University of Leicester

DEPARTMENT OF ECONOMICS

MECHANISM DESIGN AND COMMUNICATION NETWORKS

Ludovic Renou, University of Leicester, UK

Tristan Tomala, HEC School of Management, France

Working Paper No. 08/35
September 2008
Updated January 2010

Mechanism Design and Communication Networks [*]

Ludovic Renou[†] & Tristan Tomala[‡]

January 5, 2010

[*] We thank Dirk Bergemann, Subir Bose, Martin Cripps, Gianni De Fraja, Johannes Hörner, Stefano Lovo, Claudio Mezzetti, Jérôme Renault, Karl H. Schlag, Sylvain Sorin, Nicolas Vieille, Yannick Viossat, Piercarlo Zanchettin and seminar participants at: the First Northwestern-PSE Transatlantic Theory Workshop, Warwick CRETA Workshop, Cowles Summer Conference in Economic Theory, UCL Theory seminar, Paris Game Theory Seminar, for helpful discussions and comments. Anonymous referees and the editor are gratefully acknowledged for suggestions improving the paper. We owe this piece of work to a discussion Murali Agastya and one of the author had few years ago. Renou thanks the hospitality of Fuqua Business School at Duke University.

[†] Department of Economics, University of Leicester, Leicester LE1 7RH, United Kingdom. lr78@le.ac.uk

[‡] Department of Economics and Decision Sciences, HEC Paris, 78351 Jouy-en-Josas Cedex, France. tomala@hec.fr

Abstract

This paper studies a mechanism design model where the players and the designer are nodes in a communication network. We characterize the communication networks (directed graphs) for which, in any environment (utilities and beliefs), every incentive compatible social choice function is implementable. We show that any incentive compatible social choice function is implementable on a given communication network, in all environments with *either* common independent beliefs and private values *or* a worst outcome, if and only if the network is strongly connected and *weakly 2-connected.* A network is strongly connected if for each player, there exists a directed path to the designer. It is weakly 2-connected if each player is either directly connected to the designer or indirectly connected to the designer through two disjoint paths, *not necessarily directed.* We couple encryption techniques together with appropriate incentives to secure the transmission of each player's private information to the designer.

Keywords: Mechanism design, incentives, Bayesian equilibrium, communication networks, encryption, secure transmission.

JEL Classification Numbers: C72, D82.

1 Introduction

1.1 Overview and motivations

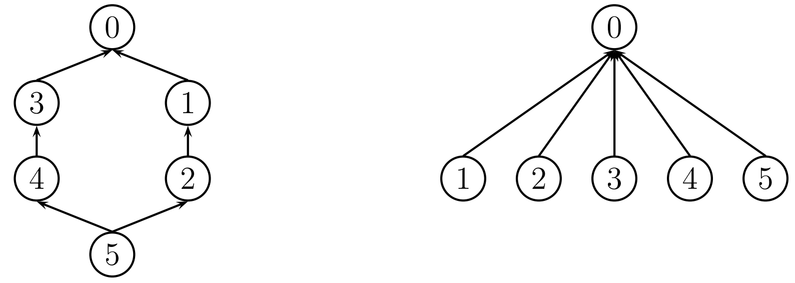
The *revelation principle* is the cornerstone of mechanism design and its applications. It asserts that the outcome of any communication system can be replicated by a direct revelation mechanism, in which agents directly and privately communicate with a designer, and truthfully report all their information (Gibbard (1973), Dasgupta, Hammond and Maskin (1979), Myerson (1979), Harris and Townsend (1981), Myerson (1982)). As a technical result, the revelation principle is a blessing. It allows to abstract away from the very details of communication systems and to focus on the social choice functions to be implemented. At the same time, it is slightly disturbing, as it implies that no decentralized communication system, however sophisticated, can dominate the centralized (direct) communication system. Yet, real-world organizations seldom take the form of centralized communication systems. The aim of this paper is to characterize the communication systems which replicate the incentive properties of centralized communication and thus, to show that incentive considerations *alone* can already explain the existence of a large variety of real-world organizations.[1]

Communication systems are naturally modeled as directed networks (graphs), in which the nodes represent the players and the designer. A player can directly communicate with another player if there exists a directed edge from that player to the other. We then associate communication networks with social environments representing the preferences and beliefs of the players, and characterize the topology of communication networks for which, in any environment, *every* incentive compatible social choice function is implementable.

The connectivity of communication networks is at the center of our analysis. A directed network is *strongly 1-connected* if for each player, there exists a directed path from this player to the designer. This is a minimal requirement that ensures that the designer may receive information from each player. A directed network is *weakly 2-connected* if each player i is either directly connected to the designer or has two disjoint

[1] There is a recent literature labeled as *algorithmic mechanism design*, that focuses on communication complexity and mechanism design (see Nisan et al. (2007) for an excellent exposition and Nisan and Segal (2006) and Van Zandt (2007) for economic applications.) Unlike this literature, we abstract from complexity considerations and entirely focus on incentives.

paths to the designer in the associated *undirected* graph. Figure 1 gives two examples of weakly 2-connected networks with a hierarchical structure. Our analysis shows that in a large class of environments, both networks have the very same incentive properties. Thus, other features, e.g. span of control, are needed to discriminate among them. We discuss some of these features in Section 1.3.



3-layer hierarchy

1-layer hierarchy

Figure 1: Two communication networks (hierarchies)

Our first main result states that any incentive compatible social choice function is implementable on a given communication network, in all environments with common independent beliefs and private values, if and only if the network is weakly 2-connected and strongly 1-connected.(In the sequel, we omit the condition of strong 1-connectedness.) The intuition for this result is as follows. A social choice function is (Bayesian) incentive compatible if, when each player expects the others to tell the truth, then no player has an incentive to lie about his own private information. Importantly, players use their prior beliefs to form their expectations. However, in a general communication network, players receive messages from their neighbors and thus, their incentives to tell the truth may be altered (since their posterior beliefs may differ from their prior beliefs). To circumvent this problem, we couple encryption techniques and incentives to transfer "securely" each player's private information to the designer through the network. Our encoding technique guarantees that no player learns anything about the types of the other players and therefore, posterior beliefs are equal to prior beliefs. To illustrate, assume that the network is *strongly* 2-connected, that is, each player is either directly connected to the designer or has two disjoint *directed* paths of communication to the designer. A player can thus send a private "encoding" key to the designer through one path and his type encoded with the key, a "cypher-type," through the other (disjoint) path. However, this is not sufficient: players must also have an incentive to truthfully forward the messages they receive. Our technique guarantees furthermore that a player's expected payoff is independent of the messages he forwards and therefore, has an incentive to truthfully forward the messages of their neighbors. Incentive compatibility ensures that players also have an incentive to truthfully report their own private information.

In this construction, the encoding technique is tailored to a given common and independent prior and cannot be generalized without this assumption. To accommodate more general beliefs, we resort to a different encoding technique which "authenticates" the player's messages in that if a player does not truthfully forward the messages of his neighbors, the deviation is detected with arbitrarily high probability. In environments with a worst outcome, the threat to be punished upon detection of a false report deters players from lying about the messages of their neighbors. Again, incentive compatibility ensures that players also have an incentive to truthfully report their own private information. This is our second main result.[2]

We now offer some motivations for our study. Firstly, as in Bolton and Dewatripont (1994), we implicitly assume that the communication network (the internal organization of the firm) is established in a prior stage and that it is relatively costly to modify. Consequently, if the designer is uncertain about which incentive compatible social choice functions he will actually have to implement, it is optimal to choose a network in the class of weakly 2-connected networks. Alternatively, we can think of our study as a worst-case analysis: If the communication network is not weakly 2-connected, there exist incentive compatible social choice functions that cannot be implemented on that network. Secondly, the previous discussion suggests that the cost of forming a link between any two agents is an important determinant in choosing among different networks (organizations). How costly is it to form such a link? To answer this question, we need to carefully interpret what a link is in our model. A link between two agents is a perfectly secure channel of communication, i.e., no other agent can eavesdrop, alter or intercept messages sent over the link, and any message sent is received with certainty. Private face-to-face communication is probably the closest instance of such perfectly secure communication in real life.[3] Such links are

[2] To the best of our knowledge, the authentication method we use is new.

[3] E-mails, phone calls or text messages are not examples of perfectly secure and reliable channels

relatively costly to establish as argued by computer scientists, see e.g., Beimel and Franklin (1999). Furthermore, Friebel and Raith (2004) argue that even if it were possible to create at no cost such perfectly secure communication links between each agent and the designer in an organization, it may not be optimal to do so. In their words, *"requiring intra-firm communication to pass through a "chain of command" can be an effective way of securing the incentives for superiors to recruit and develop the best possible subordinates."*

Related literature. The computer science literature on secure transmission of messages is closely related to this paper. Section 3.3 provides an in-depth discussion of this literature and its relationships to our study. The paper most closely related to our work is Monderer and Tennenholtz (1999), who study a similar problem to ours. Our paper substantially generalizes their results in several dimensions. Firstly, these authors consider *undirected* networks and environments with a worst outcome, common independent beliefs and private values. They show that 2-connectedness of the network is a sufficient condition for the implementation of all incentive compatible social choice functions. Crucially, in their model, edges are not directed and thus can be used to communicate in both directions. It follows that the 2-connectedness of the undirected network guarantees the existence of directed sub-networks that are strongly 2-connected. Their protocol (mechanisms and strategies) heavily exploits this fact and indeed breaks down if the undirected network does not have an underlying strongly 2-connected network. We show that in environments with common independent beliefs and private values, weak 2-connectedness, a weaker requirement than strong 2-connectedness, is a necessary and sufficient condition (the assumption of a worst outcome is superfluous). Secondly, we show that in environments with a worst outcome, weak 2-connectedness is again a necessary and sufficient condition; no further assumption on the environment is needed. We need to resort to different encryption techniques than the ones used in Monderer and Tennenholtz (1999), which would fail without common independent beliefs even on strongly 2-connected networks. Furthermore, with the very same techniques, we show that strong 3-connectedness is a sufficient condition for the implementation of all incentive compatible social choice functions in all environments. Again, the techniques of Monderer and Tennenholtz (1999) would

of communication as the recent scandal News of the World demonstrates (Guardian, 14 July 2009). In fact, if they were, there would be no need for encryption devices.

fail here.

1.2 A simple example

We now illustrate our main results within the context of a simple example. There are three players, labeled 1, 2 and 3, two types for player 2, labeled θ and θ', and two alternatives a and b. Player 2's preferences over these alternatives depend on his type (in all examples, preferences are strict). Player 2 prefers a to b if his type is θ and prefers b to a if his type is θ'. Player 1 always prefers a to b, while player 3 always prefers b to a. The designer aims at implementing the social choice function f^* that selects the preferred alternative of player 2 for each of his type: player 2 is dictatorial.

If player 2 can securely and directly communicate with the designer, f^* is clearly implementable: the designer can simply ask player 2 to directly report his preferred alternative. Suppose now that player 2 cannot directly communicate with the designer and consider the communication network $\mathcal{N}_2$ in Figure 2 (player 0 is the designer).

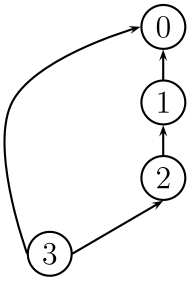


Figure 2: Communication network $\mathcal{N}_2$

With the communication network $\mathcal{N}_2$, player 2 can indirectly communicate with the designer through player 1. Moreover, player 3 has two disjoint paths of communication to the designer with player 2 on one of them. Consequently, player 2 has two disjoint paths to the designer, but one of them is not directed. The network $\mathcal{N}_2$ is thus *weakly 2-connected.* The idea is then to use the two disjoint paths from 3 to 0 to secure the communication of player 2's type to the designer, without revealing information to the other players. So, suppose that players 1 and 3 believe that player 2's type is θ with probability 1/3, independently of their own types. The goal is to design a mechanism and an equilibrium such that the designer implements a in state θ and b in state θ'.

The mechanism allows player 3 to send a real number in $[0, 1)$ to player 2 and another real number in $[0, 1)$ to player 0. Similarly, player 2 (resp., player 1) can send a real number in $[0,1)$ to player 1 (resp., player 0). An informal description of the strategies is as follows. Independently of his type, player 3 draws an "encoding key" y uniformly on $[0,1)$ and sends it to both players 0 and 2. Player 2 of type θ (resp., θ') draws a "pseudo-type" $\tilde{x}$ uniformly on $[0,1/3)$ (resp., $[1/3,1)$). The pseudo-type thus "reveals" θ, but its unconditional distribution is uniform on $[0,1)$.[4] Then, player 2 encodes his pseudo-type $\tilde{x}$ with the encoding key y received from player 3 to obtain the "cypher-type" $x = (\tilde{x} + y)\,\mathsf{mod}_{0,1}$.[5] Player 2 sends x to player 1. Player 1 has to correctly forward the message of player 2 to the designer. Let $(\hat{x}, \hat{y})$ be a pair of messages received by the designer. The allocation rule is the following: If $(\hat{x} - \hat{y})\,\mathsf{mod}_{0,1} \in [0,1/3)$, the designer implements a and implements b, otherwise.

If the players follow the prescribed strategies, $\hat{y} = y$, $\hat{x} = x$ and $(\hat{x} - \hat{y})\,\mathsf{mod}_{0,1} = \tilde{x}$. Thus, the designer correctly learns player 2's type and implements the desired social choice function f^*. In particular, players 1 and 3 expect the designer to implement a with probability 1/3 and b with probability 2/3. We now show that the players do not have an incentive to deviate from the prescribed strategies. Suppose that player 1 deviates and sends a message $\hat{x}$ to the designer instead of x. The designer implements the alternative a if $(\hat{x} - y)\,\mathsf{mod}_{0,1} \in [0,1/3)$ and b, otherwise. Since y is uniformly distributed, so is $(\hat{x} - y)\,\mathsf{mod}_{0,1}$ (see Lemma 2 in Appendix). Accordingly, player 1 expects the designer to implement a with probability 1/3 and b with probability 2/3. Thus, player 1's expected payoff does not depend on the message $\hat{x}$ that he sends. Player 1 has therefore no incentive to deviate. A similar argument applies to player 3. As for player 2, he has no incentive to deviate since f^* is incentive compatible.

It is worth stressing that the essential feature of the network is its weak 2-connectedness. For instance, if in addition to the links shown in Figure 2, player 3 has a link to player 1, the result remains valid (the network remains weakly 2-connected). Indeed, we can construct a "babbling equilibrium" in which player 3 sends an uninformative message to player 1, and player 1 plays independently of player 3's message. Alternatively, and more simply, we may let the message space from player 3 to player 1 be a singleton. In effect, we show that the weak 2-connectedness of the network is a necessary and sufficient condition for the implementation of any incentive compatible social choice

[4] More precisely, denote $\mathcal{U}_{[0,1/3)}$ (resp., $\mathcal{U}_{[1/3,1)}$) the uniform distribution on $[0,1/3)$ (resp., $[1/3,1)$). The unconditional distribution of $\tilde{x}$ is $\frac{1}{3}\mathcal{U}_{[0,1/3)} + \frac{2}{3}\mathcal{U}_{[1/3,1)} = \mathcal{U}_{[0,1)}$, the uniform distribution on $[0,1)$.

[5] For a real number r, $r\,\mathsf{mod}_{0,1} = r - \lfloor r \rfloor$, with $\lfloor r \rfloor$ the highest integer less or equal to r.

functions in environments with independent common beliefs and private values.

A further and important feature of the proposed mechanism and strategies is that players 1 and 3 learn nothing about player 2's type. This is clearly true for player 3 as he does not receive a message from player 2. As for player 1, we prove that the message x (the cypher-type) he receives is uniformly distributed on $[0,1)$ and independent of player 2's type. This feature is crucial for the implementation of incentive compatible social choice functions which depend on the private information of all players. It guarantees that posterior beliefs are equal to prior beliefs and consequently, that players' incentives to truthfully reveal their own private information are not altered.

Another important aspect is that the mechanism and strategies are tailored to environments with common independent beliefs and private values. Firstly, the partition of $[0,1)$ into $\{[0,1/3),[1/3,1)\}$ is such that the Lebesgue measure of each subset exactly matches the *common* prior beliefs of players 1 and 3. This allows the unconditional distribution of the pseudo-type to be uniform. By contrast, suppose that player 1 believes that player 2's type is θ with probability 2/3. With the above strategies, player 1 expects the designer to decode player 3's type as being θ with probability 1/3, which is different from his prior belief 2/3. Consequently, player 1's incentive to truthfully report his private information might be altered.[6] Secondly, to understand the importance of the private value assumption, suppose that player 1 prefers b to a when player 2's type is θ and a to b when player 2's type is θ' (interdependent values). If player 1 truthfully forwards the message x he received from player 2, the alternative a is implemented if and only if player 2's type is θ and the alternative b is implemented if and only if player 2's type is θ'. However, if he sends a message $\hat{x}$ independently of the message received from 2, both alternatives a and b are implemented with positive probability, regardless of player 2's type, a profitable deviation for player 1. In sum, the problem with more general environments is not only to guarantee that no information is revealed, but to provide players with incentives to truthfully communicate their private information and the messages they receive.

[6] For instance, take $\Theta_1 = \Theta_2 = \{\theta, \theta'\}$, three alternatives a, b, c, and $u_1(a,\theta) = 3/2$, $u_1(b,\theta) = 1$ and $u_1(c,\theta) = 0$. Consider the social choice function f which depends only on players 1 and 2's types with $f(\theta,\theta) = a$, $f(\theta',\theta) = f(\theta,\theta') = c$ and $f(\theta',\theta') = b$. This is incentive compatible for player 1 at state θ when he believes that player 2's type is θ with probability 2/3, but not when he believes that player 2's type is θ with probability 1/3.

With more elaborated encryption techniques, our result remain valid in environments with a worst alternative (Theorem 2). The intuition is as follows. Consider again the network $\mathcal{N}_2$. Player 3 draws a large number of independent encoding keys $y_1, \ldots, y_\eta$ and send them to players 0 and 2. Player 2 privately chooses one of these keys (with equiprobability) and uses it to encrypt his type. He then sends to player 1 the encrypted type and the unused keys, *without telling him which key was used for coding.* Player 1 has to correctly forward player 2's message to the designer. The designer compares the two vectors he receives. If these vectors differ by exactly one component η^*, he infers that the key y_{η^*} transmitted by player 3 was used for coding, and decodes player 2's type accordingly. Otherwise, the designer implements the worst alternative. This encoding technique guarantees that players 1 and 3 learn nothing about player 2's type and allows the designer to detect unilateral deviations with arbitrarily high probability, since the index η^* is the private information of player 2. In turn, the threat to implement the worst alternative upon detection of a deviation deters players from deviating.

To conclude, we preview some secondary aspects of our analysis. Firstly, the use of probabilistic coding implies that our equilibria are in mixed strategies. This point is crucial as the social choice function f^* in our example is not implementable in pure equilibria. Section 4.2 elaborates on this issue. Secondly, unlike the computer science literature, our encoding technique relies on transmitting real numbers, which may not have a finite binary expansion. This choice is well in accordance with the implementation literature: a social choice function is implementable if there exists a mechanism, possibly with continuous action spaces, and an equilibrium which corresponds to the social choice function at each state. With some modifications, our main results can be obtained with finite message spaces (see Section 3).

1.3 Applications

Our study has several implications for the optimal design of organizations (see Mookherjee (2009) for a survey). Indeed, we show that all weakly 2-connected networks have the very same incentive properties and thus, cannot be discriminated according to that property. So, which of the several weakly 2-connected networks (or hierarchies) should the designer choose?

Following Williamson (1967) and Calvo and Wellisz (1978), we may discriminate hierarchies according to their *span of control*, i.e., the number of subordinates that any manager can supervise. In a 1-layer hierarchy (the direct communication network), the designer has to directly supervise all of his employees, i.e., his span of control is the entire workforce. Alternatively, within the class of weakly 2-connected networks, there exists a network whereby the designer supervises two employees and each employee supervises a single other employee. The span of control of each manager is thus at most two. We call this network the n^*-layer hierarchy.[7] See Figure 1 for an example with $n^* = 3$. Now, if the span of control to any manager is limited (due to limited time, attention, or resources), control losses arise: information transmitted through the hierarchy might be distorted. However, as both networks have the very same incentive properties (i.e., any information the designer would have received in a 1-layer hierarchy is also received in a n^*-layer hierarchy), the n^*-layer hierarchy dominates the 1-layer hierarchy in that it minimizes the span of control of each manager, while retaining the same incentive properties.

Another distinctive mark of hierarchies is their costs of communication and information processing. For instance, in Bolton and Dewatripont (1994), the internal organization of the firm is seen as a communication network resulting from the trade-off between specialization in information processing and the cost of communication required to coordinate the activities of the firm.[8] They show that the efficient networks are pyramidal hierarchies, where each agent sends his information to at most one other agent and a unique agent (the designer) receives all the processed information. (See Radner (1993) for similar results.) Within the class of weakly 2-connected networks, only the 1-layer hierarchy is a pyramidal hierarchy. The workload of the designer in a 1-layer hierarchy is abyssal, however: he has to read the reports of all his subordinates and process all their informational contents. So, if the designer has limited resources to devote to this task, this is likely to be infeasible in practice. While there is no other pyramidal hierarchy in the class of weakly 2-connected networks, the n^*-layer hierarchy is almost pyramidal: all agents but one (the agent at the bottom) send their information to a unique superior and the designer has to process the information of only two subordinates. However, our construction requires the designer

[7] $n^* = (n+1)/2$ if n is odd and $n^* = (n+2)/2$ if n is even.

[8] Note that they abstract away from incentive considerations.

to process twice many messages. As in Bolton and Dewatripont (1994), suppose that there is a fixed cost λ to connect the designer with a subordinate and a variable cost a per message processed. The total communication cost to the designer is then $n(\lambda + a)$ with the 1-layer hierarchy and $2\lambda + (2n - 2)a$ with the n^*-layer hierarchy, where n is the number of employees. Consequently, if the cost of processing information is smaller than the cost of linking a subordinate with the designer ($a < \lambda$), the n^*-layer hierarchy is more cost efficient than the 1-layer hierarchy. This reasoning naturally focuses on the cost of linking and communicating with the designer and assumes that it is sufficiently higher than the cost of linking agents to each other. Anecdotal evidence suggests that this reasoning is adopted in practice.[9]

2 Definitions

The primitives of the model consist of two essential ingredients: social environments (players, outcomes and preferences) and communication networks.

A ***social environment*** $\mathcal{E}$ is a tuple $\langle N, A, (\Theta_i, P_i, u_i)_{i \in N} \rangle$ where $N := \{1, \ldots, n\}$ is the set of players, A the finite set of alternatives, and Θ_i the finite set of types of player $i \in N$.[10] Let $\Theta := \times_{i \in N} \Theta_i$ and $\Theta_{-i} := \times_{j \in N \setminus \{i\}} \Theta_j$, with generic elements θ and θ_{-i}, respectively. Each player knows his own type and player i of type θ_i holds a probabilistic belief $P_i(\cdot|\theta_i)$ over Θ_{-i}. Throughout the paper, we assume $P_i(\theta_{-i}|\theta_i) > 0$ for all $(\theta_i, \theta_{-i}) \in \Theta$ and for all $i \in N$. Each player has a preference relation over alternatives, which is representable by the type-dependent utility function $u_i : A \times \Theta \rightarrow \mathbb{R}$. Players are expected utility maximizers. Three properties of an environment are of particular importance to our analysis:

- The environment has a *common prior* if there exists a probability distribution P on Θ such that $P_i(\theta_{-i}|\theta_i)$ is the conditional distribution of θ_{-i} given θ_i derived from P. The common prior is *independent* if P is the product of its marginal

[9] For instance, since August 2009, the University of Leicester has adopted a new internal organization, whereby departments are now subordinated to four newly created colleges. Previously, the departments were directly subordinated to the office of the Vice-Chancellor. The main rationale evoked for this change of organization was the need to reduce the burden of communication to the Vice-Chancellor of the previous and more direct organization.

[10] In Section 4, we extend our analysis to environments with infinite type spaces.

distributions.

- The environment has *private values* if for each player i, his utility function does not depend on the types θ_{-i} of his opponents.

- The environment has a *worst outcome* if there exists an alternative $\underline{a} \in A$ such that for each player i, each type profile θ and each alternative $a \in A \setminus \{\underline{a}\}$, $u_i(\underline{a}, \theta) < u_i(a, \theta)$.

A social choice function $f : \Theta \to A$ associates with each type profile θ an alternative $f(\theta) \in A$. A social choice function is *incentive compatible* if for each player $i \in N$, for each pair of types (θ_i, θ_i') of player i, we have

$$\sum_{\theta_{-i}} u_i(f(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}) P_i(\theta_{-i}|\theta_i) \geq \sum_{\theta_{-i}} u_i(f(\theta_i', \theta_{-i}), \theta_i, \theta_{-i}) P_i(\theta_{-i}|\theta_i).$$

Note that our definition of a worst outcome is stronger than actually required; it would be enough to consider an alternative worse than any alternative in the range of the social choice function we aim to implement. Exchange economies with free disposal are examples of environments with worst outcome: the zero allocation is a worst outcome if preferences are strictly monotonic and the social choice function selects positive vectors of goods. Similarly, in quasi-linear environments, the assumption of a worst outcome is natural.

A ***communication network*** captures the possibilities of communication between the players and the designer. A communication network is a *directed* graph with $n + 1$ vertices representing the n players and the designer (henceforth, player 0). There is a directed edge from player i to player j, denoted ij, if i can send a message to j. Formally, the network, denoted by $\mathcal{N}$, is defined as a set of edges $\mathcal{N} \subseteq (N \cup \{0\}) \times (N \cup \{0\})$. We denote $C(i) = \{j \in N \cup \{0\} : ij \in \mathcal{N}\}$ the set of players to whom player i can directly send a message. Similarly, we denote $D(i) = \{j \in N \cup \{0\} : ji \in \mathcal{N}\}$ the set of players who can directly send a message to player i. A *directed path* in $\mathcal{N}$ is a finite sequence of vertices $(i_1, \ldots, i_m)$ such that $i_k i_{k+1} \in \mathcal{N}$ for each $k = 1, \ldots, m-1$. A communication network $\mathcal{N}$ is *strongly m-connected* if for each player $i \in N \setminus D(0)$, there exist m disjoint directed paths (i.e., having no common vertex except i and 0) from player i to the designer. By convention, the communication network is strongly n-connected if $N \setminus D(0) = \emptyset$. A network of particular importance is the star network

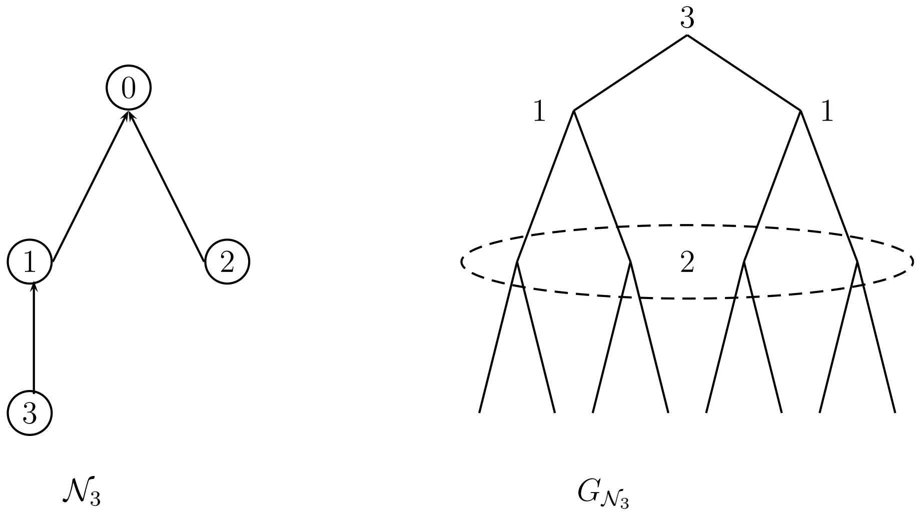
$\mathcal{N}^\star$ with the designer as the center and $D(i) = \emptyset$, $C(i) = \{0\}$ for all player $i \in N$. With the star network, each player communicates directly and privately with the designer; the star network is n-connected.

We make the following assumptions on the network throughout the paper. Firstly, we assume that networks are strongly 1-connected: for each player $i \in N$, there exists a directed path from i to 0. This assumption ensures that the designer may receive information from each player. Secondly, we assume that the graph is acyclic, that is, for each $i \in N \cup \{0\}$, there is no path from i to himself. In particular, these two assumptions imply that $C(0) = \emptyset$, i.e., the designer cannot send messages to the players. In other words, as in the classical model of mechanism design, the designer does not communicate with the players: he merely collects information and implements outcomes accordingly.

Now, we describe the interaction between a social environment and a communication network. The important feature of our model is that players can only send messages to players they are directly connected to. The interaction (the extensive-form) unfolds as follows.

- Each player i "reads" the messages he receives from players in $D(i)$. Then, he sends messages to players in $C(i)$ (he may send different messages to different players).
- The designer "reads" the messages he receives from players in $D(0)$ and selects an alternative.

Note that if $\mathcal{N} = \mathcal{N}^\star$, this corresponds to the classical model where each player communicates directly and privately with the designer. Acyclicity and strong 1-connectedness of the graph implies that the interaction as described above gives rise to a well-defined extensive-form. With acyclicity, the communication rule stating that "*a player sends his messages after having received all his messages*" generates a well-defined timing structure, where each player i is assigned a stage $t(i)$ at which he sends his messages. This statement is proved in Appendix. For instance, in Figure 3, player 3 can directly communicate with player 1, but not with player 2 and the designer. In the associated extensive-form, player 3 communicates first with player 1, and after observing player 3's message, player 1 communicates with the designer. The assumptions of directed networks, inactive designer and acyclicity (i.e., each player "speaks" only once) make our problem of implementation the hardest. Section 4 discusses several extensions. In particular, we discuss an extension of the model where the designer may send messages and show how to adapt our results.



$\mathcal{N}_3$ $G_{\mathcal{N}_3}$

Figure 3: Network $\mathcal{N}_3$ and a consistent extensive-form $G_{\mathcal{N}_3}$

A mechanism is a pair $\langle (M_{ij})_{ij\in\mathcal{N}}, g\rangle$ where for each edge ij, M_{ij} is the set of messages that player i can send to player j, and $g : \times_{i\in D(0)} M_{i0} \to A$ is the allocation rule. Note that the allocation rule depends only on the messages the designer can receive. The next step is to define the Bayesian game induced by a mechanism, a communication network and an environment.

Fix an environment $\langle N, A, (\Theta_i, P_i, u_i)_{i\in N}\rangle$, a communication network $\mathcal{N}$ and a mechanism $\langle (M_{ij})_{ij\in\mathcal{N}}, g\rangle$. Define $M_{D(i)} := \times_{j\in D(i)} M_{ji}$ as the set of messages that player i can receive and $M_{C(i)} := \times_{j\in C(i)} M_{ij}$ as the set of messages that player i can send. A *pure strategy* s_i for player i is a mapping from $M_{D(i)} \times \Theta_i$ to $M_{C(i)}$. We denote by S_i the set of player i's pure strategies and by $s_{ij}(m_{D(i)}, \theta_i)$ the message player i sends to player $j \in C(i)$ conditional on receiving the messages $m_{D(i)}$ and being of type θ_i. A *behavioral strategy* σ_i for player i maps $M_{D(i)} \times \Theta_i$ to $\Delta(M_{C(i)})$, the set of probability distributions over $M_{C(i)}$[11]. We denote by $\mathbb{P}_{\sigma,\theta}$ the probability distribution over profiles of messages (i.e., over $\times_{ij\in\mathcal{N}} M_{ij}$) induced by the strategy profile $\sigma = (\sigma_i)_{i\in N}$ at state

[11] We also find it convenient to view a behavioral strategy as a measurable mapping from $M_{D(i)} \times$

θ. The Bayesian game $G_{\mathcal{N}}$ induced by an environment, a mechanism and a network is defined as follows:

- The set of players is N, the set of player i's types is Θ_i and his beliefs are given by P_i.
- The set of strategies of player i is S_i.
- The payoff of player i is his expected utility conditional on his type and given that the outcomes are selected by the allocation rule g.

Definition 1 *The social choice function f is partially implementable on the communication network $\mathcal{N}$ if there exist a mechanism $\langle (M_{ij})_{ij \in \mathcal{N}}, g \rangle$ and a Bayesian-Nash equilibrium σ^* of $G_{\mathcal{N}}$ such that for all $\theta \in \Theta$, $g((m^*_{i0})_{i \in D(0)}) = f(\theta)$ for all profiles of messages $(m^*_{i0})_{i \in D(0)}$ received by the designer in the support of $\mathbb{P}_{\sigma^*,\theta}$.*

Denote $F_{\mathcal{N}}(\mathcal{E})$ the set of social choice functions partially implementable on the communication network $\mathcal{N}$ when the environment is $\mathcal{E}$. From the revelation principle, $F_{\mathcal{N}}(\mathcal{E}) \subseteq F_{\mathcal{N}^\star}(\mathcal{E})$ for every environment $\mathcal{E}$, and $F_{\mathcal{N}^\star}(\mathcal{E})$ is precisely the set of incentive compatible social choice functions. The aim of this paper is to characterize the communication networks $\mathcal{N}$ for which $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$ for every environment $\mathcal{E}$.

3 The main results

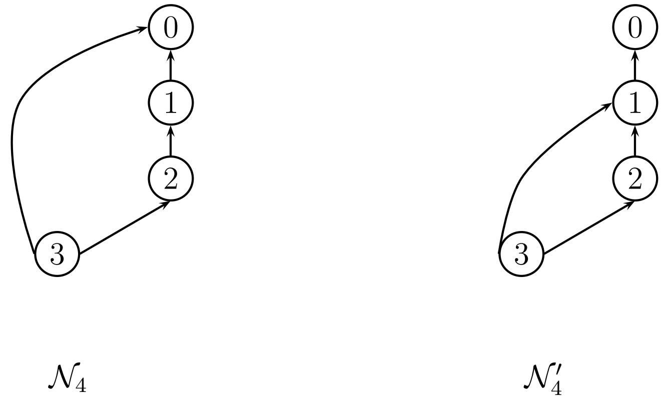
This section presents our main results regarding the partial implementation of social choice functions on communication networks. We introduce our main connectivity condition. Recall that we consider strongly 1-connected and acyclic networks. An *undirected path* in $\mathcal{N}$ is a finite sequence of vertices $(i_1, \ldots, i_m)$ such that for each $k = 1, \ldots, m-1$, either $i_k i_{k+1} \in \mathcal{N}$ or $i_{k+1} i_k \in \mathcal{N}$.

Definition 2 *The communication network $\mathcal{N}$ is* weakly 2-connected *if for each player $i \in N \setminus D(0)$, there exist two disjoint undirected paths from player i to the designer.*

$\Theta_i \times Y_i$ to $M_{C(i)}$, where $(Y_i, \mathcal{Y}_i, \mu_i)$ is a probability space independent of types and messages, i.e., a private randomization device.

In words, a network is weakly 2-connected if for each player not directly connected to the designer, there exist two disjoint paths, directed or undirected, from this player to the designer. For instance, in Figure 4, the network $\mathcal{N}_4$ is weakly 2-connected while the network $\mathcal{N}_4'$ is not. Note that in both networks, player 2 has a unique directed path to the designer and therefore, neither network is strongly 2-connected.

Importantly, if a network is not weakly 2-connected, there exists two players, i and i^*, such that all paths, directed or undirected, from player i to the designer go through player i^*. As a consequence, for each player $j \neq i$, who has a path (directed or undirected) to i, all paths (directed or undirected) from j to the designer go through player i^*. Player i^* thus "controls" all the possible messages that player i can use to communicate his private information. Player i^* even controls the messages of all players which are connected, directly or indirectly, to player i. For instance, on the network $\mathcal{N}_4'$, player 1 controls all messages that player 2 and 3 can send. These simple observations suggest that there is no hope to implement all incentive compatible social choice functions on a network which is not weakly 2-connected. We show that it is indeed the case.



$\mathcal{N}_4$

$\mathcal{N}_4'$

Figure 4: $\mathcal{N}_4$ is weakly 2-connected, $\mathcal{N}_4'$ is not

3.1 Common independent beliefs and private values

We first consider environments with common independent beliefs and private values. This assumption is common in several applications of the theory of mechanism design, e.g., auction theory (Krishna (2002)) or contract theory (Salanie (2000)). Our first result states that any incentive compatible social choice function is implementable on a network $\mathcal{N}$ for all such environments if and only if $\mathcal{N}$ is weakly 2-connected.

Theorem 1 *For all environments $\mathcal{E}$ with common independent beliefs and private values, $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$ if and only if $\mathcal{N}$ is weakly 2-connected.*

Theorem 1 extends the work of Monderer and Tennenholtz (1999) in several dimensions. Monderer and Tennenholtz consider environments and communication networks with the following properties: 1) types are independently and identically distributed, 2) a player's payoff does not depend on the private information of others (private values), 3) there exists a worst outcome (to abort the protocol) and 4) networks are undirected and repeated communication is allowed, so that each edge is directed in both ways and players may get feedback on the messages they sent. With these assumptions, they show that the 2-connectedness of the communication network is a sufficient condition for the implementation of any incentive compatible social choice function. Firstly, we show that their result extends to weakly 2-connected directed networks and that this condition is necessary. This result requires the construction of a substantially more elaborated protocol (mechanisms and strategies) than the one in Monderer and Tennenholtz (1999). Indeed, their construction relies on the existence of an underlying directed subgraph that is strongly 2-connected, so that a player can send his encrypted type on one directed path and the encryption key on the other disjoint directed path. Unlike Monderer and Tennenholtz, our assumption of weakly 2-connected networks does not guarantee the existence of two disjoint directed paths from each player to the designer. Secondly, we show that the crucial assumptions to extend their result are common independent beliefs and private values. Neither the existence of a worst outcome nor the possibility of multiple rounds of messages is essential. By contrast, Theorem 2 below shows that in environments with a worst outcome, there is no need to assume common and independent beliefs and private values. Moreover, it is important to note that the mechanism and strategies for Theorem 2 are quite different from the ones for Theorem 1. Indeed, the mechanism and strategies for Theorem 1 do not work in more general environments.

The intuition for Theorem 1 is as follows. We consider the network $\mathcal{N}_5$ in Figure 5 and show how to implement the dictatorial social choice function of player 2. Note that player 2 has a directed path of communication to the designer (through player 1) and two disjoint undirected paths of communication to the designer. However, unlike the network $\mathcal{N}_2$ in Figure 2, there is no player with a directed path to player 2 and two disjoint directed paths to the designer. This feature is essential and makes the proof of Theorem 1 quite involved for general weakly 2-connected networks (see the appendix for the general case).

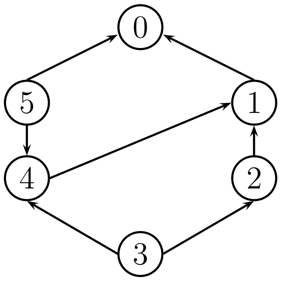


Figure 5: Communication network $\mathcal{N}_5$

As in Section 1.2, there are two alternatives a and b and two types θ and θ' for player 2. Player 2 prefers a to b if his type is θ and b to a if his type is θ'. Suppose that players 1, 3, 4 and 5 share a common prior and believes that player 2's type is θ with probability 1/3. The designer aims at implementing the dictatorial social choice function f^* of player 2.

An informal description of the strategies to implement f^* is as follows. Player 3 draws an encoding key y uniformly on $[0, 1)$ and sends it to players 2 and 4. Simultaneously, player 5 draws another encoding key z uniformly on $[0, 1)$ and sends it to the designer (player 0) and player 4. Then, player 4 encrypts the key y received from player 3 with the key z received from player 5 to obtain $w = (z + y) \,\mathsf{mod}_{0,1}$ and sends w to player 1. Player 2 of type θ (resp., θ') draws a pseudo-type $\tilde{x}$ uniformly in $[0, 1/3)$ (resp., $[1/3, 1)$) and sends the encrypted type $x = (\tilde{x} + y) \,\mathsf{mod}_{0,1}$ to player 1. Thus, player 1 receives the encrypted type x from player 2 and the modified key w from player 4. Lastly, player 1 transfers $u = (w - x) \,\mathsf{mod}_{0,1}$ to the designer. Let $(\hat{u}, \hat{z})$ be a pair of messages received by the designer. The allocation rule is the following: If $(\hat{z} - \hat{u}) \,\mathsf{mod}_{0,1} \in [0, 1/3)$, the designer implements a and otherwise, implements b.

If the players follow the prescribed strategies, then $w = (z + y) \,\mathsf{mod}_{0,1}$ and $u = (w - x) \,\mathsf{mod}_{0,1} = ((z + y) - (\tilde{x} + y)) \,\mathsf{mod}_{0,1} = (z - \tilde{x}) \,\mathsf{mod}_{0,1}$. The designer thus receives $\hat{u} = u = (z - \tilde{x}) \,\mathsf{mod}_{0,1}$ from player 1 and $\hat{z} = z$ from player 5. It follows that

$(\hat{z} - \hat{u})\,\mathsf{mod}_{0,1} = \tilde{x}$ and the designer correctly learns player 2's type and implements the desired social choice function f^*. In particular, all players but player 2 expect the designer to implement a with probability $1/3$ and b with probability $2/3$.

We now show that players do not have an incentive to deviate from the prescribed strategies and focus on player 1. From the point of view of player 1, $\tilde{x}$, y and z are mutually independent and uniformly distributed. It follows that the two messages $(z + y)\,\mathsf{mod}_{0,1}$ and $(\tilde{x} + y)\,\mathsf{mod}_{0,1}$ received by player 1 are independent and uniformly distributed (see Lemma 2 in Appendix) and convey no information about z and $\tilde{x}$. Suppose that player 1 deviates and sends the message $\hat{u}$ to the designer instead of $u = (z - \tilde{x})\,\mathsf{mod}_{0,1}$. The designer implements the alternative a if $(z - \hat{u})\,\mathsf{mod}_{0,1} \in [0, 1/3)$ and b otherwise. Since, conditionally on player 1's information, z is uniformly distributed, so is $(z - \hat{u})\,\mathsf{mod}_{0,1}$ (see again Lemma 2 in Appendix). Accordingly, player 1 expects the designer to implement a with probability $1/3$ and b with probability $2/3$. It follows that player 1's expected payoff does not depend on the message $\hat{u}$ he sends and that player 1 has no incentive to deviate. Similar arguments apply to players 3, 4 and 5. As for player 2, he has no incentive to deviate since f^* is incentive compatible.

The essential difference with the simpler example of Section 1.2 is that player 3 does not have two disjoint directed paths of communication to the designer. Thus, player 3 cannot give an encryption key to player 2 and send this key to the designer, without player 1 learning both the encryption key and player 2's encrypted type. This is precisely at this point that the protocol of Monderer and Tennenholtz fail. The novel idea is then to let player 4 encrypt the encryption key that player 3 sends to player 2, with the key received from player 5. Accordingly, player 1 receives an encrypted encryption key from player 4 and therefore, learns nothing about the type of player 2.

The proof of Theorem 1 extends these arguments to any weakly 2-connected networks (all proofs are relegated in Appendix).[12] In particular, we show that if the network is strongly 1-connected and weakly 2-connected, then there exists a protocol such that if all players abide by the protocol, the designer correctly learns the players' types and no player gets additional information about the types of his opponents. In the language of computer science, we construct a protocol for the *secret transmission*

[12] Note that the protocol (mechanism and strategies) of Monderer and Tennenholtz (1999) for undirected networks do not work in general; there is a need for encrypting encryption keys. Their protocol works only if the directed network is strongly 2-connected.

of messages. We then show that the existence of such protocol guarantees the existence of mechanism and strategies such that players are indifferent between correctly forwarding the messages they receive or lying. Thus, they indeed have an incentive to abide by the protocol. In the language of computer science, our protocol is *reliable*.

Theorem 1 also states that the weak 2-connectedness is a necessary condition to implement *all* incentive compatible social choice functions. To get some intuition for this result, let us consider a simple example. There are two players, 1 and 2, two alternatives, a and b, and two types, θ and θ' for each player. Regardless of his type, player 1 prefers a over b, player 2 of type θ prefers a over b, while player 2 of type θ' prefers b over a. Consider the social choice function f for which player 2 is dictatorial and the communication network $\mathcal{N}_6$ in Figure 6. The issue with this network, and more generally with any communication network that is not weakly 2-connected, is that player 1 controls all the information sent by player 2, and there is no way for the designer to detect a false report by player 1.

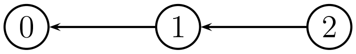


Figure 6: Communication network $\mathcal{N}_6$ is strongly 1-connected

Clearly, f is implementable on the star network $\mathcal{N}^\star$, but not on $\mathcal{N}_6$. By contradiction, suppose that f is implementable on $\mathcal{N}_6$ by the mechanism $\langle M_1, M_2, g\rangle$. There must exist an equilibrium message $m_1 \in M_1$ such that $g(m_1) = b$. However, regardless of his type and message received, player 1 has no incentives to send any message m_1 with $g(m_1) = b$, so that f cannot be implemented. The proof of Theorem 1 generalizes this argument to any network that is not weakly 2-connected.

Two further remarks are worth making. Firstly, our encoding technique extends to environments with continuous type spaces (see Subsection 4.4). Secondly, the strategies we consider are behavioral strategies. In Subsection 4.2, we prove that our result does not hold if we restrict ourself to pure equilibria, a frequently used solution concept in the mechanism design literature.

Before going further, it is worth stressing again that the encoding technique used in the proof of Theorem 1 is tailored to environments with common independent beliefs and does not apply to more general environments (even with private values). See the example in Section 1.2 for some intuition. With general beliefs, different encoding techniques have to be used: this is the object of the next section.

3.2 Worst outcome

In many concrete applications of the theory of mechanism design, players hold different and correlated beliefs about states of the world either because they have received different signals (information) or on purely subjective grounds. Moreover, the payoff of a player often depends on the private information of others. For instance, in auction models, bidders often have different information about the value of the goods for sale (e.g., mineral or oil rights) and the private information of all players influence the valuation for the good of each player. To handle these more general beliefs and payoff functions, we resort to a different encoding technique. Our new technique consists in coding the type of each player such that no information is revealed to the other players, and if a player does not truthfully forward the messages he receives, the designer detects it with arbitrarily high probability.

Theorem 2 *For all environments $\mathcal{E}$ with a worst outcome, $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$ if and only if $\mathcal{N}$ is weakly 2-connected.*

The main insight provided by Theorem 2 is that assuming a worst outcome allows to dispense with the assumptions of common independent beliefs and private values.

The intuition for Theorem 2 is as follows. We construct a mechanism such that the true type of player i is transmitted to the designer, no player $j \neq i$ gets information about the type of player i and a false report by player j is detected with arbitrarily high probability. Consider again the network $\mathcal{N}_5$ and the dictatorial social choice function of player 2.

An informal description of the strategies is the following. Player 3 sends a large number of encoding keys, all uniformly and independently drawn from $[0, 1)$ to players 2 and 4. Simultaneously, player 5 sends another large number of encoding keys all uniformly and independently drawn from $[0, 1)$ to player 4 and the designer. Player 4 thus receives a large number of keys both from player 3 and from player 5. He adds them one-by-one (addition is modulo $[0, 1)$) and sends the resulting vector of keys to player 1. Simultaneously, player 2 selects at random one of the keys received from player 3 and encrypts his type with this key. He then substitutes the selected key by the cypher-type and sends it to player 1 along with all the other keys (without telling player 1 which hey was used to encrypt his type). Lastly, player 1 received a large vector of encrypted encryption keys from player 4 and a large vector of encryption keys and the encrypted type from player 2. Player 1 then subtracts these two vectors (subtraction is component-wise modulo $[0, 1)$) and forwards the resulting vector to the designer. The designer can then detect a false report by comparing the two vectors of messages received from players 1 and 3. Namely, if player 1 truthfully forwards the message he receives, the two vectors should differ by exactly one component. In such a case, the designer decodes the type of player 2 according to this component and implements the appropriate outcome. Otherwise, the designer implements the worst outcome. By construction, only player 2 knows the key selected to encrypt his type. Thus, any deviation by players 1, 3, 4 and 5 induces the worst outcome with arbitrarily high probability: this deters them from lying.

An essential feature of Theorem 2 is the possibility to punish a detected deviation with a worst outcome. It is worth stressing, however, that our definition of a worst outcome is stronger than necessary since it does not depend on the social choice function we aim to implement. It would be enough to find an outcome worse than any outcome in the range of the social choice function.[13]

If such a worst outcome does not exist, the main difficulty for the designer is the choice of an appropriate alternative to implement whenever a false report is detected. A characterization of networks that allow to implement all incentive compatible social choice functions in all environments is left as an open problem. Yet, we provide sufficient conditions in Section 4.1. Naturally, weak 2-connectedness remains a necessary condition.

[13] It is also worth noting that Theorem 2 remains true if we consider environments with a bad outcome, i.e., an outcome $\underline{a}$ such that $u_i(f(\theta), \theta) \geq u_i(\underline{a}, \theta)$ for all $i \in N$, for all $\theta \in \Theta$. For completeness, the proof is in the appendix, Corollary 2.

3.3 Connections with computer science

An essential feature of our results is the use of encryption techniques to secure the transmission of messages from players to the designer. As already alluded in the introduction, our work is closely related to the computer science literature on secure transmission of messages, which we now review. We first discuss two important notions of security, commonly found in the computer science literature.

Message security. Informally, the transmission of a message from a sender A to a receiver B is *reliable* if A can communicate with B and no adversary, i.e., a potentially malicious third party (a hacker), can tamper with the content of the message. The transmission of a message is *secret* if no adversary finds out the content of the message sent. Information transmission is said to be *secure* if it is both reliable and secret. To discuss more precisely the notion of secrecy, let us assume that A and B have a reliable channel of communication. There are two main approaches to message security in computer science: cryptographic and information-theoretic security.

A message transmission is cryptographically secure if it is *computationally very hard* (typically NP-hard) for an adversary to find out the content of the message. This approach assumes that the adversary is computationally limited, that is, has no more computational power than a Turing Machine. The reader is referred to the seminal papers of Diffie and Hellman (1976) and Rivest et al., RSA, (1978). In particular, classical encryption techniques with *public and private keys* adopt this notion of security. For instance, the RSA encryption scheme with public keys rests on the idea that computing two large prime numbers p and q knowing their product $n = pq$ is computationally very hard.

By contrast, information-theoretic security considers adversaries with unbounded computational power and requires pieces of communication between A and B, which may be eavesdropped, to be probabilistically independent of the content of the message. This concept was originally introduced by Shannon (1949) (see also among others, Shafi and Goldwasser, 1984, Dolev et al. 1993). A simple method to achieve information-theoretic security is to map the message m to be sent to a number in, say, $\{1, \ldots, n\}$, and to add (modulo n) a uniformly distributed random key X. The encrypted message $(X + m) \bmod n$ is then uniformly distributed and independent of m: it can be publicly disclosed without harming security. The probability of guessing m correctly is $1/n$ and thus can be made arbitrarily small. Our encryption method (Lemma 2) is a continuous version of this method such that the probability of guessing correctly is zero.

As a game-theoretic model, our work follows the latter approach: the agents we consider are unboundedly rational players. These are very similar to the Byzantine adversaries considered in computer science, i.e., malicious players with unbounded computational power. The key difference, however, is that rational players respond to incentives: they do not behave maliciously if it is not optimal for them to do so.

Security in networks. Assume now that the sender A and the receiver B are some distant nodes in a network, so that there is no secure channel of communication between them. The natural question is then to characterize the networks, which guarantee the secure transmission of messages from A to B in the presence of Byzantine adversaries. This is the object of the computer science literature on secure transmission of messages. A seminal contribution is Dolev et al. (1993), who show that if the adversary controls at most t nodes, then $(2t+1)$-connectedness of the network is a necessary and sufficient condition for the secure transmission of messages from A to B. Dolev et al. assume unicast communication, i.e., a node can send different messages to its neighbors. Alternatively, Franklin and Wright (2000) study broadcast communication: any message sent by a node is automatically sent to all his neighbors. They show that $(2t+1)$-connectedness is again a necessary and sufficient for perfect security.[14]

Unlike our approach, all these results assume undirected graphs, and crucially use the possibility of messages going back and forth from the sender to the receiver (repeated communication). Dolev et al. (1993) show that in 1-way problems, i.e., if the information flows only from the sender to the receiver, a sufficient and necessary condition for the secure transmission of messages is the $(3t+1)$-connectedness of the network. Considering directed networks, Desmedt and Wang (2002) show how this bound can be lowered if there are channels of communication from the receiver to the sender. Namely, they show that if for $u \leq t$, there are $2t+1-u$ disjoint directed paths from the sender to the receiver and u disjoint directed paths from the receiver to the sender (these u paths are also disjoint from the $2t+1$ paths from the sender to the receiver),

[14] Franklin and Wright (2000) also consider a weaker notion of security: security is almost perfect when the adversary has an arbitrarily small probability of modifying the message content and to learn the content of the message. They show that $(t+1)$-connectedness is necessary and sufficient for almost-perfect security (see also Renault and Tomala, 2008).

then secure transmission of messages is possible.

Our contribution to information security. The above discussion suggests a reinterpretation of our results in the language of computer science. Starting from a communication network, a social environment and an incentive compatible social choice function f, we construct a mechanism, which implements f as a Bayesian-Nash equilibrium of the induced game. A necessary condition for this result is the possibility to construct a communication protocol with the following properties: i) the designer correctly learns the profile of types, ii) no player gets information beyond his own type, and iii) no player has an incentive to mis-execute the communication protocol. Part (ii) correspond to the computer science requirement of secrecy, while parts (i) and (iii) are the counterparts of reliability.

Before proceeding, it is worth emphasizing that the concept of Bayesian-Nash equilibrium implies that the adversary is a single potential deviant player. Such adversary has unbounded computational power, responds to incentives and controls at most one node ($t = 1$). Our main results are then reinterpreted as information transmission against this class of adversaries.

In Theorem 1, we assume common independent belief and private values, and construct a mechanism such that each player forwards the messages he receives, gets the same expected payoff regardless of the messages he forwards (see Section 1.2 and the proof of Theorem 1). With this in mind, our implementation problem is rephrased as the following problem of information transmission:

P1: *Characterize the networks for which there exists a communication protocol such that if all players abide by the protocol, the designer correctly learns the entire profile of types and no player gets additional information.*

In the presence of a worst outcome, the designer has the possibility to punish all players if he detects a deviation and we construct a protocol such that any tampering with a message is detected with arbitrarily high probability by the designer (see Section 1.2 and the proof of Theorem 2). The implementation problem gives thus rise to the following problem of information transmission:

P2: *Characterize the networks for which there exists a communication protocol such that no player gets additional information and if all but at most one player abide by the protocol, then the designer either correctly learns the entire profile of types or detects* *a deviation with arbitrarily high probability.*

Our main contribution to the literature on secure transmission of messages in networks is thus to solve problems P1 and P2 for directed graphs and 1-way problems: *the solutions are the weakly-2-connected graphs.* Compared with the computer science literature cited above, our approach through incentives allows to get a much weaker connectivity requirement. This statement is a by-product of the proofs of our main results, which are structured as follows. We first show that on any weakly-2-connected graph, there exists a communication protocol such that if all players abide by the protocol, the designer correctly learns the entire profile of types and no player gets additional information. Theorem 1 then easily follows: we use the common prior to make players indifferent between all the messages they may forward.[15] The proof of Theorem 2 uses a multiple key technique, akin to authentication schemes (see, e.g., Rabin and Ben-Or (1989)), but requires no prior knowledge of any public or private key. To the best of our knowledge, this technique is new.

To conclude, let us remark that while consistent with mechanism design theory, the use of continuous message spaces is unappealing from a computer science perspective. Yet, Theorem 1 (resp., Theorem 2) remains valid with finite message spaces, provided that prior beliefs are rational numbers (resp., that a worst outcome exists).

4 Extensions and Robustness

This section discusses various aspects of our problem and offers some generalizations.

4.1 All environments

We give sufficient conditions on the network for implementing all incentive compatible social choice functions, regardless of the environments.

Theorem 3 *If the communication network $\mathcal{N}$ is strongly 3-connected, then $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$ for all environments $\mathcal{E}$.*

The intuition is the following. Since the network is strongly 3-connected, for each player $i \in N \setminus D(0)$, there exist three disjoint directed paths from player $i \in N \setminus D(0)$

[15] We thank an anonymous referee for suggesting this structure of proof.

to the designer. For each pair of such paths, we construct a "sub-mechanism" such that any false report of messages is detected with probability 1 and which guarantees that no information about the type of player i is revealed (the construction is in the appendix, Lemma 7). A simple "majority" argument then ensures that no player has an incentive to lie. More precisely, for any unilateral deviation of player $j \neq i$, there is a pair of path from player i to the designer to which player j does not belong, and no deviation is detected on that pair of path. The designer can then correctly decode the type of player i according to the messages received from that pair of paths.

Two further remarks are worth making. Firstly, with this construction, we need each player to both draw encoding keys and to encode his type with these encoding keys. Consequently, this technique cannot be used on weakly 2-connected networks as keys might have to come from other players. Secondly, although there are three paths of communication from each player to the designer, a classical majority argument does not work in general. A player must not truthfully reveal his private information on the three paths. Simply, if a player were to do so, he would change the incentives of other players to truthfully reveal their own private information.

4.2 Pure equilibria

With the notable exception of Serrano and Vohra (2009), the literature on implementation in Bayesian environments has entirely focused on the implementation of social choice functions in *pure* equilibria (see Jackson (2001) for a survey). By contrast, the recourse to equilibria in mixed strategies is essential for our results. In effect, to transmit securely their types to the designer, it is essential for the players to encrypt their types with randomly generated keys (mixing). Although the use of randomly generated keys seems natural in our context, and indeed used in daily life (internet banking, online shopping, etc.), we might legitimately wonder whether similar results hold in environments where only pure equilibria are considered. The next theorem states that the set of social choice functions partially implementable on $\mathcal{N}$ in pure equilibria coincides with the set of incentive compatible social choice functions, irrespective of the utility functions, if and only if every player is directly connected to the designer. There is a sharp divide between implementation in pure equilibria and mixed equilibria. Denote $F^{pure}_{\mathcal{N}}(\mathcal{E})$ the set of social choice functions (partially) implementable on $\mathcal{N}$ in pure equilibria when the environment is $\mathcal{E}$.

Theorem 4 $F_{\mathcal{N}}^{pure}(\mathcal{E}) = F_{\mathcal{N}^\star}^{pure}(\mathcal{E})$ *for all environments* $\mathcal{E}$ *with common independent beliefs and private values or a worst outcome if and only if each player is directly connected to the designer i.e.,* $D(0) = N$.

The intuition is simple. If player i is not directly connected to the designer and if the social choice function depends on his type, then he must send an informative message to at least one other player, say player j. Given his updated beliefs, player j might then have no incentive to truthfully report his own private information. This reasoning is valid regardless of how many disjoint paths there are from player i to the designer.

While intuitive, Theorem 4 has remarkable implications for the topology of communication networks and implementation in pure equilibria. All but one player, say player 1, might be directly connected to the designer, player 1 might have $n-1$ disjoint paths of communication to the designer and yet, there exist incentive compatible social choice functions, which are not implementable on that network in pure equilibria. While some theorists might feel uncomfortable with equilibria in mixed strategies, the mixing through encoding techniques, as considered in this paper, seems quite natural.

4.3 Direct mechanisms

A central feature of our results is the use of encryption technique to secure the transmission of messages from the players to the designer. While the previous section shows that this is largely inescapable if we want to implement *all* incentive-compatible social choice functions, "direct" mechanisms –where players simply announce their types to their neighbors and forward messages– might suffice if we restrict attention to specific environments or to some specific incentive compatible social choice functions. For instance, consider the set of *ex-post* incentive compatible social choice functions. A social choice function f is ex-post incentive compatible if for all $i \in N$ and $\theta \in \Theta$, $u_i(f(\theta), \theta) \geq u_i(f(\theta_i', \theta_{-i}), \theta)$ for all $\theta_i' \in \Theta_i$.[16]

[16] Bergemann and Morris (2005) show that a social choice function is implementable on all type spaces if and only if it is ex-post incentive compatible.

Proposition 1 *If the communication network $\mathcal{N}$ is strongly 3-connected, then any ex-post incentive compatible social choice function is implementable on $\mathcal{N}$ by a direct mechanism.*

The intuition for Proposition 1 is simple. If a social choice function f is ex-post incentive compatible, then every player has the incentive to truthfully reveal his private information, even if he were to know the private information of some other players (e.g., his neighbors). There is therefore no particular need for encryption techniques: players can simply truthfully report their types on all paths to the designer. In the computer science terminology, secrecy is not an issue. Yet, it remains the issue of reliability: players must have the incentive to truthfully forward the messages they receive. However, with three disjoint directed paths of communication from each player $i \in N \setminus D(0)$ to the designer, a simple majority argument guarantees that no player has an incentive to misreport the messages he receives.

Furthermore, it is clear that not all ex-post incentive compatible social choice functions are implementable by direct mechanisms on weakly 2-connected networks, even in environments with common independent beliefs and private values or a worst outcome. For a counter-example, we refer the reader to the example in Section 1.2. So, weak 2-connectedness is not a sufficient condition.

In some environments, however, some ex-post incentive compatible social choice functions can be implemented by direct mechanisms, even on strongly 2-connected networks. We illustrate this possibility with the help of two important economic examples: a second-price auction and the provision of a public good.

Consider an auction with three bidders, labeled 1, 2, and 3. There is a single object to be allocated, bidder i values the object at θ_i, and bidder i's payoff is $\theta_i - x_i$ if he is allocated the object at price x_i and zero, otherwise. Consider the strongly 2-connected network $\mathcal{N}_7$ in Figure 7.

The designer aims at allocating the object to the bidder with the highest valuation (if there are several such bidders, choose one randomly). A simple and direct mechanism to implement the social choice function is as follows. Bidder 3 is required to truthfully report his valuation θ_3 to both bidders 1 and 2. Bidder 1 (resp., bidder 2) has to truthfully report his valuation θ_1 (resp., θ_2) along with bidder 3's valuation θ_3 to the designer. Let $((\hat{\theta}_1, \hat{\theta}_3^1), (\hat{\theta}_2, \hat{\theta}_3^2))$ be a profile of messages received by the designer.

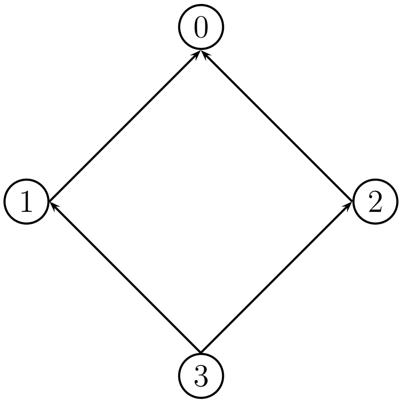


Figure 7: Communication network $\mathcal{N}_7$

The designer computes the bid-profile $(\hat{\theta}_1, \hat{\theta}_2, \max(\hat{\theta}_3^1, \hat{\theta}_3^2))$ and allocates the object to the highest bidder and charges a price equal to the second-highest bid: a second-price auction.

Since a second-price auction implements the efficient allocation in weakly dominant strategies (on the star network), no bidder has an incentive to misreport his own valuation, regardless of the reports of the other bidders. We now argue that bidder 1 has no incentive to misreport bidder 3's valuation. (A symmetric reasoning holds for bidder 2.) Clearly, if bidder 1 reports $\hat{\theta}_3^1 < \theta_3$, he does not affect the outcome since $\max(\hat{\theta}_3^1, \theta_3) = \theta_3$. Alternatively, if bidder 1 reports $\hat{\theta}_3^1 > \theta_3$, he does affect the outcome of the auction. However, this is not a profitable deviation: it not only decreases his likelihood of winning the object, but also increases the price paid if he wins.

The second example is about the provision of a public good and is adapted from Bergemann and Morris (2009). Assume that there are three players and that $\Theta_i \subseteq [0,1)$ for each player $i \in \{1,2,3\}$. The utility to player i is $(\theta_i + \gamma \sum_{j \neq i} \theta_j) x_0 + x_i$, where x_0 is the level of public good provided and x_i the monetary transfer to player i ($\gamma \geq 0$). The cost of providing the level of public good x_0 is $(1/2)(x_0)^2$. The designer aims at implementing the efficient level of public good, i.e., $(1+2\gamma)(\theta_1 + \theta_2 + \theta_3)$, at the type profile $(\theta_1, \theta_2, \theta_3)$. Again, consider the network $\mathcal{N}_7$ in Figure 7. As in the previous example, the players are required to truthfully report their types along with any message they might have received. Let $((\hat{\theta}_1, \hat{\theta}_3^1), (\hat{\theta}_2, \hat{\theta}_3^2))$ be a profile of messages received by the designer. The designer then computes the type-profile $(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)$ with $\hat{\theta}_3 := \min(\hat{\theta}_3^1, \hat{\theta}_3^2)$, produces the level $x_0 = (1+2\gamma)(\hat{\theta}_1 + \hat{\theta}_2 + \hat{\theta}_3)$ of public good and establishes the transfer $x_i = -(1+2\gamma)[\gamma\hat{\theta}_i \sum_{j \neq i} \hat{\theta}_j + (1/2)\hat{\theta}_i^2 - 2\gamma \sum_{j \neq i} \hat{\theta}_j]$ to each player i. Note that up to the term $(1+2\gamma)2\gamma \sum_{j \neq i} \hat{\theta}_j$ independent of player i's type, the transfers are identical to the generalized Vickrey-Clarke-Groves transfers of Bergemann and Morris (2009). In particular, they guarantee that the social choice function is ex-post incentive compatible (on the star network). However, and unlike the first example, the mechanism does not implement the social choice function in dominant strategies, even on the star network (unless $\gamma = 0$). Player 1 (resp., player 2) might therefore have an incentive to misreport his own type, whenever his report of player 3's type leads to $\hat{\theta}_3$ being different from player 3's true type.[17] We argue nonetheless that no player has an incentive to misreport in that example. To do so, we compute the difference $\delta_1((\hat{\theta}_1, \hat{\theta}_3^1)|\theta)$ in player 1's ex-post payoff between a truthful report (θ_1, θ_3) and the report $(\hat{\theta}_1, \hat{\theta}_3^1)$ at the type profile θ:

$$\delta_1((\hat{\theta}_1, \hat{\theta}_3^1)|\theta) = \frac{1}{2}(\theta_1 - \hat{\theta}_1)^2 + [\theta_1 + \gamma(\theta_2 + \theta_3 - \hat{\theta}_1) + 2\gamma](\theta_3 - \hat{\theta}_3),$$

with $\hat{\theta}_3 := \min(\hat{\theta}_3^1, \theta_3)$, the minimum between player 1's report about player 3's type and player 2's (true) report about player 3's type. Since $\hat{\theta}_3 \leq \theta_3$ and $\theta \in [0,1)^3$, $\delta_1((\hat{\theta}_1, \hat{\theta}_3^1)|\theta) \geq 0$ for all θ, and thus player 1 has no profitable deviation. A similar reasoning applies to player 2. As for player 3, he clearly has no profitable deviation since the social choice function is ex-post incentive compatible.

Both examples generalizes to any number of players provided that the communication network is strongly 2-connected. Lastly, note that a common feature of both examples is the existence of a "sufficient statistic" to aggregate conflicting reports about player 3's type, with the additional property that this aggregate statistic deters players 1 and 2 from lying about player 3's type. We suspect that this property can be generalized and leave it as an open issue.

4.4 Continuous type spaces

Many applications of mechanism design theory e.g., contract theory and auction theory, assume a continuous type space. While we have casted our results in environments with

[17] Remember that ex-post incentive compatibility guarantees that no player has an incentive to misreport his own type for all *truthful* reports of his opponents (but not necessarily for all reports of his opponents).

finite type spaces, they naturally extend to continuous type spaces.[18]

We now explain how to extend Theorem 1. A key feature of the proof of Theorem 1 is that player i transforms his type θ_i into a pseudo-type $\tilde{x}_i$, which reveals his type and is unconditionally uniformly distributed in $[0, 1)$. The pseudo-type is then transmitted through the network by a communication protocol. It is thus enough to show how to construct the pseudo-type in the continuous setup. Let each player's type space Θ_i be a subset of $[0, 1)$ and let types be independently distributed. Let P be the common prior and G_i be the cumulative distribution function of the marginal P^i over Θ_i. Assume that G_i is continuous. The key observation to make is that $G_i(\theta_i)$ is uniformly distributed on $[0, 1)$ and therefore, can be used as a "pseudo-type." If G_i has atoms, let θ_i^* be an atom of G_i, i.e., $\lim_{\theta_i \uparrow \theta_i^*} G_i(\theta_i) := G_i^-(\theta_i^*) < G_i^+(\theta_i^*) =: \lim_{\theta_i \downarrow \theta_i^*} G_i(\theta_i)$. Let $\hat{G}_i(\theta_i^*)$ be the realization of a uniform draw on $[G_i^-(\theta_i^*), G_i^+(\theta_i^*)]$. Let $\hat{G}_i(\theta_i) = G_i(\theta_i)$ if θ_i is not an atom. Then, $\hat{G}_i(\theta_i)$ is uniformly distributed (unconditionally on θ_i) and reveals the value of θ_i, thus is a valid pseudo-type. The mechanism construction of Theorem 1 then extends verbatim.

As for Theorem 2, it extends straightforwardly to continuous type spaces. In sum, all our constructions naturally extend to the continuous case.

4.5 Active designer

A salient feature of our model is that the designer is not active in the communication. However, in some situations, it is natural to assume that the designer can communicate with the players. For instance, a CEO has the possibility to communicate with his employees either publicly or privately.

So, let us assume that the designer can communicate with some players, so that $C(0) \neq \emptyset$. An important consequence of assuming an active designer is that the network may then contain cycles. We therefore need to relax the assumption of acyclicity. Clearly, the conditions of strong 1-connectedness and weak 2-connectedness remain necessary for the implementation of all incentive compatible social choice functions. The main insight is that these conditions are also sufficient. In other words, our results extend naturally to networks with cycles.

[18] Appropriate measurability and integrability assumptions have to be made.

Theorem 5 *For all environments $\mathcal{E}$ with common independent beliefs and private values or with a worst outcome, $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$ if and only if $\mathcal{N}$ is weakly 2-connected.*

To get an intuition for this result, consider the network $\mathcal{N}_8$ in Figure 8.

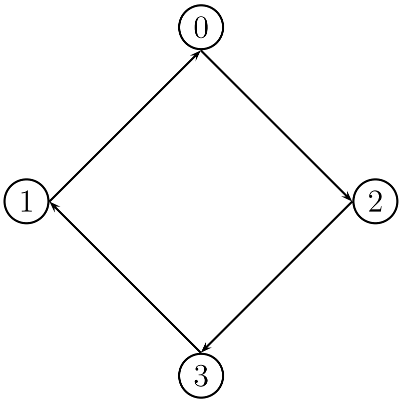


Figure 8: Communication network $\mathcal{N}_8$

The idea is simply to let the designer play the role of a provider of keys, as in the proof of Theorem 1 or Theorem 2. To be more specific, let us consider the transmission of player 3's private information in the network $\mathcal{N}_8$, when there is a worst outcome. The designer draws a large number of encoding keys and sends them to player 2. Player 2 forwards the encoding keys to player 3, who selects one key at random and uses it to encode his type. He then sends the unused keys and the encoded type to player 1, who should forward this message to the designer. Lastly, the designer compares the vector of keys he sent to player 2 and the vector of keys he receives from 1, and decodes the type of player 3 accordingly. As in the proof of Theorem 2, any deviation by player 1 or player 2 is detected with arbitrarily large probability, no information about player 3's type is revealed and the designer correctly learns the type of player 3.

Finally, let us mention that the assumption of an active designer is important in generalized principal-agents models (Myerson (1982)), where players also have to take an action, thus creating a moral hazard problem in addition to the adverse selection problem. In such models, the designer has to "securely recommend" an action to each player. We believe that our results extend to this more general framework. Indeed, if the designer has two disjoint paths of communication to each player (directed or undirected), then he can follow our protocols to privately and reliably make a recommendation to each player. A careful analysis of this issue awaits future research.

5 Conclusion

This paper completely characterizes the communication networks for which, in any environments (utilities and beliefs) with either common independent priors and private values, or with a worst outcome, every incentive compatible social choice function is (partially) implementable. We show that any weakly 2-connected communication network can replicate the incentive properties of the direct revelation mechanism. Importantly, our constructions couple encryption techniques together with incentives to secure the transmission of each player's private information to the designer.

To conclude, we believe that this paper delineates promising avenues for future research. An interesting open problem is the characterization of networks which are "equivalent" to the star network $\mathcal{N}^\star$ for *all* environments. We already know that the strong 3-connectedness of the network is sufficient, but finding necessary and sufficient conditions remains an open issue. Another interesting open issue is to consider partially known networks e.g., a model where players only known their neighbors. Other open issues include the problem of full implementation or virtual implementation on communication networks.[19]

6 Appendix

6.1 Timing Structure

In this section, we prove that the communication rule stating that "*a player sends his messages after having received all his messages*" generates a well-defined timing structure.

Lemma 1 *Let $\mathcal{N}$ be a strongly 1-connected and acyclic network. There exists an integer T and a timing function $t : N \rightarrow \{1, \ldots, T\}$ such that $t(i)$ is the stage at which player i sends his messages. Moreover, $ij \in \mathcal{N} \Rightarrow t(i) < t(j)$.*

[19] Renou (2008) is a first attempt at characterizing the social choice correspondences fully implementable in Nash equilibria on communication networks.

Proof Let $V_1 = \{i \in N : D(i) = \emptyset\}$ be the set of players who cannot receive messages. This set is clearly non-empty. For otherwise, there exists a cycle in $\mathcal{N}$. If $V_1 = N$, then $\mathcal{N} = \mathcal{N}^*$ and the proof is complete. If $V_1 \neq N$, let $V_2 = \{i : i \notin V_1 \,\text{and}\, D(i) \subseteq V_1\}$.

Claim 1 *If $V_1 \neq N$, V_2 is non-empty.*

Proof. Define $W_1 = \cup_{i \in V_1} C(i)$ as the set of players the players in V_1 can communicate to. By construction, if j is in W_1, $D(j)$ is non-empty and therefore, $j \notin V_1$. Consider then a directed path π of maximal length among the directed paths from a player in W_1 to the designer (such a path exists by strong 1-connectedness). Let j be the starting point of this directed path. We claim that j is in V_2. By contradiction, suppose that there exists $k \in D(j)$ with $k \notin V_1$. There exists then a directed path from some point m in V_1 to k, denoted $\tau = m \to l \to \cdots k \to j$. It follows that l is in W_1 and $\tau\pi$ contradicts the maximality of π. •

If $V_1 \cup V_2 = N$, the construction ends. If $V_1 \cup V_2 \neq N$, let

$$V_3 = \{i : i \notin V_1 \cup V_2 \,\text{and}\, D(i) \subseteq V_1 \cup V_2\}.$$

We continue this construction by induction. Assume that for some $k \geq 2$, the set V_s has been defined, $s \leq k$. If $\cup_{s \leq k} V_s = N$, the construction ends. If $\cup_{s \leq k} V_s \neq N$, let,

$$V_{k+1} = \{i : i \notin \cup_{s \leq k} V_s \,\text{and}\, D(i) \subseteq \cup_{s \leq k} V_s\}.$$

Claim 2 *If $\cup_{s \leq k} V_s \neq N$, V_{k+1} is non-empty.*

Proof. Let $W_{k+1} = \{j \notin \cup_{s \leq k} V_s : \exists i \in \cup_{s \leq k} V_s, j \in C(i)\}$. Since $\cup_{s \leq k} V_s \neq N$, W_{k+1} is non-empty. Consider then a directed path π of maximal length among the directed paths from a player in W_{k+1} to the designer (such a path exists by strong 1-connectedness). The starting point j of this path is in V_{k+1}. By contradiction, suppose that there exists $k \in D(j)$, $k \notin \cup_{s \leq k} V_s$. There exists then a directed path from some point m in $\cup_{s \leq k} V_s$ to k. The follower of m on this path is in W_{k+1} and this contradicts the maximality of π. •

The sequence $(\cup_{s \leq k} V_s)_k$ is a weakly increasing sequence of sets and is strictly increasing as long as $\cup_{s \leq k} V_s \neq N$. Since N is finite, there exists k such that $\cup_{s \leq k} V_s = N$. The timing function is then defined as $t(i) = s$ if $i \in V_s$. □

6.2 Probabilistic encryption

We present three important properties about the modular manipulations of real numbers in $[0,1)$. For a real number x, we denote $\lfloor x \rfloor$ the greatest integer less than or equal to x, and $x \,\mathsf{mod}_{0,1} = x - \lfloor x \rfloor$, the fractional part of x. For $(x,y) \in [0,1) \times [0,1)$, we denote $x \oplus y = (x+y) \,\mathsf{mod}_{0,1}$ and $x \ominus y = (x-y) \,\mathsf{mod}_{0,1}$.

Lemma 2 *1. For each $(x,y) \in [0,1) \times [0,1)$, $(x \oplus y) \ominus y = x$. More generally, $[0,1)$ is a commutative group for $\oplus$.*

2. Let Y be a random variable in $[0,1)$ and $x \in [0,1)$. If Y is uniformly distributed, then so are $x \oplus Y$ and $x \ominus Y$.

3. Let X, Y be independent random variables in $[0,1)$. If Y is uniformly distributed, then so are $Z = X \oplus Y$ and $W = X \ominus Y$. Furthermore, (X,Y,Z) (resp., (X,Y,W)) are pairwise-independent.

Proof of Lemma 2. (1) Consider any pair $(x,y) \in [0,1) \times [0,1)$. If $x+y \leq 1$ the statement is clear. If $x+y > 1$, $(x+y) \,\mathsf{mod}_{0,1} = x+y-1$. Thus $(x+y) \,\mathsf{mod}_{0,1} - y = x-1$ and $(x-1) \,\mathsf{mod}_{0,1} = x$.

(2) For each $z \in [0,1)$, we have

$$\begin{aligned}
\mathbb{P}(x \oplus Y \leq z) &= \mathbb{P}((x+Y) \leq z, Y \in [0, 1-x]) + \\
&\quad \mathbb{P}(x+Y-1 \leq z, Y \in (1-x, 1]) \\
&= \begin{cases} z - x + x & \text{if } z \geq x \\ z + 1 - x - (1-x) & \text{if } z < x \end{cases} \\
&= z
\end{aligned}$$

Thus, $X \oplus Y$ is uniformly distributed. Similarly, for each $z \in [0,1)$,

$$\begin{aligned}
\mathbb{P}(x \ominus Y \leq z) &= \mathbb{P}(x - Y \leq z, Y \in [0, x]) + \\
&\quad \mathbb{P}(x - Y + 1 \leq z, Y \in (x, 1]) \\
&= \begin{cases} x + 1 - (x + 1 - z) & \text{if } z \geq x \\ z + 0 & \text{if } z < x \end{cases} \\
&= z
\end{aligned}$$

Thus, $x \ominus Y$ is uniformly distributed.

(3) We only show that X and Z are independent, the rest being similar. For each $z \in [0,1)$, $\mathbb{P}(Z \leq z \mid X = x) = \mathbb{P}(x \oplus Y \leq z) = z$ from (2). □

6.3 Information transmission in weakly 2-connected network

In this section, we describe the structure of directed paths in weakly 2-connected networks and deduce that messages can be *secretly* transmitted from each player to the designer. These results are building blocks for the proofs of our main theorems.

Throughout, all networks (directed graphs) are assumed to be acyclic, strongly 1-connected and weakly 2-connected. Given a (directed) network $\mathcal{N}$, we denote $\mathcal{N}^u$ the associated undirected network: $ij \in \mathcal{N}^u$ if and only if $ij \in \mathcal{N}$ or $ji \in \mathcal{N}$.

Our definition of weakly 2-connected networks is closely related to the definition of 2-connectedness for undirected graphs. An undirected graph is 2-connected if for each pair of distinct vertices i and j, there are two disjoint paths from i to j. There are several equivalent statements for 2-connectedness of undirected graphs and the reader is referred to Bollobàs (1998, Chap. III.2). For instance, define a *cut-vertex* as a vertex i such that deleting i and all its adjacent edges yields a disconnected graph. The graph is 2-connected if and only if there is no cut-vertex. Equivalently, for each distinct vertices i, j and k, there is a path from i to j that does not contain k.

In our model, the designer (player 0) plays a special role, so that the network $\mathcal{N}$ is weakly 2-connected if and only if no player $i \in N$ is a cut-vertex of $\mathcal{N}^u$. The designer, however, can be a cut-vertex. In such case, let a *block* be a maximal 2-connected subgraph of $\mathcal{N}^u$. The undirected network $\mathcal{N}^u$ is a collection of blocks attached at 0. See Figure 9 for an example. In the sequel, we assume for simplicity that $\mathcal{N}^u$ is the only block, so that $\mathcal{N}^u$ is 2-connected. (If there are several blocks, all our arguments remain valid block-by-block.)

In the sequel, we use the letters a, b, etc. to denote nodes (players) in the network. This must not be confused with alternatives.

We define a *loop*, denoted $L(a,b)$, in $\mathcal{N}$ as a pair of directed paths with same origin a and end-point b, and no vertex in common except for the origin a and the end-point b. The loop $L(a_2, b_2)$ is a *successor* of the loop $L(a_1, b_1)$ if $a_2 \notin L(a_1, b_1)$, $b_2 \notin L(a_1, b_1)$ and the intersection $L(a_1, b_1) \cap L(a_2, b_2)$ is a path which contains at least one edge and the vertex b_1. See Figure 10 for an example.

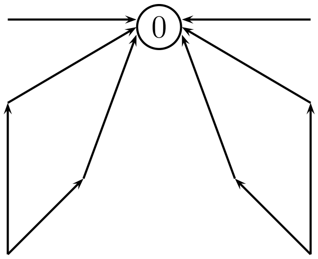


Figure 9: Blocks attached at 0

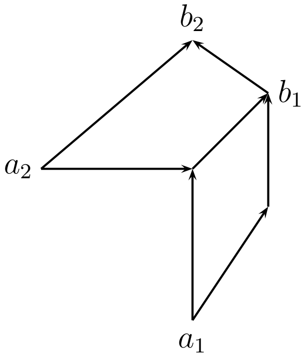


Figure 10: $L(a_2, b_2)$ is a successor of $L(a_1, b_1)$

We use the following notation: we write $i \longrightarrow k$ for a directed path $(i_0 = i, i_2, \ldots, i_R = k)$ from player i to player k and $i \longrightarrow k \longrightarrow l$ for a directed path from i to l through k, etc. We say that two directed paths $(i_0 = i, i_2, \ldots, i_R)$ and $(j_0 = i, j_2, \ldots, j_Q)$ *cross each other* if there exist r^* and q^* such that $j_{q^*} = i_{r^*}$.

To prove our main results, we use the following decomposition of directed graphs into successive loops. We assume that there are at least three player (if $n = 2$, the only strongly 1-connected and weakly 2-connected network is such that $D(0) = N$).

Proposition 2 *Let $n \geq 3$. For each $i \in N \backslash D(0)$ and each $j \in C(i)$, there exists a finite sequence of loops $L(a_1, b_1), \ldots, L(a_M, b_M)$ such that:*

1. *the edge ij belongs to $L(a_1, b_1)$,*
2. *for each $m = 1, \ldots, M-1$, $L(a_{m+1}, b_{m+1})$ is a successor of $L(a_m, b_m)$ and $a_{m+1} \notin \cup_{q \leq m} L(a_q, b_q)$, and*
3. *$b_M = 0$.*

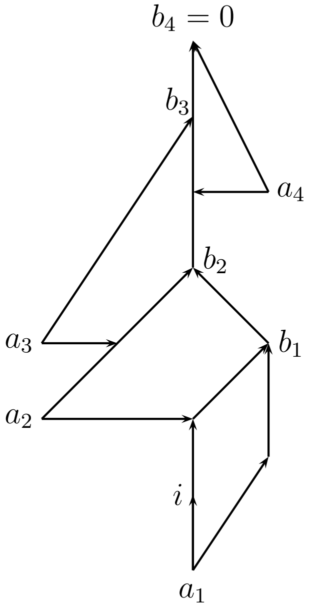


Figure 11: A sequence of loops

Proof This is trivially true if $n = 3$. Assume that $n \geq 4$. The proof rests on several lemmatas.

Lemma 3 *Let $\mathcal{N}^u$ be a 2-connected undirected graph. Let A be a non-empty set of vertices and let b and c two distinct vertices that do not belong to A. There exists $a^* \in A$ and a path from a^* to c that has no vertex in $(A \backslash \{a^*\}) \cup \{b\}$.*

Proof. Since $\mathcal{N}^u$ is 2-connected, for each $a \in A$, there exists a path from a to c that does not contain b (otherwise, b would be a cut-vertex). This path must leave the set A to reach c, thus the last point a^* in A on this path has the desired properties. •

Lemma 4 *Let $i \in N \backslash D(0)$ and $j \in C(i)$, there exists a loop that contains the edge ij.*

Proof. Remember that for each player $k \in N$, there exists a directed path from k to 0 by strong 1-connectedness and thus, $C(k) \neq \emptyset$. Consider a player $i \in N \backslash D(0)$ and $j \in C(i)$.

- Case 1. If $C(i)$ contains another player $k \neq j$, then there exists a directed path from i to 0 through the edge ij and a directed path from i to 0 through the edge

ik. These paths must cross each other (possibly at 0), thus we have found the desired loop.

- Case 2. If $C(i) = \{j\}$, denote $D_\infty(i)$ the set of players who have a directed path to i. From Lemma 3, there exists $k \in D_\infty(i)$ and an undirected path $(k_0 = k, k_1, \ldots, k_R = 0)$ from k to 0 such that no player k_r is in $D_\infty(i) \cup \{i\}$ for $r > 0$. If the edge kk_1 is directed from k to k_1, then choose a directed path from k_1 to 0 to obtain the directed path $k \rightarrow k_1 \rightarrow 0$ one the one hand and the directed path $k \rightarrow i \rightarrow j \rightarrow 0$ on the other hand. These paths must cross each other and therefore, define a loop with origin k. (The first crossing point defines the end-point of the loop.) The end-point of the loop cannot be in $D_\infty(i) \cup \{i\}$ since $k_1 \notin D_\infty(i)$. It follows that the edge ij is contained in this loop.

  If the edge kk_1 is directed from k_1 to k, then we progress along the path $(k_1, \ldots, k_R)$ until we reach a first edge $k_r k_{r+1}$ directed from k_r to k_{r+1}. Such an edge exists since, thanks to acyclicity, the edge $k_{R-1}0$ is directed from k_{R-1} to 0. Thus, there exists a directed path from k_{r+1} to 0. Consider then the directed path $k_r \rightarrow k_{r+1} \rightarrow 0$ one the one hand and the directed path $k_r \rightarrow k \rightarrow i \rightarrow j \rightarrow 0$ on the other. These paths must cross each other and thus define a loop with origin k_r. Again, the end-point of the loop cannot be in $D_\infty(i) \cup \{i\}$ since $k_r \notin D_\infty(i)$. It follows that the edge ij is contained in this loop.

•

We now construct the desired sequence of loops. We start with $i \in N \backslash D(0)$ and $j \in C(i)$.

First step. Let $L(a_1, b_1)$ be a loop containing ij and such that $t(b_1)$ is maximal among all loops that contain ij ($t(\cdot)$ is the timing function constructed in Lemma 1). (Such a loop exists by the above lemma.) If $b_1 = 0$, the construction ends. If $b_1 \neq 0$, let $c_1 \in C(b_1)$ and denote d_1 and e_1 the two predecessors of b_1 on each path of $L(a_1, b_1)$.

The construction then proceeds inductively. Assume that $L(a_1, b_1), \ldots, L(a_M, b_M)$ have been constructed for some $M \geq 1$. If $b_M = 0$, the construction ends. If $b_M \neq 0$, let $c_M \in C(b_M)$ and denote d_M and e_M the two predecessors of b_M on each of the two disjoint directed paths of $L(a_M, b_M)$.

For each subset of players N', let us denote $D_\infty(N')$ the set of players j for whom there exists a directed path from j to some player in N'. Clearly, $D_\infty(N' \cup N'') = D_\infty(N') \cup D_\infty(N'')$ and $D_\infty(D_\infty(N')) = D_\infty(N')$.

Lemma 5 *There exists a loop $L(a_{M+1}, b_{M+1})$ such that $a_{M+1} \notin \cup_{q \leq M} L(a_q, b_q) \cup D_\infty(i)$ and which contains either the path $d_M \to b_M \to c_M$ or the path $e_M \to b_M \to c_M$. Furthermore, this loop is disjoint from $\cup_{q \leq M-1} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$.*

Proof. From Lemma 3, there exists $u_M \in \cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$ and an undirected path $(\lambda_0 = u_M, \lambda_1, \ldots, \lambda_S = 0)$ from u_M to 0 disjoint from $(\cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i) \cup \{b_M\}) \backslash \{u_M\}$. Assume that $u_M \in D_\infty(L(a_M, b_M))$. There exists a directed path from u_M to b_M which goes either through d_M or through e_M. Without loss of generality, assume that this path goes through d_M. If the edge $u_M\lambda_1$ is directed from u_M to λ_1, then choose a directed path from λ_1 to 0 to obtain the directed path $u_M \to \lambda_1 \to 0$ on one hand and the directed path $u_M \to d_M \to b_M \to c_M \to 0$ on the other hand. These paths must cross each other and therefore, define a loop with origin u_M. Since $\lambda_1 \notin \cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$, the path $\lambda_1 \to 0$ cannot go through $\cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$, and thus the end-point of the loop is not in $\cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$ either. The path $d_M \to b_M \to c_M$ is thus contained in the new loop.

If the edge $u_M\lambda_1$ is directed from λ_1 to u_M, then we progress along the path $(\lambda_1, \ldots, \lambda_S)$ until we reach a first edge $\lambda_s\lambda_{s+1}$ directed from λ_s to λ_{s+1}. There must exists one such edge, because of the acyclicity of $\mathcal{N}$. Then, there is a directed path from λ_{s+1} to 0. Consider the directed path $\lambda_s \to \lambda_{s+1} \to 0$ one one hand, and the directed path $\lambda_s \to u_M \to d_M \to b_M \to c_M \to 0$ on the other. These paths must cross each other and thus define a loop with origin λ_s. As before, the end-point of the loop is not in $\cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$, thus the path $d_M \to b_M \to c_M$ is contained in this loop.

Finally, u_M cannot be in $\cup_{q \leq M-1} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$. Otherwise, the construction above provides a loop that would contradict the maximality property of b_m, for some $m < M$. That is, since $t(b_{M+1}) > t(b_m)$, the newly constructed loop would have been used at an earlier stage of the induction. Similarly, the origin a_{M+1} of the new loop cannot be in $\cup_{q \leq M} D_\infty(L(a_q, b_q)) \cup D_\infty(i)$. •

Inductive step. Let $L(a_{M+1}, b_{M+1})$ be a loop containing $d_M \to b_M \to c_M$ or $e_M \to$

$b_M \to c_M$ and such that $t(b_{M+1})$ is maximal among all loops that contain $d_M \to b_M \to c_M$ or $e_M \to b_M \to c_M$. If $b_{M+1} = 0$, the construction ends and otherwise, continues inductively.

By construction, there is a directed path from b_m to b_{m+1}, thus $t(b_m) < t(b_{m+1})$ from the definition of the timing structure. It follows that the construction stops after a finite number of iterations. This completes the proof. □

Proposition 2 is a building block for the construction of a protocol (mechanism and strategies) that allows player i to secretly send a message to the designer. Let us summarize our findings. Proposition 2 has the following implications: For each player $i \in N \setminus D(0)$ and $j \in C(i)$, there exists a finite sequence of loops $(L(a_m, b_m))_{m=1}^{M}$ such that (i) $ij \in L(a_1, b_1)$, (ii) $b_M = 0$ and (iii) the loop $L(a_{m+1}, b_{m+1})$ is a successor of the loop $L(a_m, b_m)$, $m = 1, \ldots, M-1$, with the additional property that there exists $u_m \in L(a_m, b_m) \cap L(a_{m+1}, b_{m+1})$ such that the directed path from u_m to b_m in $L(a_m, b_m)$ is part of the directed path from u_m to b_{m+1} in $L(a_{m+1}, b_{m+1})$. Moreover, the sequence of loops defines a directed path from player i to the designer through all players b_1 to b_{M-1}. To see this, note that player i belongs to the loop $L(a_1, b_1)$ from player a_1 to player b_1 and thus, belongs to one directed path to b_1. Similarly, b_1 belongs to the loop $L(a_2, b_2)$ and thus, has a directed path to b_2. Iterating this argument, we construct a directed path from i to the designer through the players b_1 to b_{M-1}. We will use this directed path to secretly transfer the private information of player i to the designer.

Proposition 3 *Let v be a random variable in $[0,1)$ privately known to player i. There exists a protocol $\mathcal{M}_i$ (i.e., a mechanism and a profile of strategies) on $\mathcal{N}$ such that whenever all players follow the prescribed strategies, the designer correctly learns the value of v. Moreover, the messages received by any player $j \neq i$ are probabilistically independent from v.*

Proof If $i \in D(0)$, this is straightforward. Fix $i \in N \backslash D(0)$ and consider the sequence of loops constructed in Proposition 2. We divide players into several categories.

- A player who belongs to one loop is *active*. All other players are inactive. Inactive players do not send or receive messages (their message sets are singletons). Let us focus now on active players.

- A player a_m who is the origin of a loop is a *provider*.

- A player b_m who is the end-point of a loop is a *lock-opener*.

- The player u_m who is the first point on the intersection of the two successive loops $L(a_m, b_m)$ and $L(a_{m+1}, b_{m+1})$ is a *lock-closer*.

- Other active players are *transmitters*.

By construction, note that a provider has no active predecessor and exactly two active successors. A lock-opener, or a lock-closer, has two active predecessors and one active successor. Transmitters have exactly one active predecessor and one active successor. Finally, player i is either a transmitter or a provider. For each loop, we label *Left (L)* the path that contains the lock-closer and *Right (R)* the other. The strategies for active players other than player i are as follows:

- Each transmitter truthfully forwards the message received from his active predecessor to his active successor.

- Each provider a_m draws an encryption key X_m uniformly in $[0, 1)$ and sends it to its two active successors.

- Each lock-closer u_m receives two numbers x_m and x_{m+1} from his two predecessors. He computes $z_m = x_m \oplus x_{m+1}$ and sends z_m to his active successor. Remark that there is no lock-closer u_{M+1} in the last loop $L(a_M, b_M)$.

- Each lock-opener b_m (with $m < M$) receives two numbers x_m^L and x_m^R from his left and right predecessors. He computes $w_m = x_m^L \ominus x_m^R$ and sends w_m to his active successor.

Player i's strategy is as follows:

- If he is a transmitter, player i receives x_1 from his active predecessor and sends $x_1 \oplus v$ to his active successor.

- If he is a provider, player i sends $X_1 \oplus v$ to his active successor on the left path and X_1 to his active successor on the right path.

See Figure 12 for a heuristic illustration of the strategies.

Firstly, we show that this protocol allows the designer to correctly learn the value of v. To this end, let us assume that these strategies are effectively played and compute the messages w_m sent by the lock-openers.

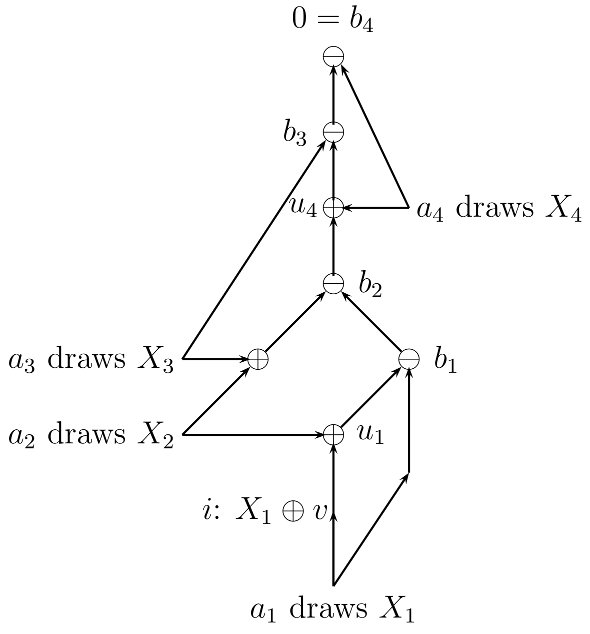


Figure 12: Providers, lock-closers $\oplus$ and lock-openers $\ominus$

The sequence of loops defines a directed path from player i to the designer. This path contains all lock-openers (b_m) and some lock-closers (u_m) and is uniquely defined if player i is a transmitter. If player i is a provider, we choose the only such path that begins with the left path of the first loop. Along this path, let us attach labels to players. All lock-openers and player i are labeled $\ominus$ and the lock-closers are labeled $\oplus$. For instance, in Figure 12, we have

$$i^{\ominus} \to u_1^{\oplus} \to b_1^{\ominus} \to b_2^{\ominus} \to u_4^{\oplus} \to b_3^{\ominus} \to b_4^{\ominus} = 0.$$

This induces a sequence in the alphabet $\{\ominus, \oplus\}$. Let $\nu(b_m)$ be the number of occurrence of two consecutive $\ominus$ appearing in the sequence before b_m (including b_m). For instance, in the example above, $\nu(b_1) = 0$, $\nu(b_2) = \nu(b_3) = 1$, $\nu(b_4) = 2$.

Lemma 6 *If the players follow the above strategies, for each $m = 1, \ldots, M-1$, we have*

$$w_m = (-1)^{\nu(b_m)} v \oplus X_{m+1}.$$

The two messages received by the designer are X_M and w_{M-1}.

Consequently, the designer can compute the value v of the private information of player i, which is $X_M \ominus w_{M-1}$ if $\nu(b_{M-1})$ is odd and $w_{M-1} \ominus X_M$ if $\nu(b_{M-1})$ is even.

Proof. We first compute w_1 and then proceed by induction. Consider the loop $L(a_1, b_1)$. Player i is either on the left path of the loop $L(a_1, b_1)$ or on the right path of $L(a_1, b_1)$. In the former case, the left path from i to b_1 is $i^{\ominus} \rightarrow u_1^{\oplus} \rightarrow b_1^{\ominus}$ and the right path is $i \rightarrow b_1$. Player b_1 thus receives $X_2 \oplus X_1 \oplus v$ from the left and X_1 from the right. It follows that $w_1 = (X_2 \oplus X_1 \oplus v) \ominus X_1 = X_2 \oplus v$. Note that in this case $\nu(b_1) = 0$. See Figure 13 for an illustration.

In the latter case, the left path is $a_1 \rightarrow u_1 \rightarrow b_1$ and the right path is $i^{\ominus} \rightarrow b_1^{\ominus}$. Player b_1 thus receives $X_2 \oplus X_1$ from the left and $X_1 \oplus v$ from the right. Thus $w_1 = (X_2 \oplus X_1) \ominus (X_1 \oplus v) = X_2 \ominus v$. Note that in this case $\nu(b_1) = 1$. See Figure 14 for an illustration. We have thus proved the lemma for $m = 1$.

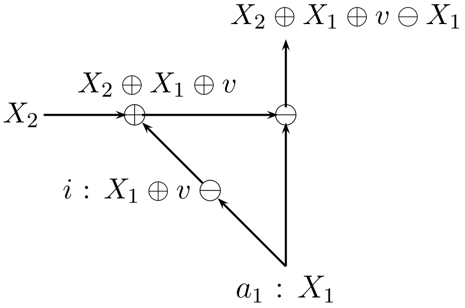

$X_2 \oplus X_1 \oplus v \ominus X_1$
$X_2 \oplus X_1 \oplus v$
X_2
$i\,:\,X_1 \oplus v$
$a_1\,:\,X_1$

Figure 13: w_1 with player i on the left path.

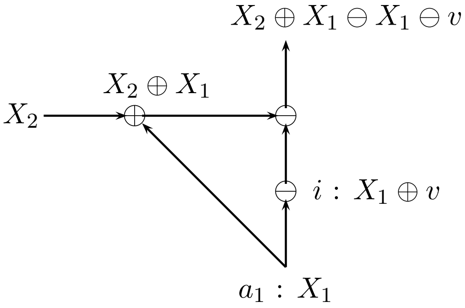


Figure 14: w_1 with player i on the right path.

We proceed now by induction. Let us assume that for some $m \leq M - 1$, $w_{m-1} = (-1)^{\nu(b_{m-1})} v \oplus X_m$ and compute w_m. Consider the loop $L(a_m, b_m)$. By construction, this loop contains b_{m-1} and u_m and the left path is the one that contains u_m. Thus, b_{m-1} is either on the left path or on the right path. In the former case, the left path of this loop is $a_m \to b^{\ominus}_{m-1} \to u^{\oplus}_m \to b^{\ominus}_m$ and the right path is $a_m \to b_m$. Since there is also the path $a_{m+1} \to u_m \to b_m$, the message received by b_m from the left is $X_{m+1} \oplus (-1)^{\nu(b_{m-1})} v \oplus X_m$ and the message received from the right is X_m. Thus,

$$w_m = (X_{m+1} \oplus (-1)^{\nu(b_{m-1})} v \oplus X_m) \ominus X_m = X_{m+1} \oplus (-1)^{\nu(b_{m-1})} v.$$

Remark that in this case $\nu(b_m) = \nu(b_{m-1})$. See Figure 15 for an illustration.

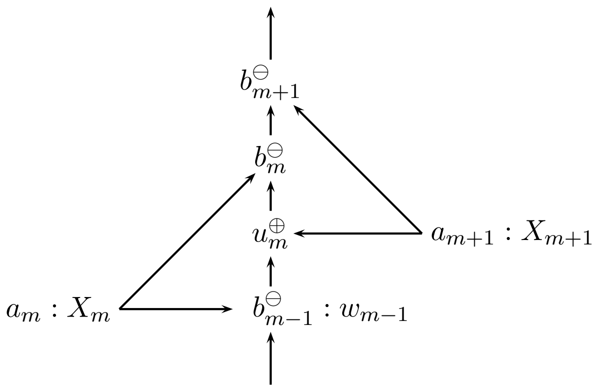


Figure 15: w_m with player b_{m-1} on the left path

In the former case, the left path is $a_m \to u_m \to b_m$ and the right path is $a_m \to b^{\ominus}_{m-1} \to b^{\ominus}_m$. Since there is also the path $a_{m+1} \to u_m \to b_m$, the message received from the left is $X_{m+1} \oplus X_m$ and the message received from the right is $(-1)^{\nu(b_{m-1})} v \oplus X_m$. Thus $w_m = (X_{m+1} \oplus X_m) \ominus ((-1)^{\nu(b_{m-1})} v) = X_{m+1} \ominus (-1)^{\nu(b_{m-1})} v$. Remark that in this case $\nu(b_m) = \nu(b_{m-1}) + 1$. See Figure 16 for an illustration.

Finally, consider the last loop $L(a_M, b_M)$, where $b_M = 0$ is the designer. By construction, this loop does not contain a lock-closer u_{M+1}. One path of this loop goes through b_{M-1}, i.e., we have $a_M \to b_{M-1} \to b_M$, and the other is $a_M \to b_M$. Other players on this loop are transmitters. The designer thus receives w_{M-1} from the first path and X_M from the other. The proof of the Lemma is thus complete. •

To complete the proof of Proposition 3, we argue that the message received by each player $j \neq i$ is probabilistically independent from v. This is clearly true for inactive players and for providers. More generally, the only messages that depend on v are those on the directed path from player i to the designer as constructed above, so the

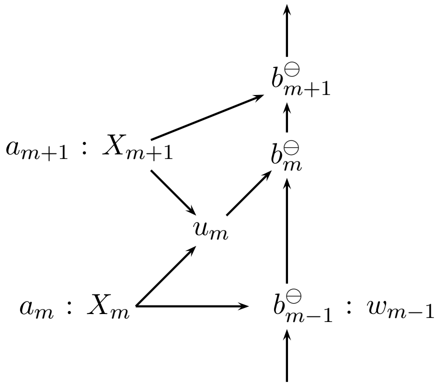


Figure 16: w_m with player b_{m-1} on the right path

statement clearly holds for players outside of this path. Transmitters on this path receive messages of the type $X \oplus v$ where X is some random variable independent from v and uniformly distributed. From Lemma 2 (iii), this is independent from v. The very same reasoning holds for lock-closers. For lock-openers, this is a consequence of the above computation: since X_m and X_{m+1} are independent and uniformly distributed, so are the two messages received by b_m. □

Corollary 1 *Let $(v_i)_{i \in N}$ be independent random variables such that v_i is known to player i only. There exists a protocol $\mathcal{M}$ on $\mathcal{N}$ such that, whenever all players abide by the protocol, the designer correctly learns the value of each v_i. Moreover, the messages received by any player j are probabilistically independent from $(v_i)_{i \neq j}$.*

Proof From Proposition 3, for each player i, there exists a protocol (mechanism and strategies) $\mathcal{M}_i$ such that player i can secretly transfer his private information v_i to the designer without revealing information to the other players. The idea is then to concatenate all these protocols "in parallel." That is, each player j plays a role in each $\mathcal{M}_i$ (inactive, provider, lock-closer, lock-opener or transmitter), and should play all the corresponding roles simultaneously. For instance, if he is transmitter in several $\mathcal{M}_i$'s, he should forward the corresponding messages on the corresponding links. Moreover, if a player is a provider in one or several $\mathcal{M}_i$'s, the random draws must be mutually independent and independent of messages received. □

6.4 Proof of Theorem 1: sufficiency

From Corollary 1, there exists a mechanism and a profile of strategies such that if all players follow the prescribed strategies, the designer correctly learns the private information of each player. We now show that, in an environment with common independent beliefs and private values, we can indeed provide the players with appropriate incentives to follow the prescribed strategies. Roughly speaking, we make sure that each player is indifferent between all the messages he may send. This is done as follows.

Fix an environment $\mathcal{E}$ with common independent beliefs and private values and an incentive compatible social choice function f. Denote P^i the marginal distribution of the common belief P on Θ_i, i.e., this is the common belief of any player $j \neq i$ on Θ_i. Without loss of generality, assume that $\Theta_i := \{1, \ldots, t_i, \ldots, T_i\}$ for each player $i \in N$ and denote $\overline{P}^i(t_i) = \sum_{\theta_i \leq t} P^i(\theta_i)$, the cumulative distribution function of P^i. Define a partition $\Pi_i = \{\Pi_i(1), \ldots, \Pi_i(T_i)\}$ of $[0,1)$ into T_i subsets with $\Pi_i(t_i) = [\overline{P}^i(t_i - 1), \overline{P}^i(t_i))$ (with $\overline{P}^i(0) = 0$). Note that if X is uniformly distributed on $[0,1)$, the event $\{X \in \Pi_i(t_i)\}$ has probability $P^i(t_i)$.

Part I. We first consider the problem of implementing the social choice function f_i^* for which player i is dictatorial, i.e., for any θ_i, define $f_i^*(\theta_i) \in \arg\max_{a \in A} u_i(a, \theta_i)$ and let $f_i^*(\theta_i, \theta_{-i}) = f_i^*(\theta_i)$ for all θ_{-i}. If $i \in D(0)$, f_i^* is clearly implementable. Assume that $i \notin D(0)$. We claim that the protocol $\mathcal{M}_i$ implies the existence of a mechanism and strategies such that player i has an incentive to truthfully reveal his type and no other active player has an incentive to manipulate the transmission of information from player i to the designer.

The mechanism and strategies are as follows:

- Player i of type t_i draws a random number v_i uniformly in $\Pi_i(t_i)$ and transmits it to the designer by the protocol $\mathcal{M}_i$.
- All other active players follow the strategies constructed in $\mathcal{M}_i$.
- Let $\hat{v}_i$ be the message decoded by the designer and denote $\hat{\theta}_i = t_i$ if $\hat{v}_i \in \Pi_i(t_i)$. (See Lemma 6.) The designer implements the alternative $f_i^*(\hat{\theta}_i)$.

Firstly, observe that the protocol $\mathcal{M}_i$ implies that each active player sends a real number in $[0,1)$. Secondly, observe that the unconditional distribution of v_i is the uniform distribution on $[0,1)$. To see this, denote $X_i^{t_i}$ a random variable uniformly distributed on $\Pi_i(t_i)$ and observe that $v_i = \sum_{t_i=1}^{T_i} \mathbf{1}_{\{\theta_i=t_i\}} X_i^{t_i}$. From Proposition 3, it follows that the designer correctly learns the type of player i if all players abide by the protocol $\mathcal{M}_i$, while no player gets additional information about the type of player i (posterior beliefs are equal to prior beliefs). So, the expected payoff of any active player $j \neq i$ of type θ_j is $\sum_{\theta_i} u_j(f_i^*(\theta_i), \theta_j) P^i(\theta_i)$.

Thirdly, we show that no active player has an incentive to deviate. This is clearly true for player i as f_i^* is incentive compatible. Consider player $j \neq i$ and suppose that j is a transmitter in the loop $L(a_m, b_m)$ for $m = 2, \ldots, M-1$. There are several cases to consider.

Case 1. Player j is on the right path of the loop $L(a_m, b_m)$ from player a_m to player b_m and moves before the lock-closer u_{m-1}. Under $\mathcal{M}_i$, he receives the message x_m. Suppose that he deviates and sends the message x'_m. It follows that the designer will receive the messages $(-1)^{\nu(b_{M-1})}(v \oplus x'_m \ominus x_m) \oplus X_M$ and X_M under the deviation, so that the decoded message is $v \oplus x'_m \ominus x_m$. Since v is uniformly distributed on $[0,1)$, it follows that the probability that $v \oplus x'_m \ominus x_m$ is in $\Pi_i(t_i)$ is $P^i(t_i)$, regardless of x'_m (see Lemma 2(ii)). Player j is thus indifferent between sending x_m and x'_m.

Case 2. Player j is on the right path of the loop $L(a_m, b_m)$ from player a_m to player b_m and moves after the lock-closer u_{m-1}, but before the lock-opener b_{m-1}. Under $\mathcal{M}_i$, player j receives the message $x_m \oplus x_{m-1}$ from the lock-closer u_{m-1}. Suppose that he deviates and sends the message x'_m. It follows that the designer will receive the messages $(-1)^{\nu(b_{M-1})}(v \oplus x'_m \ominus x_m \ominus x_{m-1}) \oplus X_M$ and X_M under the deviation. Since all random variable are uniformly distributed on $[0,1)$, so are their addition $\oplus$ or subtraction $\ominus$ (this follows from Lemma 2) and consequently, player j is indifferent between sending $x_m \oplus x_{m-1}$ and x'_m.

Case 3. Player j is on the right path of the loop $L(a_m, b_m)$ from player a_m to player b_m and moves after the lock-closer u_{m-1} and the lock-opener b_{m-1}. Under $\mathcal{M}_i$, player j receives the message $(-1)^{\nu(b_{m-1})} v \oplus x_m$. Note that j does not learn the value of x_m and believes that it is a realization of X_m. Suppose that he deviates and sends the message x'_m. It follows that the designer will receive the messages $(-1)^{\nu(b_{M-1})}(x'_m \ominus x_m) \oplus X_M$ and X_M under the deviation. Since X_m and X_M are uniformly distributed on $[0,1)$, it follows yet again that player j evaluates the probability of $\hat{v}_i = x'_m \ominus x_m \in \Pi_i(t_i)$ to be $P^i(t_i)$ and thus, is again indifferent between reporting the truth and deviating.

Case 4. Player j is on the left path of the loop $L(a_m, b_m)$ from player a_m to player b_m and moves before the lock-closer u_m. This case is similar to case 1.

Case 5. Player j is on the left path of the loop $L(a_m, b_m)$ from player a_m to player b_m and moves after the lock-closer u_m. In that case, player j also belongs to the right path of the loop $L(a_{m+1}, b_{m+1})$ and the same arguments as in case 1 apply.

Lastly, a similar reasoning applies if player j is a transmitter in the first or last loop. For instance, if player j is on the right path of the last loop $L(a_M, b_M)$ and moves before the lock-closer u_M, the same reasoning as in case 1 applies since the designer receives the message $(-1)^{\nu(b_{M-1})}v \oplus x'_M$ and X_M.

Now, suppose that player j is the provider a_m in the loop $L(a_m, b_m)$ $(m < M)$ and suppose that he sends the message x^L_m on the left path of the loop and the message x^R_m on the right path. If all other players abide by the strategies, it follows that the designer receives the messages $(-1)^{\nu(b_{M-1})}(v \oplus x^R_m \ominus x^L_m) \oplus X_M$ and X_M. Since v and X_M are uniformly and independently distributed on $[0, 1)$, it follows that the probability that the decoded type $\hat{v}_i$ is in $\Pi_i(t_i)$ is $P^i(t_i)$ and thus, player j is indifferent between following the prescribed strategy or deviating.

Similar arguments apply to the lock-closers or lock-openers, so that the prescribed strategies indeed form a Bayesian equilibrium. To summarize, incentive compatibility of the social choice function implies that player i has indeed an incentive to abide by the protocol $\mathcal{M}_i$, while all other active players have no incentive to deviate, since the protocol guarantees the same expected payoff to each active player other than player i, regardless of the message he sends.

Part II. Let f be a social choice function implementable on $\mathcal{N}^\star$, i.e., f is incentive compatible. To implement f, consider the mechanism and strategies implied by the protocol $\mathcal{M}$: each player $i \notin D(0)$ of type t_i draws a random number v_i uniformly in $\Pi_i(t_i)$ and transmits it to the designer according to the protocol $\mathcal{M}_i$, while in his role of an active player in a protocol $\mathcal{M}_j$ $(j \neq i)$, he follows the prescribed strategy.

From Corollary 1, it follows that the designer learns the true profile of types if all players abide by this protocol, while no player gets additional information about the type of his opponents. To complete the proof, note that as in part I, no player has an incentive to deviate. The expected payoff of a player i is independent of the messages he sends about his opponents (since the assumption of independent beliefs imply that we can consider each deviation as above). Incentive compatibility guarantees that player i

has indeed an incentive to abide by the sub-protocol $\mathcal{M}_i$. The proof of the sufficiency part of Theorem 1 is thus complete.

6.5 Proof of Theorem 1: necessity

Now, we prove the "only if" part of Theorem 1. The proof proceeds by contradiction. We assume that $\mathcal{N}$ is not weakly 2-connected and construct an environment with common independent belief and private values and an incentive compatible social choice function, which is not implementable on $\mathcal{N}$.

If $\mathcal{N}$ is not weakly 2 connected, there exists two distinct players i and i^* such that all paths, directed or undirected, from i to the designer go through i^*. As a consequence, for each player k that has a path to i, directed or undirected, all paths from k to 0 also go through i^*. This implies that player i^* is a cut-vertex in the network. In particular, all information regarding the players k who have a path to i, is controlled by i^*.

Let us now construct the environment and the social choice function. Assume that all players but player i have a single type and that player i has two types θ_i and θ'_i. Let a and b be two alternatives. The utilities are as follows: $u_i(a, \theta_i) = u_{i^*}(a, \cdot) = 1$, $u_i(b, \theta_i) = u_{i^*}(b, \cdot) = 0$; $u_i(a, \theta'_i) = 0$, $u_i(b, \theta'_i) = 1$. All other players are indifferent (get a utility of 0) between a and b. Any other alternative gives a utility of -1 to players i and i^* regardless of their types. The common prior is the uniform distribution on the set of types. The social choice function is the dictatorial social choice function of player i.

We claim that for every mechanism on $\mathcal{N}$, there is no equilibrium that implements this social choice function. By contradiction, assume that there exists such an equilibrium σ. Fix a profile of messages $\bar{m}_{i^*} \in M_{D(i^*)}$ for player i^* in the support of $\mathbb{P}_{\theta_i,\sigma}$, i.e., this is a message compatible with θ_i and the equilibrium strategies. Consider the deviation σ'_{i^*} for player i^* which consists in playing $\sigma_{i^*}(\bar{m}_{i^*})$ regardless of his type and messages received.

By construction of the deviation, $\sigma_{i^*}(\bar{m}_{i^*})$ is compatible with the messages sent by players who have no path to player i, i.e.,

$$\operatorname{supp} \mathbb{P}_{\theta,\sigma'_{i^*},\sigma_{-i^*}} \subseteq \operatorname{supp} \mathbb{P}_{\theta_i,\sigma} \ \ \forall \theta \in \{\theta_i, \theta'_i\}.$$

Since the strategies are assumed to implement f, it follows that the outcome is almost surely a under the deviation, regardless of the type of player i. Since player i^* prefers a to any other alternative, this deviation is profitable for player i^*.

It is worthwhile to note that weak 2-connectedness is also a necessary condition for Proposition 3 to hold. Indeed, if i^* is a cut-vertex, and if the designer learns the type of player i, then i^* must learn it as well.

6.6 Proof of Theorem 2

The proof of the "only if part" is identical to the previous one and is omitted. We turn to the "if" part and fix an environment with a worst outcome and an incentive compatible social choice function f. Without loss of generality, we assume that Θ_i is a finite subset of the open interval $(0,1)$ for each player $i \in N$. In the proof of Theorem 1, we took advantage of the environment to make players indifferent between any message they can send. This is not longer possible in environments with correlated beliefs and/or common values. We thus modify the protocol in such a way that deviations are detected with arbitrarily high probability by the designer. The threat of the worst outcome then deters the players from deviating.

Let η be a large integer. We take up the terminology and notations from Proposition 3 and modify the protocol $\mathcal{M}_i$ as follows.

- Each transmitter forwards the message received from his active predecessor to his active successor.

- Each provider a_m draws an η-vector of keys $\vec{X}_m = (X_m^1, \ldots, X_m^\eta)$ whose components are independently and uniformly distributed in $[0,1)$ and sends it to its two active successors.

- Each lock-closer u_m receives two vectors $\vec{x}_m$, $\vec{x}_{m+1}$ from his predecessors. He computes $\vec{z}_m = \vec{x}_m \vec{\oplus} \vec{x}_{m+1}$ and sends it to his active successor, where $\vec{\oplus}$ denotes component-wise addition.

- Each lock-opener b_m receives two vectors $\vec{x}_m^L$, $\vec{x}_m^R$ from his predecessors. He computes $\vec{w}_m = \vec{x}_m^L \vec{\ominus} \vec{x}_m^R$ and sends it to his active successor.

Player i behaves as follows (recall that by construction, player i is either a transmitter or a provider):

- If he is a transmitter, player i who receives $\vec{x}_1$ from his active predecessor draws uniformly a random integer η^* in $\{1, \ldots, \eta\}$, and encodes his type θ_i with the encoding key $x_1^{\eta^*}$ to obtain the cypher-type $y_1^{\eta^*}(i) = \theta_i \oplus x_1^{\eta^*}$. Player i then sends the vector $(x_1^1, \ldots, x_1^{\eta^*-1}, y_1^{\eta^*}(i), x_1^{\eta^*+1}, \ldots, x_1^{\eta})$ to his active successor.

- If he is a provider, player i draws (uniformly) a random vector $\vec{X}_1$ and a random integer η^* in $\{1, \ldots, \eta\}$ and computes $Y_1^{\eta^*}(i) = \theta_i \oplus X_1^{\eta^*}$. Player i then sends the vector $(X_1^1, \ldots, X_1^{\eta^*-1}, Y_1^{\eta^*}(i), X_1^{\eta^*+1}, \ldots, X_1^{\eta})$ to his Left active successor and $\vec{X}_1$ to his Right active successor.

The decision rule of the designer is the following. The designer receives a message $\vec{x}_M^R$ from the path $a_M \to b_{M-1} \to b_M = 0$, and a message $\vec{x}_M^L$ from the other path of the last loop $a_M \to b_M = 0$.

- If the vectors $\vec{x}_M^L$, $\vec{x}_M^R$ differ by exactly one component η^*, the designer decodes $\hat{\theta}_i = x_M^{\eta^*,R} \ominus x_M^{\eta^*,L}$ if $\nu(b_{M-1})$ is even and $\hat{\theta}_i = x_M^{\eta^*,L} \ominus x_M^{\eta^*,R}$ if $\nu(b_{M-1})$ is odd.

- Otherwise, the designer concludes that there was a deviation.

Note that no player $j \neq i$ gains information about θ_i by this modified mechanism. Indeed, player j only observes vectors of uniformly distributed numbers. If all players abide by the mechanism, then the two vectors received by the designer differ only in the component η^*, and the designer decodes correctly the type of player i from Lemma 6. The key argument is that η^* is the private information of player i. Thus, any deviation by an active player is bound to change another component with probability at least $1 - 1/\eta$.

Finally the mechanism for implementing f is the following:

- Each player i transmits his type to the designer using the modified protocol.

- If the designer concludes that there was no deviation, he implements $f(\hat{\theta}_1, \ldots, \hat{\theta}_n)$, where $\hat{\theta}_i$ is the decoded type of player i.

- Otherwise, the designer implements the worst outcome.

Let us check the equilibrium condition. The expected payoff of j under the mechanism is:

$$\sum_{\theta_{-j}} u_j(f(\theta_i), \theta_j, \theta_{-j}) P_j(\theta_{-j} \mid \theta_j) := C.$$

Assume that player j deviates in at least one sub-mechanism. His expected payoff is at most,

$$\frac{1}{\eta}W + (1-\frac{1}{\eta})\sum_{\theta_{-j}} u_j(\underline{a}, \theta_j, \theta_{-j})P_j(\theta_{-j} \mid \theta_j) := D,$$

where W is an upper bound on player j's payoff. We have

$$C - D = \frac{1}{\eta}(C-W) + (1-\frac{1}{\eta})\sum_{\theta_{-j}}(u_j(f(\theta_i), \theta_j, \theta_{-j}) - u_j(\underline{a}, \theta_j, \theta_{-j}))P_j(\theta_{-j} \mid \theta_j).$$

Letting $\varepsilon = \min_{a \neq \underline{a}, \theta}\{u_i(a,\theta) - u_i(\underline{a},\theta)\} > 0$, we find:

$$C - D \geq \frac{1}{\eta}(C-W) + (1-\frac{1}{\eta})\varepsilon.$$

Thus, for η large enough, the right-hand side is non-negative and player j has no incentive to deviate. Lastly, each player i has an incentive to transmit his true type since f is incentive compatible.

6.7 Detection with probability one

Lemma 7 *Let v be a random variable privately known by player i. If the network is weakly 2-connected, there exists a mechanism $\mathcal{M}_i$ on $\mathcal{N}$ such that, if all players abide by the mechanism, then the designer learns the value of v, whereas each player $j \neq i$ receives messages that are probabilistically independent from v. Furthermore, the designer detects deviations with probability one.*

The intuition is as follows. For each integer η, we can devise a test such that any deviation is detected with probability at least $1 - 1/\eta$. We may thus ask the players to pass *all such tests.*[20] There are several ways to construct such a test and we provide a relatively simple one. We modify our protocol M_i as follows. For simplicity, we assume that player i is not a provider.

Providers. Each provider a_m draws two independent infinite sequences $(X_\eta^{m,H}, X_\eta^{m,T})_{\eta \geq 1}$ of independently and identically (i.i.d.) distributed random variables, with uniform distribution on $[0,1)$ and sends these sequences.

Player i. Independently of his type and of the message he receives, player i draws an infinite sequence of i.i.d. fair coins $c_\eta \in \{H, T\}$. Define $(Y_\eta^H, Y_\eta^T)_{\eta \geq 1}$ as $(Y_\eta^H, Y_\eta^T) =$

[20] We thank Sylvain Sorin for suggesting this argument.

$(X_\eta^{1,H} \oplus \theta_i, X_\eta^{1,T})$ if $c_\eta = H$, and $(Y_\eta^H, Y_\eta^T) = (X_\eta^{1,H}, X_\eta^{1,T} \oplus \theta_i)$ if $c_\eta = T$. In words, for each η, player i chooses according to the toss of a fair coin whether to encode his type θ_i with $X_\eta^{1,H}$ or with $X_\eta^{1,T}$. Player i then sends the pair of sequences $(Y_\eta^H, Y_\eta^T)_{\eta \geq 1}$ to his active successor.

Other players. The other active players (transmitters, lock-closers and lock-openers) behave as in the proof of Theorem 2, except that now, vectors are sequences.

The designer. The designer receives two pairs of sequences $(x_\eta^{L,H}, x_\eta^{L,T})_{\eta \geq 1}$ and $(x_\eta^{L,H}, x_\eta^{R,T})_{\eta \geq 1}$. If for each η, it holds true that $(x_\eta^{L,H} = x_\eta^{R,H}$ and $x_\eta^{L,T} \neq x_\eta^{R,T})$ or $(x_\eta^{L,H} \neq x_\eta^{R,H}$ and $x_\eta^{L,T} = x_\eta^{R,T})$, the designer concludes that phase 1 of the test succeeds. Then, if $x_\eta^{L,T} \neq x_\eta^{R,T}$, he computes $\hat{\theta}_i = x_M^{\eta^*,R,T} \ominus x_M^{\eta^*,L,T}$ if $\nu(b_{M-1})$ is even and $\hat{\theta}_i = x_M^{\eta^*,L,T} \ominus x_M^{\eta^*,R,T}$ if $\nu(b_{M-1})$ is odd. If $x_\eta^{L,H} \neq x_\eta^{R,H}$, he computes $\hat{\theta}_i = x_M^{\eta^*,R,H} \ominus x_M^{\eta^*,L,H}$ if $\nu(b_{M-1})$ is even and $\hat{\theta}_i = x_M^{\eta^*,L,H} \ominus x_M^{\eta^*,R,H}$ if $\nu(b_{M-1})$ is odd. If all $\hat{\theta}_i^\eta$ have the same value $\hat{\theta}_i$, the designer concludes that phase 2 of the test succeeds, and regards $\hat{\theta}_i$ as the correct type of player i. If the test does not succeed, either in phase 1 or in phase 2, the designer concludes that there was a deviation.

Under these strategies, the decoded type clearly coincides with the true type. It is also clear that no player gets information about the message of player i. The sequence of coins being privately known to player i, each other active player only observes sequences of i.i.d. uniformly distributed variables. Now, we claim that any deviation is detected almost surely. Indeed, if some active player $j \neq i$ modifies the sequence, to pass the test in phase 2 he must modify an entry of the double sequence for each η. But then, to succeed in phase 1, he should modify only the component selected by player i. Consequently, the probability of passing the test while changing the message is at most the probability of guessing correctly an infinite sequence of fair coins, which is 0. Any deviation is thus detected with probability 1.

Corollary 2 *If the network is weakly 2-connected and if the environment has a bad outcome, i.e. an outcome $\underline{a}$ such that $u_i(a, \theta) \geq u_i(\underline{a}, \theta)$ for all $i \in N$, for all $a \in A$, for all $\theta \in \Theta$, then $F_{\mathcal{N}}(\mathcal{E}) = F_{\mathcal{N}^\star}(\mathcal{E})$.*

The proof consists in adapting the construction of Theorem 2. Using the above lemma, any deviation brings the bad outcome almost surely and is therefore not profitable.

6.8 Proof of Theorem 4

The "if" part being clear, we prove the "only if" part. Assume that there exists a player $i \notin D(0)$. We construct a profile of utility and an incentive-compatible social choice function $f : \Theta \to X$, which is not implementable on $\mathcal{N}$. The main feature of our construction is that when player j on a path from player i to the designer learns the type of player i, he has an incentive to misreport his own type.

Up to a relabeling of players, assume that player $1 \notin D(0)$ and $D(1) = \emptyset$, i.e., player 1 receives no messages and thus sends his messages at time 1, $t(1) = 1$.

Fix two alternatives a and b, and consider a social choice function with range $\{a, b\}$: $f : \times_{i=1}^{n} \Theta_i \to \{a, b\}$. With each type θ_i of player i, we associate a number in $\{0, 1\}$, i.e., we fix an onto mapping $\varphi_i : \Theta_i \to \{0, 1\}$. The social choice function we construct depends on types through these numbers only. Moreover, for each player i, there exists a unique $\theta_i^0 \in \Theta_i$ such that $\varphi_i(\theta_i^0) = 0$. Given a type profile $(\theta_1, \theta_2, \ldots, \theta_n)$, we denote $S = \sum_{j=2}^{n} \varphi_j(\theta_j)$ and $S_{-i} = S - \varphi_i(\theta_i)$. For convenience of language, we call $\varphi_i(\theta_i)$ the pseudo-type of player i.

The social choice function. We define the following social choice function:

$$f(\theta_1, \theta_{-1}) = \mathbf{1}_{\{\varphi_1(\theta_1)=0\}} \left[a\mathbf{1}_{\{S \le \alpha\}} + b\mathbf{1}_{\{S > \alpha\}}\right] + \mathbf{1}_{\{\varphi_1(\theta_1)=1\}} \left[b\mathbf{1}_{\{S \le \alpha\}} + a\mathbf{1}_{\{S > \alpha\}}\right],$$

where α is a fixed integer and $\mathbf{1}_E$ is the indicator function on the event E. In words, when $\varphi_1(\theta_1) = 0$, f chooses a if a large proportion of players $i = 2, \ldots, n$ are of pseudo-type 0. When $\varphi_1(\theta_1) = 1$, f chooses a if a small proportion of players are pseudo-type 0.

Next, we show that for a suitable choice of α, and for a class of utility functions, f is incentive compatible.

The utility functions. The utilities are as follows. Regardless of his type, player 1 is indifferent between a and b.

Player $i = 2, \ldots, n$ prefers a when he is of pseudo-type 0 and b when he is of pseudo-type 1. Further, his utility depends on his type and on the pseudo-type of player 1. The utility function is represented below:

where for each θ_i, $t_i(\theta_i), u_i(\theta_i), v_i(\theta_i), w_i(\theta_i)$ are positive numbers. We first show that f is incentive compatible and therefore, implementable on $\mathcal{N}^\star$ in pure strategies.

	a	b
$\theta_i : \varphi_i(\theta_i) = 0$	$t_i(\theta_i)$	0
$\theta_i : \varphi_i(\theta_i) = 1$	0	$u_i(\theta_i)$

$\theta_1 : \varphi_1(\theta_1) = 0$

	a	b
$\theta_i : \varphi_i(\theta_i) = 0$	$v_i(\theta_i)$	0
$\theta_i : \varphi_i(\theta_i) = 1$	0	$w_i(\theta_i)$

$\theta_1 : \varphi_1(\theta_1) = 1$

Claim 1. *For $\alpha = n - 2$ and suitable choices of $(t_i(\theta_i), u_i(\theta_i), v_i(\theta_i), w_i(\theta_i))_{i=2}^n$, f is incentive compatible.*

Consider the incentive constraints of player $i = 2, \dots, n$. Since f depends on pseudo-types only, the incentive constraints reduce to the incentive constraints over pseudo-types. If player i is of type θ_i such that $\varphi_i(\theta_i) = 0$, his expected payoff of announcing 0 is

$$t_i(\theta_i)P_i(S_{-i} \leq \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) + v_i(\theta_i)P_i(S_{-i} > \alpha, \varphi_1(\theta_1) = 1 \mid \theta_i)$$

If he announces 1, his expected utility is:

$$t_i(\theta_i)P_i(S_{-i} + 1 \leq \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) + v_i(\theta_i)P_i(S_{-i} + 1 > \alpha, \varphi_1(\theta_1) = 1 \mid \theta_i)$$

The associated incentive constraint says that the former is no less than the latter. This amounts to:

$$t_i(\theta_i)P_i(S_{-i} = \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) \geq v_i(\theta_i)P_i(S_{-i} = \alpha, \varphi_1(\theta_1) = 1 \mid \theta_i). \tag{1}$$

Because of the full-support assumption, both sides are positive and for each θ_i such that $\varphi_i(\theta_i) = 0$, one can find $(t_i(\theta_i), v_i(\theta_i))$ such that (1) holds.

Similarly, if player i is of type θ_i such that $\varphi_i(\theta_i) = 1$, his expected payoff of announcing 1 is:

$$u_i(\theta_i)P_i(S_{-i} + 1 > \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) + w_i(\theta_i)P_i(S_{-i} + 1 \leq \alpha, , \varphi_1(\theta_1) = 1 \mid \theta_i)$$

If he announces 0, his expected utility is:

$$u_i(\theta_i)P_i(S_{-i} > \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) + w_i(\theta_i)P_i(S_{-i} + 1 \leq \alpha, \varphi_1(\theta_1) = 1 \mid \theta_i)$$

The associated incentive constraint amounts to:

$$u_i(\theta_i)P_i(S_{-i} = \alpha, \varphi_1(\theta_1) = 0 \mid \theta_i) \geq w_i(\theta_i)P_i(S_{-i} = \alpha, \varphi_1(\theta_1) = 1 \mid \theta_i). \tag{2}$$

Both sides are positive and for each θ_i such that $\varphi_i(\theta_i) = 1$, one can find $(u_i(\theta_i), w_i(\theta_i))$ such that (2) holds.

To complete the proof of Theorem 4, we now show that the social choice function f is not partially implementable on $\mathcal{N}$.

Assume by contradiction that there exists a mechanism (M, g) on $\mathcal{N}$ and a pure strategy Bayesian-Nash equilibrium s of the induced game such that $f = g \circ s$.

Since player 1 receives no messages ($D(1) = \emptyset$), the strategy of player 1 of type θ_1 is a tuple of messages $s_1(\theta_1) = (m_{1j}(\theta_1))_{j \in C(1)}$, where $m_{1j}(\theta_1)$ is the message that player 1 sends to player $j \in C(1)$ when he is of type θ_1.

Note that, for every type profile of the other players, the pseudo-type of player 1 determines the alternative chosen by f: $\forall \theta_{-1}$,

$$f(0, (\varphi_j(\theta_j))_{j \neq 1}) \neq f(1, (\varphi_j(\theta_j))_{j \neq 1}).$$

It follows that the tuple of messages sent by player 1 of type θ_1 s.t. $\varphi_1(\theta_1) = 0$ is different from the tuple of messages sent by player 1 of type θ_1 s.t. $\varphi_1(\theta_1) = 1$. Recall that there is only one type θ_1^0 of player 1 such that $\varphi_1(\theta_1) = 0$. We thus have,

$$\forall \theta_1 \,\text{s.t.}\, \varphi_1(\theta_1) = 1, (m_{1j}(\theta_1^0))_{j \in C(1)} \neq (m_{1j}(\theta_1))_{j \in C(1)}$$

and therefore, for each θ_1 such that $\varphi_1(\theta_1) = 1$, there exists a player $i \in C(1)$ for whom $m_{1i}(\theta_1^0) \neq m_{1i}(\theta_1)$. We conclude that, for each $\theta_1 \neq \theta_1^0$, when player 1 is of type θ_1, there exists a player $i \in C(1)$ who learns from the messages that player 1's type is not θ_1^0. In particular, player i learns that the pseudo-type of player 1 is 1. We claim that this player has an incentive to deviate after receiving a message different from $m_{1i}(\theta_1^0)$.

Claim 2. *Player $i \in C(1)$, of type θ_i^0, receiving a message $m_{1i} \neq m_{1i}(\theta_1^0)$ from player 1, has an incentive to deviate from s.*

Proof. Let us fix θ_1^* such that $\varphi_1(\theta_1^*) = 1$ and a player $i \in C(1)$ of type θ_i^0, such that $m_{1i}(\theta_1^*) \neq m_{1i}(\theta_1^0)$. Consider also a profile of types $(\theta_k^*)_{k \neq 1, k \neq i}$ such that for each k, $\varphi_k(\theta_k^*) = 1$. For this type profile, $S_{-i} = n - 2$. Since $f = g \circ s$, if player i announces 0, i.e. plays what s recommends for type θ_i^0, then $S = n - 2$ and y is chosen. If player i announces 1, i.e. plays what s recommends for a type $\theta_i \neq \theta_i^0$, then $S = n - 1$ and x is chosen. Thus, if player i *knew* that the pseudo-type is 1 for every other player, he would have a clear incentive to play as if he were *not* of type θ_i^0.

Let m_i^* be the tuple of messages received by player i (under s) when the types of the other players are $(\theta_1^*, (\theta_k^*)_{k\neq 1, k\neq i})$. This tuple of messages occurs with positive probability. From the above discussion, player i deduces the pseudo-type of player 1 from m_i^*:

$$P_i(\varphi_1(\theta_1) = 1 \mid m_i^*, \theta_i^0) = 1.$$

Further, player i attributes a positive posterior probability to $(\theta_1^*, (\theta_k^*)_{k\neq 1, k\neq i})$:

$$P_i(\theta_1^*, (\theta_k^*)_{k\neq 1, k\neq i} \mid m_i^*, \theta_i^0) > 0.$$

We have thus exhibited a situation (i.e. messages, or an information set in the extensive game) where player i of type θ_i^0 knows that $\varphi_1(\theta_1) = 1$ and influences the selected alternative with positive probability. He faces thus the same kind of incentive problem as in the direct mechanism, except for the beliefs (priors are replaced by posteriors).

The expected utility of player i of type θ_i^0, conditional on m_i^* and if plays according to $s_i(\theta_0^i)$ is:

$$v_i(\theta_0^i) P(S_{-i} > n-2 \mid m_i^*, \theta_i^0) := \underline{v}.$$

If he "announces" 1, that is if he plays according to $s_i(\theta_i)$ with $\varphi_i(\theta_i) = 1$, his expected utility is:

$$v_i(\theta_0^i) P(S_{-i} + 1 > n-2 \mid m_i^*, \theta_i^0) := \overline{v}.$$

One has,

$$\overline{v} - \underline{v} = v_i(\theta_0^i) P(S_{-i} = n-2 \mid m_i^*, \theta_i^0) \geq v_i(\theta_0^i) P_i(\theta_1^*, (\theta_k^*)_{k\neq 1, k\neq i} \mid m_i^*, \theta_i^0) > 0.$$

This gives the desired contradiction. □

6.9 Proof of Theorem 5

The proof is very similar to the proofs of Theorems 1 and 2. The proof that the condition is necessary is the same. For sufficiency, the main task is to extend Proposition 3 to weakly 2-connected networks with cycles. Once this is established, Theorem 5 follows, similarly as for Theorems 1 and 2 and this part of the proof is omitted.

We now explain how to extend Proposition 3. A important remark is the following. Since the network has cycles, the existence of the timing structure is no longer guaranteed, in fact it simply fails. To define a mechanism, one has to specify a timing structure, i.e., who speaks first, who speaks second, and so on. To avoid this difficulty, we associate to the network $\mathcal{N}$, an augmented network $\mathcal{N}^A$, which is strongly 1-connected, weakly 2-connected and acyclic. Thus, Proposition 3 holds true on $\mathcal{N}^A$. Then, we show how the protocol on $\mathcal{N}^A$ induces the desired protocol on $\mathcal{N}$.

Let us fix a strongly 1-connected and weakly 2-connected network $\mathcal{N}$ (but not necessarily acyclic). Recall that a network is a set of edges. A sub-network is thus a subset of edges.

Lemma 8 *There exists an acyclic and strongly 1-connected sub-network $\mathcal{N}^a$ of $\mathcal{N}$.*

Proof. For each $i \in N$, consider a shortest directed path from i to 0 in $\mathcal{N}$. Such a shortest directed path exists since $\mathcal{N}$ is strongly 1-connected. Let $\mathcal{N}^a$ be the collection of all these paths. We claim that $\mathcal{N}^a$ has the required properties. By construction, it is strongly 1-connected. Let us show that it is acyclic. By contradiction, assume that $\mathcal{N}^a$ contains the cycle $i_1 \to i_2 \to \ldots \to i_K \to i_1$. By construction, $\mathcal{N}^a$ is such that $C(0) = \emptyset$, i.e., there is no edge $0i$ for some $i \in N$ in $\mathcal{N}^a$. It follows that the cycle does not contain the designer (player 0). It then follows that there exists $k \in \{2, \ldots, K\}$ such that the shortest path from i_k to 0 does not follow the cycle (otherwise, 0 cannot be reached, a contradiction with 1-strong connectedness). Thus, the edge $i_k i_{k+1}$ is not on a shortest path from any player j to 0, contradicting the construction of $\mathcal{N}^a$. •

With a slight abuse of notation, let $\mathcal{N}^a$ be a maximal acyclic and strongly 1-connected sub-network of $\mathcal{N}$ (it exists by the preceding lemma) and let $\mathcal{C} = \mathcal{N} \backslash \mathcal{N}^a$ be the set of edges of $\mathcal{N}$ that do not belong to $\mathcal{N}^a$. Note that every edge of $\mathcal{C}$ belongs to a cycle of $\mathcal{N}$ and that every cycle of $\mathcal{N}$ contains an edge in $\mathcal{C}$. Let $\mathcal{N}^A$ be the network obtain from $\mathcal{N}$ by replacing each edge ij in $\mathcal{C}$ by two edges: $i(j)i$ and $i(j)j$, where $i(j)$ is a fictitious player who is a duplicate of player i. That is, if ij in $\mathcal{C}$:

$$i \to j \text{ is replaced by } i \leftarrow i(j) \to j.$$

The edges of $\mathcal{N}^a$ are unchanged. See Figure 17 for an example.

Claim 3 *$\mathcal{N}^A$ is strongly 1-connected, weakly 2-connected and acyclic.*

Proof. Each "regular" player i has a directed path to 0 in $\mathcal{N}^a$ by construction. Since the fictitious player $i(j)$ is directly connected to i, he also has a path to the designer by strong 1-connectedness of $\mathcal{N}$. Weak 2-connectedness is clearly preserved by the

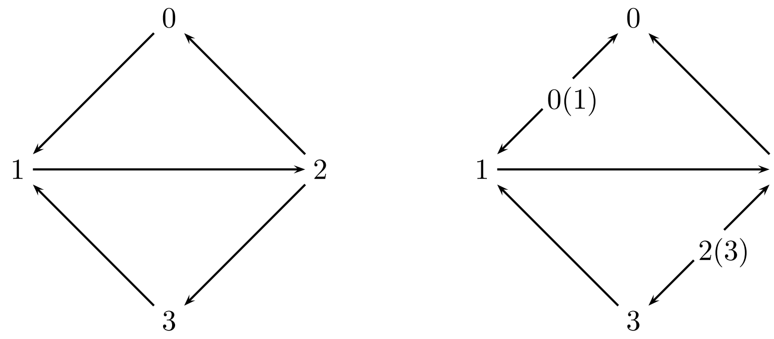


Figure 17: A cyclic network $\mathcal{N}$ and the associated acyclic $\mathcal{N}^A$

transformation. Let us show that $\mathcal{N}^A$ is acyclic. Assume that $\mathcal{N}^A$ contains a cycle. By our construction, each fictitious player has only out-going edges, thus cannot belong to a cycle. This implies that the cycle was already a cycle in $\mathcal{N}$ and therefore, it should contain an edge which belongs to $\mathcal{C}$. This is a contradiction because edges in $\mathcal{C}$ no longer appear in $\mathcal{N}^A$. •

Now, we claim that Proposition 3 extends to strongly 1-connected, weakly 2-connected networks with cycles. First, on the network $\mathcal{N}^A$, for each player i, there exists a protocol with the desired property by Proposition 3. We assume that each fictitious player has no type and a constant payoff function. Second, on the network $\mathcal{N}$, the players can replicate this protocol. The timing of the protocol is the one given by the timing structure of $\mathcal{N}^A$, which is well-defined since $\mathcal{N}^A$ is acyclic and strongly 1-connected. In particular, each duplicated player i plays only twice: he plays as the fictitious player $i(j)$ the first time and as player i the second time.

Thus, Proposition 3 extends and Theorem 5 follows, similarly as for Theorems 1 and 2.

References

[1] Amos Beimel and Matthew Franklin, Reliable Communication over Partially Authenticated Networks, Theoretical Computer Science, 1999, 220, pp. 185-210.

[2] Béla Bollobàs, Modern Graph Theory, 1998, Springer Verlag.

[3] Patrick Bolton and Mathias Dewatripont, The Firm as a Communication Network, Quaterly Journal of Economics, 1994, 109, pp. 809-839.

[4] Dirk Bergemann and Stephen Morris, Robust Mechanism Design, Econometrica, 2005, 73, pp. 1771-1813.

[5] Dirk Bergemann and Stephen Morris, Robust Implementation in Direct Mechanisms, Review of Economic Studies, 2009, 76, pp. 1175-1204.

[6] Guillermo A. Calvo and Stanislaw Wellisz, Supervision, Loss of Control and the Optimal Size of the Firm, Journal of Political Economy, 1978, 86, pp. 943-952.

[7] Partha Dasgupta, Peter Hammond and Eric Maskin, The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility, The Review of Economic Studies, 1979, 46, pp. 185-216.

[8] Yvo Desmedt and Yongge Wang, Perfectly Secure Message Transmission Revisited, Lecture Notes in Computer Science, Advances in Cryptology EUROCRYPT 2002, 2002, Volume 2332/2002, pp. 502-517.

[9] Bailey Diffie and Martin Hellman, New directions in cryptography, IEEE transactions on information theory, 1976, 22, pp. 644654.

[10] Danny Dolev, Cynthia Dwork, Orli Waarts and Moti Yung, Perfectly Secure Message Transmission, Journal of the ACM, 1993, 40, pp. 17-47.

[11] Matthew K. Franklin and Rebecca N. Wright, Secure Communication in Minimal Connectivity Models, Journal of Cryptology, 2000, 13, pp. 9-30.

[12] Guido Friebel and Michael Raith, Abuse of Authority and Hierarchical Communication, The RAND Journal of Economics, 35, 2004, pp. 224-244.

[13] Allan Gibbard, Manipulation of Voting Schemes: A General Result, Econometrica, 1973, 41, pp. 587-601.

[14] Shafi Goldwasser and Silvio Micali, Probabilistic Encryption, Journal of Computer and Systems Sciences, 1984, 28, pp. 270-299.

[15] Milton Harris and Robert M. Townsend, Resource Allocation Under Asymmetric Information, Econometrica, 1981, 49, pp. 33-64.

[16] Matthew O. Jackson, Bayesian Implementation, Econometrica, 1991, 59, pp. 461-477.

[17] Matthew O. Jackson, A Crash Course in Implementation Theory, Social Choice and Welfare, 2001, 18, pp. 655-708.

[18] Vijay Krishna, Auction Theory, Academic Press, 2002.

[19] Eric Maskin and Tomas Sjöström, Implementation Theory, Chapter 5 in Handbook of Social Choice and Welfare, Volume 1. Eds. K.J Arrow, A.K. Sen, and K. Suzumura, Elsevier 2002.

[20] Dov Monderer and Moshe Tennenholtz, Distributed Games: From Mechanisms to Protocols, Sixteenth National Conference on Artificial Intelligence, 1999, pp. 32-37.

[21] Dilip Moockerjee, Incentives in Hierarchies, 2009, forthcoming in Handbook of Organizational Economics, R. Gibbons and J. Roberts, Ed.

[22] Roger B. Myerson, Incentive Compatibility and the Bargaining Problem, Econometrica, 1979, 47, pp. 61-73.

[23] Roger B. Myerson, Optimal Coordination Mechanisms in Generalized Principal-Agent Problems, Journal of Mathematical Economics, 1982, 10, pp. 67-81.

[24] Noam Nisan and Ilya Segal, The Communication Complexity of Efficient Allocation and Supporting Prices, Journal of Economic Theory, 2006, 129, pp. 192-224.

[25] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani (Eds.), Algorithmic Game Theory, 2007, Cambridge University Press.

[26] Tal Rabin and Michael Ben-Or, Verifiable Secret Sharing and Multiparty Protocols with Honest Majority, Proceedings of the 21st Symposium on the Theory of Computing, 1989, pp 73-85.

[27] Roy Radner, The Organization of Decentralized Information Processing, Econometrica, 1993, 61, pp. 1109-1146.

[28] Ronald Rivest, Adi Shamir and Leonard Adleman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, Communications of the ACM, 1978, 21, pp. 120-126.

[29] Jérôme Renault and Tristan Tomala, Probabilistic Reliability and Privacy of Communication Using Multicast in General Neighbor Networks, Journal of Cryptology, 2008, 21, pp. 250-279.

[30] Ludovic Renou, Nash Implementation and Communication Networks, 2008, Mimeo, University of Leicester.

[31] Bernard Salanie, The Economics of Contracts - A Primer. Cambridge University Press, 2000.

[32] Claude Shannon, Communication Theory of Secrecy Systems, Bell System Technical Journal, 1949, 28, pp. 656-715.

[33] Roberto Serrano and Rajiv Vohra, Multiplicity of Mixed Equilibria in Mechanisms: a Unified Approach to Exact and Approximate Implementation, Working Paper 2009-11, Department of Economics, Brown University.

[34] Timothy Van Zandt, Communication Complexity and Mechanism Design, Journal of European Economic Association, 2007, 5, pp. 543-553.

[35] Oliver Williamson, Hierarchical Control and Optimal Firm Size, Journal of Political Economy, 1967, pp. 123-138.