

# Best-Response Mechanisms

Noam Nisan<sup>1</sup> Michael Schapira<sup>2</sup> Gregory Valiant<sup>3</sup> Aviv Zohar<sup>4</sup>

<sup>1</sup>School of Eng. and Computer Science, The Hebrew University of Jerusalem

<sup>2</sup>Computer Science Dept., Princeton University

<sup>3</sup>Dept. of Computer Science, UC Berkeley <sup>4</sup>Microsoft Research, Silicon Valley

noam@cs.huji.ac.il michael.schapira@yale.edu gvaliant@eecs.berkeley.edu avivz@microsoft.com

**Abstract:** Under many protocols—in computerized settings and in economics settings—participants repeatedly “best respond” to each others’ actions until the system “converges” to an equilibrium point. We ask when does such myopic “local rationality” imply “global rationality”, *i.e.*, when is it best for a player, given that the others are repeatedly best-responding, to also repeatedly best-respond? We exhibit a class of games where this is indeed the case. We identify several environments of interest that fall within our class: models of the Border Gateway Protocol (BGP) [7], that handles routing on the Internet, and of the Transmission Control Protocol (TCP) [5], and also stable-roommates [3] and cost-sharing [9, 10], that have been extensively studied in economic theory.

**Keywords:**best response, mechanism design, incentive compatible dynamics.

## 1 Introduction

### 1.1 Motivation: when is it best to best-respond?

The basic object of study in game theory and in economics is the *equilibrium*: a “stable” state from which none of the players wish to deviate. Equilibrium is a static concept that often abstracts away the question of how it is reached. Once we start looking at dynamics, or at algorithms for finding equilibria, we cannot escape questions of the form “*How is an equilibrium reached?*”. While there can be different formalizations of this question, in most cases, a truly satisfactory answer would have each player performing only simple “locally rational” actions and yet, mysteriously, the system would reach a global equilibrium. The simplest example of such phenomena is *repeated best-response dynamics*: each player selects the best (locally optimal) response to what others are currently doing, and this process goes on “for a while” until it “converges” to what must be a (pure Nash) equilibrium. Convergence of repeated best-response is, unfortunately, not guaranteed in general, and is the subject of much research, as is the convergence of more sophisticated “locally-rational” dynamics, *e.g.*, fictitious play or regret minimization.

Our focus in this paper is on a different question that has received little attention so far: “*Is such locally rational behavior really rational?*”. Specifically, we consider games in which repeated best-response

dynamics *do* converge to an equilibrium and study the incentive properties of this process: Is it *rational* for players to *repeatedly* best-respond? Can a long-sighted player improve, in the long run, over this repeated *myopic* optimization?

These questions about incentives are best explored in the context of games with *incomplete information*. Switching our attention from games with complete information to games with uncoupled incomplete information, we see that repeated best-response exhibits another attractive trait: to best-respond each player need only know his own utility function (“type”), as his best response does not depend on other players’ utility functions, but only on their actions. Thus, we can view best-response dynamics as a natural protocol for gradual and limited sharing of information in an effort to reach an equilibrium. Indeed, in many real-life contexts the interaction between decision makers with incomplete information takes the form of best-response dynamics (*e.g.*, Internet routing [7]). When regarding best-response dynamics from this perspective, it is an indirect mechanism in the private-information mechanism-design sense. We wish to understand when such a mechanism, that dictates that all players repeatedly best-respond, is *incentive compatible*.

### 1.2 The setting

Let us begin by laying out our setting for studying and formalizing incentives for repeated best-response.

In our framework, each player holds a private utility function, and all players’ utility functions, when put together, determine a full-information *base game* with some commonly-known strategy spaces. We desire that the outcome of the dynamics be an equilibrium of this base game.

**Base game:** We are given an  $n$ -player (one-shot) *base game*  $G$ , with players  $1, \dots, n$ , in which each player  $i$  has strategy space  $S_i$ , and  $S = S_1 \times \dots \times S_n$ . Each player  $i$  has a *utility function*  $u_i$  such that  $(u_1, \dots, u_n) \in U \subseteq U_1 \times \dots \times U_n$ , where  $U_i \subseteq \mathbb{R}^{|S|}$  is player  $i$ ’s utility space. Each player knows only his own utility function, *i.e.*, we view  $u_i$  itself as player  $i$ ’s *type*.

**Best-response mechanisms:** We study a class of indirect mechanisms, that we term “*repeated-response mechanisms*”: players take turns selecting strategies; at each (discrete) time step  $t$ , some player  $i_t$  selects and announces strategy  $s_{i_t}^t \in S_{i_t}$ . Observe that one course of action available to each player in a repeated-response mechanism is to always choose a best-response to the most recently announced strategies of the others, that is, *repeated-best-response*. We call a repeated-response mechanism in which the *prescribed behavior* for each player is to repeatedly best-respond a “*best-response mechanism*”. To fully-specify a best-response mechanism we must specify (1) the starting state; (2) the order of player activations (which player is “active” when); and (3) for each player, a rule for breaking ties among multiple best responses. All of our results hold regardless of the initial state and of the order of players’ activations (so long as it is “long enough”), and, in fact, even in more general settings.<sup>1</sup> We discuss tie-breaking rules below.

**Goal:** Our general aim is to identify interesting classes of (base) games for which best-response mechanisms are *incentive-compatible*. Intuitively, a best-response mechanism is incentive-compatible if, when all other players are repeatedly best-responding, then a player is incentivized to do the same. Defining incentive compatibility in our setting involves many intricacies. We opt to focus here on a very general notion of incentive compatibility that, we believe, captures es-

entially any variant that the reader may desire; in a companion paper [11], we present several more games (auctions) where only strictly weaker notions of incentive compatibility can be obtained. Our notion of incentive compatibility here captures the two following distinct but complementary points of view: a mechanism design perspective and a learning equilibrium [1, 2] perspective.

**Mechanism design perspective** (in a prior-free non-Bayesian setting): This point of view is natural when analyzing finite-time protocols in computerized and economic settings. We are given a game with incomplete information  $G$ , where each player’s utility function is private, and we wish to implement a pure Nash equilibrium (PNE) of  $G$ . We point out that this uncommon objective—implementing an equilibrium—proves to be a natural implementation goal in many contexts (see Section 3, where we show that desirable outcomes can be regarded as “stable states”). Best-response mechanisms are incentive compatible, from this perspective, if the desired outcomes are implemented in the ex-post Nash sense<sup>2</sup>. Importantly, from this point of view, no actual play happens during the process of best-response dynamics and players merely announce strategies as their communication with the mechanism; each player only cares about maximizing his benefit from the *final outcome* of the mechanism, that is expected to terminate after some *finite* predetermined number of time steps.

**Learning equilibrium perspective:** This point of view is natural when analyzing environments such as Internet protocols and global financial transactions, where players repeatedly interact with *each other* and there is no “final turn”. Now, the players are actually involved in *infinite* repeated play of the incomplete-information game  $G$  and each player has a rule for selecting his next strategy based on the history of play. We are interested in the natural rule that dictates that a player simply always best-respond to others’ most-recent strategies. In this context, each player wishes to maximize his *long-term* payoff, that we model to be the lim sup of his stage utilities in this infinitely-played game<sup>3</sup>. Best-response mechanisms are incentive compatible, from this perspective, if the “best-response” rules are themselves in equilibrium in this infinite

<sup>1</sup>Our results actually hold even for (1) *asynchronous* player activation orders in which multiple players can best-respond *simultaneously* or based on outdated information (as studied in [12]); (2) adaptive player activation orders that can change based on the history of play; and also when (3) the mechanism terminates as soon as all players “pass”, that is, each player repeats his last strategy.

<sup>2</sup>The Revelation Principle then implies that the direct revelation mechanism is truthful (in the ex-post-nash sense).

<sup>3</sup>In all our results, at equilibrium the lim sup is actually the limit, and thus choosing lim sup gives us the strongest and most robust results – the definition is in fact adversarial to our proofs, it potentially allows manipulators to gain utility by avoiding convergence.

game regardless of the realization of  $(u_1, \dots, u_n)$ . Using the terminology of [1, 2], this means that best-response dynamics are in “learning equilibrium”. We stress that this would not follow from the folk theorem since our players do not, in any way, punish other players for deviation. To the contrary, our incentive compatibility results establish that the natural best-response dynamics are in equilibrium without requiring players to be able to detect and penalize other players’ deviations.

**Tie-breaking rules.** When multiple best-responses exist we must specify, for each player, a *tie-breaking rule*. Importantly, this tie-breaking rule must be “uncoupled”, *i.e.*, depend solely on the player’s private information (utility function) and not on information that is unavailable to him<sup>4</sup>. Our tie-breaking rules always have the following simple form: fix, for each player  $i$ , an a-priori full order  $\prec_i$  on  $S_i$  (that can depend on  $u_i$ ), and instruct player  $i$  to break ties between multiple best-responses according to  $\prec_i$ . While this might seem innocent enough, we do get significant mileage from delicate choices of these tie-breaking rules, to the point that one may desire an intuitive justification for these choices. Roughly speaking, there are two main, conflicting, intuitions: in some cases we simply ask players to break ties so as to be “nice” to others; in other cases we break ties according to some “iterated-trembling-hand” logic.

### 1.3 Games with incentive-compatible best-response mechanisms

Our main results are identifying a class of games for which best-response mechanisms are incentive compatible, and exhibiting several interesting games that fall within this class (and thus have incentive-compatible best-response mechanisms). While at first glance, it might seem that the existence of a *unique* PNE to which best-response dynamics are guaranteed to converge implies the incentive-compatibility of best-response mechanisms, this intuition is false.

	C	D
A	2,1	0,0
B	3,0	1,2

Figure 1: A game for which best-response mechanisms are not incentive compatible.

<sup>4</sup>We note that it is also permissible for the tie-breaking rules to depend on the players’ actions, though for our purposes this was not needed.

Observe that in this game,  $(B, D)$  is the unique PNE and every sequence of best responses converges to it. Yet, consider the scenario that the starting point is the strategy profile  $(A, C)$ , and the column player repeatedly best-responds. Clearly, the row player’s *local* improvement from  $(A, C)$  to  $(B, C)$  will lead to the column player moving to  $(B, D)$ . Hence, the row player can do better by looking ahead, not moving from  $(A, C)$ , and thus “getting stuck” at  $(A, C)$ , that he strictly prefers to the unique pure Nash  $(B, D)$ . Hence, repeated best-responding is not incentive compatible in this game which is strictly-dominance-solvable, is a potential game, and has a unique and Pareto-optimal PNE.

What traits must a game have for best-response dynamics to be incentive compatible? We now present an intuitive exposition of a class of games for which this is achieved, which we term “*Never-Best-Response-Solvable (NBR-solvable) games with clear outcomes*”. In an NBR-solvable game, strategies are iteratively eliminated if a best-response never leads to them (this is slightly different from dominance-solvability and shall be defined in the following section). Intuitively, an NBR-solvable game has a *clear outcome* if when each player  $i$  considers the game after the other players have already eliminated strategies that can be eliminated regardless of what  $i$  does, he can already tell that he will not be able to do better than the outcome that is reached via repeated best-response.

Our main, and quite easy to prove, general theorem is the following. (We now state the theorem for the case that the strategy spaces are finite, though our result also holds for infinite strategy spaces.)

**Theorem** (informal): *Let  $G$  be an NBR-solvable game with a clear outcome. Then, for every starting point and every (finite or infinite) order of player activations with at least  $T = \sum_i |S_i| - n$  “rounds” (a round is a sequence of consecutive time steps in which each player is “active” at least once) it holds that:*

1. *Repeated best-response dynamics converges to a pure Nash equilibrium  $s^*$  of  $G$ .*
2. *Repeated best-response dynamics is incentive compatible.*

We prove that each of the four environments below can be formulated as a game that falls within our class of games, and that the desired outcome in each environment translates to a PNE in this formulation. Thus, the above result implies the existence of incentive-compatible best-response mechanisms that

implement the desired outcome in all the contexts below.

- **Stable-roommates.** In this classic setting [3], students must be paired for the purpose of sharing dorm rooms, and each student has a private full order over possible roommates. The objective is to find a “stable matching” where no two students prefer each other to their assigned roommates. We show that a natural mechanism, in which a student repeatedly proposes to his most preferred roommate among those that would not immediately reject him, and immediately rejects all proposers except for his most preferred proposer, is incentive compatible in well-studied environments (interns-hospitals, correlated markets).
- **Cost-sharing.** Cost-sharing arises in situations in which the cost of some public service (*e.g.*, building a bridge) must be distributed between self-interested users that can benefit from this service to different extents. We present a *distributed* mechanism that achieves this goal in an incentive-compatible manner. Our mechanism implements the outcome of the famous Moulin mechanism [9, 10] (this result can be extended to the more general class of “acyclic mechanisms” [8]).
- **Internet routing.** The Border Gateway Protocol (BGP) establishes routes between the smaller networks that make up the Internet. We abstract the results in [7] and prove that BGP is incentive compatible in realistic environments.
- **Congestion control.** The Transmission Control Protocol (TCP) handles congestion on the Internet. Building upon [5], that models key aspects of TCP, we consider behavior that is somewhat similar to TCP: increase your attempted transmission rate until encountering congestion, and then decrease the transmission rate. We show that such behavior is in equilibrium.

Our results above establish incentive compatibility of best-response mechanisms. We also consider the stronger “*collusion-proofness*” desideratum, that even a *coalition* of players not be able to deviate from repeated best-response and *all* strictly gain from doing so. We prove that in some of the above environments best-response mechanisms even achieve this stronger requirement.

## 1.4 Research agenda

We view this work as a first step towards a more

general research agenda. While convergence to equilibrium of “locally-rational” dynamics, *e.g.*, repeated best-response, fictitious play and regret minimization, has been extensively studied, little attention has been given to the question of when such locally-rational dynamics are also “globally rational”. Here, we tackle this question in the context of repeated best-response and the implementation of PNE. However, we believe that the examination of other dynamics (*e.g.*, fictitious play, regret minimization) and other kinds of equilibria (*e.g.*, mixed Nash equilibrium, correlated equilibrium) is an interesting direction for future research. Positive and negative results along these lines can help shed new light on the incentive structure of existing protocols/mechanisms (see our results for BGP and TCP and the results in [5, 7]), and provide new insights into the design of new protocols/mechanisms.

Our results for repeated best-response dynamics establish *sufficient* conditions for repeated best-response to be incentive compatible. We still lack *characterizations* of conditions that imply incentive compatibility both for general games and for specific classes of games (dominance-solvable games, potential games, *etc.*). We have thus far considered a very strong notion of incentive compatibility. We believe that considering more restrictive notions (*e.g.*, incentive compatibility in expectation) is of interest. Indeed, in a companion paper [11] we present several such results for commerce environments.

## 1.5 Organization

In the next section we formalize our model and present our general theorem. In section 3 we present our results for the four specific environments listed above. We discuss collusion-proofness in Section 4.

## 2 Incentive-compatible best-response dynamics

**Definition 2.1 (tie-breaking rules)** A tie-breaking rule (or tie-breaking order) for player  $i$  is a full order  $\prec_i$  on  $S_i$ .

When faced with a choice between multiple best-responses, player  $i$  should choose the highest (under  $\prec_i$ ) best-response. We now present the following definitions for full-information games.

**Definition 2.2 (never-best-response strategies)**  $s_i \in S_i$  is a never-best-response (NBR) under tie-

breaking order  $\prec_i$  on  $S_i$  if for all  $s_{-i}$ , there exists  $s'_i$  so that  $u_i(s_i, s_{-i}) < u_i(s'_i, s_{-i})$  OR both  $u_i(s_i, s_{-i}) = u_i(s'_i, s_{-i})$  and  $s_i \prec_i s'_i$ .

**Definition 2.3 (NBR-solvable games)** A game  $G$  is never-best-response-solvable (NBR-solvable) under tie-breaking rules  $\prec_1, \dots, \prec_n$  if there exists a sequence of eliminations of NBR strategies (under these tie breaking rules) that results in a single strategy profile.

Observe that every weakly-dominance-solvable game has a tie-breaking order under which it is NBR-solvable and every strongly-dominance-solvable game is NBR-solvable for all tie-breaking orders. Observe also that in every game that is NBR-solvable under tie-breaking rules  $\prec_1, \dots, \prec_n$  the elimination of NBR strategies (under these tie-breaking rules) has a unique order-independent outcome, that is a pure Nash equilibrium of the game. We call this outcome “the unique PNE under tie-breaking”.

**Definition 2.4 (shortest-elimination parameters)** Let  $G$  be an NBR-solvable game (under tie-breaking). Then, there exists a sequence of games  $G_0, \dots, G_r$  such that  $G = G_0$ , in  $G_r$  each player has only a single strategy, and  $\forall i \in \{0, \dots, r-1\}$ ,  $G_{i+1}$  is obtained from  $G_i$  via the removal of sets of NBR strategies (under tie-breaking). The shortest-elimination parameter  $e_G$  for  $G$  is the length of the shortest such sequence of games for  $G$ .

Observe that if, in an NBR-solvable game  $G$ , each strategy space  $S_i$  is finite, then  $e_G \leq \sum_i |S_i| - n$ . NBR solvability on its own is insufficient to guarantee incentive compatibility, and so we further restrict it.

**Definition 2.5 (globally-optimal profiles)**  $s \in S$  is globally optimal for  $i$  if  $\forall t \in S$ ,  $u_i(t) \leq u_i(s)$ .

**Definition 2.6 (clear outcomes)** Let  $G$  be an NBR-solvable game under tie breaking rules  $\prec_1, \dots, \prec_n$ . Let  $s^*$  be the unique PNE under tie-breaking of  $G$ . We say that  $G$  has a clear outcome if for every player  $i$  there exists a (player-specific) order of elimination of NBR strategies (under the given tie-breaking rules) such that  $s^*$  is globally optimal for  $i$  at the first step in the elimination sequence in which a strategy in  $S_i$  is eliminated (that is, in the game obtained after the removal of all previously-eliminated strategies from  $G$ ).

We say that an incomplete-information game  $G$  is NBR-solvable with a clear outcome (under tie-

breaking rules) if every realization of  $(u_1, \dots, u_n)$  induces a full-information game that is NBR-solvable with a clear outcome (under tie-breaking, when each player  $i$  uses the tie-breaking rule  $\prec_i$  for the realized  $u_i$ ).

Consider a best-response mechanism  $M$  for a base game  $G$ . Let  $s^t \in S$  be the players' strategies at time step  $t$ . We call  $u_i(s^t)$  player  $i$ 's stage utility at time  $t$ . If  $M$  terminates after some finite number of time steps  $T > 0$  we say that player  $i$ 's total utility is  $\Gamma_i = u_i(s^T)$  (his stage utility at the last time step of  $M$ 's execution). If  $M$  does not terminate after finite time then  $i$ 's total utility is  $\Gamma_i = \limsup_{t \rightarrow \infty} u_i(s^t)$ .  $M$  is incentive compatible if repeated best-response is a pure Nash equilibrium in this repeated game with overall utilities  $\Gamma_1, \dots, \Gamma_n$  for every realization of  $(u_1, \dots, u_n)$ . We say that  $M$  is collusion-proof if no coalition can deviate from repeated best-response and all strictly gain from doing so in this repeated game. We show that best-response mechanisms are incentive compatible for NBR-solvable games with clear outcomes.

**Theorem 2.7 (incentive-compatible mechanisms)** Let  $G$  be NBR-solvable with a clear outcome  $s^* \in S$  under tie-breaking rules  $\prec_1, \dots, \prec_n$ . Let  $M$  be a best-response mechanism for  $G$  that breaks ties as in  $\prec_1, \dots, \prec_n$ . Then, for every starting point and every (finite or infinite) order of player activations with at least  $T = e_G$  “rounds”, where a round is a sequence of consecutive time steps in which each player is “active” at least once,

1.  $M$  converges to  $s^*$ .
2.  $M$  is incentive compatible.

This holds even for (1) asynchronous player activations orders in which multiple players can best-respond simultaneously or based on outdated information (as studied in [12]); (2) adaptive player activations orders that can change based on the history of play; and also when (3) the mechanism terminates as soon as all players “pass”, that is, each player repeats his last strategy.

**Proof sketch:** Let  $G$  be an NBR-solvable game with a clear outcome (under tie-breaking). Then, there exists a sequence of games  $G_0, \dots, G_r$  with length  $r = e_G$ , such that  $G = G_0$ , in  $G_r$  each player has only a single strategy, and  $\forall i \in \{0, \dots, r-1\}$ ,  $G_{i+1}$  is obtained from  $G_i$  via the removal of sets of NBR strategies (under tie-breaking).

**Convergence:** We first show that if all players re-

peatedly best-respond then convergence to a PNE is guaranteed within  $e_G$  rounds. Consider the first round of a best-response mechanism, and consider some  $j \in [n]$  such that there exists  $s_j \in S_j$  that is NBR in  $G = G_0$ . Observe that once  $j$  is activated for the first time,  $s_j$  will never be selected thereafter. Thus, after the first round, no NBR strategy in  $G_0$  will be played ever again and hence the game is effectively equivalent to  $G_1$ . We can now use the same argument to show that after the second round the game is effectively equivalent to  $G_2$ . Thus, we mimic the elimination sequence in each strategy until we end up at  $G_r$ , whose unique strategy tuple  $s^*$  is the unique PNE under tie-breaking of  $G$ .

**Incentive compatibility:** this property follows from the fact that when each player  $i$  considers the game after the other players have already eliminated dominated strategies that can be eliminated regardless of what  $i$  does, he can already tell that he will not be able to do better than the outcome that is reached via repeated best-response.

We give the precise argument (by contradiction). Let  $i$  be a player that deviates from repeated best-response and strictly gains from doing so. The fact that  $G$  is NBR-solvable with a clear outcome (under tie-breaking) implies that there exists a (player-specific) order of elimination of NBR strategies (under the given tie-breaking rules) such that  $s^*$  is globally optimal for  $i$  at the first step in the elimination sequence in which a strategy in  $S_i$  is eliminated (that is, in the game obtained after the removal of all previously-eliminated strategies from  $G$ ). Consider this order of elimination; it induces some sequence of games  $G_0, \dots, G_l$  such that  $G = G_0$ , in  $G_l$  each player has only a single strategy, and  $\forall i \in \{0, \dots, l-1\}$ ,  $G_{i+1}$  is obtained from  $G_i$  via the removal of sets of NBR strategies (under tie-breaking) as in the  $(i+1)$ 'th step in that order. Now, let  $t_i$  be the index of the first game in the sequence in which  $i$ 's strategies are eliminated in that order. All players but  $i$  are repeatedly best-responding and in the  $t_i - 1$  first steps of the elimination sequence no strategy in  $S_i$  is eliminated. We can use the same arguments that we used to show convergence, to show that after  $t_i - 1$  rounds the game is effectively equivalent to  $G_{t_i}$ , regardless of the actions of player  $i$ . However, in that game,  $i$  can do no better than  $s^*$ —a contradiction.  $\square$

### 3 Four best-response mechanisms

We present four examples of environments that

can be formulated as games that are NBR-solvable with clear outcomes (sometimes under tie-breaking): stable-roommates games, cost-sharing games, BGP games and TCP games. This implies the existence of incentive-compatible best-response mechanisms for all these environments.

#### 3.1 Stable-roommates

This following classic setting has been extensively studied in economics, game theory and computer science.  $n$  students  $1, \dots, n$  must be paired for the purpose of sharing dorm rooms. Each student has a private *strict* ranking of the others, and prefers being matched to not being matched. The goal is to find a stable matching, *i.e.*, a matching where no two students prefer each other to their matched roommates. Unfortunately, a stable matching is not guaranteed to exist in general and, furthermore, even if a stable matching does exist (*e.g.*, in bipartite graphs), existing algorithms for reaching it are *not* incentive compatible [3]. We seek environments where a stable matching is guaranteed to exist and can be reached in an incentive compatible manner. We focus on two well-known special cases of stable roommates:

- **Intern-hospital matchings:** The “students” are divided into two disjoint sets, called *interns* and *hospitals*, and all hospitals have the same ranking of interns (*e.g.*, GPA-based).
- **Correlated markets:** The “students” are vertices in a complete graph in which every edge has a unique “weight”. The “heavier” the edge connecting a student to another student the higher that student ranks the other student.

We now show how the framework in Section 2 can be used to design natural incentive compatible mechanisms for stable-roommates. We first formulate this environment as a game and prove that this game is NBR-solvable with a clear outcome.

**Stable-roommates games:** The students are the players and each student  $i$ 's strategy space  $S_i$  is the set of all students  $j \neq i$ .  $\alpha_i(j)$  denotes student  $j$ 's rank in student  $i$ 's ranking (the *least* desired roommate's rank is 1).  $\forall s = (s_1, \dots, s_n) \in S$  (that is, choices of roommates),  $u_i(s) = \alpha_i(j)$  iff  $s_i = j$  and  $\nexists k \neq i$  such that  $s_k = j$  and  $\alpha_j(k) > \alpha_j(i)$ ; otherwise,  $u_i(s) = 0$ .<sup>5</sup> (Observe that players' utilities are correlated.)

<sup>5</sup>We note that the more natural definition of utilities that only awards utility to players that are selected by the partner they themselves choose implies a game in which all matchings are stable, and is thus not useful to us.

**Theorem 3.1** *For every stable-roommates game  $G$  it holds that in both hospital-intern matchings and correlated markets*

- $G$  is NBR-solvable.
- $G$ 's unique PNE is a stable matchings.
- $e_G \leq n$ .

**Proof sketch:** We say that a stable-roommates game is cycle-free if there is no sequence of roommates  $r_1, r_2, \dots, r_k$  of length  $k > 2$  such that each student  $r_i$  ranks student  $r_{i+1}$  higher than student  $r_{i-1}$  (where student indices are considered mod  $k$  to induce a cycle).

Any matching game that is cycle-free has an elimination sequence that can be constructed as follows: At any stage in the elimination start with some arbitrary student  $r_1$  (that has more than one strategy in the current subgame) and construct a sequence  $r_1, r_2, \dots$  of students in which  $r_{i+1}$  is the student  $r_i$  prefers the most out of the students that still have more than one possible strategy remaining (other strategies were eliminated). The number of students is finite and so the sequence must repeat. Since the game is cycle free, the cycle must be of length 2. We have thus located 2 students that desire each other the most. We can eliminate for each of the two the strategies of proposing to any other student since they are guaranteed to gain the maximal utility by proposing to each other.

All that remains is to notice that both the hospital-intern game and the correlated markets game are cycle-free. In the case of hospitals and interns, the hospitals agree about the ranking of interns and so any cycle of players will have to include a hospital that is placed after a desired intern and before a less desired one. In the case of correlated markets, any cycle of nodes in the graph must include an edge with a lower weight that appears after an edge with a higher one and therefore the preferences do not induce a cycle in the matching graph in either case.  $\square$

We observe that the following simple and computationally-efficient mechanism is a best-response mechanism for stable-roommate games, and so Theorem 2.7 implies that it implements a stable matching in an incentive-compatible manner.

**Mechanism for Stable-Roommates:**

- Go over the students in some cyclic (round-robin) order and, at each time step, allow a single student to announce another student.

- We say that a student  $i$  makes a “better offer” to another student  $j$  at time  $t$  if (1)  $i$  announces  $j$  at time  $t$ ; and (2)  $j$  prefers  $i$  to all students from whom he has “offers”, that is, all students whose last announcement was  $j$ . The mechanism prescribes that each student repeatedly check which students he can make a better offer to, and announce his most preferred student to whom he can make a better offer.
- The mechanism terminates after  $n^2$  steps and outputs all student pairs  $(i, j)$  such that  $i$ 's last announcement was  $j$  and  $j$ 's last announcement was  $i$ .

**Theorem 3.2** *The mechanism is incentive-compatible in ex-post Nash and implements a stable matching in both intern-hospital matchings and correlated markets.*

**3.2 Cost-sharing**

Cost-sharing arises in situations in which the cost of some public service (e.g., building a bridge) must be distributed between self-interested users that can benefit from this service to different extents, and is modeled as follows.  $n$  users  $1, \dots, n$  aim to share the cost of building some common infrastructure. Some cost-sharing rule specifies, for every subset of users  $S$ , and every user  $i \in S$ ,  $i$ 's “cost share”  $c_i(S)$  for building an infrastructure that only serves members of  $S$ .  $c_i(S)$  is nonnegative, monotonically non-increasing in  $S$ , and also *cross-monotonic*, that is,  $\forall i \in S \subseteq T$ ,  $c_i(S) \geq c_i(T)$ . User  $i$  gets positive (private) value  $v_i \in \mathbb{R}_{\geq 0}$  if the infrastructure serves him and 0 otherwise. The goal is to split the cost of the infrastructure between a group of users so that each user's payment is at least his cost-share, yet does not exceed his private value, that is, to find “reasonable” cost shares. Moulin [9] exhibits a centralized mechanism that achieves this (see also [10]).

We now use the framework in Section 2 to design simple and natural *distributed* incentive-compatible mechanisms that implement the same outcome as the Moulin mechanism. We present “1<sup>st</sup>-price cost-sharing games” and specific tie-breaking rules.

**1<sup>st</sup>-price cost-sharing games:** The users are the players and, for each user  $i$ ,  $S_i = \mathbb{R}_{\geq 0}$ . Given a vector of users' bids (strategies)  $\vec{b} = (b_1, \dots, b_n)$ , the “served set” for  $\vec{b}$  is the maximum-cardinality subset of users  $S$  such that  $\forall j \in S, b_j \geq c_j(S)$  (breaking ties between such sets lexicographically).  $\forall \vec{b} = (b_1, \dots, b_n)$ ,

$u_i(\vec{b}) = v_i - b_i$  if  $i$  is in the serviced set for  $\vec{b}$ ;  
 $u_i(\vec{b}) = 0$  otherwise.

**Tie-breaking rules:** Prefer bids closer to  $v_i$ , i.e.,  
 $\forall s, t \in S_i$ , if  $|s - v_i| \leq |t - v_i|$  then  $t \prec_i s$ .

**Theorem 3.3** *For every 1<sup>st</sup>-price cost-sharing game  $G$  it holds that*

- $G$  is NBR-solvable under these tie-breaking rules.
- $G$ 's unique PNE under these tie-breaking rules induces reasonable cost shares as in the outcome of the Moulin mechanism.
- $e_G \leq n$ .

**Proof sketch:** Let us show an elimination sequence for every cost sharing game. First, notice that each player can only get a non-positive utility from a bid that is above his valuation. We therefore start by eliminating these bids for all players. Next, let  $R^v$  be the set of serviced users for bids that are exactly the valuations of the players. Any player  $i \notin R^v$  will not get serviced for any set of bids that are in the remaining subgame (costs only increase as players drop out and he does not win when they all pay the maximal amount). We can therefore eliminate all strategies below  $v_i$  for any such player. For every player  $j \in R^v$ , we can eliminate all bids below  $c_j(R^v)$ , as he will only get 0 utility with those bids, and non-negative utility with higher bids. Once these are eliminated, then in the remaining subgame  $R^v$  will always be the serviced set of players and we can eliminate all bids above  $c_j(R^v)$  as well.

Note that it is also possible to perform the eliminations using a different order. Specifically, for each player  $i$  we can let all other players eliminate bids above  $v_i$ , then determine a set of serviced agents  $R_i$  for the case in which every agent  $j$  bids  $v_j$  except for agent  $i$  that bids  $\infty$ . Then, eliminate all bids for non-serviced agents (except their valuation), and check if  $c_i(R_i)$  is greater than  $v_i$ . If it is, we can eliminate bids below  $c_i(R_i)$  for agent  $i$ . Otherwise, agent  $i$  will not gain a positive utility from the service in any case and we can eliminate all his strategies except his valuation. We can then continue along the same lines as before and eliminate strategies for all other players. Either way, the elimination done by agent  $i$  leads to a subgame in which  $s^*$  is the optimal outcome for him, and so the game has a clear outcome as required.  $\square$

We observe that the following natural distributed mechanism is a best-response mechanism for 1<sup>st</sup>-price

cost-sharing games (under these tie-breaking rules), and so Theorem 2.7 implies that it implements the outcome of the Moulin mechanism in an incentive-compatible manner.

**Mechanism for cost-sharing:**

- Go over the users in some cyclic (round-robin) order and, at each time step, allow a single user to submit a bid in  $\mathcal{R}_{\geq 0}$ .
- The mechanism prescribes that each bidder  $i$  repeatedly bid as follows: submit the minimal bid  $b_i \leq v_i$  such that  $i$  is in the serviced set for the most-recently submitted bids; in the event that no such bid exists submit the bid  $b_i = v_i$ .
- The mechanism terminates after  $n^2$  time steps, outputs the serviced set  $S$  for the last-submitted bids and charges each bidder  $i \in S$  his last bid  $b_i$ .

**Theorem 3.4** *The mechanism is incentive compatible and implements reasonable cost-shares.*

This result can be extended to the class of acyclic mechanisms studied in [8].

### 3.3 Internet routing

The Border Gateway Protocol (BGP) establishes routes between the smaller networks that make up the Internet. Griffin *et al.* [6] put forth the following model for analyzing BGP dynamics. The network is an undirected graph  $G = (V, E)$  where the vertex set  $V$  consists of  $n$  source nodes and  $1, \dots, n$  a unique destination node  $d$ . Each source node has a private *strict* ranking of all simple (loop-free) routes between itself and the destination node  $d$ . Under BGP, each source node repeatedly examines its neighboring nodes' most recent route-announcements, selects to forward traffic through the neighbor whose route it likes the most, and announces its newly chosen route to all neighbors via update messages. The network is asynchronous and so nodes can select routes simultaneously and based on outdated information (update messages between nodes can be arbitrarily delayed).

BGP's convergence to a "stable" routing tree is the subject of extensive networking research. Levin *et al.* [7] observe that BGP can be regarded as best-response dynamics in a specific class of "routing games", and prove that BGP is incentive-compatible in networks for which the No Dispute Wheel [6] condition holds.

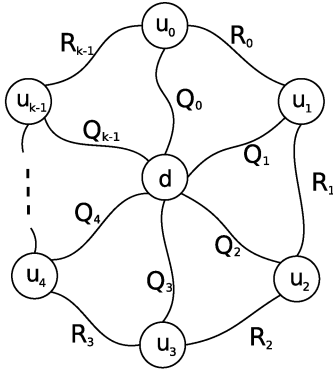
No Dispute Wheel is a generalization of the Gao-



Rexford [4] conditions, that capture common Internet routing practices. A Dispute Wheel (see Figure 2) is a 3-tuple  $(\mathcal{U}, \mathcal{R}, \mathcal{Q})$ , where  $\mathcal{U} = (u_0, u_1, \dots, u_{k-1})$  is a sequence of  $k$  vertices in  $V$ , called the ‘‘pivot nodes’’ and  $\mathcal{R} = (R_0, R_1, \dots, R_{k-1})$ ,  $\mathcal{Q} = (Q_0, Q_1, \dots, Q_{k-1})$  are two sequences of  $k$  routes, such that (indices are considered modulo  $k$ ):

- $\forall i$ ,  $Q_i$  is a simple route from  $i$  to  $d$ .
- $\forall i$ ,  $R_i$  is a simple route from  $u_i$  to  $u_{i+1}$ .
- $\forall i$ ,  $u_i$  ranks the route  $R_i Q_{i+1}$  more highly than the route  $Q_i$ .

‘‘No Dispute Wheel’’ is the condition that no Dispute Wheel exist in the network.



Each pivot node  $u_i$  would rather route clockwise through pivot node  $u_{i+1}$  than through the direct route  $Q_i$ .

Figure 2: A Dispute Wheel.

**Theorem 3.5** ([7]) *BGP is incentive-compatible in ex-post Nash in networks for which No Dispute Wheel holds.*

We now show that the class of ‘‘BGP games’’ presented in [7] falls within the category of NBR-solvable games with clear outcomes. Thus, the essence of the incentive compatibility result for BGP in [7] follows from Theorem 2.7.

**BGP games:** The source nodes are the players and, for each source node  $i$ ,  $S_i$  is the set of  $i$ ’s outgoing edges in  $E$ . Given a vector of source nodes’ traffic forwarding decisions (strategies)  $\vec{f} = (f_1, \dots, f_n)$ ,  $u_i(\vec{f} = (f_1, \dots, f_n))$  is  $i$ ’s rank for the simple route from  $i$  to  $d$  under  $\vec{f}$  (the least desired route has rank 1) if such a route exists;  $u_i(\vec{f}) = 0$  otherwise.

**Theorem 3.6** *For a BGP game  $G$  it holds that*

- $G$  is NBR-solvable.
- $G$ ’s unique PNE is a stable routing tree.
- $e_G \leq n$ .

**Proof sketch:** Let us show an elimination order in the game. At every stage in the elimination, we locate a node that can guarantee its most preferred route (in the current subgame) and eliminate all other routing actions for it. To show that such a node always exists, we begin with an arbitrary node  $a_0$  with at least 2 actions. Let  $R_0$  be  $a_0$ ’s most preferred existing route to  $d$  (a route is said to exist if all nodes along it can route accordingly in the current subgame). Let  $a_1$  be the vertex closest to  $d$  on  $R_0$ , with two available actions in the current subgame, such that  $a_1$  prefers some other route  $R_1$  to the suffix of  $R_0$  that leads from  $a_1$  to  $d$  (if no such node exists  $a_0$  can guarantee its most preferred route). Then we choose  $a_2$  to be the vertex closest to  $d$  on  $R_1$  such that  $a_2$ ’s most preferred route  $R_2$  is preferred over the suffix of  $R_1$  that leads from  $a_2$  to  $d$ . Once again if there is no such  $a_2$  we are done. We can continue to choose  $a_3, a_4, \dots$  in the same manner. Since there is a finite number of vertices, at some point some vertex will appear twice in this sequence  $(a_0, a_1, \dots)$ . This would result in the formation of a Dispute Wheel (in which the  $a_i$ s are the pivot nodes and the  $R_i$ s are the routes) which we assumed is not contained in the graph. We will therefore always be able to find a node that can guarantee its most preferred route and continue with the elimination, until there are no more nodes with several possible actions.  $\square$

### 3.4 Congestion control

Congestion control is a crucial task in communication networks. Congestion is handled via the combination of transmission-rate-adjustment protocols at the sender-receiver level (*e.g.*, TCP), and queueing management policies at the router level, that dictate how excess traffic is discarded (*e.g.*, RED). TCP is notoriously *not* incentive compatible. [5] analyzes incentives in the following TCP-inspired environment. The network is an undirected graph  $G = (V, E)$  with a given a capacity function  $c$  that specifies the capacity  $c(e)$  for each edge  $e \in E$ . The network consists of  $n$  source-target pairs of vertices  $(\alpha_i, \beta_i)$ . Every such source-target pair  $(\alpha_i, \beta_i)$  aims to send traffic along a *fixed* route  $R_i$  in  $G$ . Each source  $\alpha_i$  can select transmission rates that lie in the interval  $[0, M_i]$ , where  $M_i$  is  $\alpha_i$ ’s private information, and wishes to maximize its achieved throughput. When an edge encounters congestion, that is, the sum of incoming flows traversing it

exceeds its capacity, excess traffic must be discarded. [5] considers two capacity-allocation schemes:

- **Strict-Priority-Queueing (SPQ).**  $\forall e \in E$  there is an edge-specific order over source nodes. Capacity is shared as follows: the most highly ranked source whose route traverses the edge gets its entire flow sent along the edge (up to  $c(e)$ ); unused capacity is allocated to the second most highly ranked source whose route traverses the edge in a similar fashion, *etc.*
- **Weighted-Fair-Queueing (WFQ).**  $\forall e \in E$ , each source node  $\alpha_i$  has weight  $w_i(e)$  at  $e$ . Every source  $\alpha_i$  is then allocated capacity  $\frac{w_i}{\sum_j w_j} c(e)$ . Unused capacity is allocated in a recursive manner. The special case that  $\forall e \in E, \forall i \in [n], w_i(e) = 1$  is called “fair queueing” (FQ).

[5] considers a TCP-like protocol called Probing-Increase-Educate-Decrease (PIED) in which each source is instructed to gradually increase its transmission rate until encountering congestion and, at that point, decrease its transmission rate to its achieved throughput. [5] analyzes PIED in settings in which all edges use SPQ or all edges use WFQ, and sources priorities/weights are identical on all edges. PIED is shown to be incentive compatible in both these environments (also under asynchronous timings of rate-transmission adjustments).

**Theorem 3.7 ([5])** *PIED is incentive compatible in networks in which all edges use SPQ with coordinated priorities.*

**Theorem 3.8 ([5])** *PIED is incentive compatible in networks in which all edges use WFQ with coordinated weights (and so if all edges use FQ then PIED is incentive compatible).*

It is interesting to notice that PIED can be considered a form of better-response in a setting in which the exact available capacity is unknown. We unify the two results above for an abstracted setting by formulating the environment in [5] as a game and showing that this game is NBR-solvable with a clear outcome (under specific tie-breaking rules). Our main difference from [5] is that we allow players more knowledge about the network, while [5] uses the probing nature of PIED to learn the needed information (all that is needed is for players to be able to tell the amount of available bandwidth on their path). Thus, Theorem 2.7 implies a result that is similar in spirit to the two theorems in [5].

**TCP games:** The source nodes are the players and each source node  $i$ 's strategy space is  $S_i = [0, M_i]$ . Given a vector of source nodes' transmission rates (strategies)  $\vec{r} = (r_1, \dots, r_n)$ ,  $u_i(\vec{r})$  is  $\alpha_i$ 's achieved throughput in the unique traffic-flow equilibrium point of the network for  $\vec{r}$  ([5] shows that such a unique equilibrium point exists for the SPQ and WFQ settings with coordinated priorities/weights).

**Tie-breaking rules:**  $\forall s, t \in S_i, s \prec_i t$  iff  $s > t$ .

**Theorem 3.9** *For every TCP game  $G$  such that all edges use SPQ with coordinated priorities, or all edges use WFQ with coordinated weights, it holds that*

- $G$  is NBR-solvable under these tie-breaking rules.
- $G$ 's unique PNE under these tie-breaking is a stable flow pattern.
- $e_G \leq n$ .

For clarity of presentation we show only the proof for the case of Weighted-Fair-Queueing, with equal weights. The proof for non-equal weights and for Strict-Priority-Queueing follow similar lines.

**Proof sketch:** Let us define for each edge  $e$ , the share of each flow as  $\beta_e = c_e/k_e$  where  $k_e$  is the number of flows that traverse the edge. We construct an elimination sequence for the game as follows: Let  $e^*$  be the edge with the minimal  $\beta$ . Each flow on this edge is guaranteed  $\beta_{e^*}$  traffic through that edge, and at least that amount on all other edges. It is therefore possible to eliminate all actions of transmitting less than  $\beta_{e^*}$  for each player that goes through  $e^*$ . Now, if all flows through  $e^*$  claim their fair share, no flow can send more (no bandwidth is unclaimed). We can therefore eliminate all actions of transmitting above  $\beta_{e^*}$  for these flows. Now, we are left with a subgame with a smaller number of active players where some of the bandwidth on each edge is already used up. We can now repeat the elimination steps for the residual network graph with the remaining players.

Notice that for each bottleneck edge  $e^*$  that is found along the process there are several orders of elimination (according to ordering among players). If player  $i$  eliminates actions below  $\beta_{e^*}$  last among players that go through  $e^*$ , then he does so in a game in which the final profile is optimal for him, and so the game has a clear outcome.  $\square$

## 4 Collusion-proof best-response mechanisms

In Section 3, we establish incentive compatibility re-

sults for four environments. We are able to strengthen our results for stable-roommates (Theorem 3.1), BGP games (Theorem 3.6), and TCP games where all edges use SPQ with coordinated priorities (see Theorem 3.9). We prove that, in all these settings, best-response mechanisms are actually also collusion-proof. We observe, though, that NBR-solvability with a clear outcome does not imply collusion-proofness of best-response mechanisms in general. To see this, consider the game depicted in Figure 3 (which is simply the prisoner’s dilemma).

	C	D
C	2,2	0,3
D	3,0	1,1

Figure 3: An NBR-solvable game with a clear outcome for which best-response mechanisms are not collusion proof.

Observe that this game is indeed an NBR-solvable game with a clear outcome, yet both players prefer  $(C, C)$  to the unique equilibrium  $(D, D)$ . Thus, the two players can jointly deviate from repeated best-response and both strictly gain from doing so.

## References

- [1] R. I. Brafman and M. Tennenholtz. Efficient learning equilibrium. *Artif. Intell.*, 159(1-2):27-47, 2004.
- [2] R. I. Brafman and M. Tennenholtz. Optimal efficient learning equilibrium: Imperfect monitoring in symmetric games. In *In Proceedings of the National Conference on Artificial Intelligence (AAAI)*, pages 726-731. Press, 2005.
- [3] D. Gale and L. Shapley. College admissions and stability of marriage. *Amer. Math. Monthly*, 69(1):9-15, 1962.
- [4] L. Gao and J. Rexford. Stable Internet routing without global coordination. *IEEE/ACM Transactions on Networking*, 9(6):681-692, 2001.
- [5] P. B. Godfrey, M. Schapira, A. Zohar, and S. Shenker. Incentive compatibility and dynamics of congestion control. *SIGMETRICS Perform. Eval. Rev.*, 38(1):95-106, 2010.
- [6] T. G. Griffin, F. B. Shepherd, and G. Wilfong. The stable paths problem and interdomain routing. *IEEE/ACM Transactions on Networking*, 10(2):232-243, April 2002.
- [7] H. Levin, M. Schapira, and A. Zohar. Interdomain routing and games. In *STOC*, pages 57-66, 2008.
- [8] A. Mehta, T. Roughgarden, and M. Sundararajan. Beyond moulin mechanisms. In *EC’07: Proceedings of the 8th ACM conference on Electronic commerce*, pages 1-10, New York, NY, USA, 2007. ACM.
- [9] H. Moulin. Incremental cost sharing: Characterization by coalition strategy-proofness. *Social Choice and Welfare*, 16(2):279-320, 1999.
- [10] H. Moulin and S. Shenker. Strategyproof sharing of submodular costs: budget balance versus efficiency. *Economic Theory*, 18(3):511-533, 2001.
- [11] N. Nisan, M. Schapira, G. Valiant, and A. Zohar. Best-response auctions. *Working paper*, 2010.
- [12] N. Nisan, M. Schapira, and A. Zohar. Asynchronous best-reply dynamics. In *Proceedings of WINE*, 2008.