When Bribes are Harmless: The Power and Limits of Collusion-Resilient Mechanism Design

ARTUR GOROKH, Cornell University
SIDDHARTHA BANERJEE, Cornell University
KRISHNAMURTHY IYER, Cornell University

Collusion has long been the Achilles heel of mechanism design, as most results break down when participating agents can collude. The issue is more severe when monetary transfers (bribes) between agents are feasible, wherein it is known that truthful revelation and efficient allocation are incompatible[Schummer, 2000]. A natural relaxation that circumvents these impossibility results is that of *coalitional dominance*: replacing truthful revelation with the weaker requirement that all coalitions, whatever they may be, have dominant strategies. When a mechanism satisfies this property and is efficient, we call it *collusion resilient*. The goal of this paper is to characterize the power and limits of collusion resilient mechanisms. On the positive side, in a general allocation setting, we demonstrate a new mechanism which is collusion-resilient for *surplus-submodular* settings – a large-class of problems which includes combinatorial auctions with gross substitutes valuations. We complement this mechanism with with two impossibility results: (i) for combinatorial auctions with general submodular valuations, we show that no mechanism can be collusion-resilient, and (ii) for the problem of collective decision making, we argue that any non-trivial approximation of welfare is impossible under coalitional dominance. Finally, we make a connection between collusion resilience and false-name-proofness, and show that our impossibility theorems strengthen existing results for false-name-proof mechanisms.

1 INTRODUCTION

One of the main obstacles in employing results from theoretical mechanism design in the real world is the possibility of collusion between participating agents. An example of this is the celebrated VCG mechanism – a jewel of mechanism design theory that is rarely used in practice for this and other concerns [Ausubel and Milgrom, 2006]. The growing use of market mechanisms for solving allocation problems in a variety of settings (e.g., cloud computing, healthcare, smart grids, sharing economy platforms, decentralized currency, etc.) only lead to more opportunities for the pernicious effects of collusion to manifest themselves. Designing efficient mechanisms that are robust to collusion, and understanding their limits, is thus an pressing challenge for mechanism design.

The existing work on collusion-proof mechanisms can be roughly divided into two segments corresponding to the different choices of modeling coalitional behavior. One model assumes that agents collude only if they directly profit from it – in this case, the notion of group-strategyproofness captures sufficient conditions for mechanism to be collusion-proof. This setting admits many positive results, notable among which are the cost-sharing mechanism of Moulin [Moulin and Shenker, 2001] and the Deferred-Acceptance auctions [Milgrom and Segal, 2014].

However, the property of group-strategyproofness is only practically compelling in a transparent environment where agents have no way of compensating (i.e., *bribing*) each other so as to redistribute utility. When monetary exchanges outside of the mechanism are viable, one can assume that agents in a coalition act so as to maximize the total utility of the coalition, and then divide the reaped benefits among the coalition members via monetary transfers. Group-strategyproofness ceases to be a safeguard in this case; on the other hand, a mechanism that ensures truthful behavior even when bribes are feasible would arguably be more robust against collusion in practice.

Unfortunately, requiring truthful revelation in the presence of bribes turns out to be too demanding: for public projects, Green and Laffont [1979] showed that in the presence of coalitions, there exists no efficient, incentive-compatible direct mechanisms. Schummer [2000] extended this impossibility result to any setting with a sufficiently rich type domain (including allocation of indivisible items), even when all coalitions have at most 2 members. Only few exceptions to this impossibility are known of, in restricted single-parameter environments [Goldberg and Hartline, 2005]; any further progress, however, appears to be difficult.

In settings where the main goal is efficient allocation (i.e., the revenue generated by the principle is secondary), a promising approach for circumventing the above impossibilities is to relax truthful revelation to requiring that coalitions have dominant strategies. In other words, instead of asking for a mechanism to have a direct-revelation structure and satisfy incentive compatibility, we instead require all coalitions, whatever they might be, to have a dominant strategy, under which the resulting welfare is the maximum possible. Consider the following example.

Example 1.1 (Second Price Auction is Collusion Resilient). A single item is to be allocated via a second-price auction, among 3 agents with values $v_1 = 1, v_2 = 2, v_3 = 3$. Agent 1 bids truthfully, while agents 2 and 3 collude, behaving as to maximize their total utility. It is easy to see that the optimal strategy for the coalition $\{2, 3\}$ is to bid $b_2 = 0, b_3 = 3$; in this case agent 3 gets the item and pays the price of 1. Notice two things: first, this coalitional strategy is optimal independently of what agent 1 bids, and second, the resulting outcome is efficient. Unfortunately, single item allocation is virtually the only setting for which VCG mechanism has these two features.

The above property, which we refer to as *collusion resilience*, allows us to circumvent the impossibility of Schummer [2000], while sacrificing only the direct revelation structure, but not efficiency or incentives. This turns out to be quite a bargain, as we show that for a rather large class of settings (including *all* combinatorial auctions with gross-substitute valuations), it is possible to construct efficient collusion resilient mechanisms.

1.1 Overview of our Results

Through a sequence of results, we map out the power and limits of collusion resilience.

Positive results. In Section 2, we consider a general allocation problem with n agents who are partitioned into a set of coalitions, unknown to the designer. Coalitions then act as to maximize their total utility. In this setting we aim to construct a collusion resilient mechanism, i.e. a mechanism in which all coalitions have a dominant strategy and the allocation at the dominant strategy equilibrium is efficient.

Our main positive result is in constructing a mechanism (namely, the *Coalition-Revelation Mechanism* (CRM)) that we show satisfies collusion resilience as long as the underlying problem is *surplus-submodular* [1] – a structural condition on allocation settings which in case of multi-item allocation can be viewed as a slight generalization of gross substitutes. Because CRM is somewhat involved, we first build up to it in Section 3 by describing a simpler and more intuitive mechanism called *Expanded VCG*, which satisfies a weaker incentive guarantee. We then describe our main positive result in in Section 4.

Negative results. In Section 5 we complement our positive results with impossibility theorems that demonstrate its near-tightness. In particular, we show that no budget-balanced, IR collusion resilient mechanism exists for general submodular valuations, thereby effectively closing the question of when collusion resilience is achievable in combinatorial auctions.

For our second negative result, we consider a collective decision making problem, where agents must choose between one of two options, and monetary reparations are allowed. Here we show that not only is collusion resilience impossible, but in fact, any non-trivial constant-factor approximation

[1] A similar condition has been explored in earlier work on sybil-proof mechanisms [Yokoo et al., 2004].

of welfare is impossible under coalitional dominant strategies. Thus, in a sense, collective decision making is the worst case scenario for mechanism design in the face of collusion.

1.2 Discussion

We now briefly discuss some implications of our results, before proceeding to the technical sections.
Collusion Resilience and Surplus-Submodularity: One major upshot of our results is the leading role of surplus-submodularity in defining the limits of collusion resilient mechanisms. The critical observation is that the externality of individual agents is always larger than that of the collective in (and only in) surplus-submodular settings. To see the importance of this, note that in any mechanism, agents can misreport in two main ways: first, they may act as if being part of a larger coalition than they truly are, and second, a coalition may act as if comprised of smaller sub-coalitions. A mechanism can, through its incentive structure, prevent coalition to misreport in the first manner (as an example, in the CRM mechanism in Section 4, we use a form of the prisoners' dilemma to deter such deviations).

On the other hand, misreports of the second type are more challenging to handle, since the true underlying coalitional structure is unknown to the mechanism. Surplus-submodularity effectively eliminates the possibility that such a splitting of coalitions can be profitable, as the externality of individual agents is always larger than the collective externality. This point is crucial in driving our impossibility results in Section 5. In particular, in all our problem instances, it is the fact that a coalition may act as a group of independent agents that results in budget imbalance.
Collusion Resilience and Computation: One of the chief achievements of applying the computational lens to mechanism design was in recognizing gross substitutes as the boundary of settings where the goals of incentive compatibility, efficiency and tractability are simultaneously achievable [Leme, 2017]. A natural question to ask now is how collusion-resilient mechanisms fare in terms of computational requirements. Our positive results (in particular, the fact that our mechanism has the same complexity as that of maximizing the welfare function, and that it is collusion-resilient for all gross substitute settings) establish that *gross substitutes is also the boundary for tractable collusion-resilient mechanisms.*
Collusion Resilience and Sybil proofness The notion of collusion-resilience has strong links to the idea of sybil proofness (or false-name proofness) in mechanisms. In Appendix B, we delve further into the relation between the two concepts. Namely, we show that, in the setting of combinatorial auctions, CRM is in fact sybil-proof (under a slight modification as to impose direct revelation strategy space), and that any sybil proof mechanism can be converted into a collusion resilient mechanism via a black-box reduction. We discuss how in the light of these theorems our impossibility results can be viewed as novel contributions to the work on Sybil-proofness.

1.3 Related Work

Collusion of participants has been a notable concern since the dawn of mechanism design. Green and Laffont [1979] adopt a model of collusion that coincides with ours (wherein a coalition tries to maximize its utility) and show that, for the problem of public projects, there exists no direct mechanism that satisfies efficiency and incentive compatibility.

This negative result was later followed up by Schummer [2000], who shows a similar impossibility for any allocation problem with sufficiently rich type domain (in particular, for allocation of indivisible goods); this result was later strengthened by Mizukami [2003]. In another follow-up, Goldberg and Hartline [2005] concentrate on the single-parameter setting of allocating identical items in presence of collusion, and for that case, establish an approximately efficient and incentive compatible mechanism; this can be viewed as the first positive result in this line of work.

There has also been work on circumventing the impossibilities described above by either relaxing dominant strategy incentive constraints to Bayes-Nash equlibria [Green and Laffont, 1979, Laffont and Martimort, 2000] or assuming the ability of the mechanism designer to partially verify type reports [Penna and Ventre, 2014].

The line of work most closely related to ours is that of Chen and Micali [2012] and Deckelbaum and Micali [2017]. There, the authors introduce a notion of collusive dominant-strategy mechanisms, that is equivalent to the one we use here (see Definitions 2.3 and 2.4). They then present two results: a collusion resilient mechanism for the identical items setting, and a negative result for general combinatorial auctions. Our results can be viewed as a significant improvement over these results: the mechanism we construct is more general and satisfies collusion resilience in settings well beyond the single-parameter problem of identical items; one of the impossibility theorems we present strengthens the result proven in Deckelbaum and Micali [2017].

Another line of work closely related to our results concerns mechanisms which satisfy the property of false-name-proofness (also known as sybil-proofness) [Todo et al., 2009, Wagman and Conitzer, 2014, Yokoo et al., 2004]. There, the goal of the designer is to prevent manipulation by agents who may create multiple identities. In some settings, such a goal is identical with that of collusion resilience; in others (most notably, combinatorial auctions), they are related. We dedicate Appendix B to making a formal connection between these two notions and analyzing our contributions in light of this connection.

2 SETTING

2.1 Types, outcomes and welfare

We consider a general allocation setting with n agents who have private types $\theta = \{\theta_i\}_{i\in\mathcal{N}}, \theta_i \in \Theta_i$. The principal chooses an allocation $x \in \mathcal{X}$, whereupon an agent i with type θ_i derives value $v_i(\theta_i, x) \in \mathbb{R}$. The principal strives to choose an allocation that maximizes social welfare:

$$W(\theta, x) = \sum_{i=1}^{n} v_i(\theta_i, x).$$

A mechanism M consists of an action space [2] $\mathcal{A}_i$ for each agent i (with $\mathcal{A} \triangleq \times_i \mathcal{A}_i$), allocation rule $x(a) \in \mathcal{X}\ \forall a \in \mathcal{A}$, and transfer rule $c(a) \in \mathbb{R}^n$ for all $a \in \mathcal{A}$. Given a mechanism M, utility of agent i is given by $u_i(\theta_i, a; M) = v_i(\theta_i, x(a)) - c_i(a)$. When the employed mechanism M is clear from the context we leave this argument out for readability.

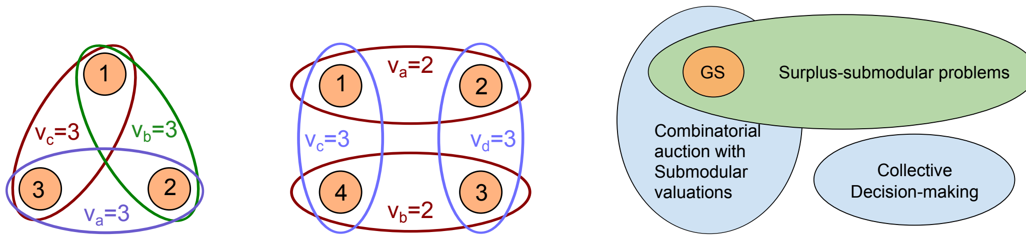
We focus on the setting where (subsets of) agents may engage in *collusion*, implying that they may coordinate their actions as to maximize the total utility. Formally, we assume that there exists an exogenously specified and unknown to the principal partition[3] $\mathcal{S} = \{S_k : k = 1, \cdots, K\}$ of the set of agents $\mathcal{N}$, with each S_k constituting a coalition of agents. We assume that agents in a coalition $S \in \mathcal{S}$ seek to maximize the *total coalitional utility*

$$U_S(\theta, a) = \sum_{i\in S} u_i(\theta_i, x(a)).$$

Such a model of coalitional behavior arises when assuming *transferable utilities*, i.e. settings where the agents in a coalition may make monetary transfers among themselves and has been studied extensively in literature [Deckelbaum and Micali, 2017, Green and Laffont, 1979, Schummer, 2000].

[2] As noted by Deckelbaum and Micali [2017], when considering collusion resilient mechanisms one cannot restrict attention to direct revelation mechanisms. Thus we have to assume general action space $\mathcal{A}$.

[3] In other words, we have $S_k \cap S_l = \emptyset$ for $k \neq l$, and $\cup_j S_j = \mathcal{N}$.



(a) Illustrating surplus-submodularity

(b) Landscape of allocation settings

Fig. 1. In Fig. 1(a), we illustrate surplus-submodularity via two examples. In the first, there are 3 items $\{1, 2, 3\}$ and 3 single-minded agents $\{a, b, c\}$ (i.e., each agent has a non-zero valuation for a particular subset) with desired sets and valuations as indicated. One can check that $W^*(\theta, S) = 3$ for any subset $S \neq \emptyset$, and hence it is surplus-submodular. Note though that the valuations are not gross substitutes or submodular. Similarly, on the right, we have 4 items $\{1, 2, 3, 4\}$ and 4 single-minded agents $\{a, b, c, d\}$. In this case, $W^*(\theta, \{a, b\}) = W^*(\theta, \{a, b, c\}) = W^*(\theta, \{a, b, d\}) = 4$ and $W^*(\theta, \{a, b, c, d\}) = 6$; since adding agent d to $\{a, b, c\}$ causes more improvement than adding to $\{a, b\}$, $W^*(\theta, \cdot)$ is not surplus-submodular.
In Fig.1(b), we depict the landscape of mechanism design settings we consider. Our positive results (Ref. Theorems 3.1, 4.2) hold for surplus-submodular settings, which are a strict superset of gross substitute valuations. On the converse side, we show an impossibility result for general submodular valuation settings (Ref. Theorem 5.2), and inapproximability for collective decision making problems (Ref. Theorem 5.6).

Several of our results rely on two additional assumptions on the underlying allocation setting. The first is a property we refer to as surplus-submodularity. We define the surplus function, which can be interpreted as best possible value that can be achieved by agents in S:

$$W^*(\theta, S) = \max_{x \in \mathcal{X}} \sum_{i \in S} v_i(\theta_i, x).$$

Also, recall the definition of a submodular function: a set function $f(\cdot)$ is submodular over domain of subsets of $\mathcal{N}$ if for any sets $S \subset T$ and any element $i \in \mathcal{N}$ we have

$$f(S \cup \{i\}) - f(S) \geq f(T \cup \{i\}) - f(T)$$

We define a setting to be surplus-submodular if it satisfies the following.

*Definition 2.1 (**Surplus-submodularity**).* The allocation problem, that is the triplet $(\Theta, \mathcal{X}, v(\cdot, \cdot))$ of type space, allocation space, and value functions, is called surplus-submodular if $W^*(t, S)$ is a submodular function over the subsets of agents S under every possible type profile $\theta \in \Theta$.

We borrow our definition of the surplus function from Gul and Stacchetti [1999]. In single-parameter settings, is easy to see that single-item auctions satisfy this condition; however, public choice settings are not surplus-submodular. Moving to multi-parameter settings, surplus-submodularity is known to hold in combinatorial auctions where agent valuations satisfy the *gross substitutes* (GS) condition [Gul and Stacchetti, 1999]; on the other hand, as we show in Figure 1(a), surplus-submodular settings extend beyond GS valuations. Finally, settings with submodular valuations are incomparable to surplus-submodularity (one exclusion follows from Figure 1(a); for the other, refer Theorem 5.2). Figure 1(b) presents a detailed comparison of different settings.

The second assumption on allocation settings we require for some of our results is that of excludability, i.e., the ability of the mechanism to exclude particular agents from the allocation, while preserving the utilities of other agents.

*Definition 2.2 (**Excludability**).* If $x \in \mathcal{X}$ is a valid outcome, then there also exists an outcome $x' \in \mathcal{X}$ such that $v_i(\theta, x') = 0$ and $v_j(\theta, x') = v_j(\theta, x)$. For such an outcome, we write $x'_i = 0$.

This is clearly satisfied by combinatorial auctions; on the other hand, it does not hold for the problem of collective decision making. For excludable settings, we let $x = \emptyset$ represent the outcome from which no agents derive any value (e.g. not allocating any items in a combinatorial auction). Note also that excludability implies that the surplus function $W^*(\cdot)$ is monotone non-decreasing (i.e., $W^*(\theta_{S,i}, S \cup \{i\}) \geq W^*(\theta_S, S)$ for all S and i).

2.2 Collusion resilience

In order to define solution concepts we use throughout the paper, we interpret a given mechanism M as a game $M_{coalition}$, in which every coalition $S \in \mathcal{S}$ is a player with type $\theta_S = \{t_i\}_{i \in S}$, utility function $U_S(\theta_S, a)$ and strategy space $\mathcal{A}_S = \{\{a_i\}_{i \in S} | \ a_i \in \mathcal{A}_i\}$ [4]. We then call a strategy profile a^{eq}, $a_i^{eq}(\theta_i) : \Theta_i \longrightarrow \mathcal{A}_i$ an $\mathcal{S}$-collusive dominant strategy equilibrium ($\mathcal{S}$-CDS) if it is a dominant strategy equilibrium in $M_{coalition}$ for any type profile Θ. Similarly, a^{eq} is $\mathcal{S}$-collusive ex-post Nash ($\mathcal{S}$-CN) equilibrium if it is a Nash equilibrium in $M_{coalition}$ for any type profile Θ.

We assume that the mechanism designer is not privy to the coalition structure $\mathcal{S}$. Given this lack of knowledge, we require for any coalition structure $\mathcal{S}$, the mechanism has an $\mathcal{S}$-CDS.

Definition 2.3. We say a mechanism M is collusive-dominant strategy compatible (resp. collusive-Nash compatible), or CDS (resp. CN) if for any coalition structure $\mathcal{S}$, there exists a collusive-dominant strategy equilibrium (resp. collusive Nash equilibrium).

We note that the above definition of CDS coincides with that proposed in Deckelbaum and Micali [2017] (who coined the term); the definition of CN is a natural extension. The above definition now leads to a natural characterization of the performance of such mechanisms.

Definition 2.4. A CDS (resp. CN) mechanism is efficient if for every coalitional structure $\mathcal{S}$, if there is an $\mathcal{S}$-CDS (resp. $\mathcal{S}$-CN) equilibrium a^{eq} at which the outcome is efficient: $W(\theta, x(a^{eq})) = \max_{x \in \mathcal{X}} W(\theta, x)$.

We henceforth refer to a mechanism M as being **collusion resilient** (resp. weakly collusion resilient), or CR, if it is CDS (resp. CN) and efficient. Moreover this definition allows for straightforward generalizations to approximate welfare approximation guarantee. The following examples provide some intuition for the definition.

Example 2.5 (Second price auction is CR). Consider the single item setting and second price auction as a mechanism. It is easy to notice that for any coalition S the dominant strategy is for the highest value agent in the coalition to submit their value and for everyone else to bid 0: submitting multiple bids can only increase the price the coalition pays if it happens to win the item. The outcome of the whole auction when every coalition employs such strategy is then clearly efficient (highest value agent always wins the item), and SP auction is thus CR.

Note that the above example demonstrates that collusion resilience does not result in any guarantees for the revenue of the mechanism, and in fact the revenue may depend on the coalitional structure $\mathcal{S}$; the welfare however remains unchanged. This property breaks down as one moves to the simplest multi-parameter setting.

Example 2.6 (VCG for unit-demand valuations is not CR). Consider the setting of unit demand valuations with 3 agents and 2 items:

[4] Note that this interpretation does not allow running VCG mechanism on coalition-players because coalitional structure $\mathcal{S}$ is unknown to the designer.

	v_{i1}	v_{i2}
Agent 1	2	0
Agent 2	0	2
Agent 3	1	1

Let x_{ij} be the indicator that agent i obtained item j, then we have $v_i(x) = \max_j v_{ij}x_{ij}$. Suppose we implement the VCG mechanism, and assume agent 3 plays truthfully, while agents 1 and 2 collude. The optimal strategy for coalition 1 & 2 then is to report (2,0) and (0,0) respectively, thus winning the first item for agent 1 at the price of 0.

To see that collusion resilience is violated here, note that the above strategy for coalition 1 & 2 is not a dominant strategy: in fact, if agent 3 reports (0,1), then 1 & 2 obtain a better utility via truthful revelation.

Moreover, what makes using VCG truly problematic in this example is not just the absence of a CDS equilibrium, but rather, the fact that under equilibrium reports, the resulting outcome is inefficient. □

Similarly, one can show that VCG fails to be collusion resilient for virtually any setting other then single item – identical items, general GS valuations, public project, etc.

2.3 Design objectives

Throughout the paper we are interested in mechanisms that satisfy constraints that have to do with participation, revenue, incentives and efficiency. The above sections lay out a strong and weak form of the incentives and efficiency objectives that we consider, namely:

(1a) Collusion-resilience
(1b) Weak Collusion-resilience

Similarly, we can define natural strong and weak design objectives for participation and transfers. Let $a^{eq}(\theta)$ be the dominant strategy or Nash equilibrium profile corresponding to type profile θ, for CR and weakly CR mechanisms respectively. Then the design objectives are as follows:

(2a) Individual Rationality $(u_i(\theta_i, x(a^{eq}(\theta))) \geq 0)\ \forall \theta \in \Theta\ \forall i \in \mathcal{N}$
(2b) Coalitional Rationality $(U_S(\theta_S, x(a^{eq}(\theta))) \geq 0)\ \forall \theta \in \Theta\ \forall S \in \mathcal{S}$
(3a) No positive transfers (NPT): $c_i(a) \geq 0 \quad \forall a \in\ \forall i \in \mathcal{N}$
(3b) Budget balance: $\sum_{i \in \mathcal{N}} c_i(a) \geq 0 \quad \forall a \in \mathcal{A}$

Our main positive result Section 4 is a mechanism that satisfies the stronger constraints (1a), (2a), (3a) for surplus-submodular settings; on the other hand, the impossibility results we present in Section 5 apply even to mechanisms that satisfy only weaker constraints (1b), (2b), (3b).

3 EXPANDED VCG

Before presenting our main result in Section 4, we first construct a simple mechanism that only satisfies the property of weak collusion resilience, in order to give the reader an intuition for how the assumption of surplus-submodularity (ref. Section 2.1) can allow the designer to ensure coalitional incentives without knowing coalitional structure $\mathcal{S}$.

The idea behind the mechanism is to expand the strategy space of the VCG mechanism such that agents report not just their individual types (i.e., values over item bundles), but in addition, values for all possible allocations – an agent can, for example, report having a value for some other agent getting an item. After collecting these reports, the mechanism then verifies whether the reported values satisfy surplus-submodularity, and allocates if this condition is met, charging agents their externalities. These two steps together result in a weakly CR mechanism.

To describe the mechanism formally, we introduce some additional notation. Let $f_i'(x)$ be the value agent i reports for allocation x, and f' the vector of all reports. The mechanism we propose takes as input an efficient allocation rule $x^*(\cdot)$, which specifies for each report vector f' and each subset S of agents, an allocation[5] $x^*(f', S)$ in S:

$$x^*(f', S) \in \operatorname*{argmax}_{x \in \mathcal{X}} \left\{ \sum_{i \in S} f_i'(x) \right\}. \tag{1}$$

Given an efficient allocation rule x^*, let $\mathrm{W}^*(f', S \mid T)$ denote the total reported value of agents in the set S under the optimal allocation for agents in T:

$$\mathrm{W}^*(f', S \mid T) = \sum_{i \in S} f_i'(x^*(f', T)). \tag{2}$$

With a small abuse of the above notation, we define the **surplus function** for a set $S \subseteq \mathcal{N}$ and given reports f' as $\mathrm{W}^*(f', S) = \mathrm{W}^*(f', S \mid S) = \max_{x \in \mathcal{X}} \sum_{i \in S} f_i'(x)$. The name comes from noting that when all coalitions report their total value truthfully (i..e., when $\sum_{i \in T} v_i(\theta_i, \cdot) = \sum_{i \in T} f_i(\cdot)$ for each coalition T), we have $\mathrm{W}^*(f', S) = W^*(\theta, S)$ for all S.

We now formally describe the Expanded VCG mechanism. For a fixed choice of an efficient allocation rule x^*, the Expanded VCG mechanism first asks every agent i to report her values $f_i'(x) \geq 0$ for every possible allocation $x \in \mathcal{X}$. The mechanism then checks if the reported surplus function $\mathrm{W}^*(f', S)$ is submodular in S, for every $S \subseteq \mathcal{N}$ [6]. If it is, the mechanism proceeds to implement the efficient allocation $x^*(f', \mathcal{N})$ and charges agents their VCG payments with the Clarke pivot rule (i.e., their externality):

$$c_i = \mathrm{W}^*(f', \mathcal{N} \setminus i \mid \mathcal{N} \setminus i) - \mathrm{W}^*(f', \mathcal{N} \setminus i \mid \mathcal{N}) \tag{3}$$

If, however, $\mathrm{W}^*(f', S \mid S)$ fails to be submodular, the mechanism does not allocate to anyone (i.e., it chooses $x = \emptyset$) [7].

For any coalition structure, we show that a particular coalitional strategy, which we refer to as the **truthful spokesman strategy**, results in a ex post collusive-Nash equilibrium. Intuitively, this involves a single representative truthfully reporting the preferences of the coalition. Formally, let $i_1^S = \operatorname{argmin}\{i | i \in S\}$ be the "representative" agent in a coalition S. Then a coalitional strategy $a_S = \{f_i\}_{i \in S}$ is a **spokesman strategy** if agents $i \in S \setminus i_1$ report $f_i'(x) = 0 \ \forall x \in \mathcal{X}$. We call a spokesman strategy truthful if $f_{i_1}'(x) = \sum_{i \in S} v_i(\theta_i, x)$.

The payment formula (3) straightforwardly guarantees NPT. In the main result of this section, we show that the Expanded VCG satisfies weak collusion resilience.

Theorem 3.1. *Assume a surplus-submodular setting that satisfies excludability. Then Expanded VCG mechanism satisfies coalitional rationality and weak collusion resilience with truthful spokesman as the equilibrium strategy profile.*

Before proceeding to the proof of this theorem, we give a couple of examples to illustrate the mechanism and its properties. In the first example we return to the setting of Example 2.6, in which we have shown that VCG fails to be CR.

[5]Although our notation $x^*(f', S)$ allows for the allocation to depend on the entire vector of reports, to avoid pathologies, we assume that the allocation $x^*(f', S)$ only depends on the sum $\sum_{i \in S} f_i'$.
[6]Note that the value of $\mathrm{W}^*(f', S)$ does not depend on a specific choice of $x^*(\cdot)$, and as a consequence the surplus-submodularity test is also independent of this choice.
[7]This step makes use of the excludability assumption we defined in Section 2.1

Example 3.2 (Expanded VCG for unit demand valuations). Consider the unit-demand setting from Example 2.6, with 3 agents and 2 items, valuations $\{(2,0),(0,2),(1,1)\}$, and where agents $\{1,2\}$ collude. Suppose we use the Expanded VCG mechanism, and consider the coalitions playing the truthful spokesman strategy, i.e., agent 3 reports truthfully, agent 1 reports having value 2 for receiving the first item and a value 2 for agent 2 receiving the second item, and agent 2 reports having value 0 for all possible allocations. In Theorem 3.1, we show that the truthful spokesman strategies forms a CN equilibrium. We now show that these reports pass the surplus-submodularity test, and that the resulting allocation is efficient.

To get intuition for the definition of surplus function given above, note first that we have $W^*(f',\{1\}) = W^*(f',\{1\} \mid \{1\}) = 4$ – this follows from observing that the optimal allocation from the point of view of agent 1 (as per her report) is to allocate the items to 1 and 2, whereupon agent 1 has a total reported value of 4. Similarly, one can check $W^*(f',\{2\}) = 0$ (as per her report), $W^*(f',\{3\}) = 1, W^*(f',\{1,3\}) = 4, W^*(f',\{1,2\}) = 4, W^*(f',\{2,3\}) = 1$, and $W^*(f',\{1,2,3\})) = 4$. It is now straightforward to verify that the surplus function is submodular.

Observe that under these reports, the mechanism allocates efficiently, giving the first item to agent 1 and the second to agent 2. Finally, using (3), note the transfers $c_1 = W^*(f',\{2,3\} \mid \{2,3\}) - W^*(f',\{2,3\} \mid \{1,2,3\}) = 1 - 0 = 1$, and $c_2 = c_3 = 0$ satisfy coalitional rationality. □

We now make a few observations about the above example. First, note that, when agent 3 reports truthfully, the above outcome (i.e., allocation coupled with payment) cannot occur under the VCG mechanism under any reports of agents 1 and 2, as both agents have to pay $c_1 = c_2 = 1$ under VCG in order to get both the items. Note also that the revenue when agents 1 and 2 are colluding is strictly lower when compared to the case of no collusion.

Finally, we can use the above example to show that Expanded VCG does not satisfy collusion resilience (as opposed to weak CR). Assume the same setting, but now with no coalitions. If agent 1 truthfully reports her value, then it is optimal for 2 to do the same. However, if instead agent 1 makes the same report as in the above example, then agent 2 is disincentivized to report truthfully, as this leads to a non-submodular surplus, whereupon the mechanism excludes all agents. On the other hand, by reporting a zero value for all outcomes, agent 2 gets a higher utility (as in the above example), thereby demonstrating that 2 now does not have a dominant strategy.

Moreover, we can also show that the surplus-submodularity test is crucial for implementing efficient outcomes, even in single-item settings.

Example 3.3 (Surplus-submodularity test is essential). There are 3 agents $\{1,2,3\}$ competing for a single item, with values $v_1 = 1$, $v_2 = 2$, $v_3 = 3$. Suppose we employ Expanded VCG, but skip the step of checking whether $W^*(f',S)$ is submodular. Now consider a case where agent 3 reports truthfully, but agents 1 & 2 collude as follows: agent 1 reports having value 4 when agent 2 receives the item, and 0 for all other allocations; agent 2 reports having value 4 for receiving the item. In absence of the surplus-submodularity test, this would allow agent 2 to obtain the item while paying a price of zero, as the externality of both agents 1 and 2 is 0 (as per formula (3)). Note however that this allocation is inefficient.

3.1 Weak Collusion Resilience of Expanded VCG

We now provide the proof of Theorem 3.1, thereby establishing the weak CR property of the Expanded VCG mechanism. Informally our proof proceeds as follows. First, we show that, whenever mechanism makes a non-trivial allocation , any coalition's payment is at least its externality. We then show that any outcome that a coalition can possibly implement is also implementable under a payment that is exactly coalition's externality, and this is done via a class of strategies we call Spokesman strategies. It then follows that when looking for an optimal strategy it is enough for coalition to only consider Spokesman strategies. We then show that a particular such strategy, Truthful Spokesman, implements the efficient outcome and forms a collusive-Nash equilibrium for any type profile and coalitional partition.

Recall that for brevity we write $\mathrm{W}^*(f', S) = \mathrm{W}^*(f', S|\, S)$. We also define the externality function:

$$E(f', S) = \mathrm{W}^*(f', \mathcal{N} \setminus S) - \mathrm{W}^*(f', \mathcal{N} \setminus S|\mathcal{N}) \tag{4}$$

With our next lemma we show that the externality function is submodular whenever surplus function is submodular. An implication of this is that in Extended VCG mechanism, every subset of agents pays at least its externality.

Lemma 3.4. *For any report f' , the surplus function $\mathrm{W}^*(f', \cdot)$ is submodular, if and only if the externality function $E(f', \cdot)$ is submodular.*

Proof. First consider the function $\hat{E}(f', S) = \mathrm{W}^*(f', S) - \mathrm{W}^*(f', S|\, \mathcal{N})$. It is a linear combination of a submodular function $\mathrm{W}^*(f', S)$ and a modular function $\mathrm{W}^*(f', S \mid \mathcal{N})$, thus is itself submodular.

Now, since $E(f', S) = \hat{E}(f', \mathcal{N} \setminus S)$, the statement of the lemma follows (if $F(S)$ is submodular then $F(\mathcal{N} \setminus S)$ is also submodular; see e.g. [Bach, 2013]). □

Now, with payment formula (3) in mind we have the following corollary.

Corollary 3.5. *In Expanded VCG, if reports f' pass the surplus-submodularity test (i.e., $\mathrm{W}^*(f', S)$ is submodular in S), then the total payment of any subset S of agents is at least its externality:*

$$C_S(f') = \sum_{i \in S} c_i(f') \geq \mathrm{W}^*(f', (\mathcal{N} \setminus S)) - \mathrm{W}^*(f', (\mathcal{N} \setminus S) \mid \mathcal{N}) \tag{5}$$

Armed with this result, we now turn to showing that Expanded VCG mechanism is weakly collusion resilient.

Proof of Theorem 3.1. We prove the theorem by first showing that in choosing an optimal policy it is enough for a coalition to only consider spokesman strategies. We then argue that, when other coalitions are playing truthful spokesman strategy, truthful spokesman is the optimal spokesman strategy for a coalition. Properties of efficiency and non-negative coalitional utility at the equilibrium then follow in a straightforward manner.

We start by demonstrating the property of coalitional rationality. If all agents play truthful spokesman, by the assumption of surplus submodularity efficient allocation is going to be made and the utility of coalition S is

$$U_S(\theta_S, f_{tr}) = W^*(\theta, \mathcal{N}) - W^*(\theta, \mathcal{N} \setminus S) \geq 0.$$

Here, the inequality follows from excludability assumption. Now, given a coalition structure $\mathcal{S}$, consider some coalition $S \in \mathcal{S}$, and assume all other coalitions play the truthful spokesman strategy. We next show for any report f'_S of the coalition S, the truthful spokesman report achieves at least as much utility for the coalition. There are two cases.

Case (1): Playing f'_S leads to $\mathrm{W}^*(f', S)$ not being submodular, which in turns leads to the empty allocation $x = \emptyset$ and zero utility. Here, using excludability and the fact that allocation rule $x^*(f')$ is efficient, it is clear the coalition can do at least as well with truthful spokesman strategy.

Case (2): The coalitional strategy f'_S passes the surplus-submodularity test, i.e., the function $\mathrm{W}^*(f', \cdot)$ is submodular. To prove the optimality of truthful spokesman in this case, we first narrow the space of reports to only spokesman reports. In particular, we first show that there is a spokesman report $\hat{f}_S = \{\hat{f}_j\}_{j \in S}$, where $\hat{f}_{i_1^S} = \sum_{i \in S} f'_i(x)$, and $\hat{f}_i = 0$ for all $i \in S \setminus i_1^S$, under which the agents in S obtain at least as much utility as $\hat{f}'$. We let $\hat{f} = (\hat{f}_S, f'_{\mathcal{N}\setminus S})$ to denote the vector of reports in which S plays the spokesman strategy $\hat{f}_S$ while the reports of other agents remain unchanged.

First, we show $\mathrm{W}^*(\hat{f}, S)$ continues to be surplus submodular. To see this, note that as only i_1^S makes a non-zero report, it is enough to check the submodularity condition is satisfied for this agent, i.e. $\mathrm{W}^*(\hat{f}, Q \cup \{i_1^S\}) - \mathrm{W}^*(\hat{f}, Q) \geq \mathrm{W}^*(\hat{f}, T \cup \{i_1^S\}) - \mathrm{W}^*(\hat{f}, T)$ whenever $Q \subset T,\ Q, T \subset \mathcal{N} \setminus S$. By definition of $\hat{f}$, we have $W^*(\hat{f}, Q \cup \{i_1\}) = W^*(f', Q \cup S)$ for all $Q \subset \mathcal{N} \setminus S$. The inequality now follows from the fact that $W^*(f', \cdot)$ is submodular.

Next, by construction of the spokesman strategy, allocation is going to be the same as for the original strategy, since the allocation only depends on the sum of all values $\sum_{i \in \mathcal{N}} f'(x)$ (ref. allocation rule (1)), but not how they are distributed between the agents. At the same time, the payment under Spokesman strategy is exactly the externality of the coalition, and by Lemma 3.5, is the minimum possible payment under the allocation $\tilde{x}$.

Next, we demonstrate that among all spokesman strategies, the truthful spokesman yields optimal surplus for S. This follows from a variant of the standard VCG argument: for coalition S playing spokesman strategy f'_S, using the payment rule (3) we get the following expression for total utility of agents in S

$$U_S(\theta, f'_S) = \sum_{i \in S} v_i(\theta_i, x^*(f', \mathcal{N})) + \sum_{i \notin S} f'_i(x^*(f', \mathcal{N})) - \mathrm{W}^*(f', \mathcal{N} \setminus S).$$

Here the last term is constant, and the first two terms are maximized when coalition S plays truthful spokesman. Thus we have shown that truthful spokesman forms a collusive-Nash equilibrium.

Finally, assuming all coalitions play truthful spokesman, the total reported value is $\sum_{i \in \mathcal{N}} f'_i(x) = \sum_{i \in \mathcal{N}} v_i(\theta_i, x) = W(\theta, x)$, and as a consequence, the welfare is maximized at this equilibrium. □

4 COALITIONAL REVELATION MECHANISM

In this section we present our main result, the *Coalitional Revelation Mechanism* (CRM) that satisfies the objectives (1a), (2a) and (3a) outlined in Section 2.3 in surplus-submodular settings.

Informally, CRM requires each agent to report her coalition in addition to her type. The mechanism checks whether the coalitional reports of the agents are consistent, and if so, implements the efficient allocation and charges each coalition its externality. To achieve its goals, two conditions must be met: (i) agents must not find it beneficial to expand their coalitions, by including agents outside their coalition in their reports (similar to the situation in Example 3.2), and (ii), a coalition must not be better off breaking into smaller coalitions. CRM explicitly creates disincentives to ensure the first, whereas the second is a consequence of surplus-submodularity.

4.1 Formal statement of the mechanism

Formally, a report of an agent i is a tuple $a_i = (\theta'_i, \mathrm{S}'_i)$, where $\theta'_i \in \Theta_i$ is the agent's reported type, and S'_i is the reported coalition that agent i is part of. We transcribe the profile of actions a as (θ', S'), where θ' denotes the profile of reported types, and S' denotes the profile of reported coalitions. Furthermore, we denote the subset of reports for a set T by including T in its subscript: thus, θ'_T denotes the type reports of agents in T, whereas S'_T denotes the coalitional reports of agents in T.

The CRM mechanism takes as input an efficient allocation function $x^*(\cdot)$, that specifies for each type-report θ' and a set T, an efficient allocation $x^*(\theta', T)$ for that set. We assume that $x^*(\theta', T)$ depends only on θ'_T, and that $x^*(\theta', T)$ excludes all the agents not in T. For sets T, U, we define the surplus functions $\mathrm{W}^*(\theta', U \mid T) = \sum_{i \in U} v_i(\theta'_i, x^*(\theta', T))$ and $W^*(\theta', T) = W^*(\theta', T \mid T)$.

Given a report profile (θ', S'), we construct a coalitional graph $G(\mathrm{S}')$ with vertices in $\mathcal{N}$ and a directed edge from i to j if $j \in S'_i$. Furthermore, given a set $T \subset \mathcal{N}$, we define $G(\mathrm{S}', T)$ to be the

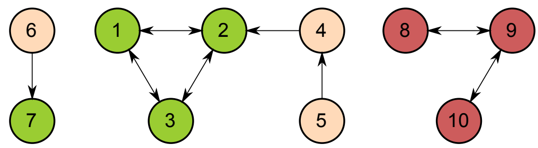


Fig. 2. In this example with 10 agents, we represent coalitional reports with arrows, e.g. the arrow from 6 to 7 denotes $S'_6 = \{7\}$, i.e., agent 6 claiming to be in coalition with 7. Here there are two consistent coalitions: $\{1, 2, 3\}$ and $\{7\}$, and these coalitions are scapegoats, because they have non-reciprocated partner claims directed towards them, from 3 and 5, respectively. Because there are 2 scapegoats, the mechanism does not allocate in this scenario. Also note that agents $\{8, 9, 10\}$ do not form a consistent coalition, and thus prevent mechanism from allocating even in the absence of scapegoats.

subgraph induced by $G(S')$ on T. We say a set $T \subset \mathcal{N}$ is a *consistent coalition*, if $S'_i = T$ for all $i \in T$ (i.e., if T is a bi-directional clique in $G(S')$). We say a consistent coalition T is a *scapegoat* if there is an agent $j \notin T$ such that $S'_j \cap T \neq \emptyset$. In such a setting, we call the agent j a *kibitzer* [8]. We illustrate these concepts in Fig. 2.

If under a report (θ', S'), the coalitional graph $G(S')$ is a disjoint union of consistent coalitions, then CRM implements the efficient allocation and charges every consistent coalition its externality. (We call such a report consistent.) The main challenge then is to create incentives for agents to report their coalitions consistently. To address this, we first discuss two ways in which inconsistent reports may arise: through unilateral or multilateral deviations.

We say an inconsistent report arises from a unilateral deviation if there is a single scapegoat. Otherwise, we say the inconsistent report arises from a multilateral deviation. (Note that a multilateral deviation may have no scapegoats, or even no consistent coalitions.)

The mechanism can be formally stated as follows:

(1) In the case of consistent reporting, an efficient allocation $x^*(\theta', \mathcal{N})$ is implemented and every consistent coalition T is charged its externality:
$$c_T = E(\theta', T) \triangleq W^*(\theta', \mathcal{N} \setminus T) - \mathrm{W}^*(\theta', \mathcal{N} \setminus T \mid \mathcal{N}). \tag{6}$$
(2) If reports are inconsistent due to a unilateral deviation, i.e., there is one scapegoat coalition T, then the allocation $x^*(\theta', T)$ (with $x^*_i(\theta', T) = \emptyset$ for all $i \in \mathcal{N} \setminus T$) is implemented, all kibitzers are charged $c_j = W^*(\theta', \mathcal{N})$, and others are charged 0.
(3) If reports are inconsistent due to multilateral deviation, empty allocation is made $x = \emptyset$ with no payments charged.

Notice that, if all agents report their types and coaltions truthfully, each coalition is charged its VCG payments with Clarke pivot rule, as per (6). Thus, it immediately follows that the mechanism satisfies coalitional rationality (CR) and no positive transfers (NPT).

The CRM mechanism does not specify how the charge c_T for a coalition is split across different members of the coalition. Note that the choice of a splitting rule does not influence collusion resilience, but may affect whether individual rationality (IR) holds. To achieve IR, one may consider a splitting rule where for each $i \in T$, we have $c_i = \gamma_i c_T$, where $\gamma_i = \frac{\mathrm{W}^*(\theta', i \mid \mathcal{N})}{\mathrm{W}^*(\theta', T \mid \mathcal{N})}$ is the fraction of the coalition's welfare that is accrued by agent i.

Thus, in demonstrating that CRM satisfies objectives outlined in Section 2.3 we are left with proving that it is collusion resilience. We address this in the next section.

[8] *Kibitzer*: [kib-it-ser] (noun) one who looks on and often offers unwanted advice or comment.

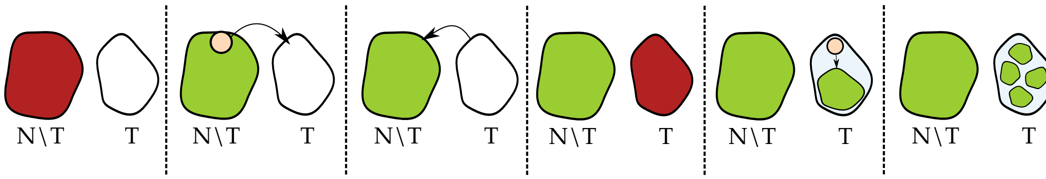


Fig. 3. An exhaustive list of non-truthful reports we consider in the proof of collusion resilience of the CRM mechanism. Here, red denotes an inconsistent report arising from a multilateral deviation, green denotes consistent reports, and orange denotes a kibitzer.

4.2 Collusion Resilience of CRM

We have the following lemma, which is an analog of Lemma 3.4.

Lemma 4.1. *Assume a surplus submodular setting. Then $E(\theta', \cdot)$ is submodular for all $\theta' \in \Theta$.*

The proof of the lemma is analogous to that of Lemma 3.4 and is left out for brevity. The main result of this section now is as follows.

Theorem 4.2. *Consider a surplus-submodular setting satisfying excludability and the non-negative values assumption. Then CRM is collusion resilient.*

Apart from satisfying stronger properties of collusion resilience and individual rationality (with a particular choice of splitting rule), CRM has a couple of important features that distinguish it from Expanded VCG. First, its computational cost is equivalent to that of VCG mechanism, in contrast to Expanded VCG where the submodularity check can be expensive. Secondly, the collusive-dominant strategy of truthful reporting does not require that agents in a coalition know each others types, thus allowing privacy for colluding agents.

Proof of Theorem 4.2. Consider a fixed coalition structure $\mathcal{S}$, and a coalition $T \in \mathcal{S}$. Our goal is to demonstrate that truthful revelation is an optimal strategy for agents in T independently of strategies employed by agents in $\mathcal{N} \setminus T$. Assume the true type profile is θ and that the reports of agents in $\mathcal{N} \setminus T$ is fixed at $\{(\theta'_i, \mathrm{S}'_i)\}_{i \in \mathcal{N} \setminus T}$. Now we consider two possible type-report profiles, $\theta' = (\theta'_T, \theta'_{\mathcal{N} \setminus T})$, and $\tilde{\theta} = (\theta_T, \theta'_{\mathcal{N} \setminus T})$, where θ_T denotes the true types of agents in T.

Suppose $G(\mathrm{S}', \mathcal{N} \setminus T)$ is inconsistent, i.e., it is not a disjoint union of bi-directional cliques (first illustration in Fig 3). In this case, the allocation chosen by the mechanism is independent of T's reports, and T's report can only affect its payment (by some agent in T becoming a kibitzer). Given this, it is optimal for the agents in T to report truthfully (and have zero payments). Hence, for the rest of the proof, we will assume that $G(\mathrm{S}', \mathcal{N} \setminus T)$ is consistent.

Given this there are two possible scenarios depending on whether or not the coalitional graph $G(\mathrm{S}')$ has a directed edge towards an agent in T from an agent not in T (a potential kibitzer). In the former case (second illustration in Fig 3), it follows immediately that the coalition T should report truthfully, since in this case, the mechanism will implement the optimal allocation $x^*(\theta', T)$ for T (since it will be targeted by a kibitzer). Thus, we are left to analyze the case where there are no potential kibitzer, i.e., there are no edges directed towards any agent in T.

Now, note that it is never optimal for any agent in T to become a kibitzer towards any consistent coalition in $\mathcal{N} \setminus T$ (third illustration in Fig 3). This is because then that agent will be charged the optimal welfare under that report, and the entire coalition will be excluded from the allocation, leading to a negative coalitional utility. Thus, the coalitional reports must be so that $G(\mathrm{S}') = G(\mathrm{S}', T) \cup G(\mathrm{S}', \mathcal{N} \setminus T)$.

This leads to three possibilities (corresponding to the last three cases illustrated in Fig. 3):
(1) $G(\mathrm{S}', T)$ has multilateral deviations leading to an internally inconsistent coalition structure;
(2) $G(\mathrm{S}', T)$ has a single scapegoat (and one or more kibitzers); and
(3) $G(\mathrm{S}', T)$ is a union of more than one consistent coalitions.
The first case is trivial, since the mechanism makes the empty allocation, which leads to lower utility for T as compared to truthful reporting.

In the second case, where there is a single scapegoat in $G(\mathrm{S}', T)$, the utility of T satisfies

$$\begin{aligned} U_T(\theta_T, (\theta', \mathrm{S}')) &\le W^*(\tilde{\theta}, T) - W^*(\theta', \mathcal{N}) \ \le\ W^*(\tilde{\theta}, T) - W^*(\theta', \mathcal{N} \setminus T) \\ &\le W^*(\tilde{\theta}, \mathcal{N}) - W^*(\theta', \mathcal{N} \setminus T) \ \le\ W^*(\tilde{\theta}, T \mid \mathcal{N}) + W^*(\tilde{\theta}, \mathcal{N} \setminus T \mid \mathcal{N}) - W^*(\tilde{\theta}, \mathcal{N} \setminus T) \\ &= W^*(\tilde{\theta}, T \mid \mathcal{N}) - E(\tilde{\theta}, T). \end{aligned}$$

The first upper bound follows from the fact that $W^*(\tilde{\theta}, T)$ is the largest value T can obtain and $W^*(\theta', \mathcal{N})$ is the payment charged to the kibitzer. The second and third inequalities follow from the fact that $W^*(\cdot)$ is monotone non-decreasiing, which in turn is implied by the excludability assumption. The expression in the last line is just the utility of truth reporting and it follows from the payment formula (6), and the last inequality also follows from the fact that $W^*(\tilde{\theta}, \mathcal{N} \setminus T) = W^*(\theta', \mathcal{N} \setminus T)$.

Finally, we focus on the third case, where $G(\mathrm{S}, T)$ is a union of more than one consistent coalitions. In other words, the agents in T report as belonging to multiple (consistent) coalitions $\{T_1, \cdots, T_K\}$. Then, the mechanism implements the allocation $x^*(\theta', \mathcal{N})$ and charges each consistent coalition T_i its externality. Thus, the total utility of the coalition T is given by

$$U_T(\theta_T, (\theta', \mathrm{S}')) = \sum_{i \in T} v_i(\theta_i, x^*(\theta', \mathcal{N})) - \sum_{k=1}^{K} E(\theta', T_k) \le \sum_{i \in T} v_i(\theta_i, x^*(\theta', \mathcal{N})) - E(\theta', T).$$

Here, in the inequality, we have used the fact that, as per Lemma 4.1, $E(\theta', \cdot)$ is submodular. Note that the expression in the right hand side is the utility the coalition receives when they report their coalition structure truthfully, and report their types to be θ'_T. In order to show that the coalition is (weakly) better off reporting their types truthfully, we make a VCG-type argument. Note that

$$\sum_{i \in T} v_i(\theta_i, x^*(\theta', \mathcal{N})) - E(\theta', T) = \sum_{i \in T} v_i(\theta_i, x^*(\theta', \mathcal{N})) + \sum_{i \notin T} v_i(\theta'_i, x^*(\theta', \mathcal{N})) - W^*(\theta', \mathcal{N} \setminus T)$$

Here the last term is constant with respect to θ'_T and the sum of first two terms is maximized when $\theta'_T = \theta_T$ since the allocation rule $x^*(\cdot, \mathcal{N})$ is efficient.

Thus we have shown that truthful revelation is a dominant strategy for CRM. We conclude the proof with noting that, when all coalitions are reporting the types and coalitions truthfully, the efficient outcome $x^*(\theta, \mathcal{N})$ is implemented and thus the mechanism is collusion resilient. □

5 IMPOSSIBILITY RESULTS

The previous sections thus demonstrate the feasibility of collusion-resilience subject to two important restrictions: *surplus-submodularity and excludability* (ref. Definitions 2.1 and 2.2). A natural question now is how critical these assumptions are for collusion-resilience; our results in this section take an important step towards understanding this.

We present impossibility results in two settings. First, we show that in combinatorial auctions, there exists no mechanism that satisfies the weaker requirements outlined in Section 2.3 (i.e., coalitional rationality, budget balance, and weak collusion resilience) for general submodular valuations. Coupled with the collusion resilience property of the CRM mechanism (ref. Theorem 4.2) for surplus submodular settings (and thereby, for gross substitutes valuations), this result essentially defines a sharp boundary for conditions under which CR mechanisms exist. Moreover, our result greatly refines the impossibility result of Deckelbaum and Micali [2017], which states that collusion resilience is impossible for general combinatorial auctions.

The second setting we consider is that of collective decision making, i.e., where agents must choose between several alternatives, with potential monetary reparations. We consider this setting for two reasons: First, the setting is one where the notions of collusion resilience, Sybil proofness and group strategy-proofness coincide, and hence our impossibility result pertains to all three. Secondly, our result suggests this setting is in a sense a worst-case scenario for collusion resilience, as we show that no non-trivial approximation of welfare is possible in the presence of collusion.

One complication for demonstrating negative results for collusion-resilience is that the standard revelation principle does not hold in this setting. Deckelbaum and Micali [2017] circumvent this via an extended revelation principle; in contrast, we adopt an alternate approach, based on the following version of the *taxation principle* applied to our setting:

LEMMA 5.1 (TAXATION PRINCIPLE). *Let M be a weakly CR mechanism, with equilibrium strategy profile a, and let S be some coalition. For any distinct type-profiles $\theta, \theta' \in \Theta^n$ such that $\theta_i = \theta'_i\ \forall i \notin S$, and $x(a(\theta)) = x(a(\theta'))$, the mechanism must charge $c_S(a(\theta)) = c_S(a(\theta'))$.*

Intuitively, the principle asserts that if all agents in the mechanism are playing an equilibrium strategy, then conditioned on identical outcomes, the payments of agents under this strategy must be independent of their types. The proof follows from standard arguments [Rochet, 1985].

5.1 Limits of Collusion Resilience in Combinatorial Auctions

For combinatorial auctions, submodular valuations are in a sense the minimal extension over gross substitutes. Our next result demonstrates that introducing *even a single agent with general submodular valuations* makes a setting non-amenable to CR mechanisms.

THEOREM 5.2 (IMPOSSIBILITY OF CR FOR SUBMODULAR VALUATIONS). *For combinatorial auction with general submodular valuations, there is no mechanism which simultaneously satisfies (i) weak coalitional rationality, (ii) budget balance, and (iii) weak collusion resilience.*

PROOF. We prove this via the contrapositive; in particular, we show that for a given setting with submodular valuations (adapted from an example in [Lehmann et al., 2001]), any mechanism M that satisfies coalitional rationality and weak collusion resilience must result in negative revenue for some instances (i.e., valuation profiles and coalition structure).

Consider the setting depicted in Table 1, in which there are 3 items and 4 agents with budget additive valuations, i.e., the value of agent i value is given by $v_i(x) = min(B_i, \sum_j x_{ij}v_{ij})$, where x_{ij} is the indicator of agent i getting the item j, $\sum_i x_{ij} \leq 1$, and $B_i \geq 0$ is a positive public parameter (in fact, in our example, agents $\{2, 3, 4\}$ have additive valuations and only agent $\{1\}$ has a true budget constraint). For the same setting, we consider four valuation profiles, denoted scenario A, B, C and D in Table 1. It is straightforward to verify that each agent has submodular valuations, but the resulting surplus function $W^*(\theta, S)$ is not submodular. We also assume the (potential) coalition structure is given by $\mathcal{S} = \{\{1, 2\}, \{3, 4\}\}$. Henceforth, we use $v_i(j)$ to denote the value agent j has for item i.

First, consider scenario A, and assume that all agents play as individuals (i.e., there is no collusion). We argue that under any mechanism M which is efficient and individually rational, agents 3 and 4 pay at most $c_3 \leq 1, c_4 \leq 1$. Indeed, note that the efficient allocation x^* does not change if agent 3 had $v_1(3) = 1$, and thus by the taxation principle (Lemma 5.1) and coalitional rationality (which for an individual agent coincides with individual rationality), we must have $c_3^A \leq 1$. Similarly, we can

Scenario A	v_1	v_2	v_3	B_i	x^*
Agent 1	3	5	3	6	2
Agent 2	1	2	0	∞	-
Agent 3	2	0	0	∞	1
Agent 4	0	0	5	∞	3

Scenario B	v_1	v_2	v_3	B_i	x^*
Agent 1	3	5	3	6	1,3
Agent 2	1	2	0	∞	2
Agent 3	2	0	0	∞	-
Agent 4	0	0	$1-\epsilon$	∞	-

Scenario C	v_1	v_2	v_3	B_i	x^*
Agent 1	$\epsilon+\delta$	0	$\epsilon+\delta$	6	1, 3
Agent 2	0	δ	0	∞	2
Agent 3	ϵ	0	0	∞	-
Agent 4	0	0	ϵ	∞	-

Scenario D	v_1	v_2	v_3	B_i	x^*
Agent 1	3	5	3	6	1,3
Agent 2	1	2	0	∞	2
Agent 3	ϵ	0	0	∞	-
Agent 4	0	0	ϵ	∞	-

Table 1. The four scenarios considered in Theorem 5.2: Note that all four correspond to the same underlying combinatorial auction with submodular valuations (with 3 items and 4 agents with budget-additive valuations.)

argue that the mechanism must charge $c_4^A \le 1$ (this follows from observing that agent 4 continues to receive item 3 in the optimal allocation as long as $v_3(4) \ge 1$). Now, since a coalition can always choose to play as individual agents would, it follows that coalition $\{3,4\}$ will pay at most $c^A_{\{3,4\}} \le 2$ in this scenario, while getting items 1 and 3.

Next consider Scenario B. Under an efficient allocation, the coalition $\{3,4\}$ should receive no items. On the other hand, since the valuations of agents 1 and 2 have not changed, we know from Scenario A that $\{3,4\}$ can misreport and receive items 1 and 3 for a cost of 2 (and thereby violate efficiency). Note now that if we demand *no positive transfers* (NPT) rather than budget balance, we immediately obtain our impossibility at this stage. More generally, however, we can observe that to disincentivize such a deviation, the mechanism must compensate agents $\{3,4\}$ via a *positive* transfer of $c_{\{3,4\}} \le -1$; this follows from the taxation principle (Lemma 5.1), using the fact that the optimal allocation does not depend on the value of $\epsilon > 0$.

Next, consider Scenario C. Here, for any $\epsilon > 0$ and $\delta > 0$, the optimal allocation x^* remains the same. Consequently, the mechanism must charge $c_{\{1,2\}} \le 2\epsilon$.

Finally, consider Scenario D. Since the allocation here is the same as in Scenario B, we have via the taxation principle that coalition $\{3,4\}$ must still get a positive transfer $c_{\{3,4\}} \le -1$. Moreover, following from Scenario C, we must have $c_{\{1,2\}} \le 2\epsilon$. Hence, for $\epsilon < 1/2$, the resulting revenue of the mechanism is negative. This completes our proof. □

5.2 Limits of Collusion Resilience in Collective decision making

Theorem 5.2 thus provides some justification for the need for the surplus-submodularity for obtaining CR mechanisms in combinatorial auctions. In this section, we consider the problem of designing CR mechanisms for collective decision making problems (described below), which are of great practical relevance, but moreover, are in a sense the furthest from our assumptions in that they violate both the surplus-submodularity and excludability conditions. Here, we demonstrate even stronger inapproximability results in the presence of collusions.

Formally, we consider the following *collective decision making* setting: There are n agents and an outcome $x \in \mathcal{X} = [0,1]$, which can represent either a fractional allocation to be chosen publicly, or a probability of picking one of the two outcomes (e.g., a political election). Every agent i has some private values $\theta_i = (v_i(0), v_i(1)) \ge 0$, corresponding to outcomes 0 and 1. [9]

[9]The impossibilities we prove in this section easily extend to the case of choosing between more than 2 options

We assume the mechanism can potentially charge agents, and thus, the utility of agent i is given by $u_i(\theta_i, x) = v_i(0)x + v_i(1)(1-x) - c_i$, where c_i is the payment charged to the agent i. Welfare $W(\theta, x)$ is then given by $W(\theta, x) = \sum_{i \in \mathcal{N}} (v_i(0)x + v_i(1)(1-x))$. Note that this setting does not satisfy excludability, as all agents experience the utility corresponding to the chosen outcome x (i.e., everyone has to live with the collective decision). Moreover, in the absence of excludability, the notion of individual rationality (i.e., participation constraint) requires a more careful definition.

Definition 5.3. A mechanism M satisfies individual rationality if for every agent i there is an action $\emptyset \in \mathcal{A}_i$ s.t. for any two action profiles a, a' such that $a'_i = a_i = \emptyset$ we have $x(a) = x(a')$ and $c_i(a) = c_i(a') = 0$.

This captures the main idea behind an IR constraint in that it models that agents have a 'non-participation' action that guarantees zero payment and does not influence the allocation.

When coalitions are not a concern, VCG mechanisms satisfies efficiency, incentive compatibility, individual rationality and NPT. However, in presence of coalitions,VCG can fail spectacularly (as any two agents can implement preferred outcome without having to pay anything). Here we show that the notion of collusion resilience does not alleviate the problem as no efficient weakly collusion resilient mechanism exists for this setting. Moreover, we argue that this impossibility result persists even if the requirement of efficiency is relaxed to demanding any non-trivial constant approximation of welfare.

Although the impossibility of approximation 5.6 supersedes the impossibility of perfect efficiency, we have decided to include both results, as the proof of the latter is simpler and is instructive in the logic of the more general and as a consequence contrived approximation result.

Theorem 5.4 (Impossibility of CR for Collective decision making). *In collective decision making settings, there is no mechanism that simultaneously satisfies* (i) *weak coalitional rationality,* (ii) *budget balance, and* (iii) *weak collusion resilience.*

The proof of this theorem is similar in its structure to that of the Theorem 5.2; for brevity we defer it to the Appendix A.

Before stating our main result, the impossibility of approximation, we first formally define the approximation factor for the described setting of collective decision making.

Definition 5.5. A mechanism M guarantees an α-approximation of welfare at the strategy profile $a(\theta)$ if for all types $\theta \in \Theta$ we have

$$\frac{W(\theta, x(a(\theta)))}{\max_{x \in \mathcal{X}} W(\theta, x)} \geq \alpha,$$

Note that an $\alpha = 0.5$ approximation can always be achieved by flipping a coin. The following result, however, shows that in the presence of collusion, this is in fact the best possible, even when allowing for positive transfers.

Theorem 5.6 (Inapproximability of Welfare in collective decision making). *In the described setting of collective decision making, there is no mechanism that simultaneously satisfies* (i) *weak coalitional rationality,* (ii) *budget balance, and* (iii) *approximates efficiency at a collusive-Nash equilibrium, with approximation factor* $\alpha > 0.5$.

Though the proof of this result follows a similar outline as our previous impossibility proofs, it is more involved. At a high level, we assume the existence of a coalitionally rational mechanism M which achieves an $\alpha = 0.5 + \beta$, for some fixed β welfare approximation, and then construct an instance with negative revenue. More precisely, for every given β, we construct a sequence of valuation profiles, and then use a descent argument to show that in at least one instance of the sequence, the 'winning' side pays an amount that is too small to cover the compensation needed to incentivize the 'losing side' to not deviate, and hence the resulting revenue is negative. The complete proof is available in the Appendix A.

5.3 Discussion

We conclude with some remarks on the implications of the impossibility results, and potential extensions. First, note that in the setting of collective decision making, collusion resilience is related to the notion of sybil-proofness, where a single agent may report to have multiple identities. This follows from the fact that a coalition S with values $\{v_i\}_{i \in S}$ can be thought of as a single agent with value $v = \sum_{i \in S} v_i$, but with $|S|$ identities. We build on this connection in Appendix B.

Second, note that in our inapproximability result, Theorem 5.6, the deviations we consider in the constructed instances increase the payoff of every agent in the coalition. This immediately implies that our impossibility results carry over to group-strategyproofness. In particular, our Theorem 5.6 implies that there is no group-strategyproof mechanism that non-trivially approximates welfare and satisfies IR and budget balance.

Finally, we note that our inapproximability results require arbitrarily large coalitions. Whether one can obtain non-trivial approximations when the size of all coalitions is bounded is an interesting open question that we leave for future work.

REFERENCES

Lawrence M Ausubel and Paul Milgrom. 2006. The lovely but lonely Vickrey auction. *Combinatorial auctions* (2006).

Francis Bach. 2013. Learning with submodular functions: A convex optimization perspective. *Foundations and Trends® in Machine Learning* (2013).

Jing Chen and Silvio Micali. 2012. Collusive dominant-strategy truthfulness. *Journal of Economic Theory* (2012).

Alan Deckelbaum and Silvio Micali. 2017. Collusion, efficiency, and dominant strategies. *Games and Economic Behavior* (2017).

Andrew V Goldberg and Jason D Hartline. 2005. Collusion-resistant mechanisms for single-parameter agents. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics.

Jerry Green and Jean-Jacques Laffont. 1979. On coalition incentive compatibility. *The Review of Economic Studies* (1979).

Faruk Gul and Ennio Stacchetti. 1999. Walrasian equilibrium with gross substitutes. *Journal of Economic theory* (1999).

Jean-Jacques Laffont and David Martimort. 2000. Mechanism design with collusion and correlation. *Econometrica* (2000).

Benny Lehmann, Daniel Lehmann, and Noam Nisan. 2001. Combinatorial auctions with decreasing marginal utilities. In *Proceedings of the 3rd ACM conference on Electronic Commerce*. ACM.

Renato Paes Leme. 2017. Gross substitutability: An algorithmic survey. *Games and Economic Behavior* (2017).

Paul Milgrom and Ilya Segal. 2014. Deferred-acceptance auctions and radio spectrum reallocation.. In *EC*.

Hideki Mizukami. 2003. On the constancy of bribe-proof solutions. *Economic Theory* (2003).

Hervé Moulin and Scott Shenker. 2001. Strategyproof sharing of submodular costs: budget balance versus efficiency. *Economic Theory* (2001).

Paolo Penna and Carmine Ventre. 2014. Optimal collusion-resistant mechanisms with verification. *Games and Economic Behavior* (2014).

J.C. Rochet. 1985. The taxation principle and multi-time Hamilton-Jacobi equations. *Journal of Mathematical Economics* 14, 2 (1985), 113 – 128. DOI: http://dx.doi.org/https://doi.org/10.1016/0304-4068(85)90015-1

James Schummer. 2000. Manipulation through bribes. *Journal of Economic Theory* (2000).

Taiki Todo, Atsushi Iwasaki, Makoto Yokoo, and Yuko Sakurai. 2009. Characterizing false-name-proof allocation rules in combinatorial auctions. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems.

Liad Wagman and Vincent Conitzer. 2014. False-name-proof voting with costs over two alternatives. *International Journal of Game Theory* (2014).

Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. 2004. The effect of false-name bids in combinatorial auctions: New fraud in Internet auctions. *Games and Economic Behavior* (2004).

A PROOFS OF IMPOSSIBILITY RESULTS

Proof of Theorem 5.4. Similar to Theorem 5.2, the proof of impossibility follows from a contrapositive argument, wherein we show that any mechanism M that satisfies coalitional rationality and weak collusion resilience in a social choice setting must result in negative revenue for some instances. This argument follows from considering a sequence of 4 instances, involving the same 3 agents competing in a social choice setting:

	$v_1(0)$	$v_2(1)$	$v_3(1)$	Coalitions	x^*	Derived bounds on payments
Scenario A	1	0	2	none	1	$c_2^A \leq 0$
Scenario B	1	2	2	none	1	$c_2^B, c_3^B \leq 0$
Scenario C	1	ϵ	ϵ	$\{2,3\}$	0	$c_3^C + c_3^C \leq -1$
Scenario D	δ	0	0	$\{2,3\}$	0	$c_2^D + c_3^D \leq -1, c_1^D \leq 0$

In the table above, we assume that agents have a positive value for one of the outcomes and value of zero for the other one, for example in Scenario A we have $v_1(0) = 1, v_1(1) = 0$.

In Scenarios A and B, the efficient allocation is $x^* = 1$. Via individual rationality applied to Scenario A, we get that $c_2^A \leq 0$. On the other hand, by applying the taxation principle (Lemma 5.1) in Scenario B, we have $c_2^B \leq 0$ (since the implemented outcome $x^* = 1$ is the same as in Scenario A); applying a symmetric argument for agent 3 instead of 2, we get $c_3^B \leq 0$.

In Scenario C, for any $\epsilon < 0.5$, we have $x^* = 0$. Since agents 2 and 3 can deviate by playing as they do in scenario B, the mechanism has to compensate them with a positive transfer of at least 2ϵ to disincentive such deviation. By the taxation principle we have $c_2^C + c_3^C \leq -1$.

Finally, consider Scenario D. By efficiency, $x^* = 0$, and by taxation principle and IR we have $c_1^D \leq 0$ (since allocation stays the same for any $\delta > 0$). On the other hand, by taxation principle applied to Scenario C, the payments of Agents 2 and 3 must satisfy $c_2^D + c_3^D \leq -1$. Thus the revenue in this instance is negative. □

Proof of Theorem 5.6. Henceforth in the proof, we assume we are given a mechanism M which is coalitionally rational, and approximately weak collusion-resilient, i.e., which is collusive-Nash compatible and guarantees an $\alpha = 0.5 + \beta$ multiplicative approximation of the optimal welfare, for some given $\beta \in (0, 0.5]$. For ease of presentation, we make an additional assumption that the mechanism M is anonymous, i.e., if an action profile A is permuted, the resulting outcome and payments stay the same. At the end of the proof we describe how the proof extends for the case when the mechanism is not anonymous.

Constructing a set of instances: We consider a collective decision making setting with $4n$ agents, and consider m instances, denoted as $\{\mathcal{I}^{(m)} : m = 1, \cdots, n\}$. In each instance $\mathcal{I}^{(m)}$, the agents $i \in \{1, \cdots, 2n\}$ form a single coalition, whereas each agent $i \in \{2n+1, \cdots, 4n\}$ is independent. The instance $\mathcal{I}^{(m)}$ is given by the following valuation profile.

$$(v_i^{(m)}(0), v_i^{(m)}(1)) = \begin{cases} \left(\frac{1}{2n}, 0\right) & \text{if } i \leq 2n & \text{(a single coalition } S = \{1, \dots, 2n\}) \\ \left(0, \frac{1}{2n}\frac{2(1-\beta)}{\beta}\right) & \text{if } 2n < i \leq 3n + m & \text{(individual agents)} \\ (0, 0) & \text{if } 3n + m < i \leq 4n & \text{(individual agents)} \end{cases}$$

In the instance $\mathcal{I}^{(m)}$, when each individual agent and the coalition play their equilibrium strategy, let $x^*_{(m)} \in [0, 1]$ denote the allocation implemented by M.

Since the agent $i = 3n + m$ can always simulate the actions of an agent i with valuation 0, we require her payoff from following the equilibrium strategy to be (weakly) higher. Note that if the agent $i = 2n + m$ acts as if having valuation 0, the allocation is $x^*_{(m-1)}$ and her payment is at most 0 (since M is IR). This implies that the payment of the agent $i = 3n + m$ is bounded by $c_i \leq v_i(1)\Delta x^*_{(m)} = v_i(1)\left(x^*_{(m)} - x^*_{(m-1)}\right)$. Due to anonymity, this bound also applies to any agent $i \in \{2n+1, \cdots 3n+m\}$.

Picking a low revenue instance: Next, we argue that there exists an instance for which the total payment of the agents $\{2n+1, 2n+2 \ldots, 4n\}$ is at most $O(1/n)$. To see this, note that there may be two possible scenarios:

Case 1. There exists an $\tilde{m}$ s.t. $\Delta x^*_{(\tilde{m})} < 0$. In this case we have $c_i \leq v_i \Delta x^*_{(\tilde{m})} < 0$ and $\sum_{i=2n+1}^{4n} c_i < 0$.

Case 2. The increments $\Delta x^*_{(m)}$ are all non-negative. In this case we have:

$$\sum_{m=1}^{n} \Delta x^*_{(m)} \leq 1,\ \Delta x^*_{(m)} \geq 0 \implies \text{there exists } \tilde{m} \in \{0, \ldots, n\}\ s.t.\ \Delta x^*_{(m)} \leq 1/n.$$

Here the first inequality follows from the fact that $x \in [0,1]$ and that $x^*_{(m)} - x^*_{(1)} = \sum_{m=1}^{n} \Delta x^*_{(m)}$. Now we have the following upper bound for the payment derived from the 'winning' individual agents $\sum_{i=2n+1}^{4n} c_i \leq \sum_{i=2n+1}^{4m} v_i(1)\Delta x^*_{(\tilde{m})} \leq \frac{2(1-\beta)}{\beta n}$.

Analyzing the deviation of a coalition: First, note that for all instances in $\mathcal{I}^{(m)}$, $m = 1 \ldots n$, by our approximation guarantee that M achieves at least a $0.5 + \beta$ fraction of the optimal welfare, we can obtain that $x^*_{(m)} \geq \frac{1}{2} + \frac{\beta}{2}$, for any $\beta \in (0, 0.5]$. Now consider the specific instance $\tilde{m}$ derived above. Notice that the coalition $S = \{1\ \ldots, 2n\}$ has $v_S(0) = 1$, and hence under equilibrium play, obtains a utility of $U_S = (1 - x^*_{(\tilde{m})}) - c_S$ with $x^*_{\tilde{m}} \geq \frac{1}{2} + \frac{\beta}{2}$.

Now consider the following sequence of deviations $\mathcal{D}(k)$ for this coalition: every agent $i \in S$ pretends to be an independent agent with value

$$(v_i^{(k)}(0), v_i^{(k)}(1)) = \begin{cases} \left(\frac{1}{2n}\left(\frac{2(1-\beta)}{\beta}\right)^2, 0\right) & if\ i \leq n+k \\ (0,0) & if n+k < i \leq 2n \end{cases}$$

Repeating the argument from the first part of the proof, we get that there exists $\tilde{k}$ s.t. the corresponding deviation leads to allocation and payments $x^{(\tilde{k})}_{(\tilde{m})}, c^{(\tilde{k})}$ that satisfy (i) total payment of the coalition is $c_S^{(\tilde{k})} \leq \left(\frac{2(1-\beta)}{\beta}\right)^2/n$, and (ii) the allocation satisfies $x^{(\tilde{k})}_{(\tilde{m})} \leq \frac{1}{2} - \frac{\beta}{2}$.

Thus, in order for this deviation to not be better than the outcome of the dominant strategy, the coalition should get a compensation of at least $c_S \leq -\beta + O(1/n)$. Since the revenue from the individual agents is $O(1/n)$, this means the mechanism yields a negative revenue for the constructed instance.

Circumventing the requirement of anonymity. Recall that in our proof we make use the assumption of anonymity. We now show how to modify the argument as to avoid using this assumption. The only place where we use anonymity is guaranteeing the uniform payment bound $c_i \leq v_i(1)\Delta x^*_{(m)}\ \forall i \in [2m+1, 3n+m]$ (and the analogous bound in constructing a picking a deviation from the constructed sequence).

In absence of anonymity, we can preserve this bound by reordering agents so that in every instance of the sequence the last agent has the largest influence on the outcome. Formally, fix instance m and let $x^*_{(m,-k)}$ represent the outcome when agent k acts as if having value 0. Then for payment of agent k we have a bound $c_k \leq v_k(x_{(m)})^* - x^*_{(m,-k)})$.

We reorder agents so that $\Delta x^*_{(m)} \geq (x_{(m)})^* - x^*_{(m,-k)})$ for all $k = [2n+1, 3n+m]$ and all $m = [1, n]$. This reordering allows us to preserve the uniform payment bound for all agents $i = [2n+1, 2n+m]$

and with it the rest of the argument remains the same as it does not use anonymity anywhere else. □

B CONNECTIONS TO SYBIL PROOFNESS

The notion of Sybil proofness, i.e., invulnaribility of a mechanism to agents who have access to multiple identities, has been investigated previously for settings of combinatorial auctions [Yokoo et al., 2004] and social choice [Wagman and Conitzer, 2014]. This property is closely related to that of collusion resilience, and in this section we investigate the connection formally.

Although the scenario in which coalitions maximize their total utility is similar to that in which individual agents have access to multiple identities, there are some important differences. In order to describe them and establish connections between the two properties, we start with formally defining Sybil proofness (sometimes called false-name-proofness), borrowing from [Yokoo et al., 2004] (Note: this is a condensed definition; for more rigorous treatment see [Yokoo et al., 2004]).

There is a set $\mathcal{N}$ of agents with types $\theta \in \Theta$, set $\mathcal{M}$, $|\mathcal{M}| = m$, of identities, and a mapping $\phi : \mathcal{N} \to 2^M$, so that $\phi(i)$ is the set of identities that belong to agent i, unknown to the designer and exclusive, $\forall i \neq j \; \phi(i) \cap \phi(j) = \emptyset$. Given a direct revelation mechanism, an agent can report a type on behalf of each of their identities $a_i \in \Theta^{|\phi(i)|}$. We also let θ' be the vector of types reported by identities, $x(\theta')$ and $c(\theta')$ respectivelty the allocation made by the mechanism and the payments charged to the identities under reports θ'.

In the setting of combinatorial auctions let $\mathcal{L}$ be the set of available items , and $x_k(\theta') \in \mathcal{L}$ be the set of items won by identity k under type reports θ', and let $c_k(\theta')$ be the payment of identity k, then utility of any agent i is

$$u_i = v_i \left(\theta_i, \cup_{k \in \phi(i)} x_k(\theta')\right) - \sum_{k \in \phi(i)} c_k(\theta') \tag{7}$$

In the collective decision making setting (Section 5.2), the utility of agent i is given by $u_i(\theta_i, \theta') = v_i x(\theta') - \sum_{j \in \phi(i)} c_j(\theta')$. We can now define the property of Sybil proofness.

Definition B.1. A direct revelation mechanism M is Sybil proof if reporting $(t_i, 0, \ldots, 0)$ (i.e., truthful reporting using a single identity) is a dominant strategy for any agent i under any partition of identities ϕ.

B.1 Comparing Sybil Proofness and Collusion Resilience

The first difference that one notices when comparing the notion of CDS (Eqn. (2.3)) with the definition above is that Sybil-proofness assumes a direct revelation mechanism, while, crucially, collusion resilience does not. However, the definition of Sybil proofness can be relaxed as to resemble Definition 2.3 without losing its essence:

Definition B.2. A mechanism M is Sybil dominant strategy compatible (in short, SDS) if any agent i with access to multiple identities $\phi(i)$ has a dominant strategy.

In the setting of social choice, it is straightforward to see that the notions of SDS and CDS coincide and thus a collusion resilient mechanism is also an efficient SDS mechanism, and vice versa. On the other hand, observe that Sybil proofness is a stronger requirement than collusion resilience. In light of this, our impossibility result, Theorem 5.6, immediately implies that there exists no sybilproof mechanism that guarantees non-trivial approximation to welfare.

The second difference we think is more substantial and has to do with the setting of combinatorial auctions. When an agent has access to multiple identities, at the end of the auction she collects the items won by all of her identities and derives the corresponding value (see (7)); her utiity is independent of which identity won which item. In contrast, under our collusion model the value derived by a coalition does depend on how the mechanism allocated the items within the coalition.

This difference, we argue, is not superficial. In many practical settings the allocations themselves are more transparent than monetary exchanges. Such is the case with e.g. allocating licenses or real estate. Employing an efficient Sybil Proof mechanism in the presence of coalitions can still lead to a non-efficient outcome, but employing a CR mechanism guarantees efficiency. This discrepancy is not just potential: for example, VCG mechanism for unit-demand setting is Sybil-proof [Yokoo et al., 2004], but is not collusion resilient.

B.2 Compatibility theorems

The aim of this section is to establish a formal connection between Sybil proof and collusion resilient mechanisms. For the reasons described above, we here concentrate on combinatorial auctions to show two complementary results. First, we show that any collusion resilient efficient mechanism is also Sybil proof and efficient. Secondly, we construct a blackbox reduction for converting an efficient Sybil proof mechanism into a collusion resilient one.

Definition B.3. A type space $\Theta = \{\Theta_i\}_{i \in \mathcal{N}}$ is downward-closed if for every i, $\theta_i \in \Theta_i$, $\lambda > 0$ there is a type $\theta_i' \in \Theta_i$ s.t. $v_i(\theta_i', x) = \lambda v_i(\theta_i, x)$ for all $x \in \mathcal{X}$.

THEOREM B.4. *Assume a downward-closed type space Θ. Let M be a mechanism that satisfies coalitional rationality, NPT and collusion resilience. Then it is also a Sybil proof mechanism that satisfies NPT, IR and efficiency.*

PROOF. Notice that agent i with access to K identities $\phi(i)$ and value $v_i(\theta_i, x)$ is equivalent to a coalition of size K with scaled valuations $v_k(\theta_k', x) = v_i(\theta_i, x)/K,\ k \in \phi(i)$ when participating in M, i.e. both the strategy space and the utility for every possible outcome coincide. The result now follows from the definition of collusion resilience. □

For our next result we need to introduce an additional assumption of closure under addition.

Definition B.5. A type space Θ is closed under addition if for any types $\theta_1, \theta_2 \in \Theta$ there exists a type $\theta_3 \in \Theta$ s.t.

$$v(\theta_3, x) = \max\{v(\theta_1, x_1) + v(\theta_2, x_2) | x_1, x_2 \in x, x_1 \cap x_2 = \emptyset\} \quad \forall\, x \in \mathcal{L}$$

For example, submodular (in items) valuations are closed under addition, but unit-demand valuations are not.

In the following theorem we aim to establish a reverse link between the notions of collusion resilient and efficient Sybil proof mechanisms.

THEOREM B.6. *In the setting of combinatorial auctions with type space Θ closed under addition, let M be an IR, budget balanced, Sybil proof mechanism (in the sense of Definition B.1) that implements an efficient allocation rule $x^*(\theta', \mathcal{N})$. Then there is a mechanism M' that satisfies coalitional rationality, budget balance and collusion resilience.*

PROOF. **Description of M'.** The main idea behind M' is to, after running M, redistribute items optimally within coalitions while preserving the dominance of truthful revelation strategy. This is done via the procedure of handling inconsistent coalition reports that we borrow from CRM (see Section 4).

More formally, a strategy of agent i in M', similarly to CRM, is a tuple (S_i', θ_i), where S_i' is the report of agent's i coalition and θ_i' is the reported type. We adopt the terminology of Section 4 and essentially repeat the coalition-verification process of CRM (steps (2) and (3) of the mechanism).

More precisely, when the reports (S', θ') are made, M' checks if all agents are a part of consistent coalition. If they are not, and there is a single scapegoat T, allocation $W^*(t', T)$ is made and all kibitzers are charged $W^*(t', \mathcal{N})$; otherwise empty allocation is made and no payments are charged.

If all agents are a part of a consistent coalition, for every coalition S the mechanism M' finds a type $\theta'_S \in \Theta$ s.t.

$$v_{i_1^S}(\theta'_S, x) = \max\{\sum_{i \in S} v_i(\theta'_i, x_i) | \cup_{i \in S} x_i = x,\ x_i \cap x_j = \emptyset\} \quad \forall x \in \mathcal{L} \tag{8}$$

for some $i_1^S \in S$. This is possible because of the assumption that θ is closed under addition. M' then creates a new type profile $\hat{\theta}$, making a substitution for every consistent coalition S: $\hat{\theta}_{i_1^S} = \theta'_S, \hat{\theta}_i = 0$ for $i \in S \setminus \{i_1^S\}$. We also let θ_S stand for the type mapped to the true types of the coalition.

Mechanism M is then run with input type profile $\hat{\theta}$. Let we let $\hat{x}_M(\hat{\theta})$ be the allocation and $\hat{c}_M(\hat{\theta})$ the payments produced by M. M' then finds an allocation x^* that re-allocates items within coalitions optimally, i.e. $\sum(v_i(\theta'_i, x_i) = v_{i_1^S}(\theta_S, \cup_{i \in S} x_i)$ (the existence of such re-allocation follows from (8). It makes the allocation x^* and charges agents in accordance to $\hat{c}_M(\hat{t})$.

M' **is Collusion resilient**. First, because M implements efficient allocation and is IR, it follows that the payments it charges to participants are VCG payments with Clarke pivot rule, i.e, their externalities.

We consider the problem of picking an optimal strategy for some coalition S. First, by the argument identical to one used in the proof of Theorem 4.2 S can restrict it coalitional reports to consistent reporting within S, i.e. either truthfully reporting $S'_i = S\ \forall i \in S$ or splitting S into several consistent coalitions.

Thus, the problem of picking an optimal report for S can be restricted to the case when mechanism M is run and the items are redistributed within the reported coalitions. But then the utility of coalition S upper bounded with the utility of agent i_1^S participating in the mechanism M (this follows from the construction of the type profile $\hat{\theta}$ and the allocation rule x^*). In mechanism M, Sybil proofness implies that it is a dominant strategy for i_1 to report $(\theta_S, 0, ..., 0)$. It then follows that it is a dominant strategy for S to report $\{(S, \theta_i)\}_{i \in S}$ when participating in M'. □

Note that in our proof we only use the assumption of efficient allocation rule once, in ruling out the possibility to deviate by creatingg an internal kibitzer. One could alternatively assume a bounded valuation type space $\Theta : v_i(\theta_i, x) \leq M\ \forall t_i,\ x \in \mathcal{X}$, and modify the mechanism as to charge kibitzers $|\mathcal{N}|M$, that is the maximum possible total welfare.

Theorem B.6 allows us to view CRM, the mechanism we construct in Section 4, as our blackbox reduction applied to VCG mechanism. In light of this theorem, our impossibility result, 5.2, can now be viewed[10] as a considerable improvement over impossibility of [Yokoo et al., 2004], which states that Sybil proof efficient mechanisms do not exist for general combinatorial auctions.

[10] This uses the fact that submodular valuations are closed under addition as per Definition B.5