Finding red balloons with split contracts: robustness to individuals' selfishness

Manuel Cebrian
MIT
cebrian@mit.edu

Lorenzo Coviello
UC San Diego
lcoviell@eng.ucsd.edu

Andrea Vattani
UC San Diego
avattani@cs.ucsd.edu

Panagiotis Voulgaris
Google Inc.
voulg@google.com

Abstract

The present work deals with the problem of information acquisition in a strategic networked environment. To study this problem, Kleinberg and Raghavan (FOCS 2005) introduced the model of *query incentive networks*, where the root of a binomial branching process wishes to retrieve an information – known by each node independently with probability $1/n$ – by investing as little as possible. The authors considered *fixed-payment contracts* in which every node strategically chooses an amount to offer its children (paid upon information retrieval) to convince them to seek the information in their subtrees. Kleinberg and Raghavan discovered that the investment needed at the root exhibits an unexpected threshold behavior that depends on the branching parameter b. For $b > 2$, the investment is linear in the expected distance to the closest information (logarithmic in n, the rarity of the information), while, for $1 < b < 2$, it becomes exponential in the same distance (i.e., polynomial in n). Arcaute et al. (EC 2007) later observed the same threshold behavior for arbitrary Galton-Watson branching processes.

The DARPA Network Challenge — retrieving the locations of ten balloons placed at undisclosed positions in the US — has recently brought practical attention to the problems of social mobilization and information acquisition in a networked environment. The MIT Media Laboratory team won the challenge by acting as the root of a query incentive network that unfolded all over the world. However, rather than adopting a *fixed-payment* strategy, the team implemented a different incentive scheme based on $1/2$-*split contracts*. Under such incentive scheme, a node u who does not possess the information can recruit a friend v through a contract stipulating that if the information is found in the subtree rooted at v, then v has to give half of her own reward back to u.

Motivated by its empirical success, we present a comprehensive theoretical study of this scheme in the game theoretical setting of query incentive networks. Our main result is that split contracts are robust — as opposed to fixed-payment contracts— to nodes' selfishness. Surprisingly, when nodes determine the splits to offer their children based on the contracts received from their recruiters, the threshold behavior observed in the previous work vanishes, and an investment linear in the expected distance to the closest information is sufficient to retrieve the information in *any arbitrary* Galton-Watson process with $b > 1$. Finally, while previous analyses considered the parameters of the branching process as constants, we are able to characterize the rate of the investment in terms of the branching process and the desired probability of success. This allows us to show improvements even in other special cases.

1 Introduction

A challenging class of crowdsourcing problems requires an interested party to provide incentives for large groups of people to contribute to the search and retrieval of rare information [20, 14, 7]. The small world problem, i.e. distributed routing of messages to unknown individuals, is the seminal example of this class and has illustrated the difficulty of the approach for almost 50 years [15, 22, 18, 4, 23]. In this class of problems, individuals in the social network act as intermediaries to create a channel between the querier and the answer. Observe that the chief difficulty of this approach is to offer incentives to the individuals to propagate the query further in the network as well as to return the answer all the way back to the querier [4]. The goal is therefore to incentivize participation of the users using some form of (possibly financial) reward. In this way, a node who does not know the answer but is offered a sufficiently high reward can act as intermediary and propagate the query by offering the neighbors a share of its reward. This setting models the social network as a marketplace of information where the users strategically act in order to maximize their utility, and raises several questions about the system's performance and the incentive propagation, the main one being: can we retrieve the answer to a difficult query when given a limited budget?

The Defense Advanced Research Projects Agency (DARPA), a research organization of the United States Department of Defense, designed a so called "Network Challenge" that conveyed a positive answer to this question.[1] The challenge consisted of locating ten moored red weather balloons placed at ten undisclosed locations in the continental United States. A single $40,000 cash prize was allocated for the first participant to submit the correct latitude and longitude (within one mile error) of all ten balloons within the contest period. In particular, the competition consisted in recruiting a team to achieve the goal. This task posed varied issues of large-scale, time-critical mobilization. In particular, in order to guarantee the participation and coordination of a large team, an adequate structure of economic incentives had to be built.

The MIT Media Laboratory team won the competition in less than 9 hours, adopting a recruitment scheme based on recursive incentives.[2] Specifically, using the $40,000 they could possibly win, they allocated an amount of $4,000 for finding each balloon. For each balloon, they would distribute the $4,000 up the chain of participants leading to successful balloon spotting, as described in their website: "[In the case we win the competition,] we're giving $2,000 per balloon to the first person to send us the correct coordinates, but that's not all – we're also giving $1000 to the person who invited them. Then we're giving $500 whoever invited the inviter, and $250 to whoever invited them, and so on...". This is equivalent to say that a node u who does not have the desired answer, can offer its friends a $1/2$-*split* contract, stipulating that if the answer is found in the subtree of a child v of u, then u will get back from v a $1/2$ fraction of whatever amount v gets. However, if u is not the querier, the total amount pocketed by u is less, as u has to give a $1/2$ fraction of its reward to its recruiter.

While the success of this strategy has been hailed as an empirical testimony to the power of incentive structures [21], the theoretical efficiency of the proven scheme has remained an open question, and motivates this work. In particular, we analyze this economic structure in the model for query incentive networks introduced By Kleinberg and Raghavan in [12]. This model considers a competitive environment where every node plays strategically. To fit the *split contracts* to this model, we generalize the splits to any fraction $0 < \rho < 1$, in the sense that any node u can offer a child v a ρ-split contract stipulating the following: if v has the answer, then v would pocket a $(1-\rho)$ of the whole reward while returning a fraction ρ to u; if v does not have the answer, then v can in turn offer some ρ'-split to its (still unrecruited) friends, and so on. Given the strategic setting, nodes will choose the splits to offer to their children so to maximize their expected payoffs; observe that contracts between different nodes can have different splits — and this is indeed the

[1] `https://networkchallenge.darpa.mil/`

[2] `http://balloon.media.mit.edu/`

case in the Nash equilibrium as our results show. The details of the original model introduced in [12] follow.

QUERY INCENTIVE NETWORKS. The scenario of interest is that of a node, the root, that is willing to invest some amount r^* to retrieve certain information from a large network in which every node plays strategically. The main goal is to characterize the tradeoff between the investment and the rarity of the information. The model, introduced by Kleinberg and Raghavan [12], is as follows: the querier node is the root of an infinite d-ary tree, where each node possesses independently the desired information with probability $1/n$, where n represents the rarity of the answer. The root offers each child u a "*fixed-payment*" contract of r^*, stipulating that the root will pay u that amount upon u providing the answer. The query propagates down the tree according to the following scheme: every node u has an integer-valued function f_u encoding its strategies; if u is offered a reward of r by its parent and does not possess the answer, then in turn it offers a reward of $1 \le f_u(r) \le r-1$ to its children. When the answer to the query is found, the root selects for payment one among the answer-holders using a fixed non-strategic rule. The payment is then propagated down through the path to that selected node, with each node along the path pocketing its share. If an intermediate node u on this path was offered r by its parent, then its overall payoff is $r - f_u(r) - 1$, where the *unit cost* is associated with the act of returning the answer[3]. The game-theoretical aspect of the model is that any node u chooses the function f_u so to maximize its payoff. To break ties, it is assumed that a node who is offered a reward of one (and does not possess the answer) will always forward the query to its children, even if its expected payoff is zero (since the unit reward would be spent when returning the answer up to its parent).

As pointed out in [12], there is a subtle deficiency with a deterministic tree: the Nash equilibria of a game played in a deterministic network tacitly assume that the nodes know the entire network. Indeed, in a Nash equilibrium, each node chooses its best strategy by knowing the strategies of every other node. However, this is unrealistic, as we want to model a setting where nodes are only aware of their neighbors. To deal with this technical issue, Kleinberg and Raghavan consider a network that can be thought as a branching process from the root. In particular, the number of children of each node is chosen independently from a binomial distribution $\mathrm{Bin}(d, q)$, where q is a constant probability of a node being present. The expected number of children of a node — i.e., the branching factor — is then $b = qd$. By classical results in the theory of branching processes, if $b < 1$ the process dies out almost surely; therefore there is no amount that the root can offer to obtain an answer with constant probability if the rarity n of the answer is large enough. Instead, for any $b > 1$, there is a constant non-zero probability that the process will generate infinitely many nodes, so that the answer is present within the first $O(\log n)$ levels of the tree with high probability. Nevertheless, Kleinberg and Raghavan show that in the Nash equilibrium the investment needed at the root can be much larger than logarithmic in n. Specifically, while an investment $r^* = O(\log n)$ is sufficient to retrieve the answer with constant probability for $b > 2$, an investment of $r^* = n^{\Theta(1)}$ is needed when $1 < b < 2$. That is, in the latter case the root must invest a reward that is exponentially larger than the expected distance from the closest answer.

Arcaute *et al*. [1] generalized the work in [12] showing that this threshold behavior at $b = 2$ still holds for arbitrary Galton-Watson branching process. They also proved that in a *ray* — a deterministic infinite path ($b = 1$, but with zero extinction probability) — the reward needed is super-exponential in the expected depth of the search tree, that is $r^* = \Omega(n!)$. Finally, they observed that this threshold behavior vanishes if the root desires to find the answer with probability tending to 1: if the desired probability is $1 - 1/n$, then

[3]As observed in [12], if nodes placed no value on this answering effort then the root could simply invest an arbitrarily small reward $\epsilon > 0$, and it would retrieve an answer because each node would have a positive payoff from participating in the game and returning the answer. To avoid this situation, a unit price is placed on the effort of returning the answer, while the cost of participating to the game is zero. This is motivated by the fact that the cost of forwarding requests to a list of friends is typically considered negligible in peer-to-peer and social-network systems [10, 24, 25] (see [12] for additional details on the motivations).

for any branching process with $b > 1$ and no extinction, the needed reward is $n^{\Theta(1)}$.

Our results. We present a theoretical study of the multi-level marketing strategies adopted by the winning team of the DARPA Network Challenge. Given the strong affinity between this challenge and the model of query incentive networks introduced in [12, 1], we frame these strategies in this model by considering split contracts as the possible offers between nodes.

Our main result is that split contracts, unlike fixed-payment contracts, are robust to a strategic environment, where every node selfishly determines the offers to its children based on the offer received from its parent. We show that for any constant $\epsilon > 0$ and Galton-Watson branching process with $b > 1$, the Nash equilibrium with split contracts uses an investment of $r^* = O(\log n)$ to retrieve the answer with probability at least $1 - \zeta - \epsilon$, where ζ is the extinction probability of the process. As the expected distance to the closest answer is $\Theta(\log n)$ and nodes pay a unit cost to return the answer, this is a constant approximation with respect to an ideal centralized non-strategic setting. In other words, the price of anarchy of the game with split contracts is constant (ignoring some pathological equilibria, see Section 4 and Appendix H).

Unlike previous work that assumed the parameters of the branching process to be held constant, we are also able to characterize the dependence of the investment with respect to the branching process and the success accuracy. This allows us to show additional improvements of split contracts over fixed-payment contracts: for example, for branching processes with no extinction, an investment of $O(n \log n)$ is enough to retrieve the answer with probability at least $1 - 1/n$, improving upon the $n^{\Theta(1)}$ investment provided in [1]. In fact, our result is even stronger since it guarantees a success probability of at least $1 - \zeta - 1/n$ in general branching processes. In the case of a ray (where the expected distance from the closest answer is n), we show that the investment needed to find the answer with constant probability is $O(n^2)$, while $\Omega(n!)$ is needed when using fixed-payment contracts [1].

Additional related work. Pickard et al. [17] described and analyzed the winning strategy of the DARPA Network Challenge. However, we distinguish ourselves from [17] in both aims and methods. The authors of [17] are mainly concerned with the motivation of the exact $1/2$-split winning strategy that was implemented by the MIT Media Laboratory, for which they show that it is in the participants' interest to recruit the highest number of friends and back the theory with an empirical analysis of the diffusion cascades. Our work considers the more general setting of split contracts in the model of query incentive networks introduced in [12] and analyzes the efficiency, in terms of investment, of the Nash equilibria.

In the context of query incentive networks with fixed-payment contracts, Kota and Narahari [13] applied the results of general branching processes from [1] to analyze the reward when the degree distribution follows a power-law and the desired success probability is at least $1 - 1/n$ and show a threshold behavior of the reward with respect to the scaling exponent. Dikshit and Yadati [3] considered the issue of the quality of the answers in query incentive networks. In particular, they define a quality conscious model of incentives and derive the same threshold behavior around the branching factor $b = 2$ found in [1, 12].

It is worth to mention additional related work that is not in the context of query incentive networks. Emek et al. [6] studied strategies of multi-level marketing, in which each individual is rewarded according to direct and indirect referrals, and show that geometric reward schemes are the only guarantee to certain desirable properties. Our setting is substantially different from [6], as the reward is based on referral rather that information retrieval. Douceur and Moscibroda [5] proposed the lottery tree as a mechanism to incentivize the adoption of a distributed systems and the solicitation of new participants. Influence in social networks is also related to our work. Kempe et al. [11] considered the algorithmic question of selecting an influential set of individuals. Jackson and Yariv [9] proposed a game-theoretic framework to model incentives in adoption processes. Hartline et al. [8] studied influence in social networks from a revenue maximization point of view. Singer [19] developed incentive-compatible mechanisms for influence maximization in several models.

2 Preliminaries

We model the network as a tree generated via a Galton-Watson branching process with offspring distribution $\{c_k\}_{k=0}^d$, that is, c_k is the probability that any node has exactly k children and $\sum_{k=0}^d c_k = 1$. We adopt the convention that the root of the tree is at level 0, its children at level 1, and so on. The probability generating function of the offspring distribution is given by

$$\Psi(x) = \sum_{k=1}^{d} c_k x^k, \qquad 0 \leq x \leq 1.$$

The branching factor of the process is defined as $b = \Psi'(1) = \sum_{k=0}^d kc_k$. A fundamental result in the theory of Galton-Watson processes states that the extinction probability ζ of a branching process is the smallest non-negative root of the equation $x = \Psi(x)$. If follows that $\zeta = 1$ if and only if $b < 1$, or $b = 1$ with $c_0 > 0$, and that $0 \leq \zeta < 1$ otherwise. For classical theory on Galton-Watson branching processes we refer to [2].

We assume that each node in the network possesses the answer to the query independently of the other nodes with probability $1/n$, where n represents the *rarity* of the answer. Note that n is the expected number of nodes to query before finding the answer. For $i \geq 0$, let ϕ_i be the probability that no node at level $j \leq i$ has the answer and $\lambda_i = \phi_{i-1} - \phi_i$ be the probability that some node at level i and no node at a lower level possesses the answer. (These probabilities are over the randomness of the branching process and of the process assigning the answers.) Moreover, conditional on the event that the branching process with probability generating function Ψ does not die out, let $h_\Psi(\epsilon, n)$ be the minimum integer i such that $\phi_i < \epsilon$. For branching factor $b > 1$, we have that $h_\Psi(\epsilon, n) = O(\log n)$ for any $\epsilon = n^{-O(1)}$, whereas in the case of $b = 1$ and $c_0 = 0$ (i.e., a ray), $h_\Psi(\epsilon, n) = n \ln \frac{1}{\epsilon}$.

Assume that r^* is the investment available at the root, which desires to retrieve the answer with probability at least $1 - \zeta - \epsilon$, for a given success accuracy $\epsilon > 0$. With the notation introduced so far, we will show that for any constants $b > 1$ and $\epsilon > 0$ an investment of $r^* = O(h_\Psi(\epsilon, n))$ suffices to propagate the query down to level $h_\Psi(\epsilon, n)$ of the tree, and hence to retrieve the answer with probability at least $1 - \zeta - \epsilon$. For ease of analysis, we assume that the root is not willing to explore the tree below level $h_\Psi(\epsilon, n)$, that is, we *truncate* the tree at that height.

SPLIT CONTRACTS. We now formalize the notion of *split contracts*. Every node including the root can offer a ρ-split contract to its children, for some $0 < \rho < 1$, stipulating the following. If the root offers a ρ-split to a child u who possesses the answer, then u receives a payment of r^* but is required to return a ρ fraction to the root, earning a total of $r^*(1 - \rho) - 1$, where we introduced a unit cost for returning the answer to the parent, as in [12, 1]. If instead u does not possess the answer then it might decide to propagate the query to its children, according to its strategy $f_u(\cdot)$, that is, offering a $f_u(\rho)$-split contract to its children. If one among u's children possesses the answer, then u receives an $f_u(\rho)$ fraction of the reward but it gives a ρ fraction back to the root and pays the unit cost to return the answer, with an overall earning of $r^*(1-\rho)f_u(\rho) - 1$. In general, consider a node u_ℓ which is reached by a query and possesses the answer, and let $u_0, u_1, \ldots, u_\ell$ be the path connecting the root to u_ℓ, where u_0 is the root. Then, if the root offered a ρ_{u_0}-split to its children, and $\rho_{u_i} = f_{u_i}(\rho_{u_{i-1}})$ is the split offered by u_i to its children for all $i < \ell$, then the root u_0 (who need not to pay the unit cost) receives a payoff of

$$r^* \cdot \rho_{u_0} \cdot f_{u_1}(\rho_{u_0}) \cdot f_{u_2}(f_{u_1}(\rho_{u_0})) \cdots f_{u_{\ell-1}}(f_{u_{\ell-2}}(\cdots)) = r^* \cdot \prod_{j=0}^{\ell-1} \rho_{u_j}.$$

Similarly, for $1 \leq i \leq \ell$, the payoff of node u_i is $\left(r^*(1 - \rho_{u_{i-1}}) \cdot \prod_{j=i}^{\ell-1} \rho_{u_j}\right) - 1$.

Without loss of generality, we assume that nodes never propose *useless* split-offers to their children, that is, ρ-split where $\rho > \rho_1 := 1 - 1/r^*$, since their children would not have incentive to play even if they possessed the answer themselves. Also, for simplicity we assume discrete domain and range for the strategy f_u of every node u, that is, $f_u : \mathcal{D}_M \to (\mathcal{D}_M \cup \perp)$, where $f_u(\rho) = \perp$ indicates that u chooses not to propagate the query, and $\mathcal{D}_M = \{\frac{\rho_1}{M}, \frac{2\rho_1}{M}, \ldots, \frac{(M-1)\rho_1}{M}, \rho_1\}$ is a discretization of the interval $(0, \rho_1]$.

Propagation of the payment. We remark that the above payoffs for the path $u_0, \ldots, u_\ell$ will turn into concrete payments only if the root selects u_ℓ among the answer-holders. Indeed, among all answer-holders reached by the query the root, will select only one for payment. In the fixed-payment model of [12, 1], this selection is made using a fixed arbitrary procedure that does not affect the strategies of the nodes (e.g., performing a random walk from the root descending down the tree; the first hit answer-holder will be paid along with its ancestors). In their setting, this choice is coherent as the root always spends a fixed investment, no matter how deep in the tree the payment is propagated. In our case this peculiarity is missing as a result of the split contract mechanism. In our model we will assume the root selects for payment one among the answer-holders (reached by the query) at *smallest depth*. This is motivated by different facts. First, if we consider some notion of time related to propagating the query one level down, then our selection mechanism better depicts the strategy adopted in the DARPA Network Challenge, where the payment was given to the first participant reporting the correct location of a balloon. Second, the actual investment of the root is in general smaller if the path to the answer is shorter. Finally, a selection mechanism based on smallest depth alleviates the false-name issue discussed in [17]. In case of multiple answer-holders at smallest depth, we assume that the root breaks ties in a way that does not affect the strategies of the nodes (e.g., performing a random walk from the root to one of the leaves of the subtree formed by all shortest paths to the answers, and selecting the corresponding answer-holder).

Difference with respect to previous work. We would like to spend a few words highlighting some of the main differences between our analysis and those in [12, 1]. One of these differences, the propagation of payments, has been already discussed above; from the technical point of view, the *smallest depth* selection mechanism introduces the hurdle that the strategy of each node does not only depend on the strategies in its subtree (as in the case of [12, 1]), but potentially on those of all nodes. We remark that the gap in efficiency of the two models is not related to the different propagation of payment. In fact, if the answer-holder were to be selected according to the smallest depth mechanism in the fixed-payment setting, then the investment needed to retrieve the answer would increase. Roughly speaking, this happens as a node further down in the tree requires higher reward to forward the query, in order to compensate for the smaller probability of having a payment candidate in its subtree.

Another salient difference between the two models concerns the *values* of the contracts: while the nature of the fixed-payment contracts of [12, 1] implies that a node being offered a reward of r can only offer an amount $r' < r$ to its children, we do not enjoy this property on the ρ's in the case of split contracts. This unfortunately precludes the inductive arguments adopted in [12, 1], making a more involved analysis necessary.

We conclude this section discussing about the gap in efficiency between split contracts, for which an investment proportional to the depth of the search tree suffices for any branching factor $b > 1$, and the results in [12, 1], for which the investment becomes exponential in the depth of the search tree when the branching factor drops below 2. In the setting of [12], the additional amount of reward δ_j that the root needs in order to explore j levels of the tree (rather than stopping at level $j-1$) can be expressed as

$$\delta_{j+1} = \frac{1 - \phi_{j-1}}{\lambda_j}\delta_j + 1.$$

When the branching factor drops below 2, the ratio $\frac{1-\phi_{j-1}}{\lambda_j}$ is greater than 1, and the investment needed at the root to propagate the query down to depth $h_\Psi(\epsilon, n)$ becomes exponential in $\log n$ (hence, poly(n)).

In contrast, the dependency on the ratio $\frac{1-\phi_{j-1}}{\lambda_j}$ is softer in our setting. In the proof of Theorem 10, we show that the ρ-split a node at level ℓ needs to receive in order to propagate the query i levels down its subtree is

$$\rho_i^{\langle \ell \rangle} = 1 - \frac{1}{r^* - i(1 + O(\frac{1-\phi_{i-1}}{\lambda_i}))}.$$

Since we can show that $\frac{1-\phi_{i-1}}{\lambda_i}$ is bounded by a constant for any branching process with $b > 1$, an investment $r^* = O(h_\Psi(\epsilon, n)) = O(\log n)$ suffices for the value $\rho^{\langle 1 \rangle}_{h_\Psi(\epsilon,n)}$ offered by the root to its children to be well-defined (i.e., in $\mathcal{D}_M$), and hence for the answer to be retrieved cheaply.

ROADMAP. The rest of the paper is structured as follows. In Section 3, we derive properties that hold for any Nash equilibrium. In Section 4, we develop a condition that we call h-consistency under which we can show that a set of strategies $\mathbf{g}$ for the nodes propagates the query to the desired level and is substantially the unique Nash equilibrium. In Section 5, we derive a bound on the investment r^*, depending on quantities related to the branching process, for which h-consistency is guaranteed to hold. Finally, in Section 6, we study such quantities of the branching process to conclude that $r^* = O(h_\Psi(\epsilon, n)) = O(\log n)$. Due to space constraints, all proofs are deferred to the Appendix.

3 Properties of Nash Equilibria

In this section we present the notion of Nash equilibrium that naturally arises in the context of split-contracts, and we then derive a manageable expression that any Nash equilibrium has to maximize. Let f_v be the function representing the strategy of node v, and $\mathbf{f}$ be the set of strategies of all nodes up to level $h_\Psi(\epsilon, n)$, as we assumed that nodes in lower levels do not play.

Definition 1 (**Nash equilibrium**). *Let r^*, Ψ, ϵ, n be the parameters of the model, and $\mathbf{f}$ be a set of functions for all nodes up to level $h_\Psi(\epsilon, n)$. For any such node v, let ρ^v be the split contract offered to v by its parent under $\mathbf{f}$. Then, $\mathbf{f}$ is a Nash equilibrium if, for each node v, v does not increase its expected payoff by deviating from $f_v(\rho^v)$ when all other nodes play according to $\mathbf{f}$. The expectation is taken over the randomness of the branching process and of the process assigning answers to nodes.*

We now give a few definitions that will be useful to derive properties of any Nash equilibrium. Given a realization of the branching process, we say that a node v at level $\ell \le h_\Psi(\epsilon, n)$ is *active* if the branching process reaches v. Moreover, given a set $\mathbf{f}$ of strategies and a realization of the branching process, we say that an active node v is $\mathbf{f}$-*reachable* if $\mathbf{f}$ forwards the query down to v. Given a realization of the branching process and of the process assigning the answer to nodes, we say that an $\mathbf{f}$-reachable node v at level ℓ is an $\mathbf{f}$-*candidate* if v holds the answer and no $\mathbf{f}$-reachable node at a level $\ell' < \ell$ does. Observe that the root selects for payment one among the $\mathbf{f}$-candidates. For each node v at level $\ell \le h_\Psi(\epsilon, n)$, set $\mathbf{f}$ of strategies, and $j \ge 1$, let $\alpha_v^{\mathbf{f}}(j|\rho)$ be the probability that there is an $\mathbf{f}$-candidate in v's subtree at distance j from v, conditional on v being $\mathbf{f}$-reachable and offering a ρ-split to its children. Similarly, for $j \ge 1$, let $\beta_v^{\mathbf{f}}(j|\rho)$ be the expected payment that v receives from its children given that v offers a ρ-split to its children and there is an $\mathbf{f}$-candidate in v's subtree at distance j from v to whom the root propagates the payment.

The following lemma characterizes an expression that must be maximized by every node up to level $h_\Psi(\epsilon, n)$ in any Nash Equilibrium.

Lemma 2. *Consider any set* $\mathbf{f}$ *of strategies, and let* ρ^v *be the split contract offered to* v *by its parent under* $\mathbf{f}$*. Then,* $\mathbf{f}$ *is a Nash equilibrium if and only if, for every node* v *up to level* $h_\Psi(\epsilon, n)$*,* $f_v(\rho^v)$ *is a value of* ρ *maximizing the function*

$$\chi_v^{\mathbf{f}}(\rho; \rho^v) := \sum_{j \geq 1} \alpha_v^{\mathbf{f}}(j|\rho) \left((1 - \rho^v) \beta_v^{\mathbf{f}}(j|\rho) - 1 \right). \tag{1}$$

To break ties in case of multiple maxima for $\chi_v^{\mathbf{f}}(\cdot; \rho^v)$, we make the same assumption as in [12, 1] that nodes favor strategies that forward the query further down in the tree. Using Lemma 2, we can now prove that Nash equilibria are "leveled".

Lemma 3. *Consider any Nash equilibrium* $\mathbf{f}$*. Then for each active node* v *at level* ℓ*,* v *is* $\mathbf{f}$*-reachable if and only if every active node at level* ℓ *is.*

By means of Lemma 3, we will say that a Nash equilibrium $\mathbf{f}$ is k*-tall* if level k is $\mathbf{f}$-reachable and level $k+1$ is not. This notion is useful in decoupling the probabilities $\alpha_v^{\mathbf{f}}(j|f_v(\rho))$ from the particular equilibrium $\mathbf{f}$ and node v. To see how, assume $\mathbf{f}$ is k-tall, $k \leq h_\Psi(\epsilon, n)$. For any node v at level $\ell \leq k$ and any $j \leq k - \ell$, let $\gamma_j^{\langle \ell \rangle}$ be the probability that there exists an $\mathbf{f}$-candidate in v's subtree at distance j from v (and therefore there is no $\mathbf{f}$-candidate in the first $\ell + j - 1$ levels). Then, as $\mathbf{f}$ is k-tall, we have that for any node at level ℓ, $\gamma_j^{\langle \ell \rangle}$ depends only on ℓ and j (and not on $\mathbf{f}$ or the specific node). This observation directly yields the following result relating the probabilities $\alpha_v^{\mathbf{f}}(j|f_v(\rho^v))$ and $\gamma_j^{\langle \ell \rangle}$.

Lemma 4. *Let* $\mathbf{f}$ *be a* k*-tall Nash equilibrium,* $k \leq h_\Psi(\epsilon, n)$*. Then, for every* $\ell \leq k$ *and node* v *at level* ℓ*,*

$$\alpha_v^{\mathbf{f}}(j|f_v(\rho^v)) = \begin{cases} \gamma_j^{\langle \ell \rangle}, & \textit{for } 1 \leq j \leq k - \ell \\ 0, & \textit{for } j > k - \ell \end{cases}$$

In general, if the query is forwarded j *levels down* v*'s subtree when* v *offers a* ρ*-split to its children, then we have* $\alpha_v^{\mathbf{f}}(j|\rho) = \gamma_j^{\langle \ell \rangle}$*.*

4 The Nash Equilibrium

In this section, we derive conditions for the existence of a Nash equilibrium that forwards the query down to level $h_\Psi(\epsilon, n)$, or, equivalently, retrieves the answer with the desired probability $1 - \zeta - \epsilon$. For ease of notation, let $h = h_\Psi(\epsilon, n)$. We proceed as follows. First we define the functions $e_i^{\langle \ell \rangle}$ and the thresholds $\rho_i^{\langle \ell \rangle}$, which intuitively represent expected rewards and contracts for a special set of strategies $\mathbf{g}$. However, to define $\mathbf{g}$, we will need all $\rho_i^{\langle \ell \rangle}$ to exist and be decreasing in i, for all $\ell \leq h$, property that we will dub h-consistency. Finally, assuming h-consistency, we will show that $\mathbf{g}$ forwards the query to level h and is a Nash equilibrium (in fact with an extra property, we will say $\mathbf{g}$ is a best-interest Nash Equilibrium).

We begin by defining the aforementioned functions and values. We provide an inductive process which defines, for each $1 \leq \ell \leq h$, a sequence of functions $e_i^{\langle \ell \rangle} : [0, 1] \to \mathbb{R}, 0 \leq i \leq h - \ell$, and values $\rho_i^{\langle \ell \rangle} \in \mathcal{D}_M$, $1 \leq i \leq h - \ell + 1$. For every $0 \leq \ell \leq h$, set $e_0^{\langle \ell \rangle}(\rho) = 0$ and $\rho_1^{\langle \ell \rangle} = \rho_1 = 1 - 1/r^*$. Suppose that all $\rho_i^{\langle \ell' \rangle}$ have been defined for $\ell < \ell' \leq h$ and $1 \leq i \leq h - \ell' + 1$. Then, for all $1 \leq i \leq h - \ell$, the function $e_i^{\langle \ell \rangle}(\rho)$ is defined as

$$e_i^{\langle \ell \rangle}(\rho) = \sum_{j=1}^{i} \gamma_j^{\langle \ell \rangle} \left[(1 - \rho) r^* \left(\prod_{t=0}^{j-1} \rho_{i-t}^{\langle \ell + t + 1 \rangle} \right) - 1 \right].$$

Having defined $e_i^{\langle \ell \rangle}(\rho)$, we define

$$\rho_{i+1}^{\langle \ell \rangle} = \max\{\rho \in \mathcal{D}_M : e_i^{\langle \ell \rangle}(\rho) \geq e_{i-1}^{\langle \ell \rangle}(\rho)\},$$

if such value exists, and leave $\rho_{i+1}^{\langle \ell \rangle}$ undefined otherwise.

For a node v at level $\ell \leq h$, $e_i^{\langle \ell \rangle}$ has the intuitive meaning of the expected reward that v receives from its children when the query is propagated i levels down v's subtree (assuming the other nodes play accordingly). The value $\rho_{i+1}^{\langle \ell \rangle}$ represents the "cheapest" split to offer a node v at level $\ell \leq h$ so that v prefers to propagate the query i levels down its subtree rather than $i-1$ (recall that, to break ties, we assumed that nodes prefer to propagate the query further down the tree). To guarantee the propagation of the query to level h, we will need the values $\rho_{h-\ell}^{\langle \ell \rangle}$ to be defined.

Definition 5 (**h-consistency**)**.** *We say that* h-consistency *holds if, for all* $1 \leq \ell \leq h$ *and* $2 \leq i \leq h - \ell + 1$*, the value* $\rho_i^{\langle \ell \rangle}$ *is defined and* $\rho_i^{\langle \ell \rangle} < \rho_{i-1}^{\langle \ell \rangle}$ *(note that* $\rho_1^{\langle \ell \rangle}$ *is always defined).*

Intuitively, the ordering of the values $\rho_i^{\langle \ell \rangle}$ in the definition of h-consistency states that if a node v propagates the query i levels down its subtree when offered a ρ-split by its father, then, in order to propagate the query $i+1$ levels down, it must be that v is offered a split not greater than ρ. This property is at the basis of the following definition of the set of strategies $\mathbf{g}$, which we will then show to be a Nash equilibrium. Note how, under $\mathbf{g}$, nodes at the same level play the same strategy.

Definition 6 (**Strategy $\mathbf{g}$**)**.** *Assume h-consistency holds. For each $1 \leq \ell \leq h$, consider the function $t^{\langle \ell \rangle}(\rho) : [0,1] \to \mathcal{D}_M \cup \{\perp\}$ defined by $t^{\langle \ell \rangle}(\rho) = \rho_{i-1}^{\langle \ell+1 \rangle}$ for the unique i such that $\rho_{i+1}^{\langle \ell \rangle} < \rho \leq \rho_i^{\langle \ell \rangle}$ (such i exists under h-consistency), where we assume $\rho_{h-\ell+2}^{\langle \ell \rangle} = 0$ and $\rho_0^{\langle \ell+1 \rangle} = \perp$. The set of strategies $\mathbf{g}$ is defined by setting $g_v(\rho) = t^{\langle \ell \rangle}(\rho)$ to every node v at level ℓ, for each $1 \leq \ell \leq h$, and letting the root play $\rho_h^{\langle 1 \rangle}$.*

It follows that, under $\mathbf{g}$, the root (at level zero) offers a $\rho_h^{\langle 1 \rangle}$-split to its children, who in turn offer $\rho_{h-1}^{\langle 2 \rangle}$-split contracts to their own children, and so on, until the nodes at level h, who do not forward the query (they play $t^{\langle h \rangle}(\rho_1^{\langle h \rangle}) = \perp$). Observe that $\mathbf{g}$ is h-tall, as all nodes up to level h are $\mathbf{g}$-reachable. The following theorem states that $\mathbf{g}$ is a Nash equilibrium.

Theorem 7 (**Nash equilibrium**)**.** *Assuming h-consistency, the set of strategies* $\mathbf{g}$ *is a Nash equilibrium.*

The key fact in the proof is to show that, for any node v at level $1 \leq \ell \leq h$ (which under $\mathbf{g}$ receives a $\rho_{h-\ell+1}^{\langle \ell \rangle}$-split from its parent and in turn offers a $\rho_{h-\ell}^{\langle \ell+1 \rangle}$-split to its children), $\chi_v^{\mathbf{g}}(\rho_j^{\langle \ell+1 \rangle}; \rho) = e_j^{\langle \ell \rangle}(\rho)$ for all $j \leq h - \ell$, and that $\rho_{h-\ell}^{\langle \ell+1 \rangle}$ is the only maximizer of (1) that propagates the query to level h. Then the theorem follows by Lemma 2.

Even though $\mathbf{g}$ is not the only Nash equilibrium, the proof of Theorem 7 shows that $\mathbf{g}$ enjoys the additional property that, for each node v and $\rho \in \mathcal{D}_M$,

$$g_v(\rho) = \arg\max_{\rho'}\{\chi_v^{\mathbf{g}}(\rho'; \rho)\}.$$

We call any equilibrium enjoying such property a *best-interest* equilibrium, as nodes choose their best option in any scenario. The following theorem shows that $\mathbf{g}$ is substantially the only best-interest equilibrium, meaning that every other best-interest equilibrium $\mathbf{f}$ coincides with $\mathbf{g}$ on all the split-offers that are actually offered to nodes under $\mathbf{f}$. As a remark, we observe that even the game with fixed-payment contracts in [12, 1] admits multiple equilibria, although the authors claim uniqueness (a counter-example is presented in Appendix H). On the positive side, the equilibrium analyzed in [12, 1] is the unique best-interest Nash equilibrium of their game.

Theorem 8 (**Uniqueness**). *Assume h-consistency. Let $\mathbf{f}$ be any ℓ-tall best-interest Nash equilibrium, for some $1 \leq \ell \leq h$, and, for each node v up to level ℓ, let ρ^v be the split contract offered to v by its parent under $\mathbf{f}$. Then, for every node v up to level ℓ, $f_v(\rho^v) = g_v(\rho^v)$.*

Theorem 8 implies that every best-interest equilibrium $\mathbf{f}$ in which the root offers a $\rho_h^{\langle 1\rangle}$-split to its children has to be h-tall, as $\mathbf{f}$ agrees with $\mathbf{g}$ on all split-offers made in $\mathbf{g}$. As h-tall equilibria retrieve the answer with the desired probability, the root has incentive to play $\rho_h^{\langle 1\rangle}$ as its strategy and would have incentive to deviate to $\rho_h^{\langle 1\rangle}$ if playing a different strategy. The following result is then implied.

Corollary 9. *Under the assumption of h-consistency, all best-interest equilibria retrieve the answer with the desired probability.*

5 Guaranteeing h-consistency

Until now, we assumed h-consistency both in the definition of $\mathbf{g}$ and in the proof that $\mathbf{g}$ is a Nash equilibrium. It therefore remains to derive conditions that ensure h-consistency. In the following theorem, we provide a lower bound on the reward r^* above which h-consistency is guaranteed. The bound reads in terms of the probabilities $\gamma_i^{\langle \ell\rangle}$ through the quantities $\Gamma_i^{\langle \ell\rangle} = \frac{1}{\gamma_i^{\langle \ell\rangle}} \sum_{j=1}^{i-1} \gamma_j^{\langle \ell\rangle}$, which, for all $1 \leq \ell \leq h$ and $1 \leq i < h - \ell$, intuitively represent the ratio between the probability that a node at level ℓ has a candidate at depth $j < i$ in its subtree versus the probability that it has one at depth i.

Theorem 10. *Suppose the discretization parameter M is large enough, say $M = \Theta(r^{*2})$, and that*

$$r^* \geq 4 \cdot h \cdot \max\left\{1, \max_{\substack{1\leq \ell \leq h \\ 1 \leq i < h-\ell}} \Gamma_i^{\langle \ell\rangle}\right\}. \tag{2}$$

Then h-consistency holds. In particular, for all $1 \leq \ell \leq h$ and $1 \leq i \leq h - \ell$, $\rho_i^{\langle \ell\rangle}$ is defined and satisfies

$$1 - \frac{1}{r^* - i} < \rho_i^{\langle \ell\rangle} \leq 1 - \frac{1}{r^* - i + 1}. \tag{3}$$

The main idea to prove the theorem is to derive tight upper and lower bounds on $\rho_{i+1}^{\langle \ell-1\rangle}$ and then proceed by induction on both ℓ and i. It can also be proven that, for a fixed i, $\rho_i^{\langle \ell\rangle}$ is decreasing in ℓ for $0 \leq \ell \leq h-i$. The intuition for this property is that a node further down in the tree is willing to give a smaller fraction of its reward back to its parent, in order to compensate the smaller probability of having a candidate in its subtree. However, we do not need this property to ensure h-consistency.

Theorem 10 along with Corollary 9 directly yields the following pivotal result, which relates the quantities $\Gamma_j^{\langle \ell\rangle}$ to the investment that is sufficient at the root to retrieve the answer with the desired probability.

Corollary 11. *Suppose condition* (2) *holds. Then, in any best-interest Nash equilibrium, the query reaches all nodes at level $h = h_\Psi(\epsilon, n)$ of the tree. That is, an answer is retrieved with probability at least $1 - \zeta - \epsilon$.*

6 Efficiency

In the previous section, we derived a lower bound on the investment r^* as a function of the values $\Gamma_i^{\langle \ell \rangle}$, for $1 \leq \ell \leq h$ and $1 \leq i \leq h - \ell$. In this section, we show our main result by relating these values to the branching process and the desired success probability. The following lemma bounds these quantities in terms of the probabilities λ_i and ϕ_i of the branching process. Recall that, for each $i \geq 0$ we defined ϕ_i as the probability that no node at level $j \leq i$ possesses the answer, and $\lambda_i = \phi_{i-1} - \phi_i$ as the probability that a node at level i possesses the answer and no node at level $j < i$ does.

Lemma 12. *For every $1 \leq \ell \leq h$ and $1 \leq i \leq h - \ell$, it holds that* $\Gamma_i^{\langle \ell \rangle} \leq \dfrac{1}{\phi_{\ell+i-1}} \dfrac{1 - \phi_{i-1}}{\lambda_i}$.

The key in proving Lemma 12 is to express $\gamma_i^{\langle \ell \rangle}$ in terms of the probabilities ϕ_j and λ_j defined above, and then to bound $\Gamma_i^{\langle \ell \rangle}$ exploiting the memory-less property of the branching process and of the process assigning the answer to the nodes.

The following technical lemma provides an upper bound to $\frac{1-\phi_{i-1}}{\lambda_i}$. In particular, for any fixed branching process with $b > 1$, this ratio is bounded by a constant, as long as ϕ_i is bounded away from the extinction probability ζ. The lemma characterizes the bound with respect to the branching process and the gap $\phi_i - \zeta$, and its proof builds on the mathematical properties of the probability generating function of the branching process. Recall that the desired success probability is $1 - \zeta - \epsilon$.

Lemma 13. *Consider any Galton-Watson branching process with branching factor $b > 1$. Then, for every i such that $\zeta + \epsilon \leq \phi_i \leq 1$, it holds that*

$$\frac{1 - \phi_i}{\lambda_{i+1}} \leq \max \left\{ \frac{1}{b-1}, \frac{1}{\epsilon} \cdot \frac{1}{1 - \Psi'(\zeta)} \right\}.$$

Our main result directly follows by combining Corollary 11, Lemma 12 and Lemma 13, along with the observation that $\phi_{\ell+i-1} > \epsilon$ (as $\phi_{\ell+i-1} \geq \phi_{h_\Psi(\epsilon,n)-1} > \epsilon$). For the case of a ray, the bound can be obtained observing that $\phi_i = (1 - 1/n)^i$ and $\lambda_{i+1} = \frac{\phi_i}{n}$, which implies $\Gamma_i^{\langle \ell \rangle} \leq \Gamma_h^{\langle 1 \rangle} \leq \epsilon^{-2} n$.

Theorem 14 (Efficiency). *Consider any Galton-Watson branching process with $b > 1$. Then, the root retrieves the answer with probability at least $\sigma = 1 - \zeta - \epsilon$ provided an investment of*

$$r^* = \frac{4}{\epsilon} \cdot \max \left\{ \frac{1}{b-1}, \frac{1}{\epsilon} \cdot \frac{1}{1 - \Psi'(\zeta)} \right\} \cdot h_\Psi(\epsilon, n).$$

In the case of a ray, with $b = 1$ and $c_0 = \zeta = 0$, an investment of $r^* = 4 \cdot \frac{n}{\epsilon^2} \cdot h_\Psi(\epsilon, n) = 4 \cdot \frac{n^2}{\epsilon^2} \ln \frac{1}{\epsilon}$ suffices.

Observe that an investment of $h_\Psi(\epsilon, n)$ is necessary even in a centralized (non-strategic) setting, where the root decides the strategies of all nodes while only guaranteeing a non-negative payoff to them (each node pays a unit cost when returning the answer). In line with intuition, the investment grows as b tends to 1 (in the limit, when the tree becomes a ray, the investment is polynomial in n), and when the accuracy ϵ approaches 0. The term $\frac{1}{1-\Psi'(\zeta)}$ can be crudely bounded by $\frac{1}{c_0}$. However, when c_0 tends to zero, so does the extinction probability ζ, which implies $\frac{1}{1-\Psi'(\zeta)} \approx \frac{1}{1-c_1}$, also suggesting a more expensive investment when the tree tends to a ray (i.e., when c_1 approaches 1).

References

[1] E. Arcaute, A. Kirsch, R. Kumar, D. Liben-Nowell, and S. Vassilvitskii. On threshold behavior in query incentive networks. In *Proceedings of the 8th ACM conference on Electronic commerce*, pages 66–74, 2007.

[2] K. B. Athreya and P. Ney. *Branching processes*. Springer-Verlag, New York, 2004.

[3] D. Dikshit and N. Yadati. Truthful and quality conscious query incentive networks. In *Internet and Network Economics*, pages 386–397. Springer, 2009.

[4] P. Dodds, R. Muhamad, and D. Watts. An experimental study of search in global social networks. *Science*, 301(5634):827, 2003.

[5] J. R. Douceur and T. Moscibroda. Lottery trees: motivational deployment of networked systems. In *SIGCOMM*, pages 121–132, 2007.

[6] Y. Emek, R. Karidi, M. Tennenholtz, and A. Zohar. Mechanisms for multi-level marketing. In *ACM Conference on Electronic Commerce*, pages 209–218, 2011.

[7] E. Hand. Citizen science: People power. *Nature*, 466(7307):685, 2010.

[8] J. D. Hartline, V. S. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. In *WWW*, pages 189–198, 2008.

[9] M. O. Jackson and L. Yariv. Diffusion of behavior and equilibrium properties in network games. *American Economic Review*, 97(2):92–98, 2007.

[10] H. Kautz, B. Selman, and M. Shah. Referral web: combining social networks and collaborative filtering. *Communications of the ACM*, 40:63–65, March 1997.

[11] D. Kempe, J. M. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *KDD*, pages 137–146, 2003.

[12] J. M. Kleinberg and P. Raghavan. Query incentive networks. In *FOCS*, pages 132–141, 2005.

[13] N. Kota and Y. Narahari. Threshold behavior of incentives in social networks. In *ACM international conference on Information and knowledge management*, pages 1461–1464, 2010.

[14] J. Lorenz, H. Rauhut, F. Schweitzer, and D. Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22):9020, 2011.

[15] S. Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

[16] V. Naroditskiy, I. Rahwan, M. Cebrian, and N. Jenning. Verification in referral-based crowdsourcing. *Working paper*, 2011.

[17] G. Pickard, W. Pan, I. Rahwan, M. Cebrián, R. Crane, A. Madan, and A. Pentland. Time critical social mobilization. *Science*, 334(6055):509–512, 2011.

[18] S. Schnettler. A structured overview of 50 years of small-world research. *Social Networks*, 31(3):165–178, 2009.

[19] Y. Singer. How to win friends and influence people, truthfully: Influence maximization mechanisms for social networks. In *WSDM*, 2011.

[20] J. Surowiecki. *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. Doubleday Books, 2004.

[21] J. C. Tang, M. Cebrián, N. A. Giacobe, H.-W. Kim, T. Kim, and D. B. Wickert. Reflecting on the darpa red balloon challenge. *Commun. ACM*, 54(4):78–85, 2011.

[22] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.

[23] D. Watts, P. Dodds, and M. Newman. Identity and search in social networks. *Science*, 296(5571):1302, 2002.

[24] B. Yu and M. P. Singh. Searching social networks. In *AAMAS*, pages 65–72, 2003.

[25] Zhang, Jun and V. Alstyne, Marshall. SWIM: fostering social network based information search. In *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, volume 2, page 1568, 2004.

A Proof of Lemma 2

Fix the available investment r^*, a set $\mathbf{f}$ of strategies, and a node v at level $\ell \leq h_\Psi(\epsilon, n)$. We need to define the following events. Let A denote the event that the root propagates the payment down through v, that is, the root selects for payment an $\mathbf{f}$-candidate in v's subtree. For each $0 \leq j \leq h_\Psi(\epsilon, n) - \ell$, let B_j denote the event that the $\mathbf{f}$-candidates are at level $\ell + j$. Finally let C denote the event that v is $\mathbf{f}$-reachable and D denote the event that there is an $\mathbf{f}$-candidate in v's subtree. Observe that the co-occurrence of B_0 and D means that v itself is an $\mathbf{f}$-candidate. Given r^* and $\mathbf{f}$, let $Y^v_{\mathbf{f},r^*}$ be the random variable denoting the payment assigned to v.

We have that

$$\begin{aligned} E[Y^v_{\mathbf{f},r^*}] &= \sum_{j \geq 0} E[Y_{\mathbf{f},r^*} | A, B_j, D, C] \Pr(A, B_j, D, C) \\ &= \Pr(A|D) \Pr(C) \sum_{j \geq 0} E[Y_{\mathbf{f},r^*} | A, B_j, D, C] \Pr(B_j, D|C). \end{aligned}$$

The first equality follows from the law of total probability together with the observation that $\Pr(A, \overline{C}) = \Pr(A, \overline{D}) = 0$ and $E[Y^v_{\mathbf{f},r^*} | \overline{A}] = 0$. The second equality follows from the chain rule of probability and the fact that $\Pr(A|B_j, D, C) = \Pr(A|D)$ for all $j \geq 0$. Observe that the term corresponding to $j = 0$ (i.e., v is the $\mathbf{f}$-candidate selected for payment) does not depend on f_v since $E[Y_{\mathbf{f},r^*}|A, B_0, D, C] = (1 - \rho^v)r^* - 1$ and $\Pr(B_0, D|C)$ only depends on the strategies of the nodes that are ancestors of v. Similarly, f_v affects neither $\Pr(C)$, which depends on the strategies of v's ancestors only, nor $\Pr(A|D)$, which is only based on the root's choice of whom to propagate the payment to. Finally, note that if v offers a ρ-split to its children, then, for $j \geq 1$, $E[Y_{\mathbf{f},r^*}|A, B_j, S, C] = (1 - \rho^v)\beta^{\mathbf{f}}_v(j|\rho) - 1$ and $\Pr(B_j, D|C) = \alpha^{\mathbf{f}}_v(j|\rho)$. Therefore, $\mathbf{f}$ is a Nash equilibrium if and only if, for every node v up to level $h_\Psi(\epsilon, n)$, $f_v(\rho^v)$ is a value ρ maximizing $\chi^{\mathbf{f}}_v(\rho; \rho^v)$. □

B Proof of Lemma 3

Let $\mathbf{f}$ be a Nash equilibrium. Fix a node v and let

$$\rho_2 = \max\{\rho \in \mathcal{D}_M : \chi^{\mathbf{f}}_v(\rho_1; \rho) \geq \chi^{\mathbf{f}}_v(\perp; \rho)\}$$

be the maximum split v's father can ask v so that v will in turn prefer to offer a ρ_1-split to their children rather than just participating in the game without propagating the query. We first argue that ρ_2 does not depend on the chosen node v, then show that $f_v(\rho) = \perp$ for every node v and $\rho_2 < \rho \leq \rho_1 = 1 - \frac{1}{r^*}$, and finally use this fact to prove the lemma.

To see that ρ_2 does not depend on v, observe that $\chi^{\mathbf{f}}_v(\perp; \rho) = 0$ as $\alpha^{\mathbf{f}}_v(j|\perp) = 0$ for $j \geq 1$, and that $\chi^{\mathbf{f}}_v(\rho_1; \rho) = \alpha^{\mathbf{f}}_v(1|\rho_1)((1 - \rho)\beta^{\mathbf{f}}_v(1|\rho_1) - 1) = \alpha^{\mathbf{f}}_v(1|\rho_1)((1 - \rho)r^*\rho_1 - 1)$.

We now show that $f_v(\rho) = \perp$, for every node v and $\rho_2 < \rho \leq \rho_1$. By contradiction, suppose $f_v(\rho) = \rho'$, for some v, $\rho_2 < \rho \leq \rho_1$, $\rho' \in \mathcal{D}_M$. On the one hand, as $\mathbf{f}$ is a Nash equilibrium, Lemma 2 implies that ρ' maximizes $\chi^{\mathbf{f}}_v(\rho'; \rho)$, and thus

$$\chi^{\mathbf{f}}_v(\rho'; \rho) \geq \chi^{\mathbf{f}}_v(\perp; \rho) = 0.$$

On the other hand, we have that

$$\chi_v^{\mathbf{f}}(\rho';\rho) = \sum_{j\geq 1} \alpha_v^{\mathbf{f}}(j|\rho')\left((1-\rho)\beta_v^{\mathbf{f}}(j|\rho') - 1\right)$$
$$\leq \sum_{j\geq 1} \alpha_v^{\mathbf{f}}(j|\rho')((1-\rho)r^*\rho_1 - 1),$$

where the last inequality follows from $\beta_v^{\mathbf{f}}(j|\rho') \leq r^*\rho_1$ for all $j \geq 1$, as ρ' must be at most ρ_1 for v's children to participate to the game. By definition of ρ_2, it must be that $\chi_v^{\mathbf{f}}(\rho_1;\rho) < \chi_v^{\mathbf{f}}(\bot;\rho) = 0$, which implies $((1-\rho)r^*\rho_1 - 1) < 0$ and thus $\chi_v^{\mathbf{f}}(\rho';\rho) < 0$, generating a contradiction.

We are now ready to prove the lemma. By contradiction, suppose the statement of the lemma does not hold. Then there must be two sibling nodes u and v (at some level $\ell < h_\Psi(\epsilon, n)$) and a value $\rho = \rho^u = \rho^v$ such that $f_u(\rho) = \rho' \neq \bot$ and $f_v(\rho) = \bot$, that is, such that u forwards the query when offered a ρ-split by its parent while v does not. By the claim above, $f_u(\rho) = \rho'$ implies that $\rho \leq \rho_2$ and therefore, by definition of ρ_2, v would have incentive to deviate from f_v, offering a ρ_1-split to their children than just participating to the game without propagating the query, contradicting that $\mathbf{f}$ is a Nash equilibrium. □

C Proof of Theorem 7

Under h-consistency, for all $\ell \leq h$ and $2 \leq i \leq h-\ell+1$, $\rho_i^{\langle\ell\rangle}$ is defined and $\rho_i^{\langle\ell\rangle} < \rho_{i-1}^{\langle\ell\rangle}$ (recall that $\rho_1^{\langle\ell\rangle}$ is defined for all $\ell \leq h$). In the proof of the theorem, we make use of the following technical lemma, that is a consequence of h-consistency.

Claim 15. *Assume h-consistency. Then, for every $1 \leq \ell \leq h$, $1 \leq i \leq h-\ell$, and $\rho_{i+2}^{\langle\ell\rangle} < \rho \leq \rho_{i+1}^{\langle\ell\rangle}$, we have that*

$$e_i^{\langle\ell\rangle}(\rho) > e_{i+1}^{\langle\ell\rangle}(\rho) > \cdots > e_{h-\ell}^{\langle\ell\rangle}(\rho),$$

and

$$e_i^{\langle\ell\rangle}(\rho) \geq e_{i-1}^{\langle\ell\rangle}(\rho) \geq \cdots \geq e_0^{\langle\ell\rangle}(\rho),$$

where we assume $\rho_{h+1}^{\langle 1\rangle} = 0$.

Proof. Consider any $i+1 \leq j \leq h-\ell$, and observe that, by definition,

$$\rho_{j+1}^{\langle\ell\rangle} = \max\{\rho' \in \mathcal{D}_M : e_j^{\langle\ell\rangle}(\rho') \geq e_{j-1}^{\langle\ell\rangle}(\rho')\},$$

and, by h-consistency (as $j+1 \geq i+2$), $\rho_{j+1}^{\langle\ell\rangle} \leq \rho_{i+2}^{\langle\ell\rangle} < \rho$. It follows that $e_j^{\langle\ell\rangle}(\rho) < e_{j-1}^{\langle\ell\rangle}(\rho)$ for all $i+1 \leq j \leq h-\ell$, which implies that

$$e_{h-\ell}^{\langle\ell\rangle}(\rho) < e_{h-\ell-1}^{\langle\ell\rangle}(\rho) < \cdots < e_i^{\langle\ell\rangle}(\rho),$$

proving the first chain of inequalities in the lemma. Now consider any $2 \leq m \leq i+1$, and observe that, by definition of $\rho_m^{\langle\ell\rangle}$,

$$e_{m-1}^{\langle\ell\rangle}(\rho_m^{\langle\ell\rangle}) \geq e_{m-2}^{\langle\ell\rangle}(\rho_m^{\langle\ell\rangle})$$

and, by h-consistency (as $i+1 \geq m$), $\rho \leq \rho_m^{\langle\ell\rangle}$. This implies that $e_{m-1}^{\langle\ell\rangle}(\rho) \geq e_{m-2}^{\langle\ell\rangle}(\rho)$ for all $2 \leq m \leq i+1$. It follows that

$$e_i^{\langle\ell\rangle}(\rho) \geq e_{i-1}^{\langle\ell\rangle}(\rho) \ldots \geq e_0^{\langle\ell\rangle}(\rho),$$

which proves the second chain of inequalities in the lemma. □

To show that $\mathbf{g}$ is a Nash equilibrium, by Lemma 2, it suffices to prove that, for every node v at level up to h, $g_v(\rho^v)$ is the value that maximizes $\chi_v^{\mathbf{g}}(\cdot;\rho^v)$, where ρ^v is the split offer v receives from its parent. Let $0 \leq i \leq h-1$, and fix a node v at level $\ell = h - i$. Under $\mathbf{g}$, v receives a $\rho_{i+1}^{\langle \ell \rangle}$-split from its parent and in turn offers a $t^{\langle \ell \rangle}(\rho_{i+1}^{\langle \ell \rangle}) = \rho_i^{\langle \ell+1 \rangle}$-split to its children. Therefore, it suffices to show that

$$\rho_i^{\langle \ell+1 \rangle} = \arg\max_{\rho'}\{\chi_v^{\mathbf{g}}(\rho';\rho_{i+1}^{\langle \ell \rangle})\}.$$

We will in fact prove something stronger, that is, for all $\rho \in \mathcal{D}_M$,

$$g_v(\rho) = t^{\langle \ell \rangle}(\rho) = \arg\max_{\rho'}\{\chi_v^{\mathbf{g}}(\rho';\rho)\}. \tag{4}$$

Fix any $\rho \in \mathcal{D}_M$. A few observations allow to prove condition (4) for the chosen ρ. First, by h-consistency, there exists unique k such that $\rho_{k+2}^{\langle \ell \rangle} < \rho \leq \rho_{k+1}^{\langle \ell \rangle}$, where we assume $\rho_{h+1}^{\langle \ell \rangle} = 0$. Second, by definition of $\mathbf{g}$ and $\chi_v^{\mathbf{g}}(\cdot;\cdot)$, node v has an incentive to play a given $\rho' \in \mathcal{D}_M$ only if there is no $\hat{\rho} > \rho'$ such that v's children would play exactly the same split contract if either offered a $\hat{\rho}$-split or a ρ'-split (otherwise, the query would propagate the same number of levels down the tree, but with v earning more if offering a $\hat{\rho}$-split to its children). This implies that if ρ' maximizes $\chi_v^{\mathbf{g}}(.,\rho)$, then $\rho' = \rho_j^{\langle \ell+1 \rangle}$ for some $0 \leq j \leq i$ (recall that node v is at level $h-i$). Third, by definition of $\mathbf{g}$, $e_j^{\langle \ell \rangle}(\cdot)$ and $\chi_v^{\mathbf{g}}(\cdot;\cdot)$, and by Claim 4, we have that $\chi_v^{\mathbf{g}}(\rho_j^{\langle \ell+1 \rangle};\rho) = e_j^{\langle \ell \rangle}(\rho)$ for all $0 \leq j \leq h-\ell$. Finally, by Lemma 15, as $\rho_{k+2}^{\langle \ell \rangle} < \rho \leq \rho_{k+1}^{\langle \ell \rangle}$, we have that $e_k^{\langle \ell \rangle}(\rho) > e_j^{\langle \ell \rangle}(\rho)$ for all $k < j < h-\ell$ and $e_k^{\langle \ell \rangle}(\rho) \geq e_j^{\langle \ell \rangle}(\rho)$ for all $0 \leq j < k$. We have that $\rho_k^{\langle \ell+1 \rangle}$ is the only maximizers of $\chi_v^{\mathbf{g}}(\cdot;\rho)$ which forwards the query to level h, while any other maximizer forwards the query to some level $\ell' < h$. Therefore, as we assumed that nodes break ties preferring to propagate the query further down the tree, $\rho_k^{\langle \ell+1 \rangle} = t^{\langle \ell \rangle}(\rho)$ is the preferred strategy of node v when offered a ρ-split from its parent. Considering all $\rho \in \mathcal{D}_M$, condition (4) follows and the theorem is proven. □

D Proof of Theorem 8

Under h-consistency, for all $\ell \leq h$, $\rho_i^{\langle \ell \rangle}$ is defined and $\rho_i^{\langle \ell \rangle} < \rho_{i-1}^{\langle \ell \rangle}$ for all $2 \leq i \leq h - \ell + 1$. Ler $\mathbf{f}$ be a best-interest Nash equilibrium that is ℓ-tall for some $\ell \leq h$. As $\mathbf{f}$ is best-interest, for every node v up to level ℓ,

$$f_v(\rho) = \arg\max_{\rho'}\{\chi_v^{\mathbf{f}}(\rho';\rho)\}, \quad \forall \rho \in \mathcal{D}_M.$$

We prove a stronger claim than the one in the theorem, that is, for every node v up to level ℓ,

$$f_v(\rho) = g_v(\rho), \quad \forall \rho^v \leq \rho \leq \rho_1, \tag{5}$$

where ρ^v is the split offered to v by its parent under $\mathbf{f}$, $\rho_1 = 1 - 1/r^*$ and $\mathbf{g}$ is the best-interest Nash equilibrium from Definition 6.

We proceed by induction on the levels of the tree, starting from level ℓ and going backwards. In particular we prove by induction that (5) holds for every node at level ℓ, for every level $\ell' \leq \ell$. Consider any node v at level ℓ. As $\mathbf{f}$ is ℓ-tall (i.e., level ℓ is $\mathbf{f}$-reachable, while level $\ell+1$ is not), node v plays $\perp$. Therefore, v's parent (at level $\ell - 1$) has incentive to offer v a ρ_1-split (the maximum split such that v has incentive to forward the answer to its parent). It follows that $\rho^v = \rho_1$ and $f_v(\rho_1) = \perp = g_v(\rho_1)$, and (5) holds for level ℓ.

Fix $0 \le i < \ell$, and suppose (5) holds for every node at level $\ell - i$. Let $\ell' = \ell - i - 1$, and consider any node v at level ℓ'. In the proof of Theorem 7, we showed that, for every $\rho \in \mathcal{D}_M$ and $1 \le j \le i$,

$$\chi_v^{\mathbf{g}}(\rho_j^{\langle \ell'+1 \rangle}; \rho) = e_j^{\langle \ell' \rangle}(\rho).$$

By the inductive hypothesis on level $\ell' + 1 = \ell - i$ and the fact that both $\mathbf{f}$ and $\mathbf{g}$ are best-interest, we have that, for every $\rho \in \mathcal{D}_M$ and $1 \le j \le i$,

$$\chi_v^{\mathbf{f}}(\rho_j^{\langle \ell'+1 \rangle}; \rho) = \chi_v^{\mathbf{g}}(\rho_j^{\langle \ell'+1 \rangle}; \rho).$$

The last two observations imply that, for every $\rho \in \mathcal{D}_M$ and $1 \le j \le i$,

$$\chi_v^{\mathbf{f}}(\rho_j^{\langle \ell'+1 \rangle}; \rho) = e_j^{\langle \ell' \rangle}(\rho). \tag{6}$$

Lemma 15, togethet with (6), implies that

(i) for every $j < i$ and $\rho_{j+2}^{\langle \ell' \rangle} < \rho' \le \rho_{j+1}^{\langle \ell' \rangle}$, node v has incentive to play $\rho_j^{\langle \ell'+1 \rangle}$ among all $\rho_i^{\langle \ell'+1 \rangle} \le \rho \le \rho_1$, and

(ii) for $\rho' = \rho_{i+1}^{\langle \ell' \rangle}$, node v has incentive to play $\rho_i^{\langle \ell'+1 \rangle}$ among all $\rho_i^{\langle \ell'+1 \rangle} \le \rho \le \rho_1$.

We need the following technical result in order to proceed with the proof.

Claim 16. *Let v be a node at level $\ell' = \ell - i - 1$. Suppose that v receives a ρ'-split from its parent, with $\rho_{j+2}^{\langle \ell' \rangle} < \rho' \le \rho_{j+1}^{\langle \ell' \rangle}$ for some $j \le i$, and that v forwards the query exactly to level $\hat{\ell} \le \ell$. Moreover, assume that* (5) *holds for every node below v. Then, $\hat{\ell} = \ell' + j + 1$ and $f_v(\rho') = \rho_j^{\langle \ell'+1 \rangle}$.*

Proof. Let $m = \hat{\ell} - \ell' - 1$. First we show that $f_v(\rho') \le \rho_m^{\langle \ell'+1 \rangle}$, and then we argue that equality must hold. To show that $f_v(\rho') \le \rho_m^{\langle \ell'+1 \rangle}$, suppose by contradiction that $f_v(\rho') > \rho_m^{\langle \ell'+1 \rangle}$, that is, there exists $k < m$ such that $\rho_{k+1}^{\langle \ell'+1 \rangle} < f_v(\rho') \le \rho_k^{\langle \ell'+1 \rangle}$. Then, the query would only be forwarded to level $\ell'+1+j < \ell'+1+m = \hat{\ell}$, as we assumed that (5) holds for all nodes below v. This generates a contradiction, and, therefore, it must be $f_v(\rho') \le \rho_m^{\langle \ell'+1 \rangle}$.

We now show that $f_v(\rho') = \rho_m^{\langle \ell'+1 \rangle}$. As $f_v(\rho') \le \rho_m^{\langle \ell'+1 \rangle}$, we have that $\beta_v^{\mathbf{f}}(k | f_v(\rho')) \le \beta_v^{\mathbf{f}}(k | \rho_m^{\langle \ell'+1 \rangle})$ for all $1 \le k \le m$, with equality if and only if $f_v(\rho') = \rho_m^{\langle \ell'+1 \rangle}$. This yields $\chi_v^{\mathbf{f}}(f_v(\rho'); \rho') < \chi_v^{\mathbf{f}}(\rho_m^{\langle \ell'+1 \rangle}; \rho')$, for $f_v(\rho') < \rho_m^{\langle \ell'+1 \rangle}$, which implies $f_v(\rho') = \rho_m^{\langle \ell'+1 \rangle}$. By (i), it must be $m = j$, which gives $\hat{\ell} = \ell'+m+1 = \ell' + j + 1$. □

We now proceed with the proof. As $\mathbf{f}$ is ℓ-tall, $f_v(\rho^v)$ must forward the query exactly to level ℓ. We first show that $\rho^v = \rho_{i+1}^{\langle \ell' \rangle}$ and $f_v(\rho_{i+1}^{\langle \ell' \rangle}) = \rho_i^{\langle \ell'+1 \rangle}$, and then we show that (5) holds for v. Note that the claim above implies that if v receives a ρ'-split from its parent with $\rho' > \rho_{i+1}^{\langle \ell' \rangle}$, then $f_v(\rho')$ does not forward the query to level exactly ℓ. Therefore, it suffices to show that $f_v(\rho_{i+1}^{\langle \ell' \rangle}) = \rho_i^{\langle \ell'+1 \rangle}$. Indeed, this would imply that $\rho^v = \rho_{i+1}^{\langle \ell' \rangle}$, as no better (larger) split forwards the query to level exactly ℓ. By contradiction, suppose v plays $f_v(\rho_{i+1}^{\langle \ell' \rangle}) = \hat{\rho} \ne \rho_i^{\langle \ell'+1 \rangle}$. As we are assuming v is offered a $\rho_{i+1}^{\langle \ell' \rangle}$-split, and (ii) implies that v prefers to play $\rho_i^{\langle \ell'+1 \rangle}$ among all $\rho > \rho_i^{\langle \ell'+1 \rangle}$, it must be $\hat{\rho} < \rho_i^{\langle \ell'+1 \rangle}$. Moreover, it must be the case that $\hat{\rho}$ forwards the query below level ℓ, otherwise v would prefer to play $\rho_i^{\langle \ell'+1 \rangle}$. However, if it was the case, v would prefer to play $\hat{\rho}$ over $\rho_i^{\langle \ell'+1\rangle}$ when offered any ρ'-split with $\rho' < \rho_{i+1}^{\langle \ell' \rangle}$. This contradicts the assumption that $\mathbf{f}$ is ℓ-tall, for which there exists $\rho' = \rho^v$ such that $f_v(\rho')$ forwards the query exactly to level ℓ.

We showed that $\rho_v = \rho_{i+1}^{\langle \ell' \rangle}$ and $f_v(\rho_v) = g_v(\rho_v)$. To complete the inductive step, we need to prove that $f_v(\rho') = g_v(\rho')$ for all $\rho^v \leq \rho' \leq \rho_1$. Fix any $\rho^v \leq \rho' \leq \rho_1$. We already proved that $f_v(\rho')$ does not forward the query exactly to level ℓ. Moreover, $f_v(\rho')$ cannot forward the the query below level $\ell' > \ell$, as otherwise v would prefer this strategy even when offered a ρ_v-split. Thus, $f_v(\rho')$ must forward the query to some level $\hat{\ell} < \ell$, and Claim 16 concludes the proof. □

E Proof of Theorem 10

Suppose condition (2) holds, that is,

$$r^* \geq 4 \cdot h \cdot \max \left\{ 1, \max_{\substack{1 \leq \ell \leq h \\ 1 \leq i < h-\ell}} \Gamma_i^{\langle \ell \rangle} \right\}.$$

We show by induction that, if the discretization parameter M is large enough, for all $1 \leq \ell \leq h$ and $1 \leq i \leq h - \ell + 1$, $\rho_i^{\langle \ell \rangle}$ is defined and satisfies

$$1 - \frac{1}{r^* - i} < \rho_i^{\langle \ell \rangle} \leq 1 - \frac{1}{r^* - i + 1}, \tag{7}$$

that is, h-consistency holds.

By definition we have $\rho_1^{\langle \ell \rangle} = \rho_1 = 1 - 1/r^*$, for all $1 \leq \ell \leq h$. Therefore (7) holds for all $1 \leq \ell \leq h$ and $i = 1$. Fix $\ell \leq h$ and suppose the claim holds for all $\ell \leq \ell' \leq h$ and $1 \leq i \leq h - \ell'$. We recall that $\rho_{i+1}^{\langle \ell-1 \rangle}$ is defined as

$$\rho_{i+1}^{\langle \ell-1 \rangle} = \max\{\rho \in \mathcal{D}_M : e_i^{\langle \ell-1 \rangle}(\rho) \geq e_{i-1}^{\langle \ell-1 \rangle}(\rho)\},$$

where

$$e_i^{\langle \ell-1 \rangle}(\rho) = \sum_{j=1}^{i} \gamma_j^{\langle \ell-1 \rangle} \left[(1-\rho) r^* \left(\prod_{t=0}^{j-1} \rho_{i-t}^{\langle (\ell-1)+t+1 \rangle} \right) - 1 \right].$$

By definition of $\rho_{i+1}^{\langle \ell-1 \rangle}$, it must be that

$$1 - \frac{1}{r^* \Delta_i} - \frac{1}{M} \leq \rho_{i+1}^{\langle \ell-1 \rangle} \leq 1 - \frac{1}{r^* \Delta_i},$$

where

$$\Delta_i = \prod_{j=0}^{i-1} \rho_{i-j}^{\langle \ell+j \rangle} - \sum_{j=1}^{i-1} \frac{\gamma_j^{\langle \ell \rangle}}{\gamma_i^{\langle \ell \rangle}} \left[\prod_{t=0}^{j-1} \rho_{i-t-1}^{\langle \ell+t \rangle} - \prod_{t=0}^{j-1} \rho_{i-t}^{\langle \ell+t \rangle} \right],$$

and M is the discretization parameter of the domain $\mathcal{D}_M$. To see this, compute the difference $e_i^{\langle \ell-1 \rangle}(\rho_{i+1}^{\langle \ell-1 \rangle}) - e_{i-1}^{\langle \ell-1 \rangle}(\rho_{i+1}^{\langle \ell-1 \rangle})$, and argue that $1 \leq (1 - \rho_{i+1}^{\langle \ell-1 \rangle}) r^* \Delta_i \leq 1 + r^* \Delta_i / M$.

We find lower and upper bounds to the term between brackets in the expression for Δ_i. First, by the inductive hypothesis, $\rho_{i-t-1}^{\langle \ell+t \rangle} > \rho_{i-t}^{\langle \ell+t \rangle}$ for all $0 \leq t \leq h - \ell$ and $0 \leq t \leq j - 1$ (with $j < i$). Therefore, we have

$$\prod_{t=0}^{j-1} \rho_{i-t-1}^{\langle \ell+t \rangle} - \prod_{t=0}^{j-1} \rho_{i-t}^{\langle \ell+t \rangle} > 0.$$

Also by induction, we have

$$\prod_{t=0}^{j-1} \rho_{i-t-1}^{\langle \ell+t \rangle} - \prod_{t=0}^{j-1} \rho_{i-t}^{\langle \ell+t \rangle} < \prod_{t=0}^{j-1} \frac{r^* - i + t + 1}{r^* - i + t + 2} - \prod_{t=0}^{j-1} \frac{r^* - i + t - 1}{r^* - i + t}$$
$$= \frac{r^* - i + 1}{r^* - i + j + 1} - \frac{r^* - i - 1}{r^* - i + j - 1}$$
$$= \frac{2j}{(r - i + j + 1)(r - i + j - 1)} < \frac{2i}{(r^*)^2},$$

as $j < i$ in the expression of Δ_i. Therefore, again by induction, we have

$$\frac{r^* - i - 1}{r^* - 1} - \frac{2i}{(r^*)^2} \Gamma_i^{\langle \ell \rangle} < \Delta_i < \frac{r^* - i}{r^*}.$$

The upper bound on $\rho_{i+1}^{\langle \ell-1 \rangle}$ follows immediately. For the lower bound, it suffices to show that $r^* \cdot \Delta_i > (r^* - i - 1)(1 + r^*/M)$. Also, note that this would imply that $\Delta_i > 0$, and so that $\rho_{i+1}^{\langle \ell-1 \rangle}$ is defined. Rearranging the terms, it suffices to show that

$$\frac{2i}{r^*(r^* - i - 1)} \Gamma_i^{\langle \ell \rangle} < \frac{1}{r^* - 1} - \frac{r^*}{M}. \tag{8}$$

By (2), we have that

$$i \le \frac{r^*}{4} \min\{1, 1/\Gamma_i^{\langle \ell \rangle}\} - 1.$$

Then, (8) holds if

$$\frac{1/2}{1 - \frac{1}{4}\min\{1, 1/\Gamma_i\}} < 1 - \frac{(r^*)^2}{M},$$

which is satisfied for M large enough. □

F Proof of Lemma 12

Recall that, for each $i \geq 0$ we defined ϕ_i as the probability that no node at level $j \leq i$ possesses the answer, and $\lambda_i = \phi_{i-1} - \phi_i$ as the probability that a node at level i possesses the answer while no node at level $j < i$ does. Also, for every $0 \leq \ell \leq h$ and $0 \leq i \leq \ell$, we defined

$$\Gamma_i^{\langle \ell \rangle} = \frac{\sum_{j=1}^{i-1} \gamma_j^{\langle \ell \rangle}}{\gamma_i^{\langle \ell \rangle}},$$

where $\gamma_i^{\langle \ell \rangle}$ is the probability that, fixed any node v at level ℓ, there is a $\mathbf{g}$-candidate u in v's subtree at distance i from v, given that v is active. We recall that a node u at level ℓ' is a $\mathbf{g}$-candidate if, under strategy $\mathbf{g}$, u is an active answer-holder and there is no active answer-holder in the first $\ell' - 1$ levels. Let L_j be the event that there is an answer holder at level j of the tree, and F_j be the event that no event L_k happens for all $k \leq j$. Observe that $\Pr(L_j, F_{j-1}) = \lambda_j$ and $\Pr(F_j) = \phi_j$. Fix a node v at level $\ell < h$. Let L_j^v be the event that there is an answer holder in v's subtree at distance j from v, and F_j^v be the event that no L_k^v happens for all $k \leq j$. Also, let A_v be the event that v is active. We have

$$\begin{aligned}
\gamma_j^{\langle \ell \rangle} &= \Pr(L_j^v, F_{\ell+j-1} | A^v) \\
&= \Pr(L_j^v | A^v, F_{\ell+j-1}) \Pr(F_{\ell+j-1} | A^v) \\
&= \Pr(L_j | F_{j-1}) \Pr(F_{\ell+j-1} | A^v) \\
&= \frac{\Pr(L_j, F_{j-1})}{\Pr(F_{j-1})} \Pr(F_{\ell+j-1} | A^v) \\
&= \frac{\Pr(L_j, F_{j-1})}{\Pr(F_{j-1})} \frac{\Pr(A^v | F_{\ell+j-1}) \Pr(F_{\ell+j-1})}{\Pr(A^v)},
\end{aligned}$$

where the third equality follows by the fact that the branching process is memory-less, and the last equality follows by Bayes' rule. Observe that the probability that v is active only depends on the existence of answer-holders on the path from the root to v or in the subtree rooted at v. Therefore, letting P^v be the event that there is no answer-holder in the path from the root to v, we can write

$$\begin{aligned}
\Pr(A^v | F_{\ell+j-1}) &= \Pr(A^v | P^v, F_{j-1}^v) \\
&= \frac{\Pr(F_{j-1}^v | A^v, P^v) \Pr(A^v | P^v)}{\Pr(F_{j-1}^v | P^v)} \\
&= \frac{\Pr(F_{j-1}) \Pr(A^v | P^v)}{\Pr(F_{j-1}^v | P^v)},
\end{aligned}$$

where the second equality follows by Bayes' rule, and the third equality by the memory-less property of the branching factor. It follows that, for all $\ell \leq h$ and $0 \leq j \leq h - \ell$,

$$\gamma_j^{\langle \ell \rangle} = \frac{\Pr(L_j, F_{j-1}) \Pr(F_{\ell+j-1}) \Pr(A^v | P^v)}{\Pr(F_{j-1}^v | P^v) \Pr(A^v)}.$$

Plugging the last expression into the definition of $\Gamma_i^{\langle \ell \rangle}$, we get

$$\begin{aligned}
\Gamma_i^{\langle \ell \rangle} &= \frac{1}{\Pr(L_i, F_{i-1}) \Pr(F_{\ell+i-1})} \sum_{j=1}^{i-1} \Pr(L_j, F_{j-1}) \Pr(F_{\ell+j-1}) \frac{\Pr(F_{i-1}^v | P^v)}{\Pr(F_{j-1}^v | P^v)} \\
&= \frac{1}{\lambda_i \phi_{\ell+i-1}} \sum_{j=1}^{i-1} \lambda_j \phi_{\ell+j-1} \frac{\Pr(F_{i-1}^v | P^v)}{\Pr(F_{j-1}^v | P^v)}.
\end{aligned}$$

As $\Pr(F_{i-1}^v | P^v) \leq \Pr(F_{j-1}^v | P^v)$ for $j \leq i$, and $\phi_{\ell+j-1} \leq 1$, we have that

$$\Gamma_i^{\langle \ell \rangle} \leq \frac{1}{\lambda_i \phi_{\ell+i-1}} \sum_{j=1}^{i-1} \lambda_j < \frac{1}{\phi_{\ell+i-1}} \frac{1 - \phi_{i-1}}{\lambda_i}.$$

□

G Proof of Lemma 13

For all $i \geq 0$, let $\hat{\phi}_i = \phi_i/p$ be the probability that for all levels up to i no node has the answer given that the root (at level zero) does not. Observe that no node up to level $i+1$ has the answer given that the root does not if and only if the root's children and their subtrees up to depth i do not have the answer. Therefore, we have that $\hat{\phi}_{i+1} = \Psi(p \cdot \hat{\phi}_i)$, where $\Psi(x), 0 \leq x \leq 1$ is the probability generating function of the branching process. It follows that

$$\begin{aligned} \lambda_{i+1} = \phi_i - \phi_{i+1} = \phi_i p \cdot \hat{\phi}_{i+1} &= \phi_i - p \cdot \sum_{k=0}^{d} c_k \hat{\phi}_i^k p^k \\ &= \phi_i - p \sum_{k=0}^{d} c_k \phi_i^k > \phi_i - \sum_{k=0}^{d} c_k \phi_i^k = \phi_i - \Psi(\phi_i). \end{aligned} \tag{9}$$

For $0 < \epsilon \leq 1 - \zeta$ and $0 \leq z < 1 - \zeta$, let

$$a(\epsilon) = \max \left\{ \frac{1}{b-1}, \frac{1}{\epsilon} \frac{1}{1 - \Psi'(\zeta)} \right\},$$

and

$$t(z, \epsilon) = a(\epsilon) \cdot (1 - z - \Psi(1 - z)) - z.$$

We need to show that, for any $0 < \epsilon \leq 1 - \zeta$,

$$\frac{1 - \phi_i}{\lambda_{i+1}} \leq a(\epsilon).$$

Observe that, by inequality (9),

$$\frac{1 - \phi_i}{\lambda_{i+1}} \leq \frac{1 - \phi_i}{\phi_i - \Psi(\phi_i)}$$

and therefore it suffices to prove that, for every $\epsilon > 0$,

$$t(1 - \phi_i, \epsilon) = a(\epsilon) \left(\phi_i - \Psi(\phi_i) \right) - (1 - \phi_i) \geq 0.$$

First, observe that, for every $\epsilon > 0$, we have $t(0, \epsilon) = 0$, since $\Psi(1) = 1$ (see [2]). Also note that

$$\left. \frac{\partial}{\partial z} t(z, \epsilon) \right|_{z=0} = a(\epsilon) \cdot \big(\Psi'(1) - 1 \big) - 1 = a(\epsilon) \cdot (b - 1) - 1,$$

which is non-negative since $a(\epsilon) \geq 1/(b-1)$. Also, observe that $\frac{\partial^2}{\partial z^2} t(z, \epsilon) < 0$ and $\frac{\partial}{\partial \epsilon} t(z, \epsilon) > 0$, for all z and ϵ in their respective domains. Therefore, since the function $t(z, \epsilon)$ is continuous, it suffices to check that $\lim_{\epsilon \to 0} t(1 - \zeta - \epsilon, \epsilon) \geq 0$. As $(1 - b)^{-1} \leq \epsilon^{-1}(1 - \Psi'(\zeta))^{-1}$ for ϵ small enough, we have that

$$\lim_{\epsilon \to 0} t(1 - \zeta - \epsilon, \epsilon) > \lim_{\epsilon \to 0} \left[\frac{1}{1 - \Psi'(\zeta)} \frac{1}{\epsilon} \left(\zeta + \epsilon - \Psi(\zeta + \epsilon) \right) \right] - 1.$$

Since $\zeta = \Psi(\zeta)$, by l'Hôpital's rule, we conclude that $\lim_{\epsilon \to 0} t(1 - \zeta - \epsilon, \epsilon) > 0$. □

H Non-uniqueness of the Nash equilibrium

In this section, we discuss the existence of multiple Nash equilibria both in the game with fixed-payment contract of [12, 1] and in the game with split contracts presented in this work.

First we recall the setting of [12, 1]. Each node has an *integer-valued* function f_v; if v is offered a reward of $r \geq 1$ by its parent, and v does not possess the answer to the query, then v offers in turn a reward of $f_v(r) < r$ to its children. Also, by definition, $f_v(1) = 0$. Kleinberg and Raghavan [12] show that a set of strategies $\mathbf{f}$ is a Nash equilibrium if and only if, for every node v, $f_v(r^v)$ is the value x maximizing the function

$$h_v(x; r^v) = (r^v - x - 1)p_v(\mathbf{f}, x).$$

Here r^v is the reward offered to v by its parent under $\mathbf{f}$, and $p_v(\mathbf{f}, x)$ is the probability that the subtree below v yields the answer, given that v does not possess the answer and offers reward x to its children. This characterization of the Nash equilibria for the game with fixed-payment contract is analogous to our result of Lemma 2 for split contracts, where the optimization is with respect to the function $\chi_v^{\mathbf{f}}(\cdot, \rho^v)$.

Using the functions $h_v(x; r^v)$, it is possible to construct a set of strategies $\mathbf{g}^{\text{fixed}}$ which optimizes $h_v(x; r^v)$ for every node v and is therefore a Nash equilibrium of the game with fixed-payment contracts. Theorem 2.2 in [12] claims that $\mathbf{g}^{\text{fixed}}$ is the unique equilibrium, in the sense that any other Nash equilibrium $\mathbf{f}$ in which $f_v(2) = 1$ is such that for all nodes v and rewards r that are reachable at v with respect to f, $f_v(r) = g_v^{\text{fixed}}(r)$. Note that this claim would imply that all equilibria have the same efficiency, in that the query is forwarded to the same levels in every equilibrium.

Unfortunately, this claim can be showed to hold true only when restricted to *best-interest* equilibria (as in our setting, see Theorem 8), that is, when considering only equilibria where $f_v(r')$ is the value x maximizing $h_v(x; r')$, for *every* r'. Note that in a *best-interest* equilibrium, nodes choose their strategies to optimize their payoff for any possible offer they *may* receive. This suggests that equilibria that are *not* best-interest are somewhat pathological, as contain nodes who do not consider their payoff globally. It is possible to show that both games admit (non-best-interest) equilibria that can be very inefficient in the sense that the query is only forwarded to a constant number of levels in the tree no matter how large the available investment r^* is. We present one of these equilibria for the case of fixed-payment contracts (the case with split contracts is similar). Consider the set of strategies $\mathbf{f}$ in which all nodes at level 1 play $f_1(r)$, all nodes at level 2 play $f_2(r)$, and all nodes below play $f_3(r)$ (recall that the root is at level zero). For a parameter $r' \geq 4$, the functions are defined as follows.

$$f_1(r) = \begin{cases} 0, & \text{if } r = 1 \\ 1, & \text{if } r = 2 \\ 2, & \text{if } r \geq 3 \text{ and } (r - r' - 1)(\lambda_1 + \lambda_2 + \lambda_3) < (r - 2 - 1)(\lambda_1 + \lambda_2) \\ r', & \text{if } r \geq 3 \text{ and } (r - r' - 1)(\lambda_1 + \lambda_2 + \lambda_3) \geq (r - 2 - 1)(\lambda_1 + \lambda_2) \end{cases}$$

$$f_2(r) = \begin{cases} 0, & \text{if } r = 1 \\ 1, & \text{if } 2 \leq r < r' \\ 2, & \text{if } r \geq r' \end{cases}$$

$$f_3(r) = \begin{cases} 0, & \text{if } r = 1 \\ 1, & \text{if } r \geq 2 \end{cases}$$

It can be verified that $\mathbf{f}$ is a Nash equilibrium, which thus forwards the query to level at most 3, regardless of the reward r^* offered by the root to the nodes at level 1. The bottleneck in the equilibrium is created by the nodes at level 3 or more, who cannot forward the query more than a single level as they never offer their children more than 1; in light of this, the nodes at level 2 are not going to offer their children more than 2 (and they do so when receiving at least r'), and in turn the nodes at level 1 do not offer more than r'. This causes the query not to be forwarded efficiently. This phenomenon cannot happen in a best-interest equilibrium as, roughly speaking, the nodes at level 3 (or more) would consider the scenario in which they get offered an amount larger than 2 and realize that it is more convenient to offer their children an amount larger than 1 (assuming the nodes below reason similarly), therefore forwarding the query deeper down the tree.