Shannon meets Myerson
Information extraction from a strategic sender

Vora, Anuj S.; Kulkarni, Ankur A.

DOI
10.1016/j.mathsocsci.2024.07.002

Publication date
2024

Document Version
Final published version

Published in
Mathematical Social Sciences

Mathematical Social Sciences 131 (2024) 48–66

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/mss

Shannon meets Myerson: Information extraction from a strategic sender ☆

Anuj S. Vora [a], Ankur A. Kulkarni [b,*]

[a] Delft University of Technology, 2628CD Delft, Netherlands

[b] Indian Institute of Technology Bombay, Mumbai 400076, India

ARTICLE INFO

Keywords:
Mechanism design
Information theory
Stackelberg game
Questionnaires
Screening

ABSTRACT

We study a setting where a receiver must design a questionnaire to recover a sequence of symbols known to a strategic sender, whose utility may not be incentive compatible. We allow the receiver the possibility of selecting the alternatives presented in the questionnaire, and thereby linking decisions across the components of the sequence. We show that, despite the strategic sender and the noise in the channel, the receiver can recover exponentially many sequences, but also that exponentially many sequences are unrecoverable even by the best strategy. We define the growth rate of the number of recovered sequences as the information extraction capacity. A generalization of the Shannon capacity, it characterizes the optimal amount of communication resources required while communicating with a strategic sender. We derive bounds leading to an exact evaluation of the information extraction capacity in many cases. Our results form the building blocks of a novel, non-cooperative regime of communication involving a strategic sender.

1. Introduction

Consider the following situation that arose during the early days of the Covid-19 pandemic. Travellers arrived at airports with varied travel histories and health inspectors had to screen these travellers based on responses to standardized questionnaires. Travellers arriving from unsafe locations were hesitant to reveal their true travel histories due to inconvenience of quarantine protocols and stigma associated with the disease, while those that arrived from safe locations wanted their true travel histories to be recorded. Some travellers had complex journeys where for some days they were at safe locations, while for other days they were at unsafe ones, and were perhaps inclined to selective misreporting. Detailed travel histories, in addition to an indication of susceptibility to infection, were useful for auxiliary studies such as contact tracing, inferring susceptibility of locations, stopping transport and more generally studying the spread of the infection. These studies evidently require the entire travel history of passengers and the health inspector's challenge was to design a questionnaire that recovered as many true travel histories as possible.

We model the above situation as follows. A *receiver* (health inspector) wishes to recover information privately known to a *sender* (traveller) over a possibly imperfect communication medium. The private information of the sender, i.e., its type, is a sequence of n symbols (locations) each drawn from a finite set of size q. Here n is the length of the travel history measured in some unit (say, "days") and comprises of locations travelled to on consecutive days. A questionnaire is characterized by a set of *alternatives* and a *selection mode*. Each alternative is drawn from the set of all possible sequences of symbols and the mode demands that the sender select any *one* alternative as its reported sequence. On viewing the sequence reported by the sender, the receiver applies a "decoding" function or an interpretation, mapping it to a *decoded* sequence. The sender is a non-cooperative agent and wishes to maximize the average *utility* $\frac{1}{n}\sum_i \mathcal{U}(x'_i, x_i)$ obtained when the true sequence $x = (x_1, \ldots, x_n)$ and the sequence $x' = (x'_1, \ldots, x'_n)$ is decoded by the receiver. This utility function dictates the sender's response to the questionnaire, and maximizing the utility may not align with the interests of the receiver.

☆ The results in this paper were presented in part the IEEE International Conference on Signal Processing and Communications (SPCOM) (Vora and Kulkarni, 2020d) held virtually in July 2020 and at the IEEE Conference on Decision and Control held virtually in December 2020 (Vora and Kulkarni, 2020c). The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This work was done when Anuj S. Vora was a doctoral student at the Systems and Control Engineering group. This research was supported by the grant CRG/2019/002975 of the Science and Engineering Research Board, Department of Science and Technology, India.

* Corresponding author.
E-mail addresses: a.vora@tudelft.nl (A.S. Vora), kulkarni.ankur@iitb.ac.in (A.A. Kulkarni).

https://doi.org/10.1016/j.mathsocsci.2024.07.002
Received 21 August 2023; Received in revised form 10 July 2024; Accepted 21 July 2024
Available online 25 July 2024

Depending on the realized sequence of symbols, for certain symbols, the sender may prefer that the receiver decodes the symbol incorrectly, whereas for other symbols it may want the receiver to know the truth. The receiver, on the other hand, is interested in maximizing the number of true sequences recovered correctly.[1] We ask the following question: given the sender's tendency to misreport its type and the possibly imperfect medium of communication, how must the receiver strategize to recover as much true information as possible? And what is the maximum *amount* of information that the receiver can extract from the sender? We call this the problem of *information extraction* from a *strategic sender*.

Notice that the receiver's options are rather limited. The receiver must publish a questionnaire before the sender's type is realized, and hence the questionnaire must be the same for all sender types. There is neither any scope for incentivizing truthful reporting by using transfers, nor do we assume incentive compatibility for the sender. How can the receiver then get any meaningful information from the sender? We allow the receiver the option to *eliminate* certain alternatives from the questionnaire. If there are q possible symbols, an exhaustive (and naive) questionnaire would have all q^n sequences as possible alternatives for the sender to choose from. Instead, we allow the receiver to select only of a subset of these sequences and publish only those sequences as available alternatives.[2] This effectively constrains the signal space of communication between the sender and the receiver, and becomes a key design tool for extracting non-trivial information from a strategic sender. Another tool with the receiver is the possibility of enforcing a linking of decisions (Jackson and Sonnenschein, 2007) across components of the sequence by suitably selecting the alternatives. The effect of this linking is that, not wanting to misreport one leg of the history forces the traveller to truthfully report other legs too; this again allows the receiver to extract more information than would be trivially possible. Indeed, we find that although we consider the n-day utility as a sum over the components of the sequence, the optimal questionnaire for n-length histories is *not* a stacking of n questionnaires of 1-length histories.

Another setting where the framework of information extraction using questionnaires is applicable arises in finance. Suppose the sender is an entity that has performed a sequence of n financial transactions. Each transaction is from a universe of q possible transactions, not all of which have verifiable records. An investigating officer (receiver) must ascertain the true sequence of transactions (say, to detect money laundering, or funding of nefarious activities), whereas the sender who has made these transactions may have an incentive to selectively misreport them. Our results in this paper provide not only strategies for the receiver, but also a structural understanding of the type space that can potentially be recovered. This understanding helps in setting up monitoring mechanisms that would discourage or prevent misreporting.

Our problem bears similarity to that of the implementation of a social choice function, as in Myerson's mechanism design setting (Myerson, 1997). Indeed, if modelled that way, the social choice function for our problem would become the identity function and incentive compatibility would require the sender to be truthful about its information. In this case, the sender's goal would coincide with the receiver's goal of recovering the truth, thereby reducing the problem to classical, non-strategic communication. Indeed in that extreme, our results do reduce to those from communication theory. Thus one may view our contribution as that of a *non-cooperative theory of communication*, generalizing Shannon's communication theory to a setting more akin to Myerson's mechanism design.

1.1. Main results

We formulate this problem as a leader–follower game with the receiver as the leader and the sender as the follower and analyse it based on the Stackelberg equilibrium. The receiver's strategy comprises of two parts — it must decide the sequences to be retained as alternatives in the questionnaire, and it must decide a mapping that "interprets" the response of the sender by mapping it to a decoded sequence. Our measure of information is the number of distinct sequences that the receiver recovers in a Stackelberg equilibrium. We find that there exists a Stackelberg equilibrium in which only this subset of the q^n sequences are retained as alternatives, and the receiver applies an identity decoding mapping on them.[3] The sequences are chosen in such a way that when *restricted* to the chosen subset, the sender's utility is incentive compatible (even though it may not be so on the full set of sequences), thereby each of the sequences in the questionnaire is recovered. This is reminiscent of the revelation principle (Myerson, 1997).

We define the limiting value (with increasing length of sequences, henceforth termed as blocklength) of the exponent of the size of this set of recovered sequences as the *information extraction capacity of the sender*. This capacity, denoted by $\Xi(\mathscr{U})$, with $\mathscr{U}$ being the single-lettered utility function of the sender, can be interpreted as the growth rate of the number of alternatives that are included in the optimal questionnaire. Equivalently, the capacity captures the rate at which the *sheet of paper* on which the questionnaire is printed – which is the communication medium between sender and receiver – must grow with increasing length of the history. It is thus a measure of the optimal amount of communication resources required while communicating with a strategic sender.

We find that $\Xi(\mathscr{U})$ is the limit of the nth root of the independence number of a certain sequence of graphs whose structure is determined by $\mathscr{U}$. The quantity $\Xi(\mathscr{U})$ bears intuitive similarity to the notion of the Shannon capacity of a graph (Shannon, 1956), a well-known and notoriously hard-to-compute quantity from communication theory. In fact, we show that the information extraction capacity *generalizes* the Shannon capacity of a graph. We show that the capacity lies between two important quantities,

$$\Gamma(\mathscr{U}) \le \Xi(\mathscr{U}) \le \Theta(G_{\mathsf{s}}^{\mathsf{Sym}}). \tag{1}$$

These bounds are independent of the blocklength n and are a function of only the utility. The lower bound $\Gamma(\mathscr{U})$ is the optimal value of an optimization problem defined over a set of permutation matrices on the alphabet $\mathcal{X}$. The upper bound is the Shannon capacity of a graph $G_{\mathsf{s}}^{\mathsf{Sym}}$ induced by the symmetric part of the utility $\mathscr{U}$.

Although this result may appear purely mathematical at first glance, it has a number of economic consequences. First, barring some corner cases, we have $\Gamma(\mathscr{U}) > 1$, whereby the receiver can recover an *exponential* number of sequences, regardless of any assumption of incentive compatibility on $\mathscr{U}$. Moreover, in general $\Gamma(\mathscr{U})$ is strictly greater than the size of the questionnaire for 1-length histories. Second, the upper bound shows that $\Xi(\mathscr{U})$ is, in general, strictly *less* than q, whereby the maximum number of sequences recovered is *exponentially smaller* than the total number of sequences. Third, for the utilities where the bounds match, the capacity is exactly characterized. Examples include cases where $\mathscr{U}$ symmetric and the corresponding $G_{\mathsf{s}}^{\mathsf{Sym}}$ is a perfect graph, and when the sender has "transitive" preferences. We discuss more

[1] As another example, recall the tale of a mischievous boy (sender) who observes a source that can be two possible states: *wolf* and *no-wolf*. The villagers (receiver) want to know if there is indeed a wolf. But the boy derives a utility $\mathscr{U}(x', x)$ when the true state is x and the villagers decode it to be x'. In the classical tale, when there is no wolf, the boy wants the villagers to think there is a wolf, i.e., $\mathscr{U}(wolf,\ no-wolf) > \mathscr{U}(no-wolf,\ no-wolf)$. But when there is a wolf, he wants them to infer that there is really a wolf, i.e., $\mathscr{U}(wolf,\ wolf) > \mathscr{U}(no-wolf,\ wolf)$.

[2] Our questionnaire requires the sender to select only one alternative.

[3] In other words, the receiver believes the sender, or takes the sender's responses at face value.

such examples later in the paper. A well-known semi-definite program based upper bound on the Shannon capacity of a graph is the Lovász theta function introduced by Lovász (1979). This, along with the lower bound, together provide two computable bounds that can approximate the capacity when it is not exactly characterized. We also derive a hierarchy of lower bounds as a function of the blocklength n by generalizing the bound $\Gamma(\mathscr{U})$, which approach the capacity asymptotically as n grows large. This allows one to approximate the capacity arbitrarily closely from the left-hand side, albeit with increasing computational burden.

When the channel is noisy, the rate of information extraction is given by $\min\{\Xi(\mathscr{U}), \Theta(G_c)\}$, where G_c is the confusability graph of the channel. This result shows that as long as the zero-error capacity of the channel is greater than $\Xi(\mathscr{U})$, the receiver can extract the maximum possible information from the sender, and otherwise it is limited by the channel capacity. This result is analogous to the setting of joint source-channel coding in information theory, where the capacity of the channel is required to be larger than the entropy of the source to ensure reliable communication (Cover and Thomas, 2012, Ch. 7).

A general lesson in these results is that there are fundamental limitations to the operation of systems with self-interested agents. Problems where agents are compromised also occur in *cyber–physical systems*, where a network of sensors connected over a communication medium is tasked with the operation of safety-critical applications such as nuclear power plants. A compromised sensor may be modelled as non-cooperative sender as in our setting. Our results show on the one hand how a receiver may strategize to obtain information from such agents, and on the other that there will usually be blind spots in the knowledge of the receiver, regardless of how it strategizes. Finally, when compromised sensors are present, improving the communication bandwidth only has a limited effect (it increases $\Theta(G_c)$ above); it may help the receiver hear the sender more clearly, but not recover more truth.

1.2. Economic insights

We new present a few economic insights that emerge from our analysis.

The key idea in the Stackelberg equilibrium strategy of the receiver is to limit the options of the sender for lying by selectively decoding only a subset of the sequences. If the receiver attempts to correctly recover a larger number of sequences by including more alternatives in the questionnaire, then the sender gets greater freedom to lie about its information, and the attempt of the receiver is counter-productive. On the other hand, if the receiver chooses too few sequences to include in the questionnaire, then the sender is compelled to speak truth given the limited choices, but the number of sequences recovered is less than optimal. The information extraction capacity is the growth rate of the "optimal" set that balances these two aspects.

Since an exhaustive questionnaire is prone to misreporting by a strategic sender, questionnaires for strategic respondents must involve some coarsening of options. This may appear unfair to senders whose type has been eliminated and are now compelled to pick a type that is not their true type. But such elimination is in the "greater good". With this strategy, a receiver acts as an *enlightened totalitarian* – by constraining what can be said, the receiver ensures that more truth is spoken. Moreover, as $n \to \infty$, the number of options retained is in general a vanishing fraction of q^n.

Questionnaires with coarse and rigid options often appear in automated assistance systems, like customer care helplines, and also in immigration forms. Our results show that such coarseness is not necessarily due to poor design or constraints of the medium (such as the length of forms), but rather a strategic choice of the receiver to ensure better reporting by the sender.

Moreover, in the optimal strategy above, the receiver takes the sender's response at face value. Hence, with a well-designed questionnaire, there is no need for further "mind games" or interpretation to be attached to the sender's response. In response, all senders whose type is included in the questionnaire, find it optimal to report it truthfully. This is analogous to the revelation principle in mechanism design.

Our results show that exponentially more sequences can be extracted by a questionnaire that employs strategic linking. Moreover, under a special structure of $\mathscr{U}$, day-wise questionnaires are optimal and in this case there is no benefit to linking.

Finally, our definition of a questionnaire is one where the sender is required to "select any one" from a set of alternatives. On a theoretical front, more general questionnaires can be considered – e.g., those where the sender can "select all that apply", or those with a "none of the above" alternative. These are fascinating directions, though beyond our present scope.

1.3. Relation to cooperative information theory

In the non-cooperative setting we consider, finite blocklength leader–follower games take the place of finite blocklength coding problems from the cooperative setting, whereas Stackelberg equilibria take the place of codes. We find that $\Xi(\mathscr{U})$ plays a role loosely analogous to that played by the entropy of a source in cooperative communication; the utility $\mathscr{U}$ of the sender is akin to the probability mass function of the source. This analogy agrees with the rate of information extraction with a noisy channel, given by $\min\{\Xi(\mathscr{U}), \Theta(G_c)\}$. Having said that, this analogy is loose because we are not concerned with a stochastic setting in this paper (recall we work in the zero-error regime). A much more complicated setting would arise when considering vanishing probability of error; our preliminary work on this line can be found in Vora and Kulkarni (2020a).

Another related viewpoint from information theory is the notion of large blocklength analysis. Information theory shows that one can, in general, communicate more information on average by exploiting structure in *sequences*, rather than in individual symbols, an idea commonly referred to as *coding*. The linking of decisions we exploit in our work is the analogue of coding in our strategic setting. In information theory it is of interest to study how the codes scale with the blocklength and thereby help quantify the improvement in accuracy of communication with increase in channel capacity. Our notion of the information extraction capacity is an attempt at capture the similar requirements in a strategic setting.

When viewed from the communication standpoint, in the noiseless case, our channel input and output spaces are both equal to the space of source sequences. In the cooperative setting this would trivially lead to recovery of all the source sequences. However, the same does not hold in our setting since the receiver chooses to selectively decode only a portion of the outputs in its optimal strategy. Thus, our results also quantify the *optimal* amount of channel resources that are required for the receiver to extract information from the sender. In a sense, this marks a shift from the communication-theoretic concept of the *capacity* of a channel to that of *capacity utilization*, as something more relevant for the non-cooperative setting.

Another subtlety here is that a strategic sender is distinct from a *defunct* sender. A defunct sender sends arbitrarily corrupted messages, whereas a strategic sender's messages being motivated by its utility have an underlying structure. The optimal questionnaire exploits this structure to obtain nontrivial information from the sender.

1.4. Related work

There have been works on strategic communication (of various flavours) in the game theory community, but to the best of our knowledge ours is the first formal information-theoretic analysis of information extraction. The first model of strategic communication was introduced by Crawford and Sobel (1982). They considered a sender and a receiver with misaligned objectives and formulated a simultaneous move game between the sender and receiver. They showed that any equilibrium involves the sender resorting to a *quantization* strategy, where the sender reports only the interval in which its information lies. Numerous variants and generalizations were subsequently studied that include the case of multi-dimensional sources (Battaglini, 2002), noisy channel medium (Sarıtaş et al., 2015), multi-sender setting with restricted state spaces (Ambrus and Takahashi, 2008), hierarchical senders (Ambrus et al., 2013). These works consider a neutral perspective between the sender and the receiver and study the Nash equilibrium strategies. The primary objective is to determine the conditions under which the equilibrium is quantized and conditions under which a fully informative equilibrium exists.

Strategic communication in control theory has been studied by Farokhi et al. (2016) and Sayin et al. (2019). Farokhi et al. (2016) studied a problem of static and dynamic estimation in the presence of strategic sensors as a game between the sensors and a receiver and characterized a class of equilibria of the game. Sayin et al. (2019) considered a dynamic signalling game between a strategic sender and a receiver. They showed that the players use linear signalling rules in equilibrium. Strategic communication has been studied from the perspective of information theory by Akyol et al. (2015, 2016) where they studied a sender-receiver game and characterized equilibria satisfying a certain rate and distortion levels. Akyol et al. (2016) also analysed the effect of side information at the receiver.

In the information design framework, there is a huge literature on the Bayesian persuasion problem first introduced by Kamenica and Gentzkow (2011). In this setting, the sender with superior information tries to influence the actions of the receiver (Bergemann and Morris, 2019). The objective in these settings is to determine conditions under which persuasion is beneficial to the sender and characterizing the equilibrium strategies and the payoff. Information-theoretic analysis of Bayesian persuasion problem was studied by Le Treust and Tomala (2019) where they studied a setting where the information of the sender is a sequence of states and it *persuades* the receiver via a noisy channel. They derived an upper bound on the payoff achieved by the sender as function of the capacity of the channel and showed that this bound can be achieved in the limit of large state sequences.

The novelty of our work and the differences with the above models are highlighted in the following points.

- The studies mentioned above consider a neutral perspective between the sender and the receiver (Battaglini, 2002; Sarıtaş et al., 2015; Ambrus and Takahashi, 2008; Ambrus et al., 2013) or consider the perspective of the sender (Farokhi et al., 2016; Sayin et al., 2019; Akyol et al., 2015, 2016; Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019; Le Treust and Tomala, 2019). In this paper, we study the problem from the perspective of the receiver and hence we formulate the game with the receiver as the leader. This perspective is scarcely studied in the literature barring a few works that we discuss ahead in this section.
- The works on cheap talk literature as well as the strategic estimation and signalling problems consider continuous state space and quadratic utilities, where the differences in the utilities of the sender and receiver are quantified by a bias. In our paper, we consider discrete state spaces with a utility for the sender that can differ arbitrarily from the objective of the receiver.
- In our setting, the information of the sender is a sequence of states, unlike the one-shot setting studied in the cheap talk literature or the Bayesian persuasion literature. Moreover, the states observed by the sender are arbitrarily generated by a probability distribution, unlike, for instance, Farokhi et al. (2016), Sayin et al. (2019) where the states are generated according to a linear dynamical system. In this manner, our work bears similarities to Le Treust and Tomala (2019) with respect to the state information observed by the sender.
- The problem of designing questionnaires for every length of travel histories boils to determining a series of static mechanisms to be designed by the receiver. This is unlike the dynamic setting in Farokhi et al. (2016), Sayin et al. (2019) where equilibrium strategies may be a function of past information and estimation and signalling strategies.
- We present a novel concept of information extraction capacity based on information-theoretic notions. This capacity characterizes the maximum "amount of information" that can be recovered from the sender. We also present a novel economic application of designing questionnaires based on this framework of information extraction.

As mentioned earlier, our setting can be viewed as a problem of implementing a social choice function that may not be incentive compatible with the preferences of the agents. A related setting is studied by Jackson and Sonnenschein (2007) where they demonstrated that these incentive constraints can be overcome by *linking* independent copies of the decision making problem. They devised a mechanism and showed that as the number of linkages grow large, the mechanism implements the social choice function *asymptotically*, i.e., the probability of the decisions on which the function cannot be implemented tends to zero. There are certain parallels between our setting and the setting considered by Jackson and Sonnenschein (2007). For instance, the linking in Jackson and Sonnenschein (2007) is akin to the block structure of our setting and the implementation of the function is analogous to information recovery by the receiver. Viewed in this manner, our inquiry can be stated as follows — how many decisions can be implemented by the social choice function exactly? Further, how does this number grow with the length of sequence of signals? Our setting is thus a *zero-error* counterpart of the implementation problem.

Our work is also related to the work of Renou and Tomala (2015) who studied the problem of approximate implementation of social choice functions where the information observed by the senders is generated according to a Markov process. They considered a setting with multiple senders and characterized the functions that can be implemented approximately in the limit of large state sequences. In contrast to our work, Renou and Tomala (2015) consider a dynamic setting where the mechanism at any time may depend on the entire past history of mechanisms, messages and the decisions of receiver. Also, in our paper, we assume that the state information (travel histories) of the sender is generated arbitrarily according to some distribution.

In the context of information theory, our problem relates to the problem of coding in presence of mismatched criteria that has been studied extensively (see Scarlett et al. (2020) for a survey). The mismatch is in the encoding and decoding criteria and is to model the inaccurate or asymmetric information about the channel or to incorporate constraints on encoding or decoding. The optimal functions of the encoder and decoder in such cases are therefore chosen only in anticipation of either a fixed encoder or decoder. Thus, they do not capture the "active" deceptive nature of the problem. The problem thematically closest to our setting is the mismatched distortion problem studied by Lapidoth (1997). In this problem, the distortion criteria of the receiver and the sender are mismatched, and the receiver aims to construct a codebook such that its own distortion is minimized. The author determines an upper bound on the distortion for a given rate of communication. A crucial difference between the setting of Lapidoth (1997) and our setting is that in the former, the objective of the sender does not depend on the sequence decoded by the receiver. This fails to capture the strategic nature of our problem where the sender is indeed affected by the actions of the receiver is therefore trying to influence the outcome by misreporting.

Our work significantly extends the results of Vora and Kulkarni (2021, 2020d). In Vora and Kulkarni (2021), we considered a situation where a health inspector designs a questionnaire to screen travellers. In this work, in contrast to the single kind of travellers, we studied a setting where the inspector encounters travellers with varying degrees of truth-telling nature and this information is only known in the form of a probabilistic belief to the inspector. We characterized the optimal questionnaires and derived preliminary bounds on the rate of information extraction. In Vora and Kulkarni (2020d) we discussed a special case of sender's utility where all the deviations from the truth contribute in equal magnitude to the benefit or loss of the sender. We showed that the information extraction capacity is bounded above by the Shannon capacity of the sender graph. This corresponds to Proposition 4.6 in this paper where the positive and negative parts of the utility are equal. We studied related strategic communication problems in Vora and Kulkarni (2020a,b) where the receiver, unlike the exact recovery in this paper, tried to achieve asymptotically vanishing probability of error. In Vora and Kulkarni (2020a), we considered a setting where the receiver allowed a certain level of distortion and devised "achievable" strategies such that the probability of error was arbitrarily low. We studied the case of binary alphabet and determined the information-theoretic rate of these achievable strategies. In Vora and Kulkarni (2020b), we considered a setting where the receiver wished to compute a function of the source. We determined sufficient and necessary conditions for reliable communication, where the incentive compatibility of the function was a sufficient condition. We also showed that when the function was incentive compatible the Stackelberg equilibrium of the game corresponds to the information-theoretic source code.

The paper is organized as follows. We formulate the problem in Section 2. In Section 3, we determine the equilibrium of the Stackelberg game with the noiseless channel. In Section 4, we define the information extraction capacity and in Sections 4.2 and 4.3, we derive lower bounds and upper bounds respectively on the capacity. Finally, we analyse the Stackelberg game with the noisy channel in Section 5. Section 6 concludes the paper.

2. Problem formulation

2.1. Notation

Random variables are denoted with upper case letters X, Y, Z and their instances are denoted as lower case letters x, y, z. Matrices are also denoted by uppercase letters. The space of scalar random variables is denoted by the calligraphic letters such as $\mathcal{X}$ and the space of n-length vector random variables is denoted as $\mathcal{X}^n$. To unclutter notation, vector random variables X, Y, Z and their instances x, y, z will be denoted without the superscript n. The set of probability distributions on a space '$\cdot$' is denoted as $\mathcal{P}(\cdot)$. The set of all conditional distributions on a set $\mathcal{Y}$ given a sequence from $\mathcal{X}$ is denoted as $\mathcal{P}(\mathcal{Y}|\mathcal{X})$. The empirical distribution of a sequence $x \in \mathcal{X}^n$ is denoted as P_x and is defined as $P_x(i) = |\{k : x_k = i\}|/n$. The joint empirical distribution of sequences (x, y) is defined similarly and is denoted as $P_{x,y}$. A graph G is denoted as $G = (V, E)$ where V is the set of vertices and E is the set of edges. When two vertices $x, y \in V$ are adjacent, we denote it either as $(x, y) \in E$ or as $x \sim y$. An independent set in G is a subset S of V such that no two vertices in S are adjacent. For a graph G, the size of the largest independent set is denoted as $\alpha(G)$. For a function or a random variable, we denote $\mathrm{supp}(\cdot)$ as its support set. For an optimization problem '$\cdot$', we denote $\mathrm{OPT}(\cdot)$ as its optimal value. Unless specified, the exp and log are with respect to the base 2. For any $n \in \mathbb{N}$, we denote $\{1, \dots, n\}$ by $[n]$.

2.2. Model

We present a model where the medium of communication is noiseless; it will be generalized later to allow for noisy communication. Let the alphabet be $\mathcal{X} = \{0, 1, \dots, q-1\}$, where $q \in \mathbb{N}$ is the alphabet size. The sender observes a sequence $X = (X_1, \dots, X_n) \in \mathcal{X}^n$, where X_i are drawn from a known distribution.[4] The sender sends a message $s_n(X) = Y \in \mathcal{X}^n$, where $s_n : \mathcal{X}^n \to \mathcal{X}^n$. The message is relayed perfectly to the receiver who decodes the message as $g_n(Y) = \hat{X}$, where $g_n : \mathcal{X}^n \to \mathcal{X}^n \cup \{\Delta\}$. Here Δ is an error symbol we introduce for convenience; we explain its meaning subsequently. Let

$$\mathcal{D}(g_n, s_n) := \left\{x \in \mathcal{X}^n \mid g_n \circ s_n(x) = x\right\} \tag{2}$$

be the set of recovered sequences when the receiver plays the strategy g_n and the sender plays the strategy s_n. We also refer to the set of recovered sequences $\mathcal{D}(g_n, s_n)$ as the set of sequences *correctly* decoded by the receiver.

We assume that the n-block utility function $\mathcal{U}_n$ takes the following form,

$$\mathcal{U}_n(\hat{x}, x) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{U}(\hat{x}_i, x_i), \tag{3}$$

where $x = (x_1, \dots, x_n) \in \mathcal{X}^n$ is the sender's type and $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n) \in \mathcal{X}^n$ is the sequence of symbols decoded by receiver, and $\mathcal{U} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the single-letter utility of the sender.

The receiver aims to maximize the size of the set $\mathcal{D}(g_n, s_n)$ by choosing an appropriate strategy g_n. The sender, on the other hand, maximizes $\mathcal{U}_n(g_n \circ s_n(x), x)$ for each x by choosing an appropriate strategy s_n. Further, we assume $\mathcal{U}_n(\Delta, x) = -\infty$ for all $x \in \mathcal{X}^n$ and $n \in \mathbb{N}$. Thus Δ is an outcome that is never preferred by the sender. We also assume that $\mathcal{U}(i, i) = 0$ for all $i \in \mathcal{X}$; this is without loss of generality as we explain below after Definition 2.1.

The operational meaning of the above model in the context of the questionnaire design is as follows. The alphabet $\mathcal{X}$ is the set of locations and n is the length of the history. The message $Y = s_n(X)$ is the alternative selected by a traveller with history $X \in \mathcal{X}^n$ and $g_n(Y)$ is the interpretation applied by the health inspector to the alternative Y. Since Δ is an outcome never preferred by the sender, if an alternative Y is mapped to Δ, it is equivalent to the alternative Y being not present in the questionnaire. Thus $C_n := \{y \in \mathcal{X}^n \mid g_n(y) \neq \Delta\}$ is the set of alternatives presented to the travellers and a traveller has to choose exactly one of the alternatives from this list. Thus if x is the true travel history of the traveller, then selecting an alternative $y \in C_n$ amounts to setting $s_n(x) = y$. The inspector then, is said to decode the response of the traveller to a travel history $\hat{x} \in \mathcal{X}^n$ if $\hat{x} = g_n(y)$. The inspector recovers the travel history x if and only if $g_n \circ s_n(x) = x$.

We formulate this problem as a leader–follower game with the receiver as the leader and the sender as the follower (Başar and Olsder, 1999).

Definition 2.1 (*Stackelberg Equilibrium*). The Stackelberg equilibrium strategy of the receiver is given as

$$g_n^* \in \arg\max_{g_n} \min_{s_n \in \mathcal{B}(g_n)} |\mathcal{D}(g_n, s_n)|, \tag{4}$$

where the set of best responses of the sender, $\mathcal{B}(g_n)$, is given as

$$\mathcal{B}(g_n) = \Big\{ s_n : \mathcal{X}^n \to \mathcal{X}^n \mid \mathcal{U}_n(g_n \circ s_n(x), x) \geq \mathcal{U}_n(g_n \circ s_n'(x), x) \quad \forall\, x \in \mathcal{X}^n, \forall\, s_n' \Big\}. \tag{5}$$

Any strategy $s_n^* \in \mathcal{B}(g_n^*)$ is said to be a Stackelberg equilibrium strategy of the sender and the pair (g_n^*, s_n^*) is said to be a Stackelberg equilibrium.

It is easy to see that the set of best responses $\mathcal{B}(g_n)$ is the same if $\mathcal{U}_n(x, x)$ is subtracted on both sides of the inequality in (5). Thus, without loss of generality, we assume $\mathcal{U}(i, i) = 0$ for all $i \in \mathcal{X}$.

In (4), we minimize over the set $\mathcal{B}(g_n)$ of the best responses of the sender because the sender may have multiple best responses and the receiver does not have control over the choice of the sender's specific best response strategy. We assume that the receiver chooses its strategy according to the worst-case over all such best responses and hence adopts a *pessimistic* viewpoint. This is also the formulation of Stackelberg equilibrium adopted in standard sources such as Başar and Olsder (1999).

[4] In the context of travellers, the term "observes" implies that the sender *arrives* with a sequence of travel locations.

Note that the problem of information extraction we study is distinct from the problem of *information design* (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019) or *information disclosure* (Akyol et al., 2016). In these cases, the problem is studied from the perspective of the sender who designs the information that is observed by the receiver so as to achieve an outcome that favours the sender. Thereby, in such a setting it is suitable to formulate the problem with the sender as the leader. On the other hand, in the case of screening of travellers, it is the role of the receiver to "ask" the agents about their information. Therefore, it is apt to study the problem with the receiver as the leader of the game.

2.3. An example

The following example illustrates some important aspects of the setting. Suppose $n = 1$ and let $\mathcal{X} = \{0, 1, 2, 3\}$ be the set of locations. Let the utility of the sender $\mathcal{U} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be given as

$$\mathcal{U} = \begin{array}{cc} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & -1 & -2 & -2 \\ 0.5 & 0 & -1 & -2 \\ -1 & 0 & 0 & -1 \\ -1 & -1 & 0.5 & 0 \end{pmatrix} \end{array}$$

Here $\mathcal{U}(i, j)$, the entry in the i^{th} row and j^{th} column, denotes the utility obtained by the sender when i is the location decoded by the receiver and j is the true location. From column 0, we can see that $\mathcal{U}(0, 0) < \mathcal{U}(1, 0)$ and $\mathcal{U}(0, 0) > \mathcal{U}(2, 0) = \mathcal{U}(3, 0)$. Hence, the sender prefers that, when the true location is 0, the receiver decodes it to be 1, but not to 2 or 3. Similar observations for column 1 shows that the sender equally prefers 1 and 2 when the true location is 1; column 2 shows that the sender prefers the location 3 when the true location is 2; the last column shows that the sender prefers 3 when the true location is 3.

Suppose the receiver declares a naive questionnaire $C = \{0, 1, 2, 3\}$ with all possible locations and announces to decode it with an identity function; i.e., the receiver's decoded location is equal to the location reported by the sender. For a true location x, suppose the sender chooses an alternative $s(x)$ from C, where $s : \mathcal{X} \to C$ is the response of the sender. A location x is recovered by the receiver if $x = s(x)$, and the set of locations recovered is $\{x \in C : s(x) = x\}$. It is easy to see that the sender has two best responses s_1, s_2, where

$$s_1(x) = \begin{cases} 1 & x = 0 \\ 1 & x = 1 \\ 3 & x = 2 \\ 3 & x = 3 \end{cases}, \quad s_2(x) = \begin{cases} 1 & x = 0 \\ 2 & x = 1 \\ 3 & x = 2 \\ 3 & x = 3 \end{cases}.$$

We have $\{x \in C : s_1(x) = x\} = \{1, 3\}$ and $\{x \in C : s_2(x) = x\} = \{3\}$. Thus, in the worst case, only one location is recovered by the receiver.

2.3.1. Elimination of alternatives improves information recovery

However, the receiver can recover more than one location by cleverly choosing its questionnaire. The main difficulty encountered above is that when all locations are included as alternatives, it creates room for lying, whereby less truth is recovered. To counter this, suppose the receiver chooses a questionnaire $\widetilde{C} = \{0, 2\}$, and an identity decoding function on $\widetilde{C}$. The (unique) best response of the sender is now a strategy $\widetilde{s}$, where

$$\widetilde{s}(i) = \begin{cases} 0 & x = 0 \\ 2 & x = 1 \\ 2 & x = 2 \\ 2 & x = 3 \end{cases}.$$

The set of recovered locations is $\{x \in \widetilde{C} : \widetilde{s}(x) = x\} = \{0, 2\}$, thereby the receiver has recovered more truth. In $\widetilde{C}$, the sender has only 0 and 2 as alternatives and 1 and 3 are left out. Given these choices, the sender is forced to report 0 truthfully, unlike earlier where it was reporting 0 as 1 when 1 was available as an alternative. Similarly, it is forced to report 2 truthfully. We thus see that a more cunningly chosen questionnaire $\widetilde{C}$ improves on the naive questionnaire C by forcing the sender to be truthful for the locations 0 and 2.

Such a questionnaire clearly leaves blind spots for the receiver — there is no hope of recovering locations 1 and 3. But including 1 in $\widetilde{C}$ would jeopardize the recovery of 0, since the sender will report 0 as 1, and including 3 will preclude the recovery of 2, since the sender will report 2 as 3. We shall see later from the main results, that the receiver cannot recover more than two symbols in the worst case over the best responses of the sender. One may ask if the receiver can recover any more locations, in the worst case, by choosing a different decoding function. We show later that this is not possible, and it suffices to consider the identity decoding function. In other words, $\widetilde{C}$ is an optimal questionnaire.

2.3.2. Linking responses can increase recovered histories on average

Now let $n = 2$, i.e., the sender has 2-length travel histories. The questionnaire will thus be composed of sequences from $\mathcal{X}^2$. From the additive nature of the utility given in (3), it is easy to see (we also show this formally) that a questionnaire $\widetilde{C}^2 := \widetilde{C} \times \widetilde{C} = \{00, 02, 20, 22\}$ recovers all the travel histories in $\widetilde{C}^2$. Thus, we have that $|\widetilde{C}^2|^{1/2} = 2$ and the receiver recovers the same amount of information *per unit length of the history*[5] as in $\widetilde{C}$. However, can the receiver do better?

Suppose the receiver declares a questionnaire $\widehat{C} = \{00, 21, 02, 23, 30\}$. Let 00 be the true sequence and consider another sequence 21 from $\widehat{C}$. We have that

$$\mathcal{U}(21, 00) = \frac{1}{2}\left(\mathcal{U}(2, 0) + \mathcal{U}(1, 0)\right) = \frac{1}{2}(-1 + 0.5) < 0,$$

whereby the sender does not prefer to report 00 as 21. Although the sender prefers the location 1 over the true location 0, the sender has to trade-off this benefit with the loss derived by reporting the location 0 as 2. Since the penalty from the latter is more than the incentive derived by misreporting, the sender prefers to report 00 over the sequence 21. This can be repeated for all pairs of sequences in $\widehat{C}$ to show that whenever the sender's true sequence is one from $\widehat{C}$, it prefers to report it truthfully. Thus, the receiver can recover all the histories from $\widehat{C}$. Notice that $|\widehat{C}|^{1/2} = 5^{1/2} > 2$. In fact, this is the largest size of questionnaire for $n = 2$. In other words, a larger value of n opens up the possibility of creating linkages across legs of the journey whereby more information can be recovered than by mere stacking of the optimal questionnaires corresponding to 1-length histories.

From these short examples, a few observations are immediately evident. It is clear that questionnaires should not be designed innocently and the receiver must *strategize* in order to extract information from the sender. In general, the receiver may be able to extract only a subset of the information from the sender. Finally, the receiver can recover more information on average when the responses of the sender are linked. We formalize these observations in this paper.

3. Information extraction from the sender in equilibrium

In this section, we characterize the size of the largest set of sequences recovered by the receiver in a Stackelberg equilibrium and its growth rate with the blocklength n. In general for a fixed receiver strategy g_n, the size of the set of recovered sequences $|\mathcal{D}(g_n, s_n)|$ could vary as the best response s_n varies over $\mathcal{B}(g_n)$. We consider the smallest such size as our notion of the number of recovered sequences. To characterize the growth rate of this size, we define the *rate* of information extraction as follows.

5 Since the total number of travel histories for any n are $|\mathcal{X}|^n$, a natural way to compare the recovery of information from questionnaires across different lengths of travel histories is to look at the nth root of the size of the questionnaire.

Definition 3.1 (*Rate of Information Extraction for a Strategy*). The number of recovered sequences by a receiver strategy g_n is defined as

$$\min_{s_n \in \mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|.$$

For any strategy g_n of the receiver, the rate of information extraction is defined as

$$R(g_n) = \min_{s_n \in \mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|^{1/n}.$$

Notice that as n increases, the number of possible sender types also increases, and it does so exponentially. It is therefore important to track the growth-rate with n of various sets of interest. The sequence $\{R(g_n)\}_{n\geq 1}$ quantifies the growth rate of the set of recovered sequences for a sequence of strategies $\{g_n\}_{n\geq 1}$. Below we characterize the number of sequences recovered in a Stackelberg equilibrium (cf. Definition 2.1) in terms of a graph induced by the utility of the sender on the space of sequences $\mathcal{X}^n$, called the *sender graph*.

Definition 3.2 (*Sender Graph*). The sender graph, denoted as $G_s^n = (\mathcal{X}^n, E_s)$, is the graph where $(x, y) \in E_s$ if either

$$\mathscr{U}_n(y, x) \geq 0 \text{ or } \mathscr{U}_n(x, y) \geq 0.$$

For $n = 1$, the graph G_s^1 is denoted as G_s and referred to as the *base graph*.

Thus, two vertices x and y are adjacent in G_s^n if the sender has an incentive to report one sequence as the other.

Remark 3.1. Two single-letter utility functions inducing the same base graph on $\mathcal{X}$ can induce two different sequences of sender graphs on $\mathcal{X}^n$ for $n > 1$. Let $\mathcal{X} = \{0, 1, 2\}$ and consider the utilities $\mathscr{U}$ and $\mathscr{U}'$ defined as

$$\mathscr{U} = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}, \quad \mathscr{U}' = \begin{pmatrix} 0 & -2.5 & -2.5 \\ 1 & 0 & -1 \\ -1.5 & 0 & 0 \end{pmatrix}.$$

It is easy to see that the graph G_s and G'_s induced on $\mathcal{X}$ by $\mathscr{U}$ and $\mathscr{U}'$ respectively is a path $0 - 1 - 2$.

Now take $n = 2$ and consider the graphs G_s^2 and $G_s'^2$ induced by $\mathscr{U}_2$ and $\mathscr{U}'_2$ respectively. It can be shown that 01 and 10 are adjacent in the graph G_s^2, but are not adjacent in the graph $G_s'^2$. Thus, although G_s and G'_s are same, G_s^2 and $G_s'^2$ are different. In short, the graphs $\{G_s^n\}_{n\geq 1}$ are defined by the utility $\mathscr{U}$ and the magnitude of the preference of the sender, rather than G_s wherein such magnitudes are suppressed. □

In the following theorem, we show that only those sequences can be recovered by the receiver that form an independent set in G_s^n. Thus, $R(g_n)$ can be at most $\alpha(G_s^n)^{1/n}$ for any strategy g_n of the receiver. Further, by choosing an appropriate strategy g_n, the receiver can recover any of the largest independent sets of G_s^n and consequently achieve the rate $\alpha(G_s^n)^{1/n}$.

Theorem 3.1. *Let $n \in \mathbb{N}$. Consider a sender with utility $\mathscr{U}$ and let G_s^n be the corresponding sender graph. For any strategy g_n of the receiver define,*

$$\mathcal{S}(g_n) = \arg\min_{s_n \in \mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|.$$

Then, for all strategies $s_n \in \mathcal{S}(g_n)$, $\mathcal{D}(g_n, s_n)$ is an independent set in G_s^n.

Furthermore, for all Stackelberg equilibrium strategies g_n^* of the receiver,

$$R(g_n^*) = \alpha(G_s^n)^{1/n}.$$

Proof. See Appendix A.1. ■

We show in the proof of the above theorem that it is sufficient for the receiver to choose a strategy g_n as

$$g_n(x) = \begin{cases} x & \text{if } x \in I_n \\ \Delta & \text{if } x \notin I_n \end{cases}, \tag{6}$$

where I_n is any largest independent set in G_s^n. Thus, the receiver decodes meaningfully only for messages in I_n. For the rest of the messages, the receiver maps them to Δ. Operationally, we interpret this strategy as a *questionnaire* in which only the sequences from I_n are retained as *alternatives* and all other sequences are dropped. The sender is required to pick one and only one alternative. Any alternative selected by the sender is taken at face value by the receiver, i.e., the receiver applies an identity decoding function. Since I_n is an independent set in G_s^n it follows from Definition 3.2 that

$$\mathscr{U}_n(x, x) > \mathscr{U}_n(y, x) \qquad \forall x, y \in I_n, x \neq y.$$

In other words $\mathscr{U}_n$ is *incentive compatible* with the identity function *when restricted to the set* I_n. Thus truth-telling is the best response for senders whose type belongs to I_n, whereby I_n is recovered by the receiver. For sequences that do not belong to I_n, the sender selects the alternative from I_n for which its utility is maximum over I_n. These latter sequences are not correctly recovered by the receiver.

One may wonder if including sequences from $\mathcal{X}^n \backslash I_n$ could lead to even more sequences recovered. This is false — inclusion of even a single additional sequence from $\mathcal{X}^n \backslash I_n$ will lead to at least one sequence from I_n not being recovered. This is because I_n is a maximum independent set, whereby every sequence in $x \in \mathcal{X}^n \backslash I_n$ is adjacent to at least one sequence in $y \in I_n$, and hence the sender has an incentive to either report x as y or y as x. In other words, at most one of x and y can be recovered. Finally, note that when $\mathscr{U}_n$ is incentive compatible on $\mathcal{X}^n$, we have $I_n = \mathcal{X}^n$ and every sequence is recovered in the equilibrium.

One may also wonder why an *undirected* graph G_s^n captures the optimal behaviour of the receiver. If two sequences x and y are adjacent in the graph G_s^n, the sender may prefer y over x or x over y (or both). Since the receiver is concerned about the number of types correctly recovered, and either type – x or y – could be the true type of the sender, the receiver must design a *common* questionnaire for either possibility. As a result the undirected graph G_s^n encapsulates the problem essence.

4. Information extraction capacity of the sender

One measure of the amount of information at any stage is the number of distinct types accessible there. In that language, the amount of information available *with* the sender is the number of distinct types that the sender can exhibit. This is clearly q^n, where recall $q = |\mathcal{X}|$. And the *growth rate* of this information is q. Theorem 3.1 shows that the maximum information the *receiver can extract* from the sender is equal to $\alpha(G_s^n)$. The limiting value of the quantity as $\alpha(G_s^n)^{1/n}$ as $n \to \infty$ is the asymptotic growth rate of the information that the receiver can extract. As such it is a fundamental limit to the amount of information that is obtainable from such a strategic sender. We call this the *information extraction capacity* of the sender.

Definition 4.1 (*Information Extraction Capacity of a Sender*). Consider a sender with utility $\mathscr{U}$ and let $\{G_s^n\}_{n\geq 1}$ be the corresponding sequence of sender graphs. The information extraction capacity of the sender is defined as

$$\Xi(\mathscr{U}) = \lim_{n\to\infty} \alpha(G_s^n)^{1/n}.$$

We show the existence of the limit in Appendix A.2. If $\Xi(\mathscr{U})$ is greater than unity, the receiver can extract an *exponentially* large number of sequences from the sender when the channel is noiseless, whereas if $\Xi(\mathscr{U}) = 1$, then asymptotically only a vanishing fraction of sequences can be recovered. If $\Xi(\mathscr{U}) = q$, it means that almost all sequences can be recovered and the number of sequences *not* recovered is asymptotically a vanishing fraction of the total number of sequences.

Let g be a Stackelberg equilibrium strategy of the receiver for $n = 1$. It is clear that $R(g) = \alpha(G_s)$. Now consider a strategy that applies g for each component of the sequence x as $\bar{g}_n(x) \equiv (g(x_1), \ldots, g(x_n))$, whereby $\bar{g}_n$ is a stacking of n copies of g. Since g corresponds to an optimal questionnaire for $n = 1$, it follows that $R(\bar{g}_n) = \alpha(G_s)$ for all n. Note that $\bar{g}_n$ is in general not a Stackelberg equilibrium, whereby $\Xi(\mathscr{U}) \geq \alpha(G_s)$. However, if $\Xi(\mathscr{U}) = \alpha(G_s)$, then it implies that $\alpha(G_s^n)^{1/n} = \alpha(G_s)$ for all n, and hence linking responses of the sender does not help the receiver recover more information. We will see later that generically, $\Xi(\mathscr{U}) > \alpha(G_s)$.

Considering that we make no assumption about incentive compatibility on $\mathscr{U}$, it is rather interesting to note that $1 < \Xi(\mathscr{U}) < q$, except in some corner cases. We show this and other properties of the information extraction capacity in the following sections.

4.1. Information extraction capacity generalizes Shannon capacity

Ideally one would like a clean and easily computable characterization of $\Xi(\mathscr{U})$. Unfortunately, such a characterization appears unlikely since $\Xi(\mathscr{U})$ generalizes some well studied computationally intractable problems, namely the Shannon capacity of a graph (Lovász, 1979). The Shannon capacity is given in terms of the strong product graph which is defined as follows.

Definition 4.2. (1) *Strong product*: Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two graphs. Then the strong product of the graphs G_1, G_2 is given by a graph $G = (V, E)$ where $V = V_1 \times V_2$. Further, two vertices $(x, x'), (y, y') \in V$, with $x, y \in V_1$ and $x', y' \in V_2$, are adjacent if and only if one of the following holds

- $x = y$ and $x' \sim y'$
- $x \sim y$ and $x' = y'$
- $x \sim y$ and $x' \sim y'$

The strong product operation is denoted as $\boxtimes$ and the product graph G is written as $G = G_1 \boxtimes G_2$.

(2) *Strong product graph*: The strong product graph denoted as $G^{\boxtimes n}$ is the graph constructed by taking the n-fold strong product of the graph G, i.e.,

$$G^{\boxtimes n} = \underbrace{G \boxtimes G \boxtimes \dots \boxtimes G}_{n}.$$

Notice that the n-fold strong product graph is constructed by using the edge relations of the base graph G_s, unlike the sender graph G_s^n which requires computing the n-block utility for determining the edges.

Definition 4.3 (*Shannon Capacity*)**.** Let G be any graph. The Shannon capacity of G is defined as

$$\Theta(G) = \lim_{n\to\infty} \alpha(G^{\boxtimes n})^{1/n},$$

where $G^{\boxtimes n}$ is the n-fold strong product given by Definition 4.2.

Example 4.1. Consider a graph G on the vertices $\{0, 1, 2\}$ given in Fig. 1(a). The strong product of the graph denoted as $G^{\boxtimes 2}$ is a graph on $\{0, 1, 2\}^2$ and is given in Fig. 1(b). It can be observed that since $\{0, 1\}$ are adjacent in G, the sequences in $\{00, 01\}$ and $\{10, 11\}$ are adjacent in $G^{\boxtimes 2}$ by the first and the second condition from Definition 4.2 respectively. The sequences $\{00, 11\}$ and $\{10, 01\}$ are adjacent due to the third condition. The sequences $\{21, 20\}$ and $\{12, 02\}$ are adjacent due to the first and second condition respectively. Since no sequence satisfies any condition with 22, it is an isolated vertex in the graph $G^{\boxtimes 2}$. □

Shannon (1956) investigated the problem of computing the maximum number of messages that can be transmitted across a noisy channel such that the receiver can recover the messages with zero probability of error. The *confusability graph* G induced by this channel is a graph with vertices as the inputs of the channel, where vertices i and j are adjacent if both can produce a common output with positive probability. In other words, i and j can be *confused* based on the output they produce (cf. Definition 5.2 for a formal discussion). Shannon showed that for any blocklength n, $\alpha(G^{\boxtimes n})$ is the maximum number of messages that can be communicated perfectly across the channel. The limit of the quantity $\alpha(G^{\boxtimes n})^{1/n}$ was termed as the *zero-error capacity* of the channel, also known as the Shannon capacity of a graph G.

There is an intuitive similarity between the above setting and our problem. The utility induces a kind of confusability at the receiver's end: if either $\mathscr{U}(i, j) \geq 0$ or $\mathscr{U}(j, i) \geq 0$, then i and j can be confused for each other by a receiver based on the sender's response. Indeed, we show that the information extraction capacity in fact generalizes the Shannon capacity of a graph.

Lemma 4.1. *Consider a graph G and let the adjacency matrix of the graph be denoted as $\mathcal{A}$. Define a utility $\mathscr{U} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ as*

$$\mathscr{U}(i, j) = \begin{cases} 0 & \text{if } i = j \text{ or } \mathcal{A}(i, j) = 1 \\ -1 & \text{if } i \neq j \text{ and } \mathcal{A}(i, j) = 0 \end{cases}.$$

Then,

$$\Xi(\mathscr{U}) = \Theta(G).$$

Proof. From the definition of $\mathscr{U}$, any pair of distinct symbols i, j are adjacent in G_s if and only if $\mathcal{A}(i, j) = 1$. Thus, distinct symbols i, j are adjacent in G_s if and only if i, j are adjacent in G and hence $G_s = G$. We now use this to show $G_s^n = G^{\boxtimes n}$.

Consider two distinct sequences $x, y \in \mathcal{X}^n$ that are adjacent in G_s^n. This can hold if and only if $\mathscr{U}(x_k, y_k) = 0$ for all k. Thus, for the adjacent sequences x and y, for all k, either '$x_k = y_k$' or 'x_k and y_k are adjacent in G_s' and hence in the graph G. This can occur if and only if x and y are adjacent in $G^{\boxtimes n}$. Thus, $G_s^n = G^{\boxtimes n}$ and hence $\Xi(\mathscr{U}) = \Theta(G)$. ■

The Shannon capacity of a graph $\Theta(G)$, is an important quantity with applications in combinatorics and computer science as well. It is, however, found to be very hard to compute barring some simple cases. For instance, the capacity of the pentagon graph was unknown for about 20 years after being first introduced by Shannon (1956). Further, the capacity of a heptagon graph is still unknown. A computable semidefinite program-based upper bound was introduced by Lovász (1979), now known as the Lovász theta number. The introduction of this quantity was a significant development and is known to lie between the NP-hard clique and chromatic numbers of a graph. However, it is known that the Lovász theta number is not a tight upper bound. For more discussion on this subject of capacity of a graph and its variations, the reader is referred to Korner and Orlitsky (1998) and the references therein.

This shows that computing the exact information extraction capacity would be hard in general and we only strive to obtain bounds on this capacity. Nevertheless, there are a few classes of utilities for which the capacity is exactly characterized and we shall discuss them in the forthcoming sections. We will also discuss cases where the capacities are equal even when the two graphs are not equal.

4.2. Lower bounds on the information extraction capacity

In this section, we present a lower bound on the information extraction capacity in terms of the optimal value of an optimization problem. This optimization problem is defined for every history length n and thus gives a series of lower bounds on the capacity. We show that this series of lower bounds is asymptotically tight. We also present bounds on the optimal value of this optimization problem. We then discuss the characteristics of the feasible region of this optimization problem.

We first define the optimization problem. Let $\mathcal{Y}_n \subseteq \mathcal{X}^n$. Let $\mathcal{Q} = \{Q^{(0)}, \dots, Q^{(|\mathcal{Q}|)}\}$, $Q^{(0)} = \mathbf{I}$ be the set of all $|\mathcal{Y}_n| \times |\mathcal{Y}_n|$ permutation matrices. For convenience, we assume that the permutation matrices are indexed by sequences from $\mathcal{Y}_n$. Consider the problem $\mathcal{O}(\mathscr{U}_n)$ as

$\mathcal{O}(\mathscr{U}_n)$:	$\max_{\mathcal{Y}_n \subseteq \mathcal{X}^n} \lvert\mathcal{Y}_n\rvert$
s.t.	$\sum_{x,z\in\mathcal{Y}_n} Q(x, z)\mathscr{U}_n(x, z) < 0 \quad \forall\, Q \in \mathcal{Q} \setminus \{\mathbf{I}\}.$

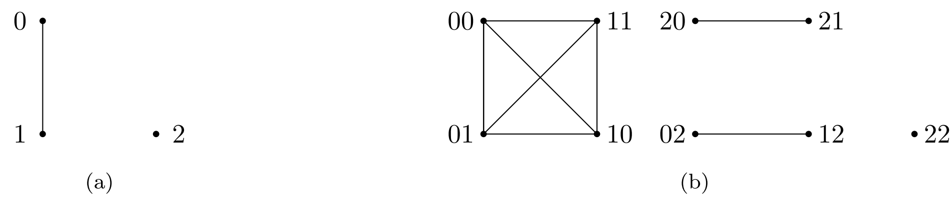


Fig. 1. (a) Graph G, (b) Graph $G^{\boxtimes n}$ (Example 4.1).

Let $\Gamma(\mathscr{U}_n) = \mathrm{OPT}(\mathcal{O}(\mathscr{U}_n))$.

Before presenting the theorem, we also discuss the symmetric part of the utility and the corresponding induced graph.

Definition 4.4 (*Symmetric Part of a Utility*). For a given utility $\mathscr{U}$, let $\mathscr{U}^{\mathsf{Sym}}$ be the symmetric part defined as

$$\mathscr{U}^{\mathsf{Sym}}(i,j) = \frac{1}{2}(\mathscr{U}(i,j) + \mathscr{U}(j,i)).$$

We denote the sender graph induced by $\mathscr{U}^{\mathsf{Sym}}$ as $G_{\mathsf{s}}^{\mathsf{Sym}}$. $\mathscr{U}^{\mathsf{Sym}}$ has a simpler structure than $\mathscr{U}$ since $\mathscr{U}^{\mathsf{Sym}}(i,j) = \mathscr{U}^{\mathsf{Sym}}(j,i)$, and thus, if a symbol i is preferred by the sender over the symbol j, then it follows that j would be preferred over i with the same extent. In fact, we later show that $(G_{\mathsf{s}}^{\mathsf{Sym}})^n$ contains only a subset of edges of G_{s}^n for all n(cf. proof of Theorem 4.5).

We now present the main result that gives a lower bound on the information extraction capacity.

Theorem 4.2. *Let $\mathscr{U}$ be any utility and G_{s}^n be the corresponding sender graph and let $\mathscr{U}^{\mathsf{Sym}}$ and $(G_{\mathsf{s}}^{\mathsf{Sym}})^n$ be the corresponding symmetric counterparts. For all n, we have $\Xi(\mathscr{U}) \geq \Gamma(\mathscr{U}_n)^{1/n}$. Further,*

$$\lim_{n\to\infty} \Gamma(\mathscr{U}_n)^{1/n} = \Xi(\mathscr{U}).$$

Moreover, $\Gamma(\mathscr{U}_n)$ is bounded as

$$\alpha(G_{\mathsf{s}}^n) \leq \Gamma(\mathscr{U}_n) \leq \alpha((G_{\mathsf{s}}^{\mathsf{Sym}})^n).$$

Proof. The proof of the lower bound on the capacity is in Appendix B.1. The proof of the bounds on $\Gamma(\mathscr{U}_n)$ is in Appendix B.2. ■

Computing $\Xi(\mathscr{U})$ requires the sequence of $\{\alpha(G_{\mathsf{s}}^n)\}_{n\geq 1}$ corresponding to the sequence of sender graphs $\{G_{\mathsf{s}}^n\}_{n\geq 1}$ induced by $\mathscr{U}$. Moreover, computation of the independence number of a generic graph can be intractable. We also show via the proof of the above theorem that for a fixed $\mathcal{Y}_n$, the problem $\mathcal{O}(\mathscr{U}_n)$ can be solved as a linear program. Thus, the above theorem provides a computable lower bound.

This theorem shows that asymptotically, the information extraction capacity is captured by sets of sequences between which the sender does not wish to deviate *on average*. It also suggests a way to approximate the capacity arbitrarily closely by taking higher values of n. In fact, improved bounds on the Shannon capacity are also obtained similarly by using the fact that $\alpha(G^{\boxtimes n}) \geq \alpha(G)^n$ (Polak and Schrijver, 2019).

For $n = 1$, the corresponding problem $\mathcal{O}(\mathscr{U})$ gives a single-letter lower bound on the capacity given by $\Gamma(\mathscr{U}) = \mathrm{OPT}(\mathcal{O}(\mathscr{U}))$. In the following, we discuss an alternate characterization of the feasible region of $\mathcal{O}(\mathscr{U})$. Although these are proved for the case of $n = 1$, they can be generalized to larger sequences as well.

For a fixed P, define

$\mathcal{T}(\mathscr{U},P)$:	$\max\limits_{P_{X,Y}\in\mathcal{P}(\mathcal{Y}^2)} \mathbb{E}_{P_{X,Y}}\mathscr{U}(X,Y)$
s.t.	$P_X = P_Y = P.$

Lemma 4.3. *Let $n = 1$. The following statements are equivalent*

1. *$\mathcal{Y}$ is feasible for $\mathcal{O}(\mathscr{U})$*
2. *For all $P \in \mathcal{P}(\mathcal{X})$ such that $\mathcal{Y} = \mathrm{supp}(P)$, the optimal value of $\mathcal{T}(\mathscr{U},P)$ is 0 and the maximum is achieved by the unique distribution $P^*_{X,Y}$ where $P^*_{X,Y}(i,i) = P(i)$ for all $i \in \mathcal{X}$.*

Proof. See Appendix B.3. ■

For the case when $\mathcal{Y}$ is the complete state space, i.e., when $\mathcal{Y} = \mathcal{X}$, the equivalence between (1) and (2) can be deduced from Lemma 1 proved by Renault et al. (2013) and from Lemma 2 proved by Renou and Tomala (2015). Our result is a straightforward generalization where P is supported on any arbitrary subset of $\mathcal{X}$. The proof follows on similar lines, but is nevertheless presented in the appendix for completeness.

Consider a set $\mathcal{Y}$ that is feasible for $\mathcal{O}(\mathscr{U})$. Let n be fixed and consider a set $\mathcal{Y}_n$ that contain symbols only from the set $\mathcal{Y}$. Further, let the sequences in $\mathcal{Y}_n$ have the same empirical distribution; that is they have the same number of each symbol, and which differ only in position. For distinct sequences $x, y \in \mathcal{Y}_n$, we can write

$$\mathscr{U}_n(x,y) = \frac{1}{n}\sum_k \mathscr{U}(x_k,y_k) = \sum_{i,j\in\mathcal{Y}} P_{x,y}(i,j)\mathscr{U}(i,j) = \mathbb{E}_{P_{x,y}}\mathscr{U}(X,Y),$$

where $P_{x,y}$ is the joint empirical distribution of the sequences x and y and $P_x = P_y$. Since x and y are distinct sequences, Lemma 4.3 states that $\mathscr{U}_n(x,y) = \mathbb{E}_{P_{x,y}}\mathscr{U}(X,Y) < 0$. Thus, the sender does not prefer the sequence x over y. This implies that, on average, the sender does not gain by *scrambling* the symbols from the set $\mathcal{Y}$ and hence the sender is "truthful on average" over the set $\mathcal{Y}$.

It can be observed that $\mathcal{T}(\mathscr{U},P)$ is an optimal transport problem (Ambrosio et al., 2008, Ch. 6) where the optimization is over joint distributions $P_{X,Y}$ with marginals equal to P. Thus, the problem $\mathcal{T}(\mathscr{U},P)$ asks for an optimal way of relatively distributing the weights $P_{X,Y}$ while keeping the marginals as P. The hypothesis of Lemma 4.3 shows that the type $P_{X,Y}$ should be such that the optimal rearrangement is $P^*_{X,Y}(i,i) = P(i)$ for all $i \in \mathcal{X}$.

Using the properties of permutations, we get the following interpretation of the feasible sets of $\mathcal{O}(\mathscr{U})$.

Remark 4.1. Let $n = 1$. A set $\mathcal{Y} \subseteq \mathcal{X}$ is feasible for $\mathcal{O}(\mathscr{U})$ if and only if any sequence of distinct symbols $i_0, i_1, \ldots, i_{K-1} \in \mathcal{Y}$ form a negative-weight chain defined as

$$\mathscr{U}(i_1,i_0) + \mathscr{U}(i_2,i_1) + \ldots + \mathscr{U}(i_0,i_{K-1}) < 0. \tag{7}$$

Notice that the chain in the above definition is a closed-chain since the first and last symbols are the same. For brevity, we describe a closed-chain as simply a chain.

To see the above assertion, take distinct symbols $i_0, i_1, \ldots, i_{K-1} \in \mathcal{Y}$ where $K \in [|\mathcal{Y}|]$, $K \geq 2$ and choose a permutation $\pi' : \mathcal{Y} \to \mathcal{Y}$ such that $i_0, i_1, \ldots, i_{K-1}$ form a cycle as $\pi'(i_k) = i_{(k+1)\mathrm{mod}K}$. The rest of the symbols are mapped to themselves. Let $Q^{\pi'}$ be the corresponding permutation matrix. Observe that $Q^{\pi'}$ is not an identity matrix. Then, it follows from feasibility of $\mathcal{Y}$ that

$$\sum_{i,j\in\mathcal{Y}} Q^{\pi'}(i,j)\mathscr{U}(i,j) = \mathscr{U}(i_1,i_0) + \mathscr{U}(i_2,i_1) + \ldots + \mathscr{U}(i_0,i_{K-1}) < 0,$$

and hence the symbols form a negative-weight chain. Since the symbols were chosen arbitrarily, this holds for all sequences of distinct symbols from the set $\mathcal{Y}$.

The expression in (7) is (upon scaling) the utility obtained by the sender when the observed sequence is $(i_0 i_1 \dots i_{K-1})$ and the decoded sequence is $(i_1 i_2 \dots i_{K-1} i_0)$, i.e., the receiver decodes the symbol i_1 in place of i_0, and i_2 in place of i_1 and so on for $i_2, \dots, i_{K-1}$. Thus these symbols form a *chain of lies* created by the sender represented as $i_0 \to i_1 \to \dots \to i_{K-1} \to i_0$. This remark shows that a set $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$ if and only if the utility obtained is negative for all possible chains of lies that can be formed from the symbols of the set $\mathcal{Y}$. We also discuss a computational approach to check for the feasibility of a set in Appendix B.4. □

We now present a result that gives another lower bound on $\Xi(\mathcal{U})$. We later use this result to show an example where even for a complete graph the capacity can be the maximum. For that we require a definition.

Definition 4.5 (*Cycle and Positive-Edges Cycle*)**.** Consider a set of distinct vertices $\{i_0, i_1, \dots, i_{K-1}\}$ from a graph. The vertices form a K-length cycle in the graph if two vertices i_l, i_m are adjacent whenever $l = (m+1) \bmod K$. Further, the cycle is a positive-edges cycle if for all m,

$$\mathcal{U}(i_l, i_m) \geq 0 \quad \text{whenever} \quad l = (m+1) \bmod K.$$

Suppose there exists a set of vertices $\mathcal{Y}$ such that no subset of symbols form such a cycle. The following proposition gives a sufficient condition for such a set to be feasible for $\mathcal{O}(\mathcal{U})$ and hence, for the size of this set to be a lower bound on $\Xi(\mathcal{U})$.

Proposition 4.4. *Consider a sender with utility $\mathcal{U}$ and let G_s be the corresponding sender graph. Suppose there exists a set $\mathcal{Y}$ such that there is no positive-edges cycle in the sub-graph induced by $\mathcal{Y}$ and*

$$\min_{i,j \in \mathcal{Y} : \mathcal{U}(i,j) < 0} |\mathcal{U}(i,j)| > (|\mathcal{Y}| - 1) \max_{i,j \in \mathcal{Y} : \mathcal{U}(i,j) \geq 0} \mathcal{U}(i,j). \tag{8}$$

Then, $\Xi(\mathcal{U}) \geq |\mathcal{Y}|$.

Proof. See Appendix B.5 ■

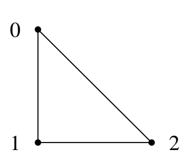
The proof of the above proposition relies on showing that $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$. Clearly, if there exists a positive edges cycle then (7) will not be satisfied. Moreover, (8) ensures that in any set of distinct symbols, there exists a pair of symbols such that the penalty for lying is large enough than the incentive. This gives that, on average, the sender does not get a better utility by misrepresenting among the symbols of $\mathcal{Y}$ and hence it is feasible for $\mathcal{O}(\mathcal{U})$.

Proposition 4.4 gives a sufficient condition in terms of the negative and positive values of $\mathcal{U}(i,j)$ for a set $\mathcal{Y}$ to be feasible for $\mathcal{O}(\mathcal{U})$. Using this result, the following example shows that $\Xi(\mathcal{U}) = q$ even when the base sender graph G_s is a complete graph.

Example 4.2. Let $\mathcal{U} : \{0,1,2\} \times \{0,1,2\} \to \mathbb{R}$ and consider the following form of $\mathcal{U}$,

$$\mathcal{U} = \begin{pmatrix} 0 & 1 & 1 \\ -4 & 0 & 1 \\ -4 & -4 & 0 \end{pmatrix}.$$

The graph induced by the utility is a 3-cycle graph and is given as



This follows since $\mathcal{U}(0,1), \mathcal{U}(0,2), \mathcal{U}(1,2) > 0$. It can be easily observed that there is no positive-edges cycle in the graph. This is because for chain $i \to j \to i$, either $\mathcal{U}(i,j) < 0$ or $\mathcal{U}(j,i) < 0$. Further, for all chains $i \to j \to k \to i$, either $\mathcal{U}(j,i) < 0$ or $\mathcal{U}(k,j) < 0$ or $\mathcal{U}(i,k) < 0$. Also, the largest weight of a chain in G_s is $-4+1+1 = -2$. Finally, $\min_{i,j:\mathcal{U}(i,j)<0} |\mathcal{U}(i,j)| = 4 > (3-1) \max_{i,j:\mathcal{U}(i,j)\geq 0} \mathcal{U}(i,j) = 2$. Thus, the conditions of Proposition 4.4 are satisfied and hence $\mathcal{X}$ is feasible for $\mathcal{O}(\mathcal{U})$ and hence $\Xi(\mathcal{U}) = q = 3$. □

We find this example to be quite surprising — even though G_s is a complete graph, the information extraction capacity is the maximum it can be. The main lesson to be drawn from it is that the *magnitude* of the gains or losses from truth-telling or lying determine the amount of information that can be extracted; the base sender graph G_s only considers the sign of these quantities.

4.3. Upper bounds on the information extraction capacity

In this section, we derive upper bounds on the information extraction capacity of the sender. We show that the Shannon capacity of the sender graph corresponding to the symmetric part of the utility is an upper bound on the capacity. We also discuss a class of utilities where the upper bound is given by the Shannon capacity of the sender graph itself.

First, consider the following upper bound for any utility $\mathcal{U}$. Recall the symmetric part of the utility from Definition 4.4.

Theorem 4.5. *Consider a utility $\mathcal{U}$ and let $\mathcal{U}^{\mathsf{Sym}}$ be its symmetric part. Let G_s^{Sym} be the sender graph corresponding to $\mathcal{U}^{\mathsf{Sym}}$. Then,*

$$\Xi(\mathcal{U}) \leq \Xi(\mathcal{U}^{\mathsf{Sym}}) \leq \Theta(G_s^{\mathsf{Sym}}).$$

Proof. See Appendix C.1 ■

As discussed in the last section, the graph G_s^{Sym} has a simpler structure compared to G_s. Moreover, in the proof of the above theorem we show that only a subset of edges of the graph G_s remain in G_s^{Sym}. Thus, the sender with the utility $\mathcal{U}^{\mathsf{Sym}}$ is more truthful about its information and as shown by the above bound, the information extraction capacity of a sender is no greater than that of a sender whose utility is equal to the symmetric part of the former's utility. In other words, ignoring the skew-symmetric part of the utility leads to an increase in the information extraction capacity. This is also because in the skew-symmetric part of the utility, for each pair of symbols $i, j \in \mathcal{X}$, we have that either i is preferred to be recovered as j or j is preferred to be recovered as i. Thus, the graph corresponding to the skew-symmetric part is a complete graph and has capacity 1.

As with the lower bound, the above theorem provides an upper bound that is independent of the blocklength n and depends only on the single letter utility $\mathcal{U}$. We can observe that the symmetric part of the utility plays a recurring role in the characterization of the bounds on the capacity.

We now present another class of utilities where the capacity is bounded above by the Shannon capacity of the sender graph.

Proposition 4.6. *Consider a sender with utility $\mathcal{U}$ given as*

$$\mathcal{U}(i,j) = \begin{cases} a & \text{if } \mathcal{U}(i,j) \geq 0 \\ -b & \text{if } \mathcal{U}(i,j) < 0 \end{cases},$$

where $a, b > 0$ and $a \geq b$. Let G_s be the corresponding sender graph. Then,

$$\Xi(\mathcal{U}) \leq \Theta(G_s).$$

Proof. See Appendix C.2 ■

The utility in the above proposition is such that the *incentive* for lying is greater than the *penalty* for lying. Intuitively, a sender with utility as above has higher tendency to lie about its information, since it can offset its penalty of lying by gaining appropriate incentive. Thus the information extraction capacity for such a sender is in general strictly less than q.

The above results characterized the utilities for which the capacity is bounded above by $\Theta(G_s)$. In the following section, we will discuss cases where the capacity is exactly equal to $\Theta(G_s)$.

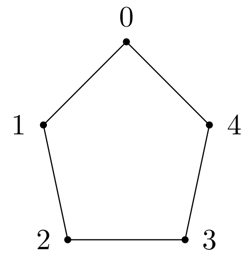


Fig. 2. Sender graph G_s on $\{0, 1, 2, 3, 4\}$ (Example 4.3).

4.4. Exact evaluation of $\Xi(\mathscr{U})$

In the earlier section, we discussed upper bounds on the information extraction capacity. A natural question that follows is that under what conditions the capacity is exactly characterized? In this section, we mention few cases of utility where the information extraction capacity is equal to the Shannon capacity of the sender graph.

Theorem 4.7. *Consider a utility $\mathscr{U}$ and let G_s be the corresponding sender graph. Then,*

$$\Xi(\mathscr{U}) = \Theta(G_s) = \alpha(G_s)$$

if any of the following hold:

1. *$\mathscr{U}$ is symmetric and G_s is a perfect graph.*
2. *$\mathscr{U}$ is of the form given in Proposition 4.6 and G_s is a perfect graph.*

Proof. See Appendix C.3 ■

Recall that a graph is *perfect* if and only if the graph and its complement do not have a odd cycle with 5 or more number of vertices (West, 2000). This means that there *does not* exist odd number of distinct symbols $\{i_0, \dots, i_{K-1}\}, K \geq 5$ such that either $\mathscr{U}(i_k, i_{k+1}) \geq 0$ or $\mathscr{U}(i_{k+1}, i_k) \geq 0$ for $k \leq K-1$ and either $\mathscr{U}(i_0, i_{K-1}) \geq 0$ or $\mathscr{U}(i_{K-1}, i_0) \geq 0$, and for all other i, j, $\mathscr{U}(i, j) < 0$. Suppose the sender has *transitive preferences*, i.e., if $\mathscr{U}(i, j) \geq 0$ and $\mathscr{U}(j, k) \geq 0$ implies $\mathscr{U}(i, k) \geq 0$ for all $i, j, k \in \mathcal{X}$. Then, if $\mathscr{U}$ is symmetric it follows that G_s is perfect. In this case, the above theorem can be used to ascertain the exact information extraction capacity.

In the above theorem, the induced sender graph is such that the capacity can be derived by just computing the independent set of the base graph G_s since the independence number of the subsequent graphs is $\alpha(G_s)^n$. It implies that in this case, linking the responses of the sender does not benefit the receiver. In the context of questionnaires, when $\Xi(\mathscr{U}) = \alpha(G_s)$, the receiver does not benefit by linking the responses over multiple days. It is optimal to just construct a *day-wise* questionnaire using the largest independent set in the base graph G_s.

The following example demonstrates how the lower bounds and the upper bounds can be used to exactly compute the capacity.

Example 4.3. Consider a sender with utility $\mathscr{U}$ on $\mathcal{X} = \{0, 1, 2, 3, 4\}$ as

$$\mathscr{U} = \begin{pmatrix} 0 & -1 & u_{02} & u_{03} & -1 \\ 1 & 0 & -1 & u_{13} & u_{14} \\ u_{20} & 1 & 0 & -1 & u_{24} \\ u_{30} & u_{31} & 1 & 0 & -1 \\ 1 & u_{41} & u_{42} & 1 & 0 \end{pmatrix},$$

where the unknown entries are such that $u_{ij} < -1$ for all $i \in \mathcal{X}, j \in \{(i+2) \bmod 5, (i+3) \bmod 5\}$. It can be observed that the base graph G_s is a pentagon and is given as in Fig. 2.

This is because i, j are adjacent if and only if $|j - i| \bmod 5 = 1$. We show that $\Xi(\mathscr{U}) = \sqrt{5}$.

Table 1
Bounds on the independence numbers of G_s^n and the rate $\alpha(G_s^n)^{1/n}$ (Example 4.3).

n	1	2	3	4	5
$\alpha(G_s^n)$	2	5	10	25	50 – 55
$\alpha(G_s^n)^{1/n}$	2	2.23	2.15	2.23	2.18 – 2.23

Notice that the graph does not contain a positive-edges cycle since $\mathscr{U}(0, 4) = -1$. Further, the graph G_s^{Sym} induced by the symmetric part $\mathscr{U}^{\mathsf{Sym}}$ is also a pentagon graph since $\mathscr{U}^{\mathsf{Sym}}(i, j) = (\mathscr{U}(i, j) + \mathscr{U}(j, i))/2 = 0$ if and only if i, j are adjacent in G_s or $i = j$. For all other i, j not adjacent in G_s^{Sym}, $\mathscr{U}^{\mathsf{Sym}}(i, j) < 0$. Since $\Theta(G_s^{\mathsf{Sym}}) = \sqrt{5}$ (Lovász, 1979), from Theorem 4.5 it follows that $\Xi(\mathscr{U}) \leq \Theta(G_s^{\mathsf{Sym}}) = \sqrt{5}$.

We now compute $\Xi(\mathscr{U})$. From the problem $\mathcal{O}(\mathscr{U})$, we only get a lower bound on $\Xi(\mathscr{U})$ since $\Gamma(\mathscr{U}) = 2$. However, consider $\mathcal{O}(\mathscr{U}_2)$ and the set $S = \{00, 12, 24, 31, 43\}$. Let x, y be a distinct pair of sequences from S. Denoting $x = (x_1, x_2), y = (y_1, y_2)$, it can be observed that if $\mathscr{U}(x_1, y_1) = 1$, then $\mathscr{U}(x_2, y_2) < -1$ and vice-versa. Thus, $\mathscr{U}_2(x, y) < 0$ for all $x, y \in S$ and hence S is an independent set in G_s^2. This gives that S is feasible for $\mathcal{O}(\mathscr{U}_2)$ and hence from Theorem 4.2 we have

$$\Xi(\mathscr{U}) \geq \Gamma(\mathscr{U}_2)^{1/2} \geq \sqrt{5}.$$

The upper and lower bounds together give that $\Xi(\mathscr{U}) = \sqrt{5}$.

We take that case of the utility where $u_{ij} = -2$ for all $i \in \mathcal{X}, j \in \{(i+2) \bmod 5, (i+3) \bmod 5\}$. For the graph G_s^n, we numerically compute the lower bounds using the integer programming formulation of the independence number of a graph (Conforti et al. (2014), Ch. 2). This helps to visualize how the size of the optimal questionnaire $\alpha(G_s^n)$ grows with the history lengths n. Table 1 shows the exact values of independence numbers for $n \leq 4$. For $n = 5$, we only have a lower bound from numerical computation and an upper bound using the fact that $\alpha(G_s^n) \leq \Xi(\mathscr{U})^n$. The variation of the independence number with increasing length of histories is given in Fig. 3(a). The green hashed line is the plot of $\Xi(\mathscr{U})^n$ and the blue solid line is the lower bound on $\alpha(G_s^n)$. The variation of the rate of information extraction is given in Fig. 3(b). Recall from the discussion after Theorem 4.7, that the receiver can construct a day-wise questionnaire by just using the largest independent sets of the graph G_s. They grow as $\alpha(G_s)^n$ and are generally suboptimal. The red dotted line in Fig. 3(a) shows the growth of the day-wise questionnaire for this utility and Fig. 3(b) shows the corresponding rate. □

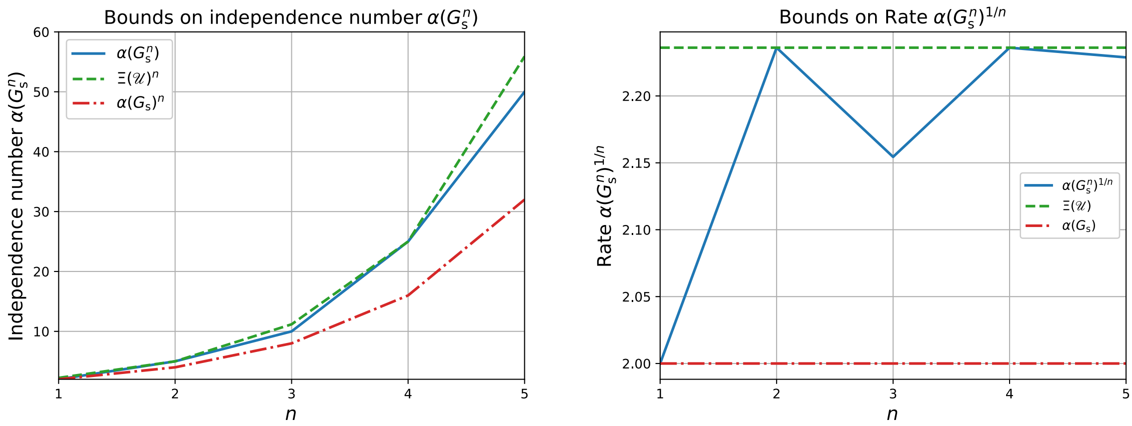
In the following, we present an example where the capacity is unknown and is approximated by lower and upper bounds.

Example 4.4. Consider a sender with utility $\mathscr{U}$ on $\mathcal{X} = \{0, 1, 2, 3, 4, 5, 6\}$ as

$$\mathscr{U} = \begin{pmatrix} 0 & -1 & -3 & -5 & -2 & -5 & -1 \\ 1 & 0 & -1 & -2 & -5 & -2 & -2 \\ -2 & 1 & 0 & -1 & -5 & -2 & -2 \\ -2 & -5 & 1 & 0 & -1 & -3 & -2 \\ -2 & -2 & -2 & 1 & 0 & -1 & -5 \\ -3 & -2 & -2 & -3 & 1 & 0 & -1 \\ 1 & -2 & -5 & -2 & -5 & 1 & 0 \end{pmatrix}.$$

The resulting graph is a cycle graph with 7 vertices as shown in Fig. 4. The graph induced by the symmetric utility $\mathscr{U}^{\mathsf{Sym}}$ is also a 7-cycle graph. The Shannon capacity of this graph is unknown. Further, the independence numbers of the strong products of the cycle graph, given by $(G_s^{\mathsf{Sym}})^{\boxtimes n}$, is exactly known only for $n = 1, 2, 3$. For $n \geq 4$ only bounds on the independence number are known (Polak and Schrijver, 2019). The bounds are given in the first row of Table 2.

We now provide numerical bounds on the independence number of G_s^n. To do so, we use the following inequality from the proof of



(a) (b)

Fig. 3. (a) Bounds on independence number with n, (b) Bounds on the rate with n (Example 4.3).

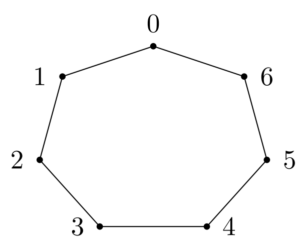


Fig. 4. Sender graph G_s on $\{0, 1, 2, 3, 4, 5, 6\}$ (Example 4.4).

Table 2
Bounds on the independence numbers of $(G_s^{\mathrm{Sym}})^{\boxtimes n}$ and G_s^n and the rate (Example 4.4).

n	1	2	3	4	5
$\alpha((G_s^{\mathrm{Sym}})^{\boxtimes n})$	3	10	33	108 – 115	367 – 401
$\alpha(G_s^n)$	3	10	31	100 – 115	310 – 401
$\alpha((G_s^{\mathrm{Sym}})^{\boxtimes n})^{1/n}$	3	3.16	3.20	3.22 – 3.27	3.25 – 3.31
$\alpha(G_s^n)^{1/n}$	3	3.16	3.14	3.16 – 3.27	3.15 – 3.31

Theorem 4.5 given in Appendix C.1 where we show that $\alpha(G_s^n)$ is bounded above as

$$\alpha(G_s^n) \le \alpha((G_s^{\mathrm{Sym}})^n) \le \alpha((G_s^{\mathrm{Sym}})^{\boxtimes n}).$$

We use numerical computations to derive lower bounds on $\alpha(G_s^n)$ and we derive the upper bounds through the bounds on $\alpha((G_s^{\mathrm{Sym}})^{\boxtimes n})$ given in Polak and Schrijver (2019).

As in Example 4.3, for the graph G_s^n, we numerically compute the lower bounds using the integer programming formulation of the independence number of a graph. These computations either yield an exact value of the independence number or a lower bound on it. Since $\Gamma(\mathcal{U}_n) \ge \alpha(G_s^n)$, we get bounds on $\Gamma(\mathcal{U}_n)$ as well. For $n \in \{1, 2, 3\}$ we have the exact value of the independence number. For $n = 4$ we only have a lower bound and for $n = 5$, we derive a lower bound by using the fact that $\alpha(G_s^5) \ge \alpha(G_s^2)\alpha(G_s^3)$. Bound for $n > 5$ can also be derived by taking such product decompositions, however, we do not present them here. Together, the bounds for history lengths till $n = 5$ are given in the second row of Table 2. The corresponding bounds on the rate are given in third and fourth row of Table 2. Computations of independence numbers for $n > 5$ were not performed due to increased computational effort.

The variation of the independence number with increasing length of histories is given in Fig. 5(a). The orange hashed line is the upper bound on $\alpha((G_s^{\mathrm{Sym}})^{\boxtimes n})$ and the blue solid line is the lower bound on $\alpha(G_s^n)$. For $n \ge 4$, the curve of $\alpha(G_s^n)$ thereby lies between the orange and the blue line. The red line is the plot of $\alpha(G_s)^n$ or equivalently, the size of day-wise questionnaire. The variation of the rate is given in Fig. 5(b) and the capacity lies between the orange and the blue line. The red dotted line is the rate of the day-wise questionnaire. □

5. Information extraction over a noisy channel

We now discuss the case where the sender and receiver communicate via a noisy channel. We determine the maximum number of sequences that can be recovered by the receiver in any equilibrium of the game. We present a notion of the asymptotic rate of information extraction and we show that it is equal to the minimum of the information extraction capacity of the sender and the zero-error capacity of the noisy channel.

5.1. Model with a noisy channel

Consider now a setting where the sender and receiver communicate via a noisy channel. As earlier, the sender observes a sequence $X \in \mathcal{X}^n$ and encodes it as $s_n(X) = Y$, where $s_n : \mathcal{X}^n \to \mathcal{Y}^n$. Here, $\mathcal{Y}$ is the input space of the channel. The message is now transmitted to the receiver via a discrete memoryless channel which generates an output $Z \in \mathcal{Z}^n$, with $\mathcal{Z}$ being the output space, according to the distribution $P_{Z|Y}$ defined as

$$P_{Z|Y}(z|y) = \prod_{i=1}^{n} P_{\mathbb{Z}|\mathbb{Y}}(z_i|y_i), \tag{9}$$

where $P_{\mathbb{Z}|\mathbb{Y}}(\cdot|\cdot) \in \mathcal{P}(\mathcal{Z}|\mathcal{Y})$. The output is decoded by the receiver as $g_n(Z) = \hat{X}$, where $g_n : \mathcal{Z}^n \to \mathcal{X}^n \cup \{\Delta\}$. Here Z is distributed according to $P_{Z|Y}(\cdot|s_n(x))$, when $s_n(x)$ is the input to the channel.

Generalizing our earlier notation in (2), let

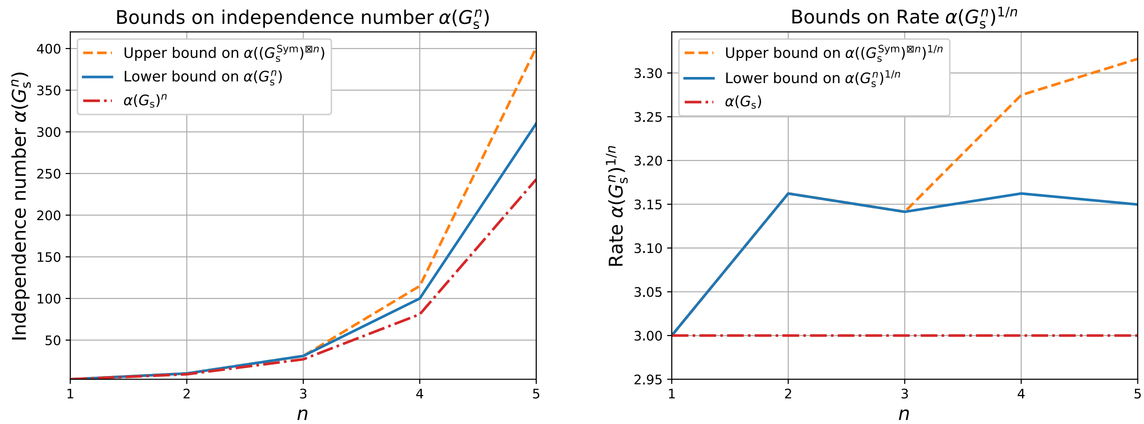
$$\mathcal{D}(g_n, s_n) := \left\{ x \in \mathcal{X}^n \mid \mathbb{P}(\hat{X} = x|X = x) = 1 \right\}, \tag{10}$$

be the set of recovered sequences when the receiver plays the strategy g_n and the sender plays the strategy s_n. The receiver tries to maximize the size of this set by choosing a strategy g_n. The sender on the other hand chooses a strategy s_n to maximize the expected utility

$$\mathbb{E}\left[\mathcal{U}_n(\hat{X}, x)\right] = \sum_{z \in \mathcal{Z}^n} P_{Z|Y}(z|s_n(x))\mathcal{U}_n(g(z), x)$$

for every $x \in \mathcal{X}^n$. The utility $\mathcal{U}_n$ is as given in (3).

As in the noiseless model, we pose the problem as a Stackelberg game.



(a) (b)

Fig. 5. (a) Bounds on independence number with n, (b) Bounds on the rate with n (Example 4.4).

Definition 5.1 (*Stackelberg Equilibrium*)**.** The optimal strategy of the receiver is given as

$$g_n^* \in \arg\max_{g_n} \min_{s_n \in \mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|. \tag{11}$$

The set of best responses of the sender $\mathscr{B}(g_n)$ is determined as

$$\mathscr{B}(g_n) = \Big\{ \ s_n : \mathcal{X}^n \to \mathcal{Y}^n \mid \mathbb{E}\Big[\mathscr{U}_n(g_n(Z), x)\Big] \geq \mathbb{E}\Big[\mathscr{U}_n(g_n(Z'), x)\Big] \\ \forall\, x \in \mathcal{X}^n, \forall\, s_n' \ \Big\}, \tag{12}$$

where Z is distributed according to $P_{Z|Y}(\cdot|s_n(x))$ and Z' is distributed according to $P_{Z|Y}(\cdot|s_n'(x))$.

Notice that as in the noiseless channel model, we adopt a *pessimistic* formulation for the receiver.

Thus, the set of best responses of the sender for a strategy g_n of the receiver is a collection of strategies, s_n, such that for all sequences x, the expected utility with respect to the distribution $P_{Z|Y}(\cdot|s_n(x))$ is the highest that can be obtained by the sender.

5.2. Stackelberg equilibrium of the game

Recall the definition of the recovered set $\mathcal{D}(g_n, s_n)$ from (10), when the receiver plays g_n and the sender plays s_n. As in Definition 3.1, the number of sequences recovered by a strategy g_n of the receiver is defined as

$$\min_{s_n \in \mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|,$$

where $\mathscr{B}(g_n)$ is as given in (12).

We now recall the definition of a graph induced by the channel, called as the confusability graph.

Definition 5.2 (*Confusability Graph*)**.** The confusability graph of a channel $P_{Z|Y}$, denoted as $G_c^n = (\mathcal{Y}^n, E_c)$, is the graph where $(y, y') \in E_c$ if there exists an output $z \in \mathcal{Z}^n$ such that

$$P_{Z|Y}(z|y)P_{Z|Y}(z|y') > 0.$$

For $n = 1$, the graph G_c^1 is denoted as G_c.

Thus, two inputs to the channel are adjacent in the confusability graph if they have a common output. The confusability graph was first introduced by Shannon (1956).

Definition 5.3 (*Asymptotic Rate of Information Extraction*)**.** Let $\{g_n^*\}_{n\geq 1}$ be a sequence of Stackelberg equilibrium strategies for the receiver and let $\{R(g_n^*)\}_{n\geq 1}$ be the corresponding sequence of rate of information extraction. The asymptotic rate of information extraction, denoted by $\mathcal{R}$, is given as

$$\mathcal{R} = \limsup_n R(g_n^*).$$

In the following theorem, we show that the receiver recovers $\min\{\alpha(G_s^n), \alpha(G_c^n)\}$ number of sequences in an equilibrium. The idea of the proof is as follows. We consider the independent sets in G_s^n and G_c^n of the size $\min\{\alpha(G_s^n), \alpha(G_c^n)\}$ denoted as I_n^s and I_n^c respectively. For every channel input sequence in I_n^c, the receiver maps the corresponding output set to a unique sequence in I_n^s. This establishes a one-to-one correspondence between the sequences in I_n^s and I_n^c. Since I_n^s is an independent set in G_s^n, the sender complies with the receiver and maps the sequences to their respective input sequences in I_n^c. The receiver is thus able to recover the set I_n^s of size $\min\{\alpha(G_s^n), \alpha(G_c^n)\}$.

Theorem 5.1. *Let $n \in \mathbb{N}$. Consider a sender with utility $\mathscr{U}$ and let G_s^n be the corresponding sender graph. Let G_c^n be the confusability graph of the channel $P_{Z|Y}$. For all Stackelberg equilibrium strategies g_n^* of the receiver,*

$$\min_{s_n \in \mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)| = \min\{\alpha(G_s^n), \alpha(G_c^n)\}.$$

Further, the asymptotic rate of information extraction is given as

$$\mathcal{R} = \min\{\Xi(\mathscr{U}), \Theta(G_c)\}.$$

Moreover,

1. *If $\Theta(G_s^{\mathrm{Sym}}) \leq \Theta(G_c)$, then $\mathcal{R} = \Xi(\mathscr{U})$.*
2. *For any n, if $\Gamma(\mathscr{U}_n)^{1/n} \geq \Theta(G_c)$, then $\mathcal{R} = \Theta(G_c)$.*

Proof. The proof of the characterization of rate is given in Appendix D.1. Part (1) follows by using Theorem 5.1 along with Theorem 4.5. Part (2) follows by using Theorem 4.2. ∎

The above result states that given a sender with information extraction capacity $\Xi(\mathscr{U})$, the zero-error capacity of the channel should be at least this number in order to extract maximum possible information from the sender. Alternatively, given the channel, the asymptotic rate of information extraction from any sender is bounded by the zero-error capacity of the channel. As long as both quantities are greater than unity, the receiver can extract exponentially large number of sequences from the sender.

The part (1) of the theorem states that if the Shannon capacity of the symmetric sender graph is less than the zero-error capacity of the channel, then the asymptotic rate of information extraction is simply the information extraction capacity of the sender. The part (2) of the theorem states that the receiver can recover an exponential number of

Table 3
Independence number of G_s^n and the rate (Example 5.1).

n	1	2	3	4	5	6
$\alpha(G_s^n)$	2	4	12	32	80	240
$\alpha(G_s^n)^{1/n}$	2	2	2.28	2.37	2.40	2.49

sequences even when the information extraction capacity of the sender is less than the zero-error capacity of the channel, provided $\Theta(G_c) > 1$.

We now present an example demonstrating the above results.

Example 5.1. Consider a sender with utility $\mathscr{U}$ on $\mathcal{X} = \{0, 1, 2, 3\}$ as

$$\mathscr{U} = \begin{pmatrix} 0 & -1 & -1 & -1 \\ 1 & 0 & -1 & -1 \\ -1 & 1 & 0 & -1 \\ 0.5 & -1 & 0.5 & 0 \end{pmatrix}.$$

It can be observed that the base graph G_s is a square graph and is shown in Fig. 6(a). We have that $\alpha(G_s) = 2$ and the largest independent sets are $\{0, 2\}$ and $\{1, 3\}$. However, it can be shown that the largest set that is feasible for $\mathcal{O}(\mathscr{U})$ is $\{0, 2, 3\}$ and hence $\Gamma(\mathscr{U}) = 3$. Moreover, $\Theta(G_s^{\mathsf{Sym}}) = 3$ and hence from Theorem 4.2 and Theorem 4.5, we get that $\Xi(\mathscr{U}) = 3$.

The independence numbers for $G_s^n, n \le 6$ is given Table 3. As in Examples 4.3 and 4.4, the independence numbers were numerically computed using the integer programming formulation. Since $\{0, 2, 3\}$ is feasible for $\mathcal{O}(\mathscr{U})$, it can be shown that $\alpha(G_s^3) \ge 3! = 12$ (cf. Lemma B.1 in Appendix B.1). Table 3 shows that in fact $\alpha(G_s^3) = 12$. Moreover, for $n = 6$, using feasibility of $\{0, 2, 3\}$, we get that $\alpha(G_s^6) \ge 6!/(2!)^3 = 90$. However, this is a weaker bound and Table 3 shows that $\alpha(G_s^6) = 240$. The second row of the table also shows the rate as a function of n. Independence numbers for $n > 6$ were not computed due to increased computational effort.

Now, consider a situation where the sender and the receiver communicate via a noisy channel. The input and the output space of the channel are $\mathcal{Y}$ and $\mathcal{Z}$ respectively, with $\mathcal{Y} = \mathcal{Z} = \{a, b, c, d\}$. The channel is defined by the probability distribution $P_{\mathbb{Z}|\mathbb{Y}}$ as

$$P_{\mathbb{Z}|\mathbb{Y}} = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \!\!\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix} \end{array},$$

where the row entries correspond to the output of the channel $\mathbb{Z}$ and the column entries correspond to the input to the channel $\mathbb{Y}$.

Consider the symbols $\{a, b\}$. Since both the symbols have a positive probability of generating the output b, i.e., $P_{\mathbb{Z}|\mathbb{Y}}(b|a)P_{\mathbb{Z}|\mathbb{Y}}(b|b) > 0$, the symbols $\{a, b\}$ can be *confused* with each other. Thus, a and b are adjacent in the confusability graph G_c induced by $P_{\mathbb{Z}|\mathbb{Y}}$ (cf. Definition 5.2). The confusability for other symbols can be deduced similarly which gives that the confusability graph G_c is a square. The confusability graph is shown in Fig. 6(b). Finally, since G_c is a perfect graph, the capacity is determined by the largest independent set of G_c, which is $\{a, c\}$ and hence $\Theta(G_c) = 2$.

For $n \in \{1, 2\}$, we have that $\alpha(G_s^n) = \alpha(G_c^{\boxtimes n})$ and hence from Theorem 5.1, the maximum information can be recovered from the sender. However, since $\Xi(\mathscr{U}) > \Theta(G_c)$, *asymptotically*, the maximum possible information cannot be recovered from the sender and hence from Theorem 5.1, we have that $\mathcal{R} = \Theta(G_c) = 2$. Table 3 shows that although the channel suffices for $n \in \{1, 2\}$, for travel histories of length greater than 2, the channel $P_{\mathbb{Z}|\mathbb{Y}}$ is not enough for maximum recovery of information.

On the other hand, consider a channel defined by the distribution $\hat{P}_{\mathbb{Z}|\mathbb{Y}}$ as

$$\hat{P}_{\mathbb{Z}|\mathbb{Y}} = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \end{array} \\ \begin{pmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 1 \end{pmatrix} \end{array}.$$

It can be observed that the symbols a and e cannot be confused with each other based on their outputs and the confusability graph induced by the channel $\hat{P}_{\mathbb{Z}|\mathbb{Y}}$, denoted as $\hat{G}_c$, is shown in Fig. 6(c).

In this case again, since $\hat{G}_c$ is a perfect graph, the capacity is $\Theta(\hat{G}_c) = 3$. Here, $\{a, c, e\}$ is the largest independent set in $\hat{G}_c$. This implies that the maximum possible information can be recovered from the sender. Also, from Theorem 5.1, we have that $\mathcal{R} = \Xi(\mathscr{U}) = 3$.

Fig. 7(a) shows the growth of independence number of G_s^n with increasing history lengths. The green hashed line is the curve of $\Xi(\mathscr{U})^n$ and the blue solid line is the curve of $\alpha(G_s^n)$. The red dotted line is the plot of $\alpha(G_s)^n$ in Fig. 7(a). It is interesting to note that although the capacity is determined at $n = 1$, it is not achieved even for history length of 6. Fig. 7(b) shows the rate of information extraction when the medium of communication is the channel given by $\hat{P}_{\mathbb{Z}|\mathbb{Y}}$. The rate monotonically increases with history lengths and approaches the maximum possible rate. The red dotted line is the rate of communication when $\mathbb{P}_{\mathbb{Z}|\mathbb{Y}}$ is the channel. □

This concludes our analysis on the topic of information extraction from a strategic sender. We have seen that the strategic setting demands a new line of analysis, that uses in part the traditional tools of information theory, but is rooted in concepts of game theory. It also leads to new concepts. Our main take away is that the *information extraction capacity of the sender*, a concept we defined and introduced in this paper, appears to be a fundamental quantity. It plays a role loosely analogous to that of the entropy of a source, characterizing the extent of information the sender can provide (or can be extracted from it). Future research will reveal the extent to which this analogy holds.

6. Conclusion

To conclude, inspired by the problem of screening of travellers with questionnaires, we considered a framework where a receiver attempted to extract information from a strategic sender. This setting was posed as a non-cooperative communication problem where the receiver (a health inspector) wishes to recover information from a misreporting sender (traveller) with zero probability of error. We considered a receiver-centric viewpoint and posed the problem as a leader–follower game with the receiver as the leader and sender as the follower. We formulated two instances of the game, with a noiseless channel, and with a noisy channel. We showed that even in the presence of the noisy channel, the receiver can extract an exponential number of sequences. To achieve this, the optimal choice of strategy for the receiver is to play a selective decoding strategy that decodes meaningfully only for a subset of sequences and deliberately induces an error on the rest of the sequences. The sequences are chosen such that the sender does not have an incentive to misreport any sequence as other, whereby, it tells the truth. In the context of designing questionnaires, this corresponds to the size of the optimal questionnaire that recovers maximum number of travel histories.

Our analysis led to new concepts: the rate of information extraction and the information extraction capacity of the sender. We showed that the maximum rate of information extraction is equal to the information extraction capacity of the sender in the noiseless channel case. In the presence of the noisy channel, the receiver can still extract information with this rate, provided the zero-error capacity of the channel is larger than the information extraction capacity of the sender. We

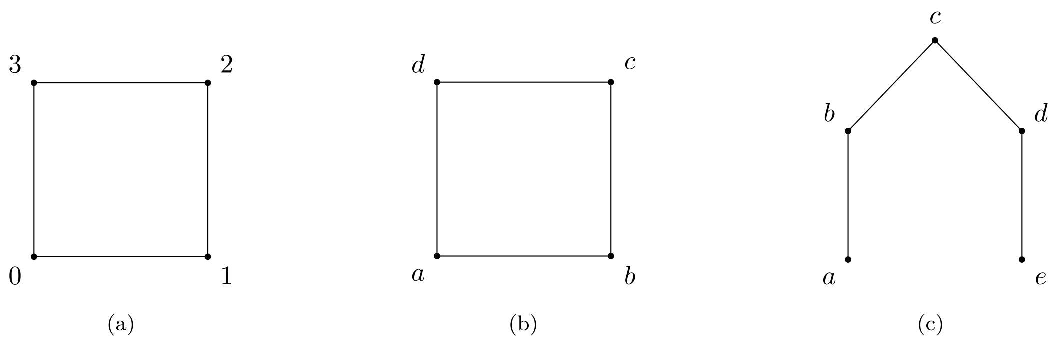


Fig. 6. (a) Sender graph G_s, (b) Confusability graph G_c, (c) Confusability graph $\widehat{G}_c$ (Examples 5.1).

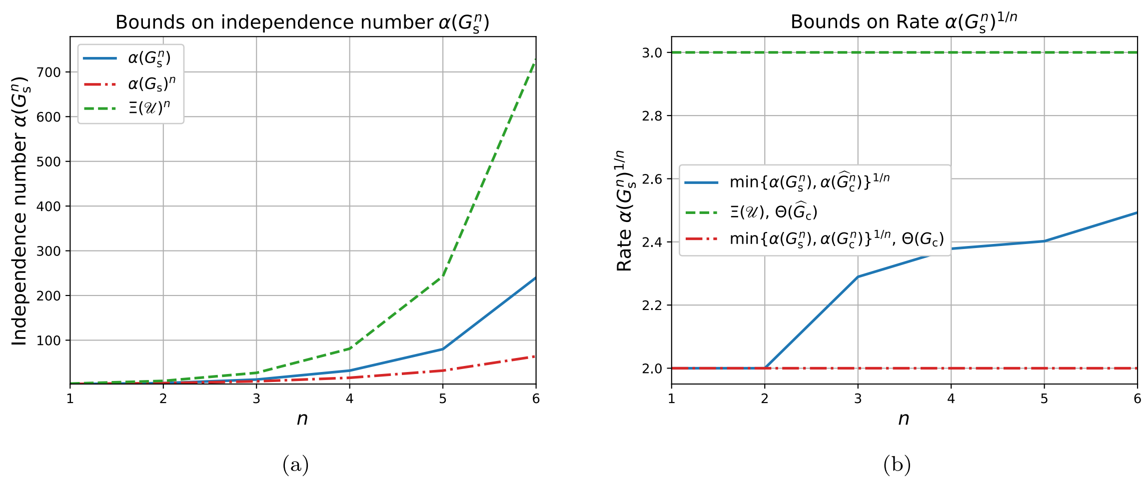


Fig. 7. (a) Bounds on independence number with n, (b) Bounds on the rate with n (Examples 5.1).

derived single-letter lower bounds and upper bounds. The lower bound is the optimal value of an optimization problem over permutation matrices. The upper bound is the Shannon capacity of the sender graph corresponding to the symmetric part of the utility. The information extraction capacity characterizes the fundamental limit to the amount of information that can be recovered with questionnaires.

CRediT authorship contribution statement

Anuj S. Vora: Writing – original draft, Methodology, Investigation, Formal analysis. **Ankur A. Kulkarni:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Data availability

No data was used for the research described in the article.

Acknowledgements

We thank the two anonymous reviewers for their evaluation of the paper. Their comments and suggestions have led to significant improvements in the exposition of the model presented in this paper. Specifically, we acknowledge the first reviewer for suggesting the re-organization of the results and the second reviewer for recommending addition of numerical examples.

This research was supported by the grant CRG/2019/002975 of the Science and Engineering Research Board, India, Department of Science and Technology, India.

Appendix A. Preliminaries

A.1. Proof of Theorem 3.1

Proof. For strategies g_n of the receiver such that $\min_{s_n \in \mathcal{S}(g_n)} |\mathcal{D}(g_n, s_n)| = 1$, the claim trivially holds. Let g_n be such that $|\mathcal{D}(g_n, s_n)| \geq 2$ for all strategies $s_n \in \mathcal{S}(g_n)$. We prove the claim by contradiction.

Suppose for some strategy $s_n \in \mathcal{S}(g_n)$, the set $\mathcal{D}(g_n, s_n)$ is not an independent set in G_s^n. Thus, there exists distinct sequences $\bar{x}, \hat{x} \in \mathcal{D}(g_n, s_n)$ such that $\mathcal{U}_n(\bar{x}, \bar{x}) \leq \mathcal{U}_n(\hat{x}, \bar{x})$. Using this, define a strategy $\bar{s}_n$ as

$$\bar{s}_n(x) = \begin{cases} s_n(x) & \forall\ x \neq \bar{x} \\ s_n(\hat{x}) & \text{for } x = \bar{x} \end{cases}. \tag{13}$$

Observe that $\bar{s}_n$ is also a best response since

$$\mathcal{U}_n(g_n \circ \bar{s}_n(x), x) = \mathcal{U}_n(g_n \circ s_n(x), x) \quad \forall\ x \neq \bar{x}$$

and for $x = \bar{x}$,

$$\begin{aligned} \mathcal{U}_n(g_n \circ \bar{s}_n(\bar{x}), \bar{x}) &= \mathcal{U}_n(g_n \circ s_n(\hat{x}), \bar{x}) \\ &= \mathcal{U}_n(\hat{x}, \bar{x}) \\ &\geq \mathcal{U}_n(\bar{x}, \bar{x}) = \mathcal{U}_n(g_n \circ s_n(\bar{x}), \bar{x}). \end{aligned} \tag{14}$$

Here (14) follows since $g_n \circ s_n(\hat{x}) = \hat{x}$, which in turn holds since $\hat{x} \in \mathcal{D}(g_n, s_n)$.

Now, for all $x \in \mathcal{D}(g_n, s_n) \backslash \{\bar{x}\}$, $g_n \circ \bar{s}_n(x) = g_n \circ s_n(x) = x$ and hence x lies in $\mathcal{D}(g_n, \bar{s}_n)$ and $\mathcal{D}(g_n, s_n)$. However, when $x = \bar{x}$, $g_n \circ \bar{s}_n(\bar{x}) = \hat{x} \neq \bar{x} = g_n \circ s_n(\bar{x})$. Thus, the sequence $\bar{x}$ lies in $\mathcal{D}(g_n, s_n)$ but is not recovered by the pair $(g_n, \bar{s}_n)$ and hence does not lie in $\mathcal{D}(g_n, \bar{s}_n)$. Thus, $|\mathcal{D}(g_n, \bar{s}_n)| < |\mathcal{D}(g_n, s_n)|$. However, this is a contradiction since $s_n \in \mathcal{S}(g_n)$. Thus, for all $s_n \in \mathcal{S}(g_n)$, the set $\mathcal{D}(g_n, s_n)$ is an independent set in G_s^n.

We now prove the second part of the result.

Since $\mathcal{D}(g_n, s_n)$ is an independent set in G_s^n for all g_n and for all $s_n \in \mathscr{B}(g_n)$, this implies that $R(g_n) \le \alpha(G_s^n)^{1/n}$ for all strategies g_n. We now show that for Stackelberg equilibrium strategies g_n^*, $R(g_n^*) = \alpha(G_s^n)^{1/n}$.

Consider an independent set I_n in G_s^n such that $|I_n| = \alpha(G_s^n)$ and define a strategy g_n for the receiver as

$$g_n(x) = \begin{cases} x & \text{if } x \in I_n \\ \Delta & \text{if } x \notin I_n \end{cases}.$$

Since Δ is never preferred by the sender, we can assume without loss of generality, that for all $s_n \in \mathscr{B}(g_n)$ and for all x, $g_n \circ s_n(x) \in I_n$ and hence $\mathcal{D}(g_n, s_n) \subseteq I_n$. We will now show that in fact the two sets are equal.

Consider an $x \in I_n$. For any $s_n \in \mathscr{B}(g_n)$, the utility of the sender is

$$\mathscr{U}_n(g_n \circ s_n(x), x) = \mathscr{U}_n(x', x)$$

for some $x' \in I_n$. Since I_n is an independent set in G_s^n, $\mathscr{U}_n(x', x) < 0$ for all $x' \in I_n, x' \neq x$. Since $x \in I_n$ was arbitrary,

$$\mathscr{U}_n(g_n \circ s_n(x), x) \le 0 \quad \forall\, x \in I_n,$$

with equality if and only if $g_n \circ s_n(x) = x$. Clearly, the optimal choice of s_n for the sender, is such that $s_n(x) = x$ for all $x \in I_n$. Specifically, all the strategies $s_n \in \mathscr{B}(g_n)$ are such that $s_n(x) = x$ for all $x \in I_n$. Thus, for all $s_n \in \mathscr{B}(g_n)$, $\mathcal{D}(g_n, s_n) = I_n$ and hence $R(g_n) = \alpha(G_s^n)^{1/n}$. It follows that for all Stackelberg equilibrium strategies g_n^* of the receiver, $R(g_n^*) = \alpha(G_s^n)^{1/n}$. ■

A.2. Proof of existence of information extraction capacity

Before proving the existence, we prove the following lemma.

Lemma A.1. *Let $m, n \in \mathbb{N}$. Consider a sender with utility $\mathscr{U}$ and the corresponding sender graph G_s^n. Then,*

$$\alpha(G_s^{m+n}) \ge \alpha(G_s^m)\alpha(G_s^n).$$

Proof. Consider an independent set I_m in G_s^m and an independent set I_n in G_s^n. The claim will follow by showing that $I_m \times I_n$ is an independent set in G_s^{m+n}.

Consider sequences $x, y \in \mathcal{X}^{m+n}$ such that $x = (w^m, w^n)$, $y = (v^m, v^n)$, where $w^m, v^m \in I_m$ and $w^n, v^n \in I_n$. Now,

$$\mathscr{U}_{m+n}(y, x) = \frac{m}{m+n}\mathscr{U}_m(v^m, w^m) + \frac{n}{m+n}\mathscr{U}_n(v^n, w^n).$$

Since I_m and I_n are independent sets and x, y are distinct, $\mathscr{U}_m(v^m, w^m) \le 0$ and $\mathscr{U}_n(v^n, w^n) \le 0$ with strict inequality in at least one the terms and hence, $\mathscr{U}_{m+n}(y, x) < 0$. This holds for all distinct sequences $x, y \in I_m \times I_n$ which shows that $I_m \times I_n$ is an independent set in G_s^{m+n}. Thus, $\alpha(G_s^{m+n}) \ge |I_m||I_n|$. Taking I_m and I_n to be the corresponding largest independent sets from G_s^m and G_s^n respectively, the claim follows. ■

Theorem A.2. *Consider a sender with utility $\mathscr{U}$ and let $\{G_s^n\}_{n\ge1}$ be the corresponding sequence of sender graphs. Then the limit in Definition 4.1 exists.*

Proof. From Lemma A.1, $\alpha(G_s^{m+n}) \ge \alpha(G_s^m)\alpha(G_s^n)$, for all $m, n \in \mathbb{N}$. Define $\beta_n = \log(\alpha(G_s^n))$ to get $\beta_{m+n} \ge \beta_m + \beta_n$. From Fekete's lemma (Schrijver, 2003), the limit of the sequence $\{\beta_n/n\}_{n\ge1}$ exists and is equal to $\sup_n \beta_n/n$. Using this and from the continuity and monotonicity of exp(.),

$$\lim_n \exp\left(\frac{\beta_n}{n}\right) = \sup_n \exp\left(\frac{\beta_n}{n}\right).$$

Substituting $\beta_n = \log(\alpha(G_s^n))$, the claim follows. ■

Appendix B. Lower bounds

B.1. Proof of Theorem 4.2

To prove Theorem 4.2, we first define a lemma. For that consider the following definition

Definition B.1. Let $K \in \mathbb{N}$ and $\mathcal{Y} \subseteq \mathcal{X}$. Define the set $T_{\mathcal{Y}}^K$ as

$$T_{\mathcal{Y}}^K = \left\{ x \in \mathcal{Y}^{K|\mathcal{Y}|} : P_x(i) = \frac{1}{|\mathcal{Y}|} \quad \forall\, i \in \mathcal{Y} \right\}.$$

Thus, the set $T_{\mathcal{Y}}^K$ is a set of all those sequences where every symbol from the set $\mathcal{Y}$ occurs exactly K times.

Using this definition, we define a set $T_{\mathcal{Y}_n}^K$,with $\mathcal{Y}_n \subseteq \mathcal{X}^n$, where $T_{\mathcal{Y}_n}^K$ consists of $nK|\mathcal{Y}_n|$-length sequences constructed by concatenating sequences from $\mathcal{Y}_n$, each appearing exactly K times. The following lemma gives a sufficient condition in terms of the optimization problem $\mathcal{O}(\mathscr{U}_n)$ for the independence of the set $T_{\mathcal{Y}_n}^K$.

Lemma B.1. *Let $n, K \in \mathbb{N}$. Consider a sender with utility $\mathscr{U}$ and let G_s^n be the corresponding sender graph. Let $\mathcal{Y}_n \subseteq \mathcal{X}^n$ be a set feasible for the problem $\mathcal{O}(\mathscr{U}_n)$. Then, $T_{\mathcal{Y}_n}^K$ is an independent set in the graph $G_s^{nK|\mathcal{Y}_n|}$.*

Proof. Fix a $K \in \mathbb{N}$. For distinct sequences $x, y \in T_{\mathcal{Y}_n}^K$,

$$\mathscr{U}_{nK|\mathcal{Y}_n|}(y, x) = \sum_{v,w\in\mathcal{Y}_n} P_{x,y}(v, w)\mathscr{U}_n(v, w) = \mathbb{E}_{P_{x,y}}\mathscr{U}_n(X, Y),$$

where the joint empirical distribution $P_{x,y}$ satisfies

$$\sum_{v'\in\mathcal{Y}_n} P_{x,y}(v', w) = \sum_{w'\in\mathcal{Y}_n} P_{x,y}(v, w') = \frac{1}{|\mathcal{Y}_n|} \quad \forall\, v, w \in \mathcal{Y}_n.$$

Define a matrix $D \in \mathbb{R}^{|\mathcal{Y}_n|\times|\mathcal{Y}_n|}$ where $D(v, v) = 1 - P(v)$, where $P(v) = 1/|\mathcal{Y}_n|$ for all $v \in \mathcal{Y}_n$ and $D(v, v') = 0$ for all $v \neq v'$. Then from Lemma 5 in Renault et al. (2013), we have that $P_{x,y} = V_{x,y} - D$, where $V_{x,y}$ is a convex combination of the permutation matrices $\mathcal{Q} = \{Q^{(0)}, \ldots, Q^{(|\mathcal{Q}|)}\}$ on $\mathcal{Y}_n$, i.e., $V_{x,y} = \sum_m \alpha_m Q^{(m)}$, with $\sum_m \alpha_m = 1$ and $\alpha_m \ge 0$ for all m. Thus,

$$\begin{aligned}\mathbb{E}_{P_{x,y}}\mathscr{U}_n(X, Y) &= \sum_{v,w\in\mathcal{Y}_n} V_{x,y}(v, w)\mathscr{U}_n(v, w) \\ &= \sum_m \alpha_m \sum_{v,w\in\mathcal{Y}_n} Q^{(m)}(v, w)\mathscr{U}_n(v, w).\end{aligned}$$

Since x and y are distinct, it follows that $\alpha_0 < 1$, where α_0 corresponds to the coefficient of the identity matrix $Q^{(0)}$. Thus,

$$\mathscr{U}_{nK|\mathcal{Y}_n|}(y, x) = \frac{1}{|\mathcal{Y}_n|} \sum_{m\in[|\mathcal{Q}|]} \alpha_m \sum_{v,w\in\mathcal{Y}_n} Q^{(m)}(v, w)\mathscr{U}_n(v, w).$$

Since $\mathcal{Y}_n$ is feasible for $\mathcal{O}(\mathscr{U})$, the term $\sum_{v,w\in\mathcal{Y}_n} Q^{(m)}(v, w)\mathscr{U}_n(v, w)$ is negative for all $m \in [|\mathcal{Q}|]$. Thus, $\mathscr{U}_{nK|\mathcal{Y}_n|}(y, x) < 0$ for all distinct $x, y \in T_{\mathcal{Y}_n}^K$ and hence, $T_{\mathcal{Y}_n}^K$ is an independent set in the graph $G_s^{nK|\mathcal{Y}_n|}$. ■

We now prove Theorem 4.2.

Proof of Theorem 4.2. From the feasibility condition of $\mathcal{O}(\mathscr{U}_n)$, it follows trivially that all independent sets of G_s^n are feasible for $\mathcal{O}(\mathscr{U}_n)$. Thus, $\Gamma(\mathscr{U}_n) \ge \alpha(G_s^n)$ and hence $\lim_{n\to\infty} \Gamma(\mathscr{U}_n)^{1/n} \ge \Xi(\mathscr{U})$.

Now we prove that $\lim_{n\to\infty} \Gamma(\mathscr{U}_n)^{1/n} \le \Xi(\mathscr{U})$. Let $K \in \mathbb{N}$. Consider a set $\mathcal{Y}_n$ that maximizes $\mathcal{O}(\mathscr{U}_n)$. Following Definition B.1, we define a set $T_{\mathcal{Y}_n}^K$ which consists of sequences constructed by concatenating all sequences from $\mathcal{Y}_n$, each appearing exactly K times. This construction gives sequences of length $nK|\mathcal{Y}_n|$. From Lemma B.1 it follows that $T_{\mathcal{Y}_n}^K$ is an independent set in $G_s^{nK|\mathcal{Y}_n|}$ for all $K \in \mathbb{N}$. Consequently, the capacity is bounded as

$$\Xi(\mathscr{U}) = \lim_{K\to\infty} \alpha(G_s^{nK|\mathcal{Y}_n|})^{\frac{1}{nK|\mathcal{Y}_n|}}$$

$$\geq \lim_{K\to\infty}\left(|T^K_{\mathcal{Y}_n}|^{\frac{1}{K|\mathcal{Y}_n|}}\right)^{1/n}.$$

We use Stirling's approximation to determine $\lim_{K\to\infty}|T^K_{\mathcal{Y}_n}|^{1/K|\mathcal{Y}_n|}$. Notice that $|T^K_{\mathcal{Y}_n}|$ is given as $|T^K_{\mathcal{Y}_n}| = (K|\mathcal{Y}_n|)!/(K!)^{|\mathcal{Y}_n|}$. Thus,

$$\Xi(\mathcal{U}) \geq \lim_{K\to\infty}\left(|T^K_{\mathcal{Y}_n}|^{\frac{1}{K|\mathcal{Y}_n|}}\right)^{1/n} = |\mathcal{Y}_n|^{1/n} = \Gamma(\mathcal{U}_n)^{1/n}.$$

Taking the limit as $n \to \infty$, the second claim follows. ■

B.2. Proof of bounds on $\Gamma(\mathcal{U}_n)$ in Theorem 4.2

The lower bound trivially follows from the feasibility condition of $\mathcal{O}(\mathcal{U}_n)$. To see the upper bound, note that if $\mathcal{Y}_n$ is feasible for $\mathcal{O}(\mathcal{U}_n)$, then

$$\mathcal{U}_n(x,y) + \mathcal{U}_n(y,x) < 0 \quad \forall\, x,y \in \mathcal{Y}_n, x \neq y.$$

This implies that for all distinct sequences $x, y \in \mathcal{Y}_n$, we have $\mathcal{U}_n^{\mathsf{Sym}}(x,y) = \mathcal{U}_n^{\mathsf{Sym}}(y,x) < 0$. Thus, $\mathcal{Y}_n$ is an independent set in $(G_{\mathsf{s}}^{\mathsf{Sym}})^n$ and the upper bound follows.

B.3. Proof of Lemma 4.3

Proof. Let (1) hold. The proof follows on the lines of Lemma 1 in Renault et al. (2013). For the sake of completeness, we repeat some of the arguments here.

Let $P \in \mathcal{P}(\mathcal{X})$ with $\mathrm{supp}(P) = \mathcal{Y}$ be any distribution and let $P_{X,Y} \in \mathcal{P}(\mathcal{Y}\times\mathcal{Y})$ be a distribution that satisfy $P_X = P_Y = P$. Repeating the arguments of the proof of Lemma B.1, we get that

$\mathbb{E}_{P_{X,Y}}\mathcal{U}(X,Y) = \sum_m \alpha_m \sum_{i,j\in\mathcal{Y}} Q^{(m)}(i,j)\mathcal{U}(i,j)$. Since $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$, the summation $\sum_{i,j\in\mathcal{Y}} Q^{(m)}(i,j)\mathcal{U}(i,j)$ is negative for all $Q^{(m)}$ that is not an identity matrix. Thus, $\mathbb{E}_{P_{X,Y}}\mathcal{U}(X,Y) \leq 0$ for all $P_{X,Y}$ and equality holds if and only if $P_{X,Y}$ is a diagonal matrix. Moreover, this holds for arbitrary $P \in \mathcal{P}(\mathcal{Y})$.

Let (2) hold. We will show that the support set of P, denoted by $\mathcal{Y}$, is a feasible set for $\mathcal{O}(\mathcal{U})$. Let Q be any $|\mathcal{Y}|\times|\mathcal{Y}|$ non-identity permutation matrix. Let $\hat{P}_{X,Y} \in \mathbb{R}^{|\mathcal{Y}|\times|\mathcal{Y}|}$ be a matrix such that $\hat{P}_{X,Y}(i,i) = P(i)$ for all $i \in \mathcal{Y}$ and $\hat{P}_{X,Y}(i,j) = 0$ for all $i \neq j$. Define a matrix $W_{X,Y}$ using Q as $W_{X,Y} = \hat{P}_{X,Y} + \beta(Q - \mathcal{I})$, where $\mathcal{I}$ is $|\mathcal{Y}| \times |\mathcal{Y}|$ identity matrix and taking $\beta \in (0, \min_{i\in\mathcal{Y}} P(i))$ ensures that the entries of $W_{X,Y}$ is non-negative. It is easy to check that the marginals of $W_{X,Y}$ are such that $W_X = P = W_Y$. Since (2) holds, we have that $\mathbb{E}_{W_{X,Y}}\mathcal{U}(X,Y) < 0$. Using the above form of $W_{X,Y}$, we get $\mathbb{E}_{W_{X,Y}}\mathcal{U}(X,Y) = \beta\sum_{i,j\in\mathcal{Y}} Q(i,j)\mathcal{U}(i,j)$. It follows that $\sum_{i,j\in\mathcal{Y}} Q(i,j)\mathcal{U}(i,j) < 0$. Since the matrix Q was arbitrary, we have the above inequality for all $|\mathcal{Y}|\times|\mathcal{Y}|$ non-identity permutation matrices Q. Thus, $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$. ■

B.4. Computational approach to check for the feasibility of a set $\mathcal{Y}$

We now discuss a computational method that can be used to check for the feasibility of a set $\mathcal{Y}$ for the problem $\mathcal{O}(\mathcal{U})$. Given a set of vertices $\mathcal{Y}$, its feasibility for $\mathcal{O}(\mathcal{U})$ can be checked using the Bellman–Ford algorithm in the following way. Define a weighted directed graph $\bar{G}_{\mathsf{s}}$ with vertices $\mathcal{X}$ induced by $\mathcal{U}$ where every vertex is connected to every other vertex and the edge from a vertex i to j has weight $\mathcal{U}(j,i)$. Define a utility $\mathcal{U}' = -\mathcal{U}$ and let the directed graph induced by $\mathcal{U}'$ be denoted as G'_{s}. Suppose we have a set $\mathcal{Y}$ that is feasible for the problem $\mathcal{O}(\mathcal{U})$. Then, from the characterization of feasible region of $\mathcal{O}(\mathcal{U})$, it implies that there is no zero or positive-weight directed cycle in the subgraph of $\bar{G}_{\mathsf{s}}$ induced by $\mathcal{Y}$. Here, weight of a cycle is the sum of the weights of the edges. Equivalently, there is no negative-weight or zero-weight directed cycle in the subgraph of G'_{s} induced by $\mathcal{Y}$. We use this observation as follows.

The Bellman–Ford algorithm (Cormen et al., 2009, Ch. 26), determines the shortest paths to all vertices in a directed graph from a given source vertex. The algorithm also detects whether the graph has a negative-weight directed cycle in the graph, in which case there may not exist a shortest path between two vertices. Suppose there is no zero-weight directed cycle in the graph G'_{s}. Then, this algorithm can be used to check for the feasibility of a given $\mathcal{Y}$ in the following manner.

- Given a set $\mathcal{Y}$ consider the subgraph of G'_{s} induced by the vertices $\mathcal{Y}$
- Apply the Bellman–Ford algorithm on this subgraph
- If the algorithm detects a negative-weight directed cycle in the subgraph, then the set $\mathcal{Y}$ is not feasible for $\mathcal{O}(\mathcal{U})$. Otherwise, $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$.

In the worst case, we need to check feasibility of all subsets of $\mathcal{X}$ in order to determine the optimal $\mathcal{Y}$. However, the total number of subsets of $\mathcal{X}$ may be very large which may make this procedure impractical. Nevertheless, the above method is still independent of the blocklength n and is thus an effective tool to derive a lower bound for the capacity.

B.5. Proof of Proposition 4.4

Recall the definition of a positive-edges cycle in Definition 4.5. We first demonstrate that it is necessary that a feasible set of $\mathcal{O}(\mathcal{U})$ does not contain a positive-edges cycle. Let $\mathcal{Y}$ be feasible for $\mathcal{O}(\mathcal{U})$. Let $i_0,\dots,i_{K-1}$ be distinct symbols from $\mathcal{Y}$ and suppose there exists a positive-edges cycle in $\mathcal{Y}$, such that $\mathcal{U}(i_l, i_m) \geq 0 \quad \forall\ l = (m+1)\mathrm{mod}\ K$. Clearly, this means that the set $\mathcal{Y}$ does not satisfy the condition in (7) and hence $\mathcal{Y}$ is not feasible for $\mathcal{O}(\mathcal{U})$.

Suppose there is no positive-edges cycle in $\mathcal{Y}$ and there is at least one pair of (i_k, i_j), $k > j$ such that $\mathcal{U}(i_k, i_j) < 0$. Then,

$$\begin{aligned}
&\mathcal{U}(i_1,i_0) + \mathcal{U}(i_2,i_1) + \ldots + \mathcal{U}(i_0,i_{K-1}) \\
&\leq \min_{i,j\in\mathcal{Y}:\mathcal{U}(i,j)<0} |\mathcal{U}(i,j)| + (|\mathcal{Y}|-1)\max_{i,j\in\mathcal{Y}:\mathcal{U}(i,j)\geq 0} \mathcal{U}(i,j) \\
&< 0.
\end{aligned}$$

Since $i_0,\dots,i_{K-1}$ were arbitrary, it follows that $\mathcal{Y}$ is feasible for $\mathcal{O}(\mathcal{U})$.

Appendix C. Upper bounds

C.1. Proof of Theorem 4.5

Proof. Let $n \in \mathbb{N}$. From Theorem 4.2, we have that $\alpha(G_{\mathsf{s}}^n) \leq \alpha((G_{\mathsf{s}}^{\mathsf{Sym}})^n)$. The claim follows by taking the limit of the nth root.

We now prove the second part of the result. Let $n \in \mathbb{N}$. We first prove that if $\mathcal{U}$ is symmetric then $\Xi(\mathcal{U}) \leq \Theta(\mathcal{U})$. Consider the graph $G_{\mathsf{s}}^{\boxtimes n}$ derived by taking the n-fold strong product of G_{s}. Consider a distinct pair of sequences $x, y \in \mathcal{X}^n$ that are adjacent in $G_{\mathsf{s}}^{\boxtimes n}$. Then, for all $k \in [n]$, either $x_k = y_k$ or x_k, y_k are adjacent in G_{s}. Thus,

$$\mathcal{U}_n(y,x) = \frac{1}{n}\sum_{y_k\neq x_k}\mathcal{U}(y_k,x_k) \geq 0.$$

Hence, x, y are adjacent in G_{s}^n as well, which gives $G_{\mathsf{s}}^{\boxtimes n}$ is a subgraph of G_{s}^n. Thus, $\alpha(G_{\mathsf{s}}^n) \leq \alpha(G_{\mathsf{s}}^{\boxtimes n})$ and taking the limit of the nth root and using the first part of the result, the claim follows. ■

C.2. Proof of Proposition 4.6

Proof. Let $n \in \mathbb{N}$. We prove this by showing that if distinct sequences $x, y \in \mathcal{X}^n$ are not adjacent in G_{s}^n, then x, y are not adjacent in $G_{\mathsf{s}}^{\boxtimes n}$.

Let x, y be not adjacent in G_{s}^n. Then, $\mathcal{U}_n(y,x) < 0$ and $\mathcal{U}_n(x,y) < 0$. Since, $b \leq a$, it follows that

$$\left|\{k : \mathcal{U}(y_k,x_k) = -b\}\right| > \frac{n}{2},$$
$$\left|\{k : \mathcal{U}(x_k,y_k) = -b\}\right| > \frac{n}{2}.$$

This implies that there is some k such that $\mathcal{U}(y_k,x_k) = \mathcal{U}(x_k,y_k) = -b$ and hence x_k, y_k are not adjacent in G_{s}. It follows that x, y are not adjacent in $G_{\mathsf{s}}^{\boxtimes n}$ as well. This gives that $\alpha(G_{\mathsf{s}}^n) \leq \alpha(G_{\mathsf{s}}^{\boxtimes n})$. The claim follows by taking the limit. ■

C.3. Proof of Theorem 4.7

Proof.

1. Follows from Theorem 4.5 and the fact that for a perfect graph, $\Theta(G_{\mathrm{s}}) = \alpha(G_{\mathrm{s}})$ (Lovász, 1979).
2. Follows from Proposition 4.6 and the fact that for a perfect graph, $\Theta(G_{\mathrm{s}}) = \alpha(G_{\mathrm{s}})$. ■

Appendix D. Noisy channel results

Before we prove the results, we state a few definitions. For any strategy g_n of the receiver, recall the worst-case best response set $\mathcal{S}(g_n)$ defined as

$$\mathcal{S}(g_n) = \arg\min_{s_n\in\mathscr{B}(g_n)} |\mathcal{D}(g_n, s_n)|.$$

Let

$$\mathcal{Z}(y) = \mathrm{supp}(P_{Z|Y}(\cdot|y)),$$

where y is an input to the channel. Note that since the output space of the channel is $\mathcal{Z}^n$, $\mathcal{Z}(y) \subseteq \mathcal{Z}^n$ for all y.

D.1. Proof of Theorem 5.1

Proof. Consider a Stackelberg equilibrium strategy g_n^* of the receiver. It is known from Theorem 3.1, that the set of recovered sequences from the sender can be at most $\alpha(G_{\mathrm{s}}^n)$ and hence $\min_{s_n\in\mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)| \leq \alpha(G_{\mathrm{s}}^n)$. Further, at most $\alpha(G_{\mathrm{c}}^n)$ sequences can be transmitted with zero error through the channel and hence for all s_n, $|\mathcal{D}(g_n^*, s_n)| \leq \alpha(G_{\mathrm{c}}^n)$. Together, it follows that

$$\min_{s_n\in\mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)| \leq \min\{\alpha(G_{\mathrm{s}}^n), \alpha(G_{\mathrm{c}}^n)\}.$$

We now show that equality holds in the above relation.

Let $d = \min\{\alpha(G_{\mathrm{s}}^n), \alpha(G_{\mathrm{c}}^n)\}$. Clearly, there exists an independent set I_n^{s} in G_{s}^n such that $|I_n^{\mathrm{s}}| = d$. Similarly, there exists an independent set I_n^{c} in G_{c}^n such that $|I_n^{\mathrm{c}}| = d$. Let the sequences in the sets I_n^{s} and I_n^{c} be denoted as x^i and y^i respectively, with $i \in [d]$ and $x^i \in \mathcal{X}^n, y^i \in \mathcal{Y}^n$. With this convention, define the strategy g_n as

$$g_n(z) = \begin{cases} x^i & \text{if } z \in \mathcal{Z}(y^i) \\ \Delta & \text{if } z \notin \bigcup_{i=1}^d \mathcal{Z}(y^i) \end{cases}. \tag{15}$$

Note that $\mathcal{Z}(y^i)$ are disjoint sets. We show that the strategy g_n of the receiver ensures that all strategies s_n in the set of best responses $\mathscr{B}(g_n)$ are such that $\mathcal{Z}(s_n(x^i)) = \mathcal{Z}(y^i)$ for all $i \in [d]$. Fix an index i, let s_n be any strategy for the sender and let $s_n(x^i) = y^* \in \mathcal{Y}^n$. Notice that if $\mathcal{Z}(y^*) \nsubseteq \bigcup_{j=1}^d \mathcal{Z}(y^j)$, then $g_n(z) = \Delta$ for some $z \in \mathcal{Z}(y^*)$, $z \notin \bigcup_{j=1}^d \mathcal{Z}(y^j)$. This gives that $\mathbb{E}\left[\mathcal{U}_n(g_n(Z), x^i)\right] = -\infty$. Thus, y^* is such that $\mathcal{Z}(y^*) \subseteq \bigcup_{j=1}^d \mathcal{Z}(y^j)$.

Writing $\mathcal{Z}(y^*) = \bigcup_{j=1}^d \mathcal{Z}(y^*) \cap \mathcal{Z}(y^j)$, it follows that

$$\begin{aligned}\mathbb{E}\left[\mathcal{U}_n(g_n(Z), x^i)\right] &= \sum_{z\in \bigcup_{j=1}^d \mathcal{Z}(y^*)\cap\mathcal{Z}(y^j)} P_{Z|Y}(z|y^*)\mathcal{U}_n(g_n(z), x^i) \\ &= \sum_{j=1}^d \sum_{z\in\mathcal{Z}(y^*)\cap\mathcal{Z}(y^j)} P_{Z|Y}(z|y^*)\mathcal{U}_n(x^j, x^i).\end{aligned}$$

The last equation follows from the definition of g_n. Since I_n^{s} is an independent set in G_{s}^n, $\mathcal{U}_n(x^j, x^i) < \mathcal{U}_n(x^i, x^i) = 0$ for all $j \neq i$ and hence

$$\sum_{j=1}^d \sum_{z\in\mathcal{Z}(y^*)\cap\mathcal{Z}(y^j)} P_{Z|Y}(z|y^*)\mathcal{U}_n(x^j, x^i) \leq 0$$

with equality if and only if $\mathcal{Z}(y^*) = \mathcal{Z}(y^i)$. Thus, the strategy y^* is such that $\mathcal{Z}(y^*) = \mathcal{Z}(y^i)$. Clearly, this holds for all $s_n \in \mathscr{B}(g_n)$ and for all $i \in [d]$.

It easy to see that the utility of the sender and the receiver do not depend on the exact choice of y^* so long as $\mathcal{Z}(y^*) = \mathcal{Z}(y^i)$. Hence, without loss of generality, we consider s_n to be such that $s_n(x^i) = y^i$ for all $i \in [d]$. Thus, when the sequence $x^i \in I_n^{\mathrm{s}}$ is observed by the sender, it encodes it as y^i. The channel generates an output z, which belongs to the support set $\mathcal{Z}(y^i)$. The receiver maps all such z to x^i thereby ensuring $\mathbb{P}(\widehat{X} = x^i|X = x^i) = 1$ for all $i \in [d]$. Thus, for all $x^i \in I_n^{\mathrm{s}}$ and $s_n \in \mathscr{B}(g_n)$, $x^i \in \mathcal{D}(g_n, s_n)$. Hence $\mathcal{D}(g_n, s_n) \supseteq I_n^{\mathrm{s}}$ for all $s_n \in \mathscr{B}(g_n)$. Since $|I_n^{\mathrm{s}}| = d$, $|\mathcal{D}(g_n, s_n)| \geq d$. Using $|\mathcal{D}(g_n, s_n)| \leq d = \min\{\alpha(G_{\mathrm{s}}^n), \alpha(G_{\mathrm{c}}^n)\}$, it follows that for all Stackelberg equilibrium strategies g_n^*, $\min_{s_n\in\mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)| = \min\{\alpha(G_{\mathrm{s}}^n), \alpha(G_{\mathrm{c}}^n)\}$.

We now determine the asymptotic rate of information extraction. It can be shown that, for a discrete memoryless channel given by (9), the graph G_{c}^n is same as the graph constructed by taking n-fold strong product of G_{c} (Shannon, 1956), i.e., $G_{\mathrm{c}}^n = G_{\mathrm{c}}^{\boxtimes n}$. Thus, $\min_{s_n\in\mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)| = \min\{\alpha(G_{\mathrm{s}}^n), \alpha(G_{\mathrm{c}}^{\boxtimes n})\}$ and hence

$$\begin{aligned}R(g_n^*) &= \min_{s_n\in\mathscr{B}(g_n^*)} |\mathcal{D}(g_n^*, s_n)|^{1/n} \\ &= \min\left\{\alpha(G_{\mathrm{s}}^n)^{1/n}, \alpha(G_{\mathrm{c}}^{\boxtimes n})^{1/n}\right\}.\end{aligned}$$

The claim follows after taking the limit. ■

Thus, from the strategy defined in (15), it can be observed that the receiver decodes meaningfully only for a subset of sequences. In particular, it chooses d number of inputs, y^i with $i \in [d]$, which can be distinguished from each other and maps the respective support sets $\mathcal{Z}(y^i)$, to d distinct sequences x^i. For the rest of the outputs from the channel, the receiver declares an error Δ. In response, the optimal strategy for the sender is (without loss of generality) such that it maps the sequences x^i to the inputs y^i. Also, for all strategies s_n and for all $x \in \mathcal{X}^n \setminus \{x^i\}_{i\in[d]}$, $s_n(x) = y^*$, where y^* is such that $\mathcal{Z}(y^*) \subseteq \bigcup_{j=1}^d \mathcal{Z}(y^j)$. This is because for any other input $y' \neq y^*$, if $\mathcal{Z}(y') \nsubseteq \bigcup_{j=1}^d \mathcal{Z}(y^j)$, then the channel output z may lie outside $\bigcup_{j=1}^d \mathcal{Z}(y^j)$ for which $g_n(z) = \Delta$ and the corresponding utility is $-\infty$.

References

Akyol, E., Langbort, C., Başar, T., 2015. Privacy constrained information processing. In: 54th IEEE Conference on Decision and Control. CDC, IEEE, pp. 4511–4516. http://dx.doi.org/10.1109/CDC.2015.7402924.

Akyol, E., Langbort, C., Başar, T., 2016. Information-theoretic approach to strategic communication as a hierarchical game. Proc. IEEE 105 (2), 205–218. http://dx.doi.org/10.1109/JPROC.2016.2575858.

Ambrosio, L., Gigli, N., Savaré, G., 2008. Gradient Flows: in Metric Spaces and in the Space of Probability Measures. Springer Science & Business Media, http://dx.doi.org/10.1007/978-3-7643-8722-8.

Ambrus, A., Azevedo, E.M., Kamada, Y., 2013. Hierarchical cheap talk. Theor. Econ. 8 (1), 233–261. http://dx.doi.org/10.3982/TE1038.

Ambrus, A., Takahashi, S., 2008. Multi-sender cheap talk with restricted state spaces. Theor. Econ. https://econtheory.org/ojs/index.php/te/article/view/20080001/0.

Başar, T., Olsder, G., 1999. Dynamic Noncooperative Game Theory, second ed. SIAM, Philadelphia, http://dx.doi.org/10.1137/1.9781611971132.

Battaglini, M., 2002. Multiple referrals and multidimensional cheap talk. Econometrica 70 (4), 1379–1401. http://dx.doi.org/10.1111/1468-0262.00336.

Bergemann, D., Morris, S., 2019. Information design: A unified perspective. J. Econ. Lit. 57 (1), 44–95. http://dx.doi.org/10.1257/jel.20181489.

Conforti, M., Cornuéjols, G., Zambelli, G., 2014. Integer Programming, vol. 271, Springer, http://dx.doi.org/10.1007/978-3-319-11008-0.

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2009. Introduction to Algorithms. MIT Press.

Cover, T.M., Thomas, J.A., 2012. Elements of Information Theory. John Wiley & Sons, http://dx.doi.org/10.1002/047174882X.

Crawford, V.P., Sobel, J., 1982. Strategic information transmission. Econometrica 1431–1451. http://dx.doi.org/10.2307/1913390.

Farokhi, F., Teixeira, A.M., Langbort, C., 2016. Estimation with strategic sensors. IEEE Trans. Autom. Control 62 (2), 724–739. http://dx.doi.org/10.1109/TAC.2016.2571779.

Jackson, M.O., Sonnenschein, H.F., 2007. Overcoming incentive constraints by linking decisions. Econometrica 75 (1), 241–257. http://dx.doi.org/10.1111/j.1468-0262.2007.00737.x.

Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Amer. Econ. Rev. 101 (6), 2590–2615. http://dx.doi.org/10.1257/aer.101.6.2590.

Korner, J., Orlitsky, A., 1998. Zero-error information theory. IEEE Trans. Inform. Theory 44 (6), 2207–2229. http://dx.doi.org/10.1109/18.720537.

Lapidoth, A., 1997. On the role of mismatch in rate distortion theory. IEEE Trans. Inform. Theory 43 (1), 38–47. http://dx.doi.org/10.1109/18.567644.

Le Treust, M., Tomala, T., 2019. Persuasion with limited communication capacity. J. Econom. Theory 184, 104940. http://dx.doi.org/10.1016/j.jet.2019.104940.

Lovász, L., 1979. On the Shannon capacity of a graph. IEEE Trans. Inform. Theory 25 (1), 1–7. http://dx.doi.org/10.1109/TIT.1979.1055985.

Myerson, R.B., 1997. Game Theory: Analysis of Conflict. Harvard University Press, http://dx.doi.org/10.2307/j.ctvjsf522.

Polak, S.C., Schrijver, A., 2019. New lower bound on the Shannon capacity of C7 from circular graphs. Inform. Process. Lett. 143, 37–40. http://dx.doi.org/10.1016/j.ipl.2018.11.006.

Renault, J., Solan, E., Vieille, N., 2013. Dynamic sender–receiver games. J. Econom. Theory 148 (2), 502–534. http://dx.doi.org/10.1016/j.jet.2012.07.006.

Renou, L., Tomala, T., 2015. Approximate implementation in Markovian environments. J. Econom. Theory 159, 401–442. http://dx.doi.org/10.1016/j.jet.2015.07.009.

Sarıtaş, S., Yüksel, S., Gezici, S., 2015. On multi-dimensional and noisy quadratic signaling games and affine equilibria. In: American Control Conference. ACC, IEEE, pp. 5390–5395. http://dx.doi.org/10.1109/ACC.2015.7172182.

Sayin, M.O., Akyol, E., Başar, T., 2019. Hierarchical multistage Gaussian signaling games in noncooperative communication and control systems. Automatica 107, 9–20. http://dx.doi.org/10.1016/j.automatica.2019.05.024.

Scarlett, J., Somekh-Baruch, A., Martinez, A., et al., 2020. Information-theoretic foundations of mismatched decoding. arXiv preprint arXiv:2003.11694.

Schrijver, A., 2003. Combinatorial Optimization: Polyhedra and Efficiency, vol. 24, Springer Science & Business Media.

Shannon, C., 1956. The zero error capacity of a noisy channel. IRE Trans. Inf. Theory 2 (3), 8–19. http://dx.doi.org/10.1109/TIT.1956.1056798.

Vora, A.S., Kulkarni, A.A., 2020a. Achievable rates for strategic communication. In: 2020 IEEE International Symposium on Information Theory. ISIT, IEEE, pp. 1379–1384. http://dx.doi.org/10.1109/ISIT44484.2020.9174307.

Vora, A.S., Kulkarni, A.A., 2020b. Communicating with a strategic sender. In: Twenty-Sixth National Conference on Communications. NCC, IEEE, pp. 1–6. http://dx.doi.org/10.1109/NCC48643.2020.9056042.

Vora, A.S., Kulkarni, A.A., 2020c. Information extraction from a strategic sender over a noisy channel. In: 59th IEEE Conference on Decision and Control. CDC, pp. 354–359. http://dx.doi.org/10.1109/CDC42340.2020.9303935.

Vora, A.S., Kulkarni, A.A., 2020d. Zero error strategic communication. In: 2020 International Conference on Signal Processing and Communications. SPCOM, IEEE, pp. 1–5. http://dx.doi.org/10.1109/SPCOM50965.2020.9179592.

Vora, A.S., Kulkarni, A.A., 2021. Optimal questionnaires for screening of strategic agents. In: IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, IEEE, pp. 8173–8177. http://dx.doi.org/10.1109/ICASSP39728.2021.9414345.

West, D.B., 2000. Introduction to Graph Theory, second ed. Prentice Hall.