new correlating device, and when coalitions plan their deviations: before (*ex-ante*) or after (*ex-post*) receiving the agreement recommendations. *Ex-ante* definitions can be found in Milgrom & Roberts (1996), Moreno & Wooders (1996), and Ray (1996), and *ex-post* definitions can be found in Einy & Peleg (1995), Ray (1998), and Bloch & Dutta (2007).

In our framework of cheap-talk pre-play communication, a deviating coalition can use communication channels to share private information and to construct a new correlating device, and can plan deviations before, after or during the process of receiving their recommendations. At first look, the equilibrium defined in definition 2.3 seems to be resistant only against joint deviations that are planned

at the *ex-ante* stage.⁵ However, it turns out that in this framework an *ex-ante* k -strong correlated equilibrium is also an *ex-post* k -strong correlated equilibrium (see Heller, 2008). The intuition is that a decision to play an S -deviation is common knowledge in S . This allows the use of interactive epistemology (Aumann, 1976) to show that one can assume w.l.o.g. that all the deviators in a profitable *ex-post* S -deviation share the same posterior belief about the distribution of a^S . Using this, it is possible to “emulate” any profitable *ex-post* S -deviation by a profitable *ex-ante* S -deviation.

Similar to the existing definitions of a strong correlated equilibrium, we assume that the deviating players are myopic: they do not take into account the possibility that there may be further deviations.⁶

Definition 2.4: Given a coalition $S \subseteq N$, we define an *S-replacing scheme* as a function $r^S : A^S \rightarrow [0,1]$. Given a strategy profile q and another strategy profile \tilde{p} (which we refer to as the “reaction” profile), we say that the profile $p \in \Delta(A)$ is an *S-replacement of the strategy profile q with a reaction profile \tilde{p}* , if there exists a replacing scheme r^S , such that for all $a \in A$, we have $p(a) = (1 - r^S(a^S))q(a) + \sum_{b^S \in A^S} q(b^S)r^S(b^S)\tilde{p}(a)$. Let $R(q, S, \tilde{p}) \subseteq \Delta(A)$ denote the *set of all S-replacements of q with \tilde{p}* .

Thus, a correlated strategy profile is an S -replacement of the profile $q \in \Delta(A)$ with a reaction profile \tilde{p} , if it is induced by an S -replacing scheme r^S in the following process:

- Given their recommendations $a^S \in A^S$ (chosen by a correlation device according to q) the players of S do a joint lottery.
- With probability $1 - r^S(a^S)$ they obey their recommendations, and everyone plays according to q .
- With probability $r^S(a^S)$ the correlation device is “replaced”, and everyone plays according to \tilde{p} .

Definition 2.5: Given a correlated strategy profile $q \in \Delta(A)$, we say that an uncorrelated strategy profile $\tilde{q} \in IA$ is a *k-strong punishing strategy profile* (for q) if for every coalition $S \subseteq N$ satisfying $|S| \leq k$ and for every S -deviation $\tilde{p} \in D(\tilde{q}, S)$ (from \tilde{q}), the following two conditions hold:

- $\exists i \in S$ s.t. $u^i(q) > u^i(\tilde{p})$.
- For every S -replacement (of q with \tilde{p}) $p \in R(q, S, \tilde{p})$, $\exists i \in S$ s.t. $u^i(q) \geq u^i(p)$.⁷

⁵ Because when considering whether the S -deviation p is profitable, one evaluates $u^i(q)$ and $u^i(p)$, without conditioning the utilities on *ex-post* information about the recommendations a^S .

⁶ Assuming otherwise leads to the concepts of a coalition-proof Nash equilibrium (Bernheim et al., 1987) and of a correlated coalition-proof equilibrium (defined in the literature that was referred to after definition 2.3)

⁷ The second condition can be omitted if simultaneous messages are allowed (As discussed in Section 6).

A profile q is a k -strong punishable strategy profile if there exists a k -strong punishing profile (for q).

In our cheap-talk protocol (as described in the next section) the players construct together a correlation device that recommends each player what to play according to a k -punishable strategy profile q . The punishing strategy profile \tilde{q} is used to "deter" a coalition S (of up to k players) from using an " S -lie" (defined in Section 3): deviate while communicating with the members of $-S$ in order to "manipulate" the correlation device and induce a new correlation. The properties of our protocol (presented in Section 3) guarantee that any use of an S -lie is detected by a non-deviating player with probability $\lambda < 1$ at the *ex-ante* stage, and with probability 1 at the *ex-post* stage. If an S -lie is detected, then all the non-deviating players play the (uncorrelated) punishing strategy profile \tilde{q} , and the deviating players deviate from \tilde{q} to an S -deviation $\tilde{p} \in D(\tilde{q}, S)$.

The first condition guarantees that if the detecting probability λ is large enough then every S -lie at the *ex-ante* stage is not profitable to at least one deviating player (as discussed after definition 2.6).

Assume that the players in S consider the use of an S -lie at the *ex-post* stage with probability $r^S(a^S)$ (which depends on their recommendations $a^S \in A^S$). Because the S -lie is detected with probability 1, the resulting correlated strategy profile that is induced by the S -lie is an S -replacement: $p \in R(q, S, \tilde{p})$. Thus the second condition guarantees that every S -lie at the *ex-post* stage is not profitable to at least one deviating player.⁸

Definition 2.6: Given a game G , a coalition S , a k -strong punishable strategy profile q with a k -strong punishing profile \tilde{q} , let the S -punishment be: $m_{q,\tilde{q}}^S = \min_{\tilde{p} \in D(\tilde{q}, S)} \max_{i \in S} (u^i(q) - u^i(\tilde{p}))$, and let the *minimal punishment* be: $m_{q,\tilde{q}} = \min_{S \subseteq N, |S| \leq k} (m_{q,\tilde{q}}^S)$. Let $w_q = \max_{i \in N, a \in A} (u^i(a) - u^i(q))$ be the *maximal profit*. Finally, let the *minimal detecting probability* $0 \leq \lambda_{q,\tilde{q}} < 1$ be a solution (in the interval $[0,1)$) to the equation $(1 - \lambda_{q,\tilde{q}}) \cdot w_q = \lambda_{q,\tilde{q}} \cdot m_{q,\tilde{q}}$.

Given a coalition S with up to k players and an S -deviation $\tilde{p} \in D(\tilde{q}, S)$ from the punishing profile \tilde{q} , one of the deviating players loses $\max_{i \in S} (u^i(q) - u^i(\tilde{p})) > 0$ if the profile q is replaced with \tilde{p} . The S -punishment $m_{q,\tilde{q}}^S > 0$ is the minimal such loss for all possible S -deviations $\tilde{p} \in D(\tilde{q}, S)$, and the minimal punishment $m_{q,\tilde{q}} > 0$ is the minimal S -punishment for all coalitions with up to k players. All those expressions are positive due to the first condition in definition 2.5 (for every S -deviation

⁸ When considering whether an S -replacement is profitable, one evaluates $u^i(q)$ and $u^i(p)$, without conditioning the utilities on *ex-post* information about a^S . This is justified due to similar reasons as those discussed after definition 2.3 and in Heller (2008).

$\tilde{p} \in D(\tilde{q}, S)$, $\exists i \in S$, $u^i(q) > u^i(\tilde{p})$). The expression $w_q = \max_{i \in N, a \in A} (u^i(a) - u^i(q)) \geq 0$ is the maximal profit a player may earn from replacing the profile q with another profile.

Assume that the players in S consider the use an S -lie at the *ex-ante* stage (before receiving their recommendations). If the S -lie, is undetected then the profit of any deviating player is at most w_q , and if it is detected, then there exists a deviating player (say player i) that loses at least $m_{q, \tilde{q}} > 0$. Let λ be the probability that the S -lie is detected. Player i 's expected profit from the use of the S -lie is at most: $(1 - \lambda) \cdot w_q - \lambda \cdot m_{q, \tilde{q}}$, and if $\lambda > \lambda_{q, \tilde{q}}$, then the equality in the definition of $\lambda_{q, \tilde{q}}$, $(1 - \lambda_{q, \tilde{q}}) \cdot w_q = \lambda_{q, \tilde{q}} \cdot m_{q, \tilde{q}}$, guarantees that player i loses (in expectation) if the S -lie is used.

Definition 2.7: Given a game G (with n players), let \bar{G} be the *cheap-talk extended game* with the following pre-play talk phase:

- The talk phase includes infinite number of turns.
- According to some commonly known round-robin order, at each turn t , a *communicating coalition* $S(t) \subseteq N$ is chosen, and each of its members simultaneously sends a message to everyone in $S(t)$.
- The messages are taken from some finite alphabet M (that contains the null message ϕ).

Playing phase: After the talk phase, each player i simultaneously chooses an action in A^i .

We have defined the cheap-talk to be infinite in the spirit of Aumann & Hart (2003) who discuss 2-player games, and show that any artificial restriction on the length of the conversation would limit the set of Nash equilibria in the extended game due to terminal effects propagating backwards (as in the finitely repeated Prisoner's Dilemma). Observe that in our definition all coalitions have substantial communication capabilities: every coalition S has an infinite number of turns in which it is the communicating coalition ($S=S(t)$) and its members can use those turns to share information, plan a deviating scheme and implement a correlating device.

Denote the set of all t_0 -period histories by $H_{t_0} = \prod_{t < t_0} \prod_{i \in S(t)} M$, of all t_0 -information sets of player i_0 by $H_{t_0}^{i_0} = \prod_{t < t_0, i_0 \in S(t)} \prod_{i \in S(t)} M$ ($H_{t_0}^{i_0}$ only includes the turns in which i_0 was in the communicating coalition), of all infinite histories by H_∞ and of all infinite information sets of player i_0 by $H_\infty^{i_0}$.

A (behavioral) i -strategy in \bar{G} is a pair of measurable⁹ functions $c^i = (f^i, g^i)$ where:

⁹ H_t, H_t^i have the discrete topology (all sets are measurable) and H_∞, H_∞^i have the usual product topology (the smallest σ -field containing all finite cylinders).

- $f^i : \bigcup_{t, i \in S(t)} H_t^i \rightarrow \Delta(M)$ - A function according to which player i chooses his messages (when he is included in the communicating coalition, i.e. $i \in S(t)$).
- $g^i : H_\infty^i \rightarrow \Delta(A^i)$ - A function according to which i chooses his action in the playing phase.

We use the term *protocol* to denote an uncorrelated strategy profile in \bar{G} (an n -tuple $c = (c^1, \dots, c^n)$). Given a protocol c and a history $h \in H$, we refer to $(g^1(h^1), \dots, g^n(h^n))$ as the protocol's action profile or as the (protocol's) recommendations.

Definition 2.8: Given a game G and a cheap-talk extension \bar{G} , we say that a protocol c is *finite* if there exists a random variable t_* with a finite expected value (which we call the protocol's *length*), such that for all i and for all $t > t_*$, $f_t^i = \phi$. Therefore if all the players follow the finite protocol c , then the functions g^i depend non-trivially only on the first t_* turns. Observe that only the players who play according to c are bounded by the protocol finiteness. A deviating coalition S can continue to send non-null messages in turns after t_* .

Definition 2.9: Given a game G and a cheap-talk extension \bar{G} , we say that a protocol c is **polite** if for every turn t , at most one player sends a message different than ϕ (the null message).

Definition 2.10: Given a game G and a cheap-talk extension \bar{G} , we say that a protocol c is *using only private 2-player channels* if for every turn t in which $|S(t)| > 2$, all the players send ϕ .

Definition 2.11: Given a game G , a cheap-talk extension \bar{G} and a protocol c , let $q_c \in \Delta(A)$ denote the unique probability according to which actions are chosen at the playing phase, if all the players play in \bar{G} according to c . Let $q \in \Delta(A)$ be a strategy profile. We say that *a protocol c implements q* if $q = q_c$.

Definition 2.12: Given a game G , a cheap-talk extension \bar{G} , a protocol c , and a coalition S , we say that another protocol c_S is an *S -protocol-deviation* (from c) if for every $i \notin S$, $c^i = c_S^i$ (i.e. everyone outside S plays according to c , and only the members of S may deviate).

3. A Universal Cheap-Talk Minority-Proof Protocol

Theorem 3: Let G be a game with n players, let $k < n/2$, let $q \in \Delta(A)$ be a k -strong correlated equilibrium, which is k -strong punishable, with rational parameters (i.e. $q(a)$ is rational $\forall a \in A$), and let \bar{G} be a cheap-talk extended game with a finite alphabet M (which depends on q). Then there exists a finite protocol c (an uncorrelated strategy profile in \bar{G}), which implements q and is a k -strong Nash equilibrium. Furthermore, c is a polite protocol that uses only 2-player private channels.

A proof of Theorem 3: We first give a detailed constructive description of our protocol c , and then prove that it implements q and that it is a k -strong Nash equilibrium in \bar{G} . For simplicity of presentation, we assume n to be odd and $n = 2 \cdot k + 1$. Our protocol's talk phase is divided into "constructing" phases and "lottery" phases.

The constructing phase is based on the k -private protocol presented in Ben-Or et al. (1988), which deals with fault-tolerant distributed computation. Their setup includes n players, each holding a secret input x_i , who compute n polynomials $(f_i(x_1, \dots, x_n))_{i \in N}$ (the outputs). Their protocol, if followed by all the players, allows each player i to obtain the value of $f_i(x_1, \dots, x_n)$, while not acquiring any information about the values of the other outputs or inputs: the conditional distribution of $(x_j)_{j \in N, j \neq i}$ and $(f_j(x_1, \dots, x_n))_{j \in N, j \neq i}$ given all the messages he received and sent (and his input x_i) is the same as the conditional distribution given only $f_i(x_1, \dots, x_n)$ and x_i . Moreover, if any coalition S (with up to k players) shares all the messages each of them received and sent (and their inputs $(x_j)_{j \in S}$), then the resulting conditional distribution of $(f_j(x_1, \dots, x_n))_{j \in -S}$ and $(x_j)_{j \in -S}$ is the same as the conditional distribution given only $(f_j(x_1, \dots, x_n))_{j \in S}$ and $(x_j)_{j \in S}$.

Let $d \in \mathbf{N}$ be the common denominator of $\{q(a)\}_{a \in A}$, let $p \in \mathbf{N}$ be a prime number larger than d and all $|A^i|$ ($p > d$, $\forall i, p > |A^i|$), let \mathbf{Z}_p be the finite field of integers modulo p , and let the alphabet of \bar{G} be $M = \mathbf{Z}_p \cup \phi$. Let $(f_i(x))_{i \in N}$ be polynomials over \mathbf{Z}_p which satisfy the following conditions:

- If x is chosen uniformly over $\{1, \dots, d\} \subseteq \mathbf{Z}_p$, then $\Pr(f_1(x), \dots, f_n(x) = (i_1, \dots, i_n)) = q(a_{i_1}, \dots, a_{i_n})$, where $a_{i_j} \in A^j$ is the i_j -th action of player j .
- If $x \in \{d+1, \dots, p-1\}$, then $f_1(x) = \dots = f_n(x) = 0$.

At the beginning of each constructing phase, each player randomly chooses a secret input $\tilde{s}_i \in \mathbf{Z}_p$, and the players communicate until each player i obtains the value of $f_i(\tilde{s}) = f_i(\tilde{s}_1 + \dots + \tilde{s}_n)$, which is interpreted to be the protocol's recommendation for player i : if $f_i(\tilde{s}) = l$ then he should play his l -th action in the playing phase.

In each lottery phase, the players make a joint lottery: with a large enough probability $\lambda < 1$, all the messages of the last constructing phase are revealed to everyone, each player verify that the recommendation he received is indeed equal to $f_i(\tilde{s})$, and a new constructing phase is played; with probability $1-p$ nothing is revealed, the talk phase ends (i.e. since that turn, everyone sends null

messages), and the players play the last recommended action. If at any time during the talk phase a deviation from the protocol is revealed, then the players play a punishing strategy profile.

In the next sub-section we describe the details of the constructing phase, and in the following one we describe the details of the lottery phase. The constructive phase is adapted from Ben-Or et al. (1988), and repeated here to make this paper self contained (some technical details are omitted, see Ben-Or et al., 1988; Goldreich et al., 1987).

The constructing phase:

1) *Choosing the secret action profile:* Each player $i \in N$ randomly chooses his secret input $\tilde{s}_i \in \mathbf{Z}_p$. Let $\tilde{s} \in \mathbf{Z}_p$ be the sum of all inputs $\tilde{s} = \tilde{s}_1 + \dots + \tilde{s}_n$.¹⁰ The random chosen action profile is the one interpreted by the values of $f(\tilde{s}) = (f_i(\tilde{s}))_{i \in N}$ (as described above).

2) *Sharing the secret:* Each player $i \in N$ shares his secret input $\tilde{s}_i \in \mathbf{Z}_p$ among the other players. We begin by presenting a definition for sharing a secret input by a polynomial:

Definition 3.1: Let $(y_1, \dots, y_n) \in \mathbf{Z}_p$ be some distinct ($y_i \neq y_j$) non-zero elements in \mathbf{Z}_p . We say that a secret input $s \in \mathbf{Z}_p$ is *shared among the players* by a polynomial $l(x)$:

- $l(0) = s$.
- Each player i receives the value of $l(y_i)$ (i 's "share" of the secret).
- For each coalition S with up to k players ($|S| \leq k$), the set of shares $(l(y_i))_{i \in S}$ is completely independent from the secret input s (i.e. the conditional distribution of s given $(l(y_i))_{i \in S}$ is the same as the prior distribution).

In that case, we say that the *polynomial* $l(x)$ *encodes the secret input* s .

We now describe how the second sub-phase is done. For each player $i \in N$:

- Player i randomly chooses k elements $(\tilde{b}_{(i,j)} \in \mathbf{Z}_p)_{j=1..k}$. Let $l_i(x)$ be the following polynomial over \mathbf{Z}_p : $l_i(x) = \tilde{s}_i + \tilde{b}_{(i,1)}x + \dots + \tilde{b}_{(i,k)}x^k$.
- Player i sends player j his *share*: $l_i(y_j)$.

At the end of this sub-phase, \tilde{s}_i is shared among the players by the polynomial $l_i(x)$: Every

¹⁰ All the operations described in this section (sums and multiplications) are operations in \mathbf{Z}_p (i.e. modulo p).

coalition S with up to k players ($i \notin S$) knows only k non-zero values of $(l_i(y_j))_{i \in S}$, which are completely independent of the secret value $\tilde{s}_i = l_i(0)$. Observe that every coalition with $k+1$ players can compute the interpolation polynomial $l_i(x)$ by sharing their $k+1$ shares (as done later in the revealing sub-phase), and evaluate \tilde{s}_i .

3) *Distributed Computations*: In this sub-phase the players compute together the values of $(f_i(\tilde{s}))_{i \in N}$, while not revealing any information about the secret inputs $(\tilde{s}_1, \dots, \tilde{s}_n)$. The value of $f_i(\tilde{s})$ can be computed from $(\tilde{s}_1, \dots, \tilde{s}_n)$ by a series of additions and multiplications. We have to show that we can share the result of each computation step (by a polynomial of degree k): adding two secret inputs, multiplying a secret input with a constant, and multiplying two secret inputs. We begin by showing it for the two simpler computation steps:

- *Addition of two secrets*: Let $s', s'' \in \mathbf{Z}_p$ be two secret inputs that are shared by the polynomials (of degree k) $l'(x), l''(x)$, i.e.: each player i knows the values of $l'(y_i), l''(y_i)$. In order to share $s' + s''$, each player i computes $l'(y_i) + l''(y_i)$. The resulting polynomial $l'(x) + l''(x)$ encodes $s' + s''$.
- *Multiplication of a secret with a known constant*: Let $s' \in \mathbf{Z}_p$ be a secret shared by the polynomial (of degree k) $l'(x)$. Let $b \in \mathbf{Z}_p$ be a known constant. In order to share $b \cdot s'$, each player computes $b \cdot l'(y_i)$. The resulting polynomial $b \cdot l'(x)$ encodes $b \cdot s'$.

As a corollary we have:

Corollary 3.2 (Matrix multiplication): Let A be a constant $n \cdot n$ matrix, and let each player i have a secret input $s_i \in \mathbf{Z}_p$. Let $\vec{s} = (s_1, \dots, s_n)$, and define $\vec{r} = (r_1, \dots, r_n)$ by: $\vec{r} = A \cdot \vec{s}$. Then the players can jointly compute \vec{r} , such that the only information obtained by each player i is the value of r_i .

Proof: Matrix multiplication is equivalent to computing n linear functionals. The players can independently compute each functional r_i (by using the two computation steps described above: addition and multiplication by a constant), and reveal the outcome to player i (by sending him their shares of the polynomial that encodes the value of r_i).

Now we show how the players can share the result of a *multiplication of two secrets* (by a polynomial of degree k). Let $s', s'' \in \mathbf{Z}_p$ be two secret inputs that are shared by the polynomials (of degree k) $l'(x), l''(x)$. The sharing of $s' \cdot s''$ is done as follows:

- Let $\hat{h}(x) = l'(x) \cdot l''(x)$. Note that the free coefficient of the polynomial $\hat{h}(x)$ is $s' \cdot s''$. There are two problems to use $\hat{h}(x)$ to encode $s' \cdot s''$: its non-free coefficients are not random (for example,

$h(x)$ cannot be irreducible) and its degree is $2 \cdot k$ instead of k . We fix those problems by the following two sub-steps.

- *Randomization*: Each player i randomly selects a polynomial $p_i(x)$ of degree $2 \cdot k$ with a zero free coefficient, and distributes its shares $p_i(y_j)$ among the players. Then each player i computes $\hat{h}(y_i) + p_1(y_i) + \dots + p_n(y_i)$. Let $h(x) = \hat{h}(x) + p_1(x) + \dots + p_n(x)$ be the resulting polynomial. Note that $h(0) = \hat{h}(0) = s' \cdot s^n$, and all other coefficient of $h(x)$ are completely random.
- *Degree reduction*: Let $h(x) = h_0 + h_1 \cdot x + \dots + h_{2k} x^{2k}$, let $\vec{h} = (h_0, h_1, \dots, h_{2k})$, let $s_i = h(y_i)$, and let $\vec{s} = (s_1, \dots, s_n)$. Let $g(x) = h_0 + h_1 \cdot x + \dots + h_k x^k$ be the truncation of $h(x)$, let $\vec{g} = (h_0, \dots, h_k, 0, \dots, 0)$, let $r_i = g(y_i)$, and let $\vec{r} = (r_1, \dots, r_n)$. Let $B = (b_{i,j})$ be the $n \cdot n$ matrix where $b_{i,j} = (y_j)^{i-1}$, and let P be the linear projection $(\alpha_0, \dots, \alpha_{2k}) \cdot P = (\alpha_0, \dots, \alpha_k, 0, \dots, 0)$. Observe that: $\vec{h} \cdot B = \vec{s}$, $\vec{h} \cdot P = \vec{g}$, and $\vec{g} \cdot B = \vec{r}$. Since B is not singular (because the y_i -s are distinct), we have: $\vec{s} \cdot (B^{-1}PB) = \vec{r}$. By corollary 3.2, the players can compute \vec{r} , such that the only information obtained by each player i is the value of r_i , and thus $s' \cdot s^n$ can be shared by $g(x)$.

4) *Revealing the recommendations*: All the players send each player i their shares of the polynomial that encodes $f_i(\vec{s})$, he interpolates the polynomial, and evaluates the value of $f_i(\vec{s})$.

The lottery phase:

Let $\tilde{q} \in IA$ be an uncorrelated punishing strategy profile (for q). Let λ be a rational number satisfying $\lambda_{q,\tilde{q}} < \lambda < 1$ (where $\lambda_{q,\tilde{q}}$ is the minimal detecting probability as defined in 2.6). During the lottery phase, the players perform a joint lottery¹¹ and with probability λ all the players send everyone all the messages they sent and received in the last constructing-phase ("the revealing sub-phase").¹² If a player finds out that there was a deviation in the constructing phase (i.e. a sender claims he sent a message $a \in M$ while a receiver claims he received another message $b \in M$, or some player i did not do what he was supposed to do, like sending a wrong number in some computation step or choosing a polynomial with a wrong free coefficient), then he tells everyone about it, and then the (non-deviating) players stop communicating (i.e. send a null message for the rest of the infinite talk phase) and play the

¹¹ Assuming $\lambda=l/m$ (such that l and m are natural), the joint lottery can be done by each player simultaneously tells everyone a random chosen $y_i \in (0, \dots, m-1)$ and comparing $y = (y_1 + \dots + y_n) \bmod n$ to l ($y < l$ with probability λ). The lottery can also be implemented by a polite protocol by encoding the secret y_i -s by polynomials and computing y .

¹² This can be done with only private 2-player channels: in each turn (according to some known round-robin order), player i tells player j a message he has sent/received in the last constructing-phase. After enough turns, all the players receive all the messages sent in the last constructing phase twice (once from the sender and once from the receiver).

punishing strategy profile \tilde{q} . With probability $1 - \lambda$, the content of the last constructing phase is not revealed, the (non-deviating) players stop communicating, and each player plays his part of the action profile of the last constructing phase.

Remarks:

- If all the players received a “0-recommendation” ($\forall i, f_i(\tilde{s}) = 0$, which happens when $\tilde{s} \in \{d + 1, \dots, p - 1\}$), then they play the revealing phase with probability 1. If some players claim they received a 0-recommendation while others not, then they play the punishing strategy profile.
- If at any time during the talk phase (not necessarily during the revealing sub-phase), a deviation from the protocol is revealed by a non-deviating player (for example, he receives a null message instead of a number), then the players play the punishing strategy profile \tilde{q} .

Proving the protocol is a k -strong Nash equilibrium:

Let S be a coalition with up to k players. We have to show that there is no profitable deviation for S , i.e. that in every S -protocol-deviation c_S there is a deviating player $i \in S$, such that $u^i(q_{c_S}) \leq u^i(q_c)$. The possible S -protocol-deviations can be divided into a few kinds: choosing numbers in \mathbf{Z}_p non-uniformly, sharing information, not following S -part of the action profile and S -lies (deviating while communicating with the members of $-S$ in order to change their recommendations). We show that none of those kinds (nor a combination of them) is profitable for S .

Choosing numbers non-uniformly: In the beginning of the constructing phase, each player should randomly choose a secret input $\tilde{s}_i \in \mathbf{Z}_p$. The fact that the action profile depends only on the sum of those \tilde{s}_i -s, guarantee that choosing \tilde{s}_i in any arbitrary way, does not affect the distribution of the action profile. Later, in the constructing phase, players should choose random coefficients for the polynomials $p_i(x)$ (in the randomization sub-step). Choosing the values of those coefficients (or of the secret inputs) in any arbitrary way might be done in order to gain information about the action profile, but this is equivalent to sharing information (which is discussed in the next paragraph).

Sharing information: In this deviation, S members follow the protocol c when communicating with members of $-S$, but deviate when communicating among themselves: in turns when the communicating coalition is $S(t) \in S$, they send messages that contain some information about their secret inputs or about messages they received or sent in earlier parts of the protocol. Such sharing can be done in "silent" turns (when the players are supposed to send null messages).¹³ We now show that such sharing

¹³ For example, the deviating players can use some of the infinite number of turns after the protocol ends (when they are supposed to send only null messages) for sharing information.

does not give the deviating players any non-trivial information about $-S$ part of the action profile: $(f_j(\tilde{s}))_{j \in -S}$. In each step of the constructing phase, each player i shares his secret information (his original secret input \tilde{s}_i and new inputs he obtains during the distributed computations sub-phase) by encoding them with random polynomials of degree k . The players of S know at most k shares of any encoding polynomial of a player in $-S$, and thus the coefficients of those polynomials, and specifically the free coefficient, are completely independent of their shares. Thus, the conditional distribution of $(f_j(\tilde{s}))_{j \in -S}$ and $(\tilde{s}_j)_{j \in -S}$ given all the messages the players of S received and sent (and the $(\tilde{s}_j)_{j \in S}$), are the same as the conditional distribution given only $(f_j(\tilde{s}))_{j \in S}$ and $(\tilde{s}_j)_{j \in S}$ (Ben-Or et al., 1988).

Not following their part of the action profile: The players of S may plan a deviation in the playing phase: playing according to an S -deviating scheme $d^S : A^S \rightarrow \Delta(A)$ instead of following the protocol's action profile. However, the fact that q is a k -strong correlated equilibrium and that $q = q_c$ guarantees that such a deviating scheme is not profitable for at least one player in S .

S-lies: We define an S -lie as an S -protocol-deviation in which the players of S deviate while communicating with the non-deviating players of $-S$, and as a result a non-deviating player $j \notin -S$ receives a different recommendation than $f_j(\tilde{s})$ (or does not receive a valid recommendation at all). Such S -lies can be used at the *ex-ante* stage (the first 3 sub-phases of the constructing phase when the players do not know their recommendation $a^N \in A^N$) or at the *ex-post* stage (the fourth sub-phase of the constructing phase).

We first deal with S -lies at the *ex-ante* stage. With probability λ a revealing sub-phase is played after the constructing phase. In the revealing sub-phase, the members of $-S$ share their $k+1$ (or more) shares $l_i(y_j)$ of each random polynomial $l_i(x)$ (of degree k) that was used to encode the secret inputs \tilde{s}_i . This allows each non-deviating player to interpolate the coefficients of all those polynomials, and evaluate the true value of all \tilde{s}_i . Thus the players in $-S$ can check whether any non-deviating player j received a different recommendation than $f_j(\tilde{s})$. Thus any *ex-ante* S -lie is detected with probability $\lambda > \lambda_{q, \bar{q}}$, and definition 2.6 (of the minimal detecting probability $\lambda_{q, \bar{q}}$) guarantees that any such S -lie is not profitable to at least one player in S .

We now discuss S -lies at the *ex-post* stage. At least $k+1$ non-deviating players in $-S$ send their true shares to each non-deviating player. Thus each player can interpolate his polynomial and evaluate $f_j(\tilde{s})$ based only on the shares received from the non-deviating players. Thus any deviation at that stage (sending a wrong share) is detected with probability 1, and definition 2.5 (of k -strong punishing

profile) guarantees that the S -lie is unprofitable to at least one deviating player. **QED (theorem 3).**

The reader should note that our protocol is not a $(k+1)$ -strong Nash equilibrium, as any coalition with $k+1$ players can share their $l_i(y_j)$ -s and know the action profile: $f(\tilde{s}) = (f_i(\tilde{s}))_{i \in N}$.

4. Non-Existence of a Universal Cheap-Talk ($n/2$)-Proof Protocol

In this section we show by way of examples that for every n , our result is tight: first for 2-player games (example 4.1) and then for general n -player games (example 4.2).

The following example (4.1) shows that there is a 2-player game \hat{G} and a (1-strong) correlated equilibrium \hat{q} which is 1-strong punishable (with rational parameters), such that for every protocol \hat{c} (an uncorrelated strategy profile in the extended cheap-talk game), which implements \hat{q} , there is a player with a profitable deviation in the playing phase (thus there is no Nash equilibrium in the extended game with the same payoffs as \hat{q}).

Example 4.1 (based on Ben-Porath, 1998): Let \hat{G} be a “chicken” game, as described in Table 4.1 (Alice is the row player and Bob is the column player). Let \hat{q} be the following correlated equilibrium, which is 1-strong punishable with a punishing profile $\tilde{q} = (d_1, d_2)$, as described in Table 4.2:

Table 4.1: “Chicken” payoffs

		Bob	
		e^2	d^2
Alice	e^1	3,3	1, 4
	d^1	4, 1	0, 0

Table 4.2: The correlated equilibrium \hat{q}

		Bob	
		e^2	d^2
Alice	e^1	$\frac{1}{3}$	$\frac{1}{3}$
	d^1	$\frac{1}{3}$	0

Assume to the contrary that there is a cheap-talk protocol c such that $\hat{q} = q_c$ and that c is a Nash equilibrium in \bar{G} (the extended game). First we would like to note that because the game has only two players then both players completely know the history of messages:¹⁴ i.e. $\forall t, H_t = H_t^1 = H_t^2$, $H_\infty = H_\infty^1 = H_\infty^2$. The protocol c determines a probability function \Pr_c over the set of infinite histories H_∞ . We show first that with probability 1 the recommendations to both players are pure:

- Let $H_I \subseteq H_\infty$ be the set of histories where both players mix their actions.¹⁵ Observe that $\Pr_c(H_I) = 0$ (because otherwise $q_c(d^1, d^2) > 0$).

¹⁴ W.l.o.g. we can ignore turns in which $|S(t)| = 1$ (i.e. turns in which a player communicates “with himself”).

¹⁵ While our “minority-proof” protocol chooses a pure action profile (given a certain history of messages), in general protocols may choose a random action profile.

- Let $H_{II} \subseteq H_\infty$ be the set of histories where one player (say Alice) mixes while the other player (say Bob) plays a pure action. In every $h \in H_{II}$, Alice has a profitable deviation in the playing phase: if Bob's action is d^2 , then Alice deviates and play e^1 , and if Bob's action is e^2 , then Alice deviates and plays d^1 . Thus, $\Pr_c(H_{II}) = 0$ (because we assume c to be a Nash equilibrium).

Let $H_{III} \subseteq H_\infty$ be the set of histories where the recommended action profile is (e^1, e^2) . Observe that $\Pr_c(H_{III}) = 1/3$ (because $\hat{q}(e^1, e^2) = 1/3$). We finish the proof by observing that both players have a profitable deviation from the protocol c in the playing phase of \bar{G} : playing d^i instead of e^i in every history in H_{III} , contradicting our assumption that c is a Nash equilibrium in \bar{G} .

Example 4.2 shows that for every n , there exists a game G with $2n$ players and a n -strong correlated equilibrium $q \in \Delta(A)$ with rational parameters, which is n -strong punishable, such that in every finite protocol c that implements q (in the extended game \bar{G}), there is a coalition with n players that has a profitable coalitional-deviation. In particular there is no n -strong Nash equilibrium in the extended game with the same payoffs as q .

Example 4.2: Let G be a game with $2n$ players: $\{A^1, \dots, A^n, B^1, \dots, B^n\}$. Each player in $A = \{A^1, \dots, A^n\}$ has two pure actions: $\{e^1, d^1\}$, and each player in $B = \{B^1, \dots, B^n\}$ has two pure actions: $\{e^2, d^2\}$. The payoff of the game is:

- If any two players in A played a different action (i.e. A^i played e^1 while A^j played d^1), or any two players in B played a different action, then all the players get 0.
- Otherwise (all players in A played the same action as well as all the players in B), the payoff is the same as in example 4.1 (as described in table 4.3).

Let q be the n -strong correlated equilibrium, which is n -strong punishable (with a punishing profile $(d^1, \dots, d^1, d^2, \dots, d^2)$), that is described in Table 4.4. Thus in q all players in A play the same action, as do all the players in B .

Table 4.3: G payoff

		$B = \{B^1, \dots, B^n\}$	
		(e^2, \dots, e^2)	(d^2, \dots, d^2)
$A = \{A^1, \dots, A^n\}$	(e^1, \dots, e^1)	3 to everyone	1 to $\{A^1, \dots, A^n\}$ 4 to $\{B^1, \dots, B^n\}$
	(d^1, \dots, d^1)	4 to $\{A^1, \dots, A^n\}$ 1 to $\{B^1, \dots, B^n\}$	0 to everyone

Table 4.4: n -strong correlated equilibrium q

		$B = \{B^1, \dots, B^n\}$	
		(e^2, \dots, e^2)	(d^2, \dots, d^2)
A	(e^1, \dots, e^1)	$1/3$	$1/3$
	(d^1, \dots, d^1)	$1/3$	0

Assume to the contrary that there is a finite cheap-talk protocol c (with length t_*) such that $q=q_c$ and that c is an n -strong Nash equilibrium in \bar{G} . We construct an "emulating" protocol in $\hat{\bar{G}}$ (the 2-player game of example 4.1), in which Alice and Bob talk by emulating what $A=\{A^1, \dots, A^n\}$ and $B=\{B^1, \dots, B^n\}$ would have talked in the protocol c in \bar{G} and their recommendations in the playing phase emulate the recommendations of A and B . Due to example 4.1, one of the players (say Alice) has a profitable deviation (in the playing phase of $\hat{\bar{G}}$). We end the proof by constructing a profitable deviation for A that includes information sharing among the members of A (after the original protocol ends), and playing a joint deviation in the playing phase that emulates Alice's profitable deviation in $\hat{\bar{G}}$.

We now formally present the constructions described above. We begin by defining the alphabet and the order of communication in $\hat{\bar{G}}$:

- The alphabet \hat{M} in $\hat{\bar{G}}$ is $\hat{M} = \prod_{i=1..n} M$ (where M is the alphabet of \bar{G}). We relate a message of Alice $\hat{m}^{Alice} \in \hat{M}$ to a vector of messages of A in \bar{G} : $\hat{m}^{Alice} = (m^{A^1}, \dots, m^{A^n})$, and similarly for Bob.
- The communicating coalition in turn t in $\hat{\bar{G}}$ is determined according to the communicating coalition in turn t in \bar{G} : $\hat{S}(t) = \begin{cases} \{Alice\} & S(t) \in \{A\} \\ \{Bob\} & S(t) \in \{B\} \\ \{Alice, Bob\} & otherwise \end{cases}$

In order to simplify notation, we relate to an information set of player i_0 in \bar{G} : $h_{t_0}^{i_0} \in H_{t_0}^{i_0}, h_{t_0}^{i_0} = \prod_{t < t_0, i_0 \in S(t)} \prod_{i \in S(t)} m_t^i$ as $h_{t_0}^{i_0} = \prod_{t < t_0, i_0 \in S(t)} (m_t^{A^1}, \dots, m_t^{A^n}, m_t^{B^1}, \dots, m_t^{B^n})$ and let $m_t^i = \phi$ for each $i \notin S(t)$ (i.e. we denote the message from a player which is not part of the communicating coalition in a given turn to be a null message), and similarly we relate to an information set in $\hat{\bar{G}}$: $\hat{h}_{t_0}^{i_0} \in \hat{H}_{t_0}^{i_0}, \hat{h}_{t_0}^{i_0} = \prod_{t < t_0, i_0 \in \hat{S}(t)} \prod_{i \in \hat{S}(t)} \hat{m}_t^i$ as $\hat{h}_{t_0}^{i_0} = \prod_{t < t_0, i_0 \in \hat{S}(t)} (\hat{m}_t^{Alice}, \hat{m}_t^{Bob})$ and let $\hat{m}_t^i = \phi$ for each $i \notin \hat{S}(t)$.

We continue by relating each information set of Alice in $\hat{\bar{G}}$ with a set of information sets for each $A^i \in A$ (and similarly for Bob and B):

- Given an information set of Alice in turn t_0 : $\hat{h}_{t_0}^{Alice} = \prod_{t < t_0, Alice \in \hat{S}(t)} (\hat{m}_t^{Alice}, \hat{m}_t^{Bob}) \in \hat{H}_{t_0}^{Alice}$, where $\hat{m}_t^{Alice} = (m_t^{A^1}, \dots, m_t^{A^n})$ and $\hat{m}_t^{Bob} = (m_t^{B^1}, \dots, m_t^{B^n})$, let the dual information set of A^j in turn t_0 be $h_{t_0}^{A^j} = \prod_{t < t_0, A^j \in S(t)} (m_t^{A^1}, \dots, m_t^{A^n}, m_t^{B^1}, \dots, m_t^{B^n})$ assuming that the following constraint holds: if $i \notin S(t)$ then

$m_t^i = \phi$. If the constraint does not hold, we say that $\hat{h}_{t_0}^{Alice}$ does not have a dual information set for player A^i in G . Similarly, given the information sets of the members of A in turn t_0 , $(h_{t_0}^{A^i})_{i=1..n}$ (where $h_{t_0}^{A^i} = \prod_{t < t_0, A^i \in S(t)} (m_t^{A^1}, \dots, m_t^{A^n}, m_t^{B^1}, \dots, m_t^{B^n})$), let $\hat{h}_{t_0}^{Alice}$ be the dual information set of Alice in turn t_0 : $\hat{h}_{t_0}^{Alice} = \prod_{t < t_0, Alice \in \hat{S}(t)} (\hat{m}_t^{Alice}, \hat{m}_t^{Bob}) \in \hat{H}_t^{Alice}$. We similarly define dual information sets of Bob and B .

- The dual information sets of the infinite histories and infinite information sets are similarly defined.

Let $(f^{A^1}, \dots, f^{A^n}, f^{B^1}, \dots, f^{B^n})$ be the functions according to which the players in \bar{G} (who follow the protocol c) talk in the talking phase ($f^i : \bigcup_{t, i \in S(t)} H_t^i \rightarrow \Delta(M)$), and let $(g^{A^1}, \dots, g^{A^n}, g^{B^1}, \dots, g^{B^n})$ be the (measurable) functions according to which they play in the playing phase ($g^i : H_\infty^i \rightarrow \Delta(A^i)$). We now construct $\hat{c} = (\hat{c}^{Alice}, \hat{c}^{Bob}) = ((\hat{f}^{Alice}, \hat{f}^{Bob}), (\hat{g}^{Alice}, \hat{g}^{Bob}))$, the ‘‘emulating’’ protocol in $\bar{\hat{G}}$:

- We define the talking function of Alice $\hat{f}^{Alice} : \bigcup_{t, Alice \in \hat{S}(t)} \hat{H}_t^i \rightarrow \Delta(\hat{M})$ in turn t (such that $Alice \in \hat{S}(t)$) to be: $\hat{f}^{Alice}(\hat{h}_t^{Alice}) = (f^{A^1}(h_t^{A^1}), \dots, f^{A^n}(h_t^{A^n}))$, where $h_t^{A^i}$ is the dual information set of A^i and if $A^i \notin S(t)$ then $f^{A^i}(h_t^{A^i})$ is defined to be null.¹⁶ We similarly define the talking function of Bob. Observe that the constraint mentioned earlier in the definition of a dual information set (if $i \notin S(t)$ then $m_t^i = \phi$) always holds (as can be shown by induction on t).

- We define the playing function of Alice $\hat{g}^{Alice} : \hat{H}_\infty^{Alice} \rightarrow \Delta(\hat{A}^{Alice})$ to be:

$$\hat{g}^{Alice}(\hat{h}_\infty^{Alice}) = \begin{cases} g^{A^i}(\hat{h}_\infty^{A^i}) & \text{if } g^{A^1}(\hat{h}_\infty^{A^1}) = \dots = g^{A^n}(\hat{h}_\infty^{A^n}), \\ d^1 & \text{otherwise} \end{cases},$$

where $h_\infty^{A^i}$ is the dual information set of A^i . We similarly define Bob’s playing function.

Observe that $\hat{q} = q_{\hat{c}}$. Thus due to example 4.1 one of the players (say Alice) has a profitable deviation from \hat{c} in the playing phase of $\bar{\hat{G}}$. Let $\tilde{g}^{Alice} : \hat{H}_\infty^{Alice} \rightarrow \Delta(\hat{A}^{Alice})$ be Alice playing function in her profitable deviation. Observe that \hat{c} is finite (and has the same length t_* as c), and therefore \tilde{g}^{Alice} depends (non-trivially) only on $\hat{H}_{t_*}^{Alice}$ (since after t_* , the players only send null messages): $\tilde{g}^{Alice} : \hat{H}_{t_*}^{Alice} \rightarrow \Delta(\hat{A}^{Alice})$.

We finish the proof by constructing a profitable A -protocol-deviation contradicting our assumption

¹⁶ Originally, $f^{A^i}(h_t^{A^i})$ is defined only for turns in which $i \in S(t)$.

that c is an n -strong Nash equilibrium in \bar{G} . In the A -protocol-deviation the players in A begin by talking according to the original protocol c until it ends (i.e. until t_*). After c is over, they use some turns in which $S(t) = A$ (when they should only send null messages according to the protocol) to fully share their sets of information (i.e. at the end of those turns each player in A knows $\{h_{t_*}^{A^1}, \dots, h_{t_*}^{A^k}\}$ which are the dual information sets of $\hat{h}_{t_*}^{Alice}$ in \bar{G}). Therefore, afterwards they can evaluate and play the A -strategy induced by $\tilde{g}^{Alice}(\hat{h}_{t_*}^{Alice})$ (i.e. they all play $\tilde{g}^{Alice}(\hat{h}_{t_*}^{Alice})$ and if it is a mixed strategy they make a joint lottery and play together according to the result). Observe that every player in A earns from the deviation the same profit that Alice earns from the deviation in \bar{G} (thus it is a profitable A -deviation).

Remarks about example 4.2:

- Observe that our requirement in the definition of cheap-talk (definition 2.7) that each coalition S has an infinite number of turns in which $S = S(t)$ is necessary for the proof of example 4.2. If in a different pre-play communication framework the players can limit the private communication channels of sub-coalitions (for example, at some point the grand coalition can decide that the talking phase ends, and the playing phase is immediately played), then the proof of example 4.2 does not hold.
- Our requirement that c is finite can be replaced by a requirement that the communication order gives the players of A and the players in B enough opportunities to secretly share information. For example, requiring that the order of the communicating coalitions $S(t)$ is such that between any two turns in which the players of A communicate with the players of B , there is a turn in which $A = S(t)$ and a turn in which $B = S(t)$.

5. An Example for Applicability – a 5-Player “Chicken” Game

In this section, we study a 5-player “chicken” game, in which the use of our "minority-proof" protocol can give a substantial gain to all players. Let G be the following game:

- Each of the 5 players has two pure actions: s ("swerve") and d ("drive straight")
- The payoff function is:
 - If all players play s , then everyone gets 4.
 - If up to 2 players play d , then those who played d get 5 while the others get 2.
 - If more than 2 players play d , then everyone gets 0.

The presence of a fair mediator allows the players to achieve the following correlated strategy profile q :

- With probability $3/8 = 37.5\%$: all the players play s .
- For each of the 10 couples (i,j) where $i \neq j$, with probability $1/16 = 6.25\%$: (i,j) play d , while the other players play s .

One can verify that:

- The profile q is a 2-strong correlated equilibrium with a symmetric payoff of 3.5 which is the best 2-strong correlated equilibrium symmetric payoff.
- The profile q is 2-strong punishable (with the punishing strategy $\tilde{q} = (d, d, d, d, d)$).
- The payoff of q is strictly better than the best symmetric payoff in the convex hull of Nash equilibria – 3.2. (achieved by choosing each of the ten couples (i,j) with probability 10%, and playing the Nash equilibrium in which (i,j) play d and the others play s .)

We now wish to compare our protocol with the existing protocols in the literature (Barany, 1992; Ben-Porath, 1998; Ben-Porath, 2003; Forges, 1990; Gerardi, 2004), when a fair mediator does not exist, and the players can only use cheap-talk. These protocols can implement q , but only as a (1-strong) Nash equilibrium. This implementation is "weak" in the sense that it is possible for two players (for example, players 1 and 2 in Ben-Porath, 2003) to jointly interfere with the protocol and guarantee a payoff of 5 for themselves (and 2 to the other players). Contrary to that, the use of our protocol gives a "stronger" implementation as a 2-strong Nash equilibrium. An analog example can be devised for any odd number of players.

6. Concluding Remarks:

1) Two possible extensions of our protocol are:

- Implementing a k -coalition-proof correlated equilibrium as a k -coalition-proof Nash equilibrium (Bernheim et al., 1987) in the extended cheap-talk game.
- Implementing a k -strong (or k -coalition-proof) correlated equilibrium of a Bayesian game.

2) **Ex-ante and ex-post strong correlated equilibria:** A related question is the relation between the various sets of strong correlated equilibria that were defined in the literature. In Heller (2008) we prove that the set of *ex-ante* strong correlated equilibria (as defined in Moreno & Wooders, 1996) is included in the set of *ex-post* strong correlated equilibria (as defined in any of the alternative definitions of Einy & Peleg, 1995; Ray, 1998; Bloch & Dutta, 2007, Heller, 2008). This is different than the coalition-proof case where it is known that there is no inclusion relationship between the sets of *ex-ante* and *ex-post* coalition-proof correlated equilibria (as discussed in the referred papers).

- 3) **A simultaneous protocol:** The implementation of a simultaneous protocol depends on the ability of players to send and receive messages exactly at the same time. This may be problematic in some real-world environments, and thus we have chosen to construct our protocol in the more robust form of a polite protocol. If simultaneous messages are possible, one can alter the last sub-phase of the constructing phase of our protocol (“revealing the recommendations”) into a simultaneous one, by sending the shares of the polynomials that encode the recommendations simultaneously to all players.¹⁷ This allows weakening definition 2.5 of a k -strong punishing strategy profile by omitting the second condition, and only requiring the first condition ($\forall \tilde{p} \in D(\tilde{q}, S), \exists i \in S, u^i(q) > u^i(\tilde{p})$). This weaker definition, allows the implementation of a larger set of k -strong correlated equilibria.
- 4) **Cryptographic protocols:** In situations in which the players are computationally restricted and one assumes the existence of “one-way” functions, it is possible to construct a protocol that implements any k -strong correlated equilibrium as a k -strong Nash equilibrium, without the restriction $k < n/2$, as discussed in Goldreich et al. (1987), Gossner (1998), Dodis et al. (2000), Urbano and Vila (2002), Lepinski et al. (2004), Abraham et al. (2006), and the references within.

References:

- Abraham I., Dolev D., Gonen R., Halpern J., 2006. Distributed Computing Meets Game Theory: Robust Mechanisms for Rational Secret Sharing and Multiparty Computation. Proceedings of the 25th annual ACM symposium on Principles of distributed computing, ACM, 53-62.
- Aumann, R., 1959. Acceptable Points in General Cooperative n -person Games, in Kuhn, H.W., Luce, R.D. (Eds.), Contributions to the Theory of Games IV. Princeton University Press, N.J, pp. 287-324.
- Aumann, R., 1974. Subjectivity and Correlation in Randomized Strategies. J. Math. Econ. 1, 67-96.
- Aumann, R., 1976. Agreeing to Disagree. Annals of Statistics, vol. 5, no.6, 1236-1239.
- Aumann, R., 1987. Correlated Equilibrium as an Expression of Bayesian Rationality. Econ. 55, 1-18.
- Aumman, R., Hart S., 2003. Long Cheap Talk. Econometrica 71 (6), 1619-1660.
- Barany, I., 1992. Fair Distribution Protocols or How the Players Replace Fortune. Mathematics of Operations Research. 17, 329-340.
- Ben-Or, M., Goldwasser S., Wigderson A., 1988. Completeness Theorems for Non-Cryptographic

¹⁷ When altering the protocol into a simultaneous one, a few minor changes are needed as well: making the joint lottery simultaneous (as described in the first part of footnote 11, and if a player does not send any y_i , then the other players should assume that $y_i = 0$), restricting the 1st remark on page 14 (playing the punishment if some, but not all, of the players claim to receive a 0-recommendation) only to the cases where a majority of the players claim to receive a 0-recommendation, and restricting the 2nd remark on page 14 (playing the punishment if a deviation is detected) only to the *ex-ante* stage.

- Fault-Tolerant Distributed Computation (extended abstract). Proceedings of the 20th ACM symposium on the theory of computing (STOC), ACM, 1-10.
- Ben-Or M., Rabin T., 1989. Verifiable Secret Sharing and Multiparty Protocol with Honest Majority. In Proc. 21st STOC, 73-85, ACM.
- Ben-Porath, E., 1998. Communication Without Mediation: Expanding the Set of Equilibrium Outcomes by “Cheap” Pre-play Procedures. *J. Econ. Theory* 80, 108-122.
- Ben-Porath, E., 2003. Cheap Talk in Games with Incomplete Information. *J. Econ. Theory* 108, 45-71.
- Bernheim, B., Peleg, B., Whinston, M., 1987. Coalition-proof Nash equilibria. *J. Econ. Theory* 42, 1-12.
- Bloch F., Dutta B., 2007. Correlated Equilibria, Incomplete Information and Coalitional Deviations. mimeo (based on The Warwick Economics Research Paper Series - paper 763).
- Chaum, D., Crepeau, C., Damgard I., 1988. Multi-Party Unconditionally Secure Protocols. Proceedings of the twentieth annual ACM symposium on Theory of computing, ACM, 11-19.
- Crawford, V., Sobel, J., 1982. Strategic Information Transmission. *Econometrica* 50, 579-594.
- Dodis, Y., Halevi, S., Rabin, T., 2000. A Cryptographic Solution to a Game Theoretic Problem, In *Advances in Cryptology - CRYPTO 2000*, volume 1880 of LNCS, pages 112–130. Springer-Verlag.
- Einy, E., Peleg B. (1995), Coalition Proof Communication Equilibria, in *Social Choice, Welfare & Ethics* (W. Barnett, H. Moulin, M. Salles and N. Schofield, eds.), Cambridge, New-York and Melbourne: Cambridge University Press.
- Eliasz K., 1999, Fault Tolerant Implementation, *Review of Economic Studies*, Vol. 69(3), 589-610.
- Forges, F., 1990. Universal Mechanisms, *Econometrica* 58, 1341-1364.
- Gerardi, D., 2004. Unmediated Communication in Games with Complete and Incomplete Information, *J. Econ. Theory* 114, 104-131.
- Goldreich O., Micali S., Wigderson A., 1987. How to play any mental game—a completeness theorem for protocols with honest majority, *Proc. of the 19th Annual STOC*, ACM, 1987, 218–229.
- Gossner, A., 1998. Secure Protocols or How communication Generates Correlation, *J. Econ. Theory* 83, 69-89.
- Heller Y., 2008, Ex-ante and Ex-post Strong Correlated Equilibria, mimeo, February 2008, <http://www.tau.ac.il/~helleryu/ex-ante-ex-post.pdf>
- Krishna R.V., 2004. Extended Conversations in Sender-Receiver Games, *Edinburgh School of Economics*, [ESE Discussion Papers](#) 126.
- Lehrer, E., Sorin S., 1997. One-Shot Public Mediated Talk, *Games Econ. Behav.* 20, 131-149.
- Lepinski, M., Micali, S., Peikert C., Shelat A., 2004. Completely Fair SFE and Coalition-Safe Cheap

- Talk. Proceedings of the 23rd annual ACM symposium on Principles of distributed computing, 1-10.
- Milgrom, P., Roberts, J., 1996. Coalition-Proofness and Correlation with Arbitrary Communication Possibilities, *Games Econ. Behav.* 17, 113-128.
- Moreno, D., Wooders, J., 1996. Coalition-Proof Equilibrium, *Games Econ. Behav.* 17, 80-113.
- Molotov-Ribbentrop Pact, 1939. Modern History sourcebook edited by Paul Halsall (1997), <http://www.fordham.edu/halsall/mod/1939pact.html> (visited in January 2008).
- Nash, J.F., 1951. Non-Cooperative Games, *Ann. Math.* 54, 286-295.
- Ray I., 1996. Coalition-Proof Correlated Equilibrium: a Definition, *Games Econ. Behav.* 17, 56-79.
- Ray, I., 1998. Correlated Equilibrium as a Stable Standard of Behavior, *Review of Economic Design*, 3, 257-269.
- Shamir A., 1979. How to Share a Secret. *Communications of the ACM*, vol. 22, 612-613.
- Urbano, A., Vila J.E, 2002. Computational Complexity and Communication: Coordination in Two-Player Games, *Econometrica* 70 (5), 1893-1927.