Autonomous Agents and Multi-agent Systems manuscript No.
(will be inserted by the editor)

Negotiation in Exploration-based Environment

Israel Sofer · David Sarne · Avinatan Hassidim

Received: date / Accepted: date

Abstract When two parties need to split some reward between them, negotiation theory can predict what offers the parties will make and how the reward will be split. When a single party needs to evaluate several alternatives and choose the best among them, optimal-stopping-rule theories guide it as to how to perform the exploration, what to explore next and when to stop. We consider a model in which party A needs to choose one alternative, but has no information and no means of acquiring information on the value of each alternative. Party B, on the other hand, has no interest in what party A chooses, but can perform (costly) exploration to learn about the different alternatives. As both negotiation and exploration take time, the common deadline and discounting factor further tie the processes together. We study the combined model, providing a comprehensive game theoretic based analysis, enabling the extraction of the payments that need to be made between agents A and B, and the social welfare. Special emphasis is placed on studying the effect of interleaving negotiation and exploration, and when is this method preferred over separating the two. In addition to

Preliminary results of this work appeared in Proceedings of the Twenty-Sixth National Conference on Artificial Intelligence (AAAI-2012).

Israel Sofer
Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel
E-mail: israel.sofer@gmail.com

David Sarne
Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel
E-mail: david.sarne@gmail.com

Avinatan Hassidim
Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel
E-mail: avinatanh@gmail.com

exploring the basic questions, we also consider the case in which one of the parties has some control over the parameters of the problem (e.g. the negotiation protocol), and show how it increases the utility of this party but decreases the overall welfare.

Keywords Negotiation · Costly Exploration

1 Introduction

Negotiation is one of the main mechanisms for reaching an agreement between individuals with conflicting goals, and is an important research area in autonomous agents and multi-agent systems [49,31]. Through the process of negotiation, agents can decide how to divide among themselves the gains achieved from cooperation, thus reaching mutually beneficial agreements [58]. As such, negotiation is common in task distribution and resource allocation domains [40], e.g., when dividing a fixed amount (pie/cake/dollar) [68,70], deciding on airplane landing rules [66] and coordinating schedules using Personal Digital Assistants (PDAs) [87].

One inherent characteristic of negotiation is that it is costly and time-consuming [40,3]. The fact that negotiation takes time is commonly modeled through the use of discounting to the agent's utilities or through constant costs incurred due to the delay in task execution [40,3]. Nevertheless, the effect of the time consumed during negotiation may be far more complex. In particular, as the negotiation progresses, the environment (and consequently the issues being negotiated) may change, as the agents learn more about the world.

Negotiation models typically assume that task execution takes place only upon reaching an agreement regarding all negotiated issues [20]. Still, in many settings, an agent may prefer to negotiate over the different issues one at a time [10]. Whenever the negotiation concerns task execution, the agents may even choose to carry out the execution of the different tasks upon which an agreement was already reached prior to continuing the negotiation over the remaining issues/tasks. Consequently, when allowing task execution throughout the negotiation, any negotiation step needs to take into consideration the possible outcomes of executing the subset of tasks currently offered for exploration and their effect on the remainder of the negotiation.

One interesting type of task for which terms of execution are often negotiated between agents is *costly exploration* [15,57,71]. Costly exploration (commonly modeled as "optimal stopping rule" problem) is crucial whenever a number of possible opportunities are available, from which only one can be exploited, and the value of each opportunity is unknown. Disambiguating the value of an opportunity may be resource consuming [29,32] or associated with direct monetary costs [4,71]. The goal of costly exploration is to maximize the overall benefit, defined as the value of the opportunity eventually picked minus the costs accumulated during the exploration process. The costly exploration problem applies to a variety of real-world situations, including: job search, technology R&D, deciding on a path to route packets and many more [51,77,48].

This paper studies a negotiation protocol in which both negotiation and the task execution take time and need to be completed prior to some pre-defined deadline. Specifically, we consider the task of costly exploration as defined above. The model consists of two fully rational and self-interested agents who use a standard alternating-offers protocol. The agents have complete information on each other's preferences and on the environment settings (e.g., the opportunities sequence, discounting factor and the negotiation deadline). Only one of the agents can carry out the exploration, whereas the other agent is the one benefiting from the values revealed through the exploration. The agents negotiate over the opportunities that the exploring agent will explore and how much the beneficiary agent will pay in exchange for said exploration/s.

Our analysis distinguishes between two important variants of the above negotiation protocol. In the first (herein denoted "non-interleaved protocol"), the negotiation and exploration are two separate processes that need to be carried out under one deadline. The agents first negotiate over the entire exploration plan and the appropriate side-payment. Then, once an agreement is reached, the exploration is executed as agreed upon within the limitations of the pre-defined deadline. In the second negotiation protocol (herein denoted "interleaved protocol"), the agents negotiate over one opportunity to be explored at a time (and the appropriate payment for exploring that opportunity). The exploration of that opportunity is then executed, and the process may be resumed by negotiating the exploration of an additional opportunity, and so on. Meaning that the negotiation and exploration are interleaved with each other. A key difference between the two variants is that in the interleaved protocol the agents do not decide beforehand how the exploration will be performed and how much the beneficiary agent will pay the exploring agent, but rather negotiate it as the exploration progresses. One key advantage of the interleaved negotiation protocol is that the agent benefiting from the exploration can make sure that the other agent executes the exploration as agreed upon. Furthermore, this form of negotiation is often more intuitive for people as it does not require agreeing on a complex exploration strategy in advance. In addition, when using this protocol the decisions are based on more accurate information (e.g., the results of the explorations executed so far), hence decreasing the risk that the agents undertake. It also enables any of the agents to opt out at any time, e.g., if external conditions have changed, as it does not bind the agents to a comprehensive exploration process. The advantage of the non-interleaved protocol is that the negotiation itself takes less time (as it considers the entire exploration plan) and therefore more time can potentially be allocated for the exploration itself, potentially resulting in greater social welfare.

1.1 Application Domains

A negotiation over exploration, where one side carries the exploration for the benefit of the other, can be found in several real-life application. For example, consider a firm that needs to fill a managerial position (e.g., replacing a highly skilled professional that is about to retire). The task, which is often beyond the abilities of the HR department, can be outsourced to a headhunter who specializes in this particular areas and can approach suitable candidates employed elsewhere. The headhunter can propose an initial list of potential candidates, each associated with some a priori uncertain worth to the firm (e.g., based on her Linkdin profile or resume). The headhunter can

disambiguate this uncertainty and provide all the information necessary for determining the "worth" of a given candidate by meeting her, interviewing her and negotiating the terms and perks on behalf of the firm. Naturally this latter process is costly, especially if the potential candidates are spread over different geographic locations. The firm and the headhunter though need to negotiate over the order according the different candidates will be interviewed and under what conditions (e.g. under what findings) the search process should be terminated and concluded, as well as the payments that will be made by the firm to the headhunter. The headhunter can potentially spend days in contacting, visiting and interviewing each candidate, hence delaying or extending the negotiation between the firm and the headhunter limit the number of professionals that could potentially be interviewed, assuming the firm needs to fill-in the position by some pre-defined deadline (e.g., due to the retirement of the current position holder).
Additional domains to which the model studied in this paper can be mapped include:

– Investments - An investor negotiating with a business advisor (or an accounting firm) the order according to which the latter will perform due-diligence checks and essential background research to several different companies the investor considers investing in.
– Research and development - A company negotiating with an external research lab over the order according to which the latter will research and evaluate different technological solutions for a manufacturing problem the company encountered.
– Oil drilling - A company that purchased the drilling rights for a certain land negotiating with a contractor, capable of drilling, over the execution of exploratory drills (assuming the company will be developing one field eventually).

1.2 Main Contributions

The main contributions of this paper are threefold: First, this paper is the first to formally present and analyze a model involving negotiation and costly exploration under a unified time constraint (deadline). The analysis provided encompasses the two common negotiation modes: interleaved and non-interleaved. It enables deriving the agents' expected utilities for each step of the negotiation and the payments that need to be made. In particular, it shows that in both cases the agents will follow the optimal exploration strategy for the underlying stand-alone exploration problem, though the expected utility of any of the agents is protocol-dependent. As part of the analysis we investigate the agents' preferences over which protocol to choose and which protocol obtains a higher social welfare (i.e. expected joint utility) for a given setting. Some of the results here are somehow counter-intuitive, e.g., an agent may prefer the interleaved protocol despite the fact that, given that all other setting parameters are similar, the non-interleaved protocol is proven to yield a greater social welfare (joint expected utility). The analysis also compares the two protocols with the legacy negotiation protocol, according to which the negotiation and exploration are not interleaved and exploration does not take time, i.e., is not bounded by the negotiation deadline or discounted (the divided surplus is thus fixed throughout the negotiation). The comparison reveals that several dependencies known for the agents' utilities in the setting parameters (e.g., in the negotiation horizon and the discounting of gains) do not generally hold in our interleaved and non-interleaved protocols.

Second, this paper deals with the complexity of solving for the interleaved protocol, which is exponential in the number of opportunities. We show that for any interleaved protocol we can present a simple equivalent non-interleaved protocol. In the equivalent non-interleaved setting the agents end up with the same expected utilities as in the original problem, however the solution complexity is of the same order of magnitude as solving the stand-alone exploration problem.

Finally, we study the case where one agent has more control over the negotiation process (e.g., the negotiation horizon or the number of opportunities which will be explored every time). We show that when allowing one of the agents to choose some of the protocol parameters, the ratio between the agent utility with the extra control and the utility without it can be unbounded. Further, we show that, non-intuitively, the ability to bound the decrease in social welfare due to such an extra control is not correlated with the amount of control awarded to the agent.

1.3 The Essence of Combining Negotiation and Exploration

Our results show that the agents will follow the optimal exploration strategy for the underlying stand-alone exploration problem. Hence, the surplus that is divided in equilibrium is fixed and can be computed regardless of the negotiation. Therefore, one may wonder why not solve each problem independently, first computing the surplus, and then negotiating over it using the legacy alternating offers protocol. However this approach completely ignores the fact that negotiation takes, time, and this time comes at the expense of the time alloted for exploration, thus reducing the surplus that can be divided between the negotiating parties. Indeed computing the utility gained by each of the participants can not be done based only on the equilibrium path, but derives directly from the exploration sequence of the off path.

The following example illustrates the difference between negotiating over the optimal sequence surplus as a self contained process compared to the model describes in this paper. Consider a case in which there are two opportunities, one which has value 8 with certainty, and the other is either 5 or 10, each with 50%. To simplify the example, there is no cost to exploring the opportunities, and there is no discounting of gains. Suppose that there are 4 negotiation periods, that exploration takes one period of time, and so is making an offer.

If the exploration and the negotiation were separate, we could say that the expected value of this sequence is 9 as we can accommodate the exploration of both opportunities (hence the second opportunity will be explored first, yielding a gain of 10 with probability of 0.5, and otherwise gaining 8 which is the value of the first opportunity). Since there is no discounting of gains, the outcome of the negotiation over the expected surplus of 9 is that the last player to offer takes over the entire surplus. However, in our model, rejecting the first offer would limit the players to exploring just one opportunity, which would reduce the surplus to 8. Therefore, the gains are split, so that one player gets 1 unit of surplus, and the other player gains 8 units.

In the following section we review related work. Section 3 presents the negotiation protocol, distinguishing between the two main variants. The analysis of the two protocols as well as of the underlying stand-alone exploration problem is given in Section 4. Section 5 illustrates the differences in the resulting individual utilities and in social welfare when using the interleaved and non-interleaved protocols, and supplies a comparison with the legacy negotiation protocol. It also supplies the bounds (whenever applicable) for the resulting decrease in the social welfare and increase in individual utilities as discussed above. Finally, we conclude with a discussion and directions for future research in Section 6.

2 Related Work

Multi-agent negotiation is an active research area that has attracted the attention of many researchers, prompting several literature reviews over the years (see [49,66, 62,13,31]) as well as international competition for designing automated negotiation agents [9,8]. The purpose of the negotiation is usually to reach an agreement about the provision of a service by one agent for another, division of resources or task allocation [74,40,81]. The negotiation problem is defined by the negotiation space, which typically includes a negotiation protocol (the set of the interaction rules between the agents), negotiation objectives (the range of issues to negotiated) and negotiation strategies (the sequences of actions that the agents plan to take in order to achieve their objectives) [22]. As such, the strategy of the agents and the agreement reached over the negotiation objectives might change according to the negotiation protocol.

The common form of negotiation involves two agents, where in each step of the negotiation one of the agents makes an offer, and the other agent decides whether to accept or reject it. The agents usually take turns in being the agent who makes the offer (often referred to as "alternating offers") and the negotiation terminates upon reaching an agreement (on the issues negotiated) [68,58]. In some cases the negotiation is also constrained by a pre-defined deadline and terminates if the deadline is reached, even if the agents failed to reach an agreement [70,50,34]. The effect of time, either with or without deadlines, is usually modeled through the use of a discounting factor over the surplus divided between the agents. The discounting factor can either be common for all agents and all items [68], different for each agent [3] or different for each item [20].

The analysis proposed in this field relates to numerous negotiation protocol variants, differing in the assumptions they make regarding the negotiation mechanism. These include assumptions related to the agents' characteristics, such as the information that negotiators have, their level of rationality and level of cooperation. The information that negotiators have, e.g., about their environment [68,56], the other negotiators' types or utility functions [69,86], can be incomplete or may change during the negotiation [28,36]. Still, significant body of work assumes the agents have full information over the environment and each other preferences [24]. The level of rationality of the negotiators ranges from fully rational [85,18] to bounded rational [47, 65].[1] Finally, the agents' level of cooperation may range from self-interested agents to fully cooperative ones (i.e., when trying to maximize the social welfare) [35].

Other assumptions that differentiate between the different models are those made over the negotiation mechanism itself. For example, the negotiation can be over a single issue [68,3] or multiple-issues [19,44,86]. Another distinguishing aspect in this context is the interaction between the participating agents: The negotiation can be limited to the use of proposals and counter proposals [20,3], or can be extended to allow argumentation as well, allowing the agents to affect the negotiation state by adding information about their beliefs [75,30,1,2].

The main differentiating factor of the model presented in this paper from others is that both the negotiation itself and the resulting exploration are constrained by the same deadline. Prior work, which consider negotiation deadlines, tends to assume that the time constraint only applies to the negotiation itself, whereas the execution of the negotiated task does not depend on the time it took to reach an agreement [40,68,19]. A negotiation deadline that does affect task execution can be found in negotiation literature, mostly in the sense of a resulting decrease in the amount of time remaining in which the negotiated resource can be used. For example, Kraus and Wilkenfeld consider a case where two agents negotiate over painting a wall, and the longer the negotiation takes the less time remains available to paint the wall [39]. Another example is in their extension of the model to the multi-agent case [38]. Here three robots that are stationed on a satellite, are instructed to move a telescope from one location to another as soon as possible. A delay in moving the telescope will reduce the number of pictures sent back to scientists on earth. Yet, these works all consider a negotiation over a single item/resource, therefore the nature of the negotiated solution does not substantially change over time. In our case, the decrease in the number of explorations that can potentially be executed due to the extension of the negotiation often results in a substantially different exploration scheme to be negotiated (see Section 4.1 for more details). More importantly, when negotiating over a single item under time constraint, the issue of partial execution throughout the negotiation (i.e., "interleaved" negotiation), which is the essence of our paper, becomes irrelevant as the task can be executed only once an agreement is reached. Even in cases where multi-issue negotiation is considered, the task execution for all issues takes place only after an agreement for all issues is reached [20,10]. Consequently, in models of the latter type, the process of disambiguating the uncertainty associated with task execution does not affect the course of negotiation. All in all, while negotiation processes are commonly recognized to be interleaved with other business processes [33], to the best of our knowledge, an analysis of an interleaved negotiation of the type presented in this paper has not been introduced to date.

The idea of interleaving negotiation and task execution under time constraints resembles, in a way, the work on temporal reasoning mechanisms for Distributed Continual Planning (DCP) systems [26]. Here, the focus is on how to control and coordinate actions of multiple agents in a shared environment, where planning is adapted throughout task execution. Various continual planning tools in which plan-

[1] Much effort has been dedicated in recent years to developing automatic negotiators that can successfully negotiate with people, usually applying machine learning mechanisms [45,27,37,46,7,6].

ning and execution are interleaved can be found, mostly integrating Real-Time and Artificial Intelligence design technologies [55, 52, 76]. Recently, much effort has been dedicated to such planning mechanisms as part of DARPA's "Coordinators" project [83, 11, 5]. Most of this body of research, however, consider cooperating agents that collaborate as teammates in order to reach a shared goal [17, 61]. Some works in this area consider agents that are not necessarily aimed at forming a good collective plan but rather at ensuring that their local goals are achieved, when viewed in a global context [25, 26]. These, however, lack the negotiation aspects, and the consensus is typically reached by simple means that do not necessarily guarantee a convergence to some agreement. Another somehow similar negotiation model is the one introduced by Larson and Sandholm [41, 42]. Here, the negotiation involves a costly exploration in the sense that the computation of the benefit in individual solutions in comparison to joint solutions to a given problem that the agents face is costly. The agents thus need to decide on any step of the process to which computation to direct their resources. In our model, the exploration is of a different nature, i.e., disambiguating the value of an opportunity.

The fact that task execution is interleaved in the negotiation substantially complicates the process of computing the agents' strategies, as discussed in Section 4. While related work commonly addresses the computational complexity aspects of negotiation protocols [16, 14, 24], the analysis given there is irrelevant to our case. This is because the increased complexity in prior works is typically attributed to the task execution itself (e.g., the optimization problem in this case is NP-complete [43]) rather than to the effect of continuing the negotiation process after an agreement is reached and executed.

Another aspect that distinguishes this paper from prior work is the uncertainty associated with the process. In prior work, the uncertainty results primarily from the incomplete information (commonly modeled through a probability distribution) that the agents have about the negotiation environment, the other agents' preferences, their prospective utilities, the discounting factor that applies to each agent and the weights of the negotiation issues [56, 69, 86, 36, 3]. The uncertainty in the negotiation model analyzed in this paper, on the other hand, is attributed to the results of task execution, which affect agents' gains.

The costly exploration problem (commonly modeled as "optimal stopping rule" problem) embedded in our negotiation protocol is standard for settings where individuals need to search for an applicable opportunity while incurring an exploration cost (see several literature reviews [77, 51, 54]).[2] Over the years, many costly exploration model variants have been considered, focusing on different aspects of the model, such as the decision horizon (finite versus infinite) [48], the presence of the recall option [51], the distribution of values and the extent to which findings remain valid throughout the process [67]. The problem in its most general form and its analysis is given by Weitzman [84]. Nevertheless, despite considering settings where agents cooperate in exploration [72, 64, 63], the optimal costly exploration literature has not addressed

[2] While optimal stopping is usually discussed in the context of models such as the "secretary problem" [21], the latter does not involve search costs and the goal is to maximize the probability of finding the best candidate rather than minimizing cost, hence the great difference between the two.

exploration as part of negotiation settings and in particular as part of interleaved settings.

The game theoretic analysis used in the paper follows the standard assumptions. In particular we assume that players are rational, and that this is common knowledge. In addition, we assume that the utility is quasi linear, and the players are risk neutral[3] (trying to maximize their expected utility). Given the extensive form game played by the players, the natural solution concept is a (pure) subgame perfect equilibrium, and we prove the existence and uniqueness of this equilibrium for all our protocols. For background on game theory, see [59, 60]

3 The Negotiation Protocol

We begin with a formal description of the negotiation protocol, specifying the players, the available actions and temporal structure, outcomes and the players' utility over the outcomes.

Players The protocol assumes a negotiation setting with two agents: Agt_1 and Agt_2. Both agents are assumed to be fully rational and self-interested. It is assumed that Agt_1 is in need for carrying out an exploration process and that Agt_2 is the one actually capable of carrying out the exploration.

Available actions and temporal structure We model the protocol as an extensive-form game with perfect information (i.e., all actions are perfectly observable), perfect recall (i.e., on each stage both agents have full recollection of the moves made so far) and finite horizon. Chance nodes relate to the uncertainty associated with the exploration process. On each step of the process, one of the agents plays the role of a "proposer" and the other plays the role of a "responder". The negotiation takes place according to an alternating-offers protocol, by which the former proposer becomes the responder and vice-versa at each negotiation step. The proposer, either being Agt_1 or Agt_2, needs to propose the exploration of a subset of opportunities out of a set of n opportunities $B = \{b_1, \ldots, b_n\}$ and the compensation that will be paid by Agt_1 to Agt_2 if the proposed exploration takes place eventually. The responder can either accept or decline the proposal. Both agents can opt out from the process, at any step.

We consider two variants of the negotiation protocol. In the non-interleaved negotiation protocol the negotiation step that takes place at time period t is based on having the proposing agent (either Agt_1 or Agt_2) offer a complete exploration plan from this point onwards, denoted S^t, that Agt_2 should follow (note that even though both agents take turns in proposing only Agt_2 can perform the exploration), for a payment $M(t)$ that will be made by Agt_1 to Agt_2 at time t. The exploration plan S^t specifies the order at which opportunities will be explored and the conditions under which the exploration will terminate at any point of the process. At this point we do not limit the way the exploration plan is specified and it can take any form (e.g., rule based, decision tree). Later on we prove that the plan will necessarily be specified as

[3] At some points in the paper we also consider the distribution of payoffs, but we do not assume this is a consideration for the players.

a sequence of reservation values. If an agreement is reached the negotiation ends and Agt_2 explores the opportunities according to the exploration strategy S^t (see Figure 1(a)). The second form of negotiation is one where negotiation and exploration steps are interleaved throughout the process: at each negotiation step the proposing agent (either Agt_1 or Agt_2) offers the exploration of a single opportunity b_i (by Agt_2) in exchange for a payment $M(t)$ (made by Agt_1 to Agt_2). If an agreement is reached Agt_2 explores the agreed opportunity and the agents advance to the next negotiation step (see Figure 1(b)). We emphasize that the sequence of opportunities over which the agents negotiate is not fixed a priori—it can be implicitly decided by the agents.

We assume that each negotiation step takes $j > 0$ time periods and each exploration step takes $k \geq 0$ time periods.[4] The process externally terminates after T time periods or when one of the agents chooses to opt out. The value of T is thus the overall deadline set for the process (negotiation and exploration).

Outcomes Each opportunity b_i encapsulates a value v_i (representing an expense, utility, etc.) unknown to the agents. The value of opportunity b_i derives from a probability distribution function, denoted $f_i(x)$. By exploring opportunity b_i, its value v_i is disambiguated (i.e., revealed to the agents). The agent executing the exploration (Agt_2) incurs a fee (cost), denoted c_i.

Players' utility over the outcomes The value v_i of an opportunity b_i is absolute and after the exploration it can be exploited only by Agt_1. While any explored opportunity is applicable for Agt_1, it is only capable of exploiting one. In the case where none of the opportunities is being exploited, Agt_1 obtains a default value v_0. Thus, given several opportunities in which values were revealed, the agent prefers exploiting the one associated with the highest or lowest value, denoted "best".[5] In this paper we will consider the overall benefit-maximization problem, hence the best value is the highest one, though the same analysis applies to the minimization problem.

The protocol assumes that gains, exploration costs and payments made are discounted using a discounting factor δ (per time period). For simplicity, the protocol assumes that only after an exploration step is completed (i.e. after k time periods) the cost c_i is incurred and the value v_i is obtained. The payment $M(t)$ is made upon acceptance, before the exploration takes place, i.e., j time periods after the proposal is made. For example, a proposal made at time t will define a payment $M(t + j)$ that will be made (if accepted) at time $t + j$. The utility of Agt_2 out of the negotiation, denoted U_2, is the discounted sum of payments it receives and the exploration fees it incurs throughout the negotiation. The utility of Agt_1, denoted U_1, is the difference between the discounted best value that was revealed by the time the negotiation terminated and the initial default value, v_0, minus the discounted sum of payments it makes to Agt_2. Given a negotiation setting $(v_0, \delta, B, T, k, j)$, the goal is to find the

[4] The case of $j = 0$ is ill-defined as it enables an infinite negotiation.

[5] The choice of minimum or maximum is application-dependent. For example, if the opportunities are different production technologies, as in the R&D example, then the company will pick the one associated with the minimum cost. If the opportunities are interviewees (potential employees in the headhunting example), the value of each opportunity represents the company's benefit from hiring her, and the firm will recruit the one associated with the maximum value among those interviewed.

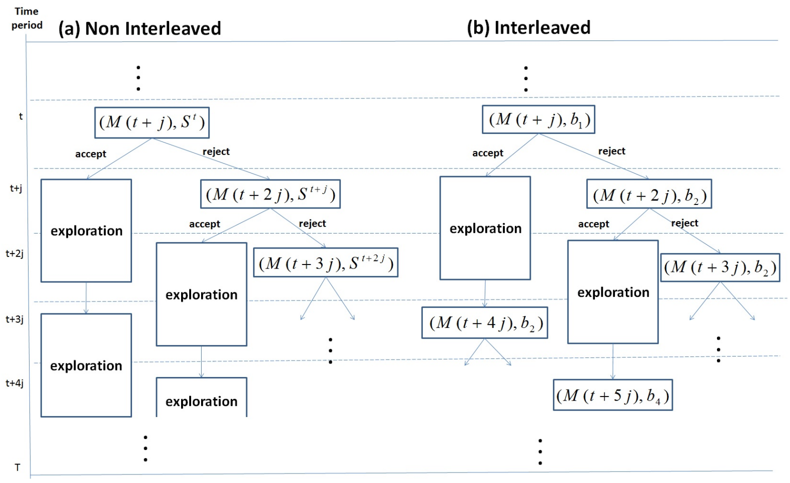


Fig. 1: Interleaved and non-interleaved protocols when an exploration takes twice as many time periods as a negotiation phase (k=2j).

proposals that will be made throughout the negotiation and the resulting utilities U_1 and U_2.

We emphasize that the setting studied is a complete information one, meaning that both agents are familiar with the distribution function, $f_i(x)$, and the cost, c_i, for each opportunity in B and the other parameters (v_0, δ, T, k, j) in the setting. The only parameter that is not a common knowledge (though symmetric for both agents) is the value, v_i, encapsulates for each opportunity and can only be revealed through exploring the opportunity.

Table 1 provides a mapping of the four applications mentioned in Subsection 1.1 to the model outlined above.

We illustrate the two protocol variants using the following toy example: assume there are two opportunities, b_1 and b_2, where b_1 is associated with a uniform distribution between 0 and 10 and b_2 is associated with a uniform distribution between 7 and 9. The exploration cost of both opportunities is $c = 0.2$, the default known value is $v_0 = 0$ and the discounting factor is $\delta = 0.9$. Finally, the process deadline is $T = 4$ and both negotiation and exploration steps take one time period (i.e., $j = 1$ and $k = 1$). Figure 2 illustrates a possible process flow for this example when using the non-interleaved protocol. Agt_1 first proposes to explore b_1 and then b_2 for a payment $M(2) = 4$. Agt_2 rejects the offer, the agents change roles and Agt_2 proposes next. Agt_2 proposes to explore b_1 and then b_2 for a payment $M(3) = 6$. Agt_1 accepts the offer and Agt_2 explores the opportunities one after the other, and in this example obtains the value 8 for b_1 and 9 for b_2. The utility of Agt_1 will be the discounted maximum value found minus the discounted payment M, i.e., $U_1 = 0.9^4 * 9 - 0.9^2 * 6 = 1.04$. The utility of Agt_2 will be the discounted sum of payments minus the discounted exploration fees, i.e., $U_2 = 0.9^2 * 6 - 0.9^3 * 0.2 - 0.9^4 * 0.2 = 4.58$.

Application	Agt_1	Agt_2	Opportunities	Opportunity value	Exploration cost	Deadline justification
Head hunting	firm that needs to fill a managerial position	specialized head-hunter	suitable candidates employed elsewhere	candidate's worth to the firm (taking into account skills, experience and requested salary and perks)	meeting the candidate, interviewing her and negotiating the terms of employment	a due-date by which the position needs to be filled (e.g., the grace period of currently resigning employee that holds this position)
Investments	investor interested in investing in a company	business advisor or an accountant	companies the investor considers investing in	true worth of the company	cost of due-diligence checks and essential background research	beginning of a new fiscal year; date by which funds become available
R&D	company seeking a technological solution for a manufacturing problem	external research lab	different production technologies	cost of implementing the proposed solution	R&D costs; materials, mokups and prototypes	due date for production
Oil drilling	company that purchased the drilling rights for a certain land	contractor with drilling equipment and expertise	different locations for exploratory drills	estimated amount of oil gallons that can be produced	cost of exploratory drills (varies according to location, terrain structure, etc.)	deadline for developing the field (according to the license)

Table 1: The mapping of the four applications mentioned in 1.1 to the studied negotiation model.

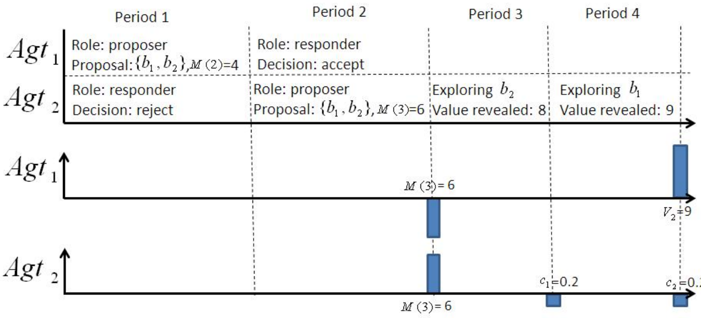


Fig. 2: A possible process flow in the non-interleaved protocol for the toy example given in the text. The upper part relates to the negotiation and exploration carried out, and the bottom part depicts the utility gains/losses.

In the interleaved protocol the agents negotiate over the payment for exploring one opportunity at a time as illustrated in Figure 3 for a specific possible process flow. Here, Agt_1 proposes to explore b_1 in the first time period for a payment $M(2) = 4$. Agt_2 accepts the offer in the first time period, thus that opportunity is explored in the second time period, revealing a value of 8. Based on the value revealed, Agt_2 proposes to explore b_2 for a payment $M(4) = 1$. Agt_1 accepts, and the exploration

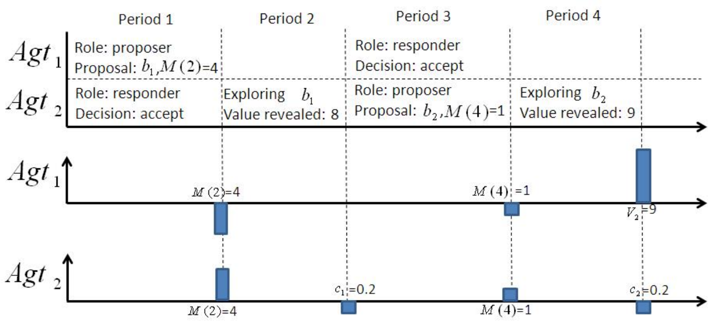


Fig. 3: A possible process flow in the interleaved protocol for the toy example given in the text. The upper part relates to the negotiation and exploration carried out, and the bottom part depicts the utility gains/losses.

takes place during the fourth time period, revealing a value of 9. The utility of Agt_1 will be the discounted maximum value found minus the discounted payments, i.e., $U_1 = 0.9^4 * 9 - 0.9 * 4 - 0.9^3 * 1 = 1.58$. The utility of Agt_2 will be the discounted sum of payments minus the discounted exploration fees, i.e., $U_2 = 0.9 * 4 - 0.9^2 * 0.2 + 0.9^3 * 1 - 0.9^4 * 0.2 = 4.04$.

4 Analysis

We first review the solution to the exploration problem if it is carried out stand-alone and then analyze the solution to the negotiation problem.

4.1 Optimal Exploration

The exploration setting embedded in the negotiation protocol augments the canonical sequential exploration problem described by Weitzman [84] when considering a finite number of exploration periods. The problem (referred to as "the stand-alone exploration problem" onwards) can be defined by the tuple (v_0, δ, B, T, k). The optimal exploration strategy in this case is the one that maximizes the discounted expected value obtained when the process terminates minus the discounted expected sum of costs incurred along the exploration. A strategy S is thus a mapping of a world state $W = (v, \mathcal{B}, T-t+1)$ to an opportunity $b_i \in \mathcal{B}$ which value should be obtained next, where v is the maximum value obtained by the agent so far, $\mathcal{B} \subseteq B$ is the set of opportunities with values still unknown and t is the current time period, hence $T-t+1$ is the number of remaining potential exploration periods.[6] ($b_i = \emptyset$ if the process is to be terminated at this point.)

The complexity of solving the stand-alone exploration problem is setting-dependent. For example, if the deadline T enables the agent to potentially explore all the opportunities in $\mathcal{B}$ if requested to do so, then the solution complexity is linear in the number

[6] While the proper representation should use a single variable to represent the number of remaining periods, we prefer the use of $T - t$ as it later coincides with the negotiation analysis.

of opportunities [84]. The optimal exploration strategy in this case is based on setting for each opportunity b_i a reservation value, denoted r_i, satisfying:[7]

$$\delta^k c_i = \delta^k \int_{x=r_i}^{\infty} (x - r_i) f_i(x) dx - (1 - \delta^k) r_i. \tag{1}$$

Given the set of reservation values $\{r_1, ..., r_n\}$, the agent should explore the opportunity associated with the highest reservation value and terminate the exploration process once the maximum value found so far is greater than the maximum reservation value among those assigned to the remaining unexplored opportunities[8]. Intuitively, the reservation value r_i can be seen as the value for which the agent is indifferent to the exploration of opportunity b_i (i.e., the exploration cost c_i equals the discounted improvement in the value it will be able to exploit). Meaning the value r_i for which the discounted cost of exploring the opportunity (left hand side) equals the discounted marginal gain from exploring b_i in case a better value is found (first term on right hand side) minus the loss in exploitation value due to postponing the exploitation for additional k periods (second term on the right hand side).

To illustrate the nature of the solution, consider the example used in Figures 2 and 3 where we have two opportunities to explore and $k = 1$. The reservation value of opportunities b_1 and b_2 according to Equation 1 are: $r_1 = 8$ and $r_2 = 7.32$. The optimal strategy is thus to first explore b_1. If the value v_1 found in b_1 is greater than 7.32 then the agent should terminate the exploration, ending up with v_1. Otherwise the agent should explore opportunity b_2, ending up with $\max(v_1, v_2)$.

Another case where the solution is relatively simple is when the opportunities are homogeneous (i.e., associated with the same distribution of values and exploration costs). Here, regardless of the deadline T, the solution is based on reservation values, and can be calculated by substituting $c_i = c$, $r_i = r$ and $f_i(x) = f(x)\ \forall i$ in Equation 1. According to this solution, opportunities are explored in a random order and the exploration resumes as long as the value obtained is below the calculated reservation value [51]. The solution complexity is constant.

In other cases, e.g., when opportunities are heterogeneous and the number of opportunities which can be explored is bounded (i.e., the exploration horizon becomes a constraint over the number of opportunities that can be explored), the optimal decision rule regarding the next opportunity to explore may not be based on reservation values, and the computational complexity of the optimal exploration strategy may increase substantially. For example, consider a case of three heterogeneous opportunities with discrete distributions of values as described in Figure 4 and assume that only two opportunities can be explored (i.e., $T = 2$ and $k = 1$) and also that

[7] This equation is different from [84] in the sense that the cost is multiplied by δ^k because it is incurred after the exploration, whereas in [84] the exploration cost is incurred before the exploration takes place. This, however, does not qualitatively change the results reported in this paper.

[8] In this sense the reservation value is just a threshold, and if the value of this threshold is too small the agent halts the exploraton

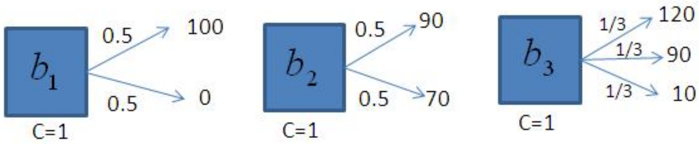


Fig. 4: An example of three heterogeneous opportunities. The optimal exploration strategy here cannot be represented as a sequence of reservation values.

$\delta = 1$. In an attempt to solve according to Equation 1[9], the resulting reservation values are $r_1 = 98$ (resulting from: $1 = 0.5(100 - r_1)$), $r_2 = 88$ (resulting from: $1 = 0.5(90 - r_2)$) and $r_3 = 117$ (resulting from: $1 = \frac{1}{3}(120 - r_3)$) for b_1, b_2 and b_3 respectively. The order of exploring the opportunities according to the reservation value rule is b_3, b_1, b_2, however in this case only two opportunities can be explored. If exploring b_3 and then b_1 according to the reservation value rule, then the expected benefit is: $\frac{1}{3} \times 120 + \frac{2}{3}[\frac{1}{2}(\frac{1}{2} \times 100 + \frac{1}{2} \times 90) + \frac{1}{2}(\frac{1}{2} \times 100 + \frac{1}{2} \times 10)] = 90$. However, consider an alternative exploration strategy according to which b_3 is explored first. If the value obtained in b_3 is $v_3 = 10$ then opportunity b_2 is explored, and if $v_3 = 90$ then opportunity b_1 is explored. The expected benefit in this case is $\frac{1}{3} \times 120 + \frac{2}{3}[\frac{1}{2}(\frac{1}{2} \times 100 + \frac{1}{2} \times 90) + \frac{1}{2}(\frac{1}{2} \times 90 + \frac{1}{2} \times 70)] = 98\frac{1}{3}$, which dominates the reservation-value-based strategy.

In this case, the problem cannot be mapped to [84]. In that paper, Weitzman conjectured that the problem of extracting a solution when the number of opportunities is greater than the number of exploration opportunities is difficult. As far as we know, the best solution to date is the trivial one, namely to evaluate, for each subset of remaining unexplored opportunities the best value obtained so far, and the benefit in continuing the exploration of each of the yet unexplored opportunities. This solution can be represented as a decision tree of the opportunities' exploration order and the potential observed values. However, the size of this tree is exponential in the number of opportunities still available for exploration and the number of values in the distribution of the different opportunities.

4.2 Non-interleaved Negotiation

We begin with the analysis of the non-interleaved negotiation protocol. The analysis uses a standard backward induction technique [78] based on the sub-game perfect equilibrium concept, starting from the last step and then going backwards [68,73]. The idea is that the payment offered in exchange for the exploration that the proposal defines will guarantee the responding agent's indifference between accepting and rejecting the proposal (and thus the offer is necessarily accepted) [80,3,40]. Once an

[9] For the discrete case the calculation of the reservation value is similar, replacing the integral by a sum and the probability distribution function with discrete probability P_i in Equation 1:

$$\delta^k c_i = \delta^k \sum_{x=r_i}^{\infty}(x - r_i)P_i(x) - (1 - \delta^k)r_i. \qquad (2)$$

agreement is reached, the negotiation terminates and the exploration is executed by Agt_2 according to the exploration strategy agreed upon. Proposition 1 suggests that both agents will never opt out of the negotiation (before the deadline), once started, until an agreement is reached.

Proposition 1 *If the proposer finds it beneficial to continue negotiating at time t, then at any time $t^* > t$, where $T - t^* + 1 \geq j + k$, the proposer (regardless of its identity) will necessarily find it beneficial to continue with the negotiation.*

Proof Assume that it is beneficial to continue negotiating at time t when the best value obtained so far is v. Since the proposer finds it beneficial to resume the negotiation, there must be at least one opportunity which expected discounted improvement of the value v is greater than its discounted exploration cost (i.e., $\delta^k \int_{x=v}^{\infty}(x - v)f_i(x)dx - (1 - \delta^k)v > c_i$). If the agents reached time $t^* > t$ and an agreement has not been yet reached (i.e., no exploration takes place in the non-interleaved protocol prior to reaching an agreement), then that opportunity is necessarily still relevant. Therefore, further exploration is beneficial and the benefit generated through such exploration will be divided through negotiation. The condition $T - t^* + 1 \geq j + k$ is required in order to guarantee that at least one exploration can be executed after the negotiation that will take place at time t^*. ■

We use $U_p(t)$ to denote the expected utility of agent p (where $p = 1, 2$) from time period t onwards and $U_{\bar{p}}(t)$ to denote the utility of the other agent. The expected benefit of agent p from the process as a whole, denoted U_p, is thus given by $U_p = U_p(0)$. Since value in our problem is generated only through exploration, the joint utility of both agents from the negotiation equals the value generated from explorations carried out in the underlying stand-alone exploration problem. We use $EV(t)$ to denote the expected discounted benefit (overall) from following the optimal strategy over the underlying stand-alone exploration problem from time step t onwards. Similarly, we use $EC(t)$ to denote the expected discounted cost that Agt_2 will incur in order to explore the opportunities according to the optimal strategy for the underlying stand-alone exploration problem from time step t onwards. The value $EC(t)$ is the cost part of $EV(t)$, hence the improvement in the exploited value for Agt_1 from time period t onwards is greater than $EV(t)$. However, its expected utility $U_1(t)$ is less than $EV(t)$ because of the side-payments. Similarly, while Agt_2 incurs $EC(t)$, its expected utility $U_2(t)$ is positive, as it takes into consideration also the side-payments obtained from Agt_1.

Theorem 1 presents the agents' utilities from the negotiation and the payment $M(t)$ in each step, for the unique subgame perfect equilibrium.

Theorem 1 *The unique sub-game perfect equilibrium in the non-interleaved model variant is to have at any time step t the proposing agent offering an exploration strategy according to the optimal exploration solution for the stand-alone problem $(v_0, \delta, B, T - t - j + 1, k)$ for a payment $M(t + j)$:*

$$M(t+j) = \begin{cases} U_2(t+j) + EC(t+j) & \textit{if proposer} = Agt_1 \\ EV(t+j) - U_1(t+j) + EC(t+j) & \textit{if proposer} = Agt_2 \end{cases} \quad (3)$$

This offer is necessarily accepted and the agents' expected utilities $U_1(t)$ and $U_2(t)$ are:

$$U_p(t) = \begin{cases} \Sigma_{i=1}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^{i+1}\delta^{ij})EV(t+ij) & \textit{if p = proposer} \\ \Sigma_{i=2}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^{i}\delta^{ij})EV(t+ij) & \textit{if p = responder} \end{cases} \quad (4)$$

Proof We first prove that the resulting exploration is identical to the optimal exploration if carried out stand-alone. Then we prove that the payment is necessarily made according to (3) and that the resulting agents' utilities are given by (4).

The proposer p, when making an offer at time t, merely needs to guarantee that the responder's utility at time $t+j$ if receiving its offer, $U_{\bar{p}}(t+j)$, is equal to its utility at that time if rejecting the current offer. Since the discounted sum of utilities at time t if the offer is accepted is given by the expected benefit of the offered exploration plan, denoted $\delta^j EV'(t+j)$, the proposer's discounted expected utility, which it attempts to maximize, is $\delta^j EV'(t+j) - \delta^j U_{\bar{p}}(t+j)$. This latter term is maximized when $EV'(t+j)$ is maximized, i.e., when offering to follow the optimal exploration strategy, resulting in $EV(t+j)$.

As for the payment $M(t+j)$, we distinguish between two cases. The first is when the proposing agent is Agt_1. In this case, it needs to make sure that Agt_2's utility at time $t+j$ from accepting the proposal equals its utility in case of rejection, $U_2(t+j)$. Since, by accepting the offer Agt_2 incurs a cost $EC(t+j)$, the payment $M(t+j)$ needs to be the sum of the last two terms, i.e., $U_2(t+j) + EC(t+j)$. Similarly, when the proposing agent is Agt_2, it needs to make sure that Agt_1's utility from accepting the proposal at time $t+j$ equals its utility in case of rejection, $U_1(t+j)$. Since Agt_1's expected benefit at time $t+j$ from the exploration carried out by Agt_2 is $EV(t+j) + EC(t+j)$ (as Agt_2 is the one incurring the costs), and Agt_2 needs to guarantee that Agt_1 receives only $U_1(t+j)$, it will request that the payment, $M(t+j)$, will be $EV(t+j) - U_1(t+j) + EC(t+j)$.

Now, we need to prove the validity of (4). The proof is inductive, showing that if the agents' utilities are given by (4) for any time $t' > t$ then their utilities at time t are also given by that equation. Since the proposer needs to guarantee the responder's discounted next step's utility, an agent's utility in time step t is given by:

$$U_p(t) = \begin{cases} \delta^j EV(t+j) - \delta^j U_{\bar{p}}(t+j) & \text{if p = proposer} \\ \delta^j U_p(t+j) & \text{if p = responder} \end{cases} \quad (5)$$

Using (5) we can validate the correctness of (4) for time t where no more than a single negotiation step can be carried out while leaving time for at least one exploration within the remaining time (i.e., for time $t = \lfloor \frac{T-k}{j} \rfloor \cdot j - j + 1$). When getting to the last negotiation step, the proposer takes over the entire exploration's expected benefit and the responder receives no expected benefit, i.e,:

$$U_p(\lfloor \frac{T-k}{j} \rfloor \cdot j - j + 1) = \begin{cases} \delta^j EV(\lfloor \frac{T-k}{j} \rfloor \cdot j + 1) & \text{if p = proposer} \\ 0 & \text{if p = responder} \end{cases} \quad (6)$$

Substituting $t = \lfloor \frac{T-k}{j} \rfloor \cdot j - j + 1$ in (4) obtains an upper index (for the sum) of $\lfloor \frac{T - (\lfloor \frac{T-k}{j} \rfloor \cdot j - j + 1) - k + 1}{j} \rfloor$, which equals 1. Therefore (4) is the same as (6).

According to the inductive assumption, the agents' utilities for any time $t' > t$ are given by (4). Therefore, substituting $U_p(t+j)$ and $U_{\bar{p}}(t+j)$ according to (4) in (5) obtains:

$$U_p(t) = \begin{cases} \delta^j EV(t+j) - \delta^j (\Sigma_{w=1}^{\lfloor \frac{T-(t+j)-k+1}{j} \rfloor}((-1)^{w+1}\delta^{wj})EV(t+j+wj)) & \text{if p = proposer} \\ \delta^j (\Sigma_{w=1}^{\lfloor \frac{T-(t+j)-k+1}{j} \rfloor}((-1)^{w+1}\delta^{wj})EV(t+j+wj)) & \text{if p = responder} \end{cases} \quad (7)$$

Substituting $i = w + 1$ in (7) obtains:

$$U_p(t) = \begin{cases} \delta^j EV(t+j) - (\Sigma_{i=2}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^i\delta^{ij})EV(t+ij)) & \text{if p = proposer} \\ \Sigma_{i=2}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^i\delta^{ij})EV(t+ij) & \text{if p = responder} \end{cases} \quad (8)$$

and since $\delta^j EV(t+j) - (\Sigma_{i=2}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^i\delta^{ij})EV(t+ij)) = \Sigma_{i=1}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^i\delta^{ij})EV(t+ij)$ we obtained (4).

The uniqueness of the solution results from the fact that in the last negotiation step the unique solution is to take over the entire exploration's expected benefit. Similarly in any of the inductive steps backward the only possible solution is to offer the other side exactly its discounted expected benefit if rejecting the proposal.[10] ■

When $k = 0$, i.e., when the exploration takes no time at all, all opportunities can be potentially explored once an agreement is reached. In this case the solution to the stand-alone problem does not depend on the remaining time $T - t + 1$ (i.e., the solution to $(v_0, \delta, B, T - t + 1, 0)$ is the same regardless of the value of t). Moreover, since all opportunities can potentially be explored, the optimal strategy can be extracted simply by calculating the appropriate reservation values according to (1). The optimal exploration will follow the order of the reservation values, terminating the exploration process once the maximum value found so far is greater than the maximum reservation value among those assigned to the remaining unexplored opportunities (see Section 4.1). For the same reasons, the value of $EV(t)$ in this case does not depend on t, i.e., $EV(t) = EV\ \forall t \leq T$. Furthermore, in case $k = 0$ and $j = 1$ (i.e., each negotiation step takes one time period), we obtain:

$$U_p(t) = \begin{cases} (\delta - \Sigma_{i=2}^{T-t+1}(-1)^i\delta^i)EV & \text{if p = proposer} \\ \Sigma_{i=2}^{T-t+1}(-1)^i\delta^i EV & \text{if p = responder} \end{cases} \quad (9)$$

This latter result is in fact the solution for the legacy negotiation protocol with no exploration as appears in [20].[11]

[10] In some degenerate cases the optimal exploration itself is not unique and there is possible more than one expected-benefit-maximizing way to explore the opportunities (e.g., when two opportunities have the same reservation value). In this case we the agents can follow an arbitrary pre-defined sequencing rule for the opportunities associated with the same reservation value when constructing their offers.

[11] See Equation 1 in [20], which defines the portion out of the δEV pie that is being divided between the two parties.

Non-interleaved Negotiation with Immediate Exploration An important variant of the non-interleaved negotiation protocol is the one where upon acceptance of a proposal the exploration process starts immediately (rather than waiting j time periods). This case is common in domains where upon a rejection of a proposal the two sides need to spend time preparing their counter proposal, whereas acceptance does not involve any delays in the execution of the exploration. Formally, we no longer use the parameter j to indicate the amount of time that the agents need to wait until a new proposal can be made or until the exploration can start executing from the time the proposal was made . Instead, we distinguish between: j_{reject}, denoting the amount of time (measured in time periods) that the agents need to wait between one proposal and the next if the first was rejected, and j_{accept}, denoting the amount of time (measured in time periods) that the agents need to wait from the time a proposal was made until exploration according to this proposal can start executing, if that proposal is accepted. The non-interleaved negotiation with immediate exploration protocol is thus defined as: $j_{accept} = 0$, $j_{reject} > 0$ and $k \geq 0$. The solution for this latter protocol is similar to the one given above for the general non-interleaved negotiation protocol: instead of using $EV(t + ij)$ in Equation 4 we use $EV(t)$ and there is no need to discount the expected benefit when the proposal is accepted (i.e., instead of using δ^{ij} in Equation 4 we use $\delta^{(i-1)j}$). The agents' utilities are thus calculated as:

$$U_p(t) = \begin{cases} \Sigma_{i=1}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^{i+1}\delta^{(i-1)j})EV(t) & \text{if p = proposer} \\ \Sigma_{i=2}^{\lfloor \frac{T-t-k+1}{j} \rfloor}((-1)^{i}\delta^{(i-1)j})EV(t) & \text{if p = responder} \end{cases} \tag{10}$$

Similarly, the appropriate modification of Equation 3 describing the payments made in this case becomes (replacing $EC(t + j)$ with $EC(t)$ and $EV(t + j)$ with $EV(t)$):[12]

$$M(t) = \begin{cases} \delta^j U_2(t+j) + EC(t) & \text{if proposer} = Agt_1 \\ EV(t) - \delta^j U_1(t+j) + EC(t) & \text{if proposer} = Agt_2 \end{cases} \tag{11}$$

A corollary of Theorem 1 is that at time period t in the non-interleaved negotiation with immediate exploration protocol, the proposing agent will offer an exploration strategy according to the optimal exploration solution for the stand-alone problem $(v_0, \delta, B, T - t + 1, k)$.

The importance of the non-interleaved with immediate exploration negotiation protocol is that, as we show later in this section, it facilitates the solution of the interleaved negotiation protocol that is given in the following paragraphs, substantially reducing its computational complexity.

4.3 Interleaved Negotiation Protocol

Next we analyze the interleaved negotiation protocol where, in each step, the proposer offers the exploration of a specific opportunity for some payment $M(t)$. In this protocol, in some situations, depending on the best value found so far, the agents will

[12] Notice that in this case the payment is made at the time of the proposal, rather than at time $t + j$.

opt out in step $t < T$ of the negotiation (in contrast to Proposition 1 which precludes such scenarios in the non-interleaved case). As with the non-interleaved protocol, the analysis of the negotiation process uses a standard backward induction technique [78]: the payment proposed for exploring opportunity b_i guarantees that the responding agent is indifferent between accepting and rejecting the proposal (and thus the offer is necessarily accepted).

Given a negotiation setting $(v_0, \delta, B, T, j, k)$ we define a negotiation step as a tuple $(t, v, \mathcal{B})$ representing reaching t with a set $\mathcal{B} \subseteq B$ of opportunities that are available to explore, and the best value found so far is v. We use $U_P(t, v, \mathcal{B})$ to denote the expected utility gain (onwards) of agent p if in state $(t, v, \mathcal{B})$. In the case of Agt_1, $U_P(t, v, \mathcal{B})$ captures the discounted expected improvement in its exploitation value due to the explorations to come minus the expected discounted payments made to Agt_2 from this step onwards. Similarly, for Agt_2, $U_P(t, v, \mathcal{B})$ represents the expected discounted payments received minus the expected discounted exploration costs incurred throughout future explorations.

Theorem 2 unfolds the unique subgame perfect equilibrium in the interleaved model variant, including the agents' utilities $U_P(t, v, \mathcal{B})$ from the negotiation and the payment $M(t)$ offered at each step.

Theorem 2 *The unique sub-game perfect equilibrium in the interleaved model variant mandates that when the agents reach state $(t, v, \mathcal{B})$, the proposer will choose to either: (a) terminate the negotiation if the optimal solution to the stand-alone exploration problem $(v, \mathcal{B}, T - t + 1, j + k)$ is to terminate the exploration, or otherwise (b) offer to explore the first opportunity according to the optimal exploration strategy for the stand-alone problem $(v, \mathcal{B}, T - t + 1, j + k)$, for a payment:*

$$M(t, v, \mathcal{B}) = \begin{cases} U_2(t+j, v, \mathcal{B}) + \delta^k c_i \\ -\delta^k \int_{y=-\infty}^{\infty} U_2(t+j+k, \max(v,y), \mathcal{B} - b_i) f_i(y) dy & \textit{if proposer} = Agt_1 \\ \delta^k \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y) dy - U_1(t+j, v, \mathcal{B}) \\ +\delta^k \int_{y=-\infty}^{\infty} U_1(t+j+k, \max(v,y), \mathcal{B} - b_i) f_i(y) dy & \textit{if proposer} = Agt_2 \end{cases} \quad (12)$$

Where $U_P(t, v, \mathcal{B})$ is given by:

$$U_1(t, v, \mathcal{B}) = \begin{cases} -\delta^j M(t+j, v, \mathcal{B}) + \delta^{j+k} \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y) dy \\ +\delta^{j+k} \int_{y=-\infty}^{\infty} U_1(t+j+k, \max(v,y), \mathcal{B} - b_i) f_i(y) dy & \textit{if } \; T - t + 1 \geq j + k \\ 0 & otherwise \end{cases} \quad (13)$$

$$U_2(t, v, \mathcal{B}) = \begin{cases} \delta^{j+k} \int_{y=-\infty}^{\infty} U_2(t+j+k, \max(v,y), \mathcal{B} - b_i) f_i(y) dy \\ +\delta^j M(t+j, v, \mathcal{B}) - \delta^{j+k} c_i & \textit{if } \; T - t + 1 \geq j + k \\ 0 & otherwise \end{cases} \quad (14)$$

This offer is necessarily accepted.

Proof We first prove that Equations 13 and 14 capture the expected utility of Agt_1 and Agt_2, respectively, if the proposer's proposal to explore opportunity b_i is accepted. Then, we prove that the payment that makes the responder indifferent between accepting or rejecting the proposal is captured by Equation 12. Finally, we prove that opportunity b_i is necessarily selected according to the optimal exploration strategy for the appropriate stand-alone exploration problem.

If the proposal is accepted and opportunity b_i is explored, then Agt_1's expected utility is the sum of the discounted payment $M(t+j, v, \mathcal{B})$ it makes to Agt_2 at time $t+j$ (i.e., discounted by δ^j), the discounted expected improvement in its current value after exploring b_i, given by $\int_{y=-\infty}^{\infty}(\max(v, y) - \frac{v}{\delta^{j+k}})f_i(y)dy$ (which becomes applicable at time $t + j + k$ thus discounted by δ^{j+k}) and its expected discounted benefit from the next negotiation (that takes place in $j + k$ periods, thus discounted by δ^{j+k}) onwards, given by $\int_{y=-\infty}^{\infty} U_1(t + j + k, \max(v, y), \mathcal{B} - b_i)f_i(y)dy$, which corresponds to Equation 13. Agt_2's expected utility in this case is the sum of the discounted payment $M(t + j, v, \mathcal{B})$ it receives from Agt_1 at time $t + j$ (thus discounted by δ^j), its discounted cost, c_i, of exploring opportunity b_i (paid at time $t + j + k$, thus discounted by δ^{j+k}) and its expected discounted benefit from the following negotiation step (starting in $j + k$ periods, thus discounted by δ^{j+k}) onwards, given by $\int_{y=-\infty}^{\infty} U_2(t + j + k, \max(v, y), \mathcal{B} - b_i)f_i(y)dy$, which corresponds to Equation 14. In both Equations 13 and 14, if no additional negotiation can be carried out while leaving time for exploration, the agents' expected utilities is zero, as there is no further benefit that can be generated and divided between the agents.

If the proposal is rejected, then the agents can negotiate again after j time periods (keeping the same v and $\mathcal{B}$ values), hence the agents' utilities are given by:

$$U_p(t, v, \mathcal{B}) = \delta^j U_p(t + j, v, \mathcal{B}) \tag{15}$$

In order to guarantee that the responding agent is indifferent between accepting and rejecting the proposal, we need to set $M(t+j, v, \mathcal{B})$ such that the responder's utility according to (13) and (14) is equal to (15). Equating (13) and (15) yields Equation 12 for the case where Agt_2 is the proposer. Similarly, Equating (14) and (15) yields Equation 12 for the case where Agt_1 is the proposer.

So far, we have proven that equations 12-14 capture the appropriate payment and the agents' utilities if the proposer requests the exploration of opportunity b_i. Now we present an inductive proof, showing that the opportunity b_i that will be included in the proposer's proposal is indeed the one that ought to be explored according to the optimal exploration strategy for the appropriate stand-alone problem $(v, \mathcal{B}, T - t + 1, j + k)$. When getting to the last negotiation iteration (i.e., at time t where $T-(2j+k)+1 < t \leq T-(j+k)+1$), the utilities $U_2(t+j+k, \max(v, y), \mathcal{B}-b_i)$ and $U_1(t + j + k, \max(v, y), \mathcal{B} - b_i)$ are zero (by definition, according to (13) and (14)). Similarly, when calculating the payment $M(t+j, v, \mathcal{B})$ according to (12), such that $T - (2j + k) + 1 < t \leq T - (j + k) + 1$ (i.e., in the call to $M(t + j, v, \mathcal{B})$ made by (13) and (14) when in the last negotiation iteration), the utilities $U_2(t + j, v, \mathcal{B})$ and $U_1(t + j, v, \mathcal{B})$ are zero. Therefore we obtain:

$$M(t+j,v,B) = \begin{cases} \delta^k c_i & \text{if proposer} = Agt_1 \\ \delta^k \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y)dy & \text{if proposer} = Agt_2 \end{cases} \quad (16)$$

$$U_p(t,v,\mathcal{B}) = \begin{cases} -\delta^j M(t+j,v,\mathcal{B}) + \delta^{j+k} \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{k+j}}) f_i(y)dy & \text{if p} = Agt_1 \\ \delta^j M(t+j,v,\mathcal{B}) - \delta^{j+k} c_i & \text{if p} = Agt_2 \end{cases} \quad (17)$$

Substituting (16) in (17) obtains:

$$U_p(t,v,\mathcal{B}) = \begin{cases} \delta^{j+k} \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y)dy - \delta^{j+k} c_i & \text{if p} = \text{proposer} \\ 0 & \text{if p} = \text{responder} \end{cases} \quad (18)$$

The term $\delta^{j+k} \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y)dy - \delta^{j+k} c_i$ in (18) is actually the expected benefit from the exploration of b_i, given that the best known value at time t is v. Therefore the proposer will prefer proposing the exploration of b_i associated with the best expected benefit, which is equivalent to the exploration rule according to the optimal solution to the stand-alone problem.

Now assume that the proposer at any negotiation step $t' > t$ offers the exploration of the next opportunity that should be explored according to the optimal stand-alone exploration strategy. We need to show that the proposer that needs to make a proposal at time t follows the same rule. We use $EV(t,v,\mathcal{B})$ to denote the expected net benefit from the exploration that is about to take place from time period t onwards when the best value found so far is v and the set of available opportunities is $\mathcal{B}$. We first show that $U_1(t,v,\mathcal{B}) + U_2(t,v,\mathcal{B}) = EV(t,v,\mathcal{B})$. The proof for this is once again inductive, showing that if $U_1(t,v,\mathcal{B}) + U_2(t,v,\mathcal{B}) = EV(t,v,\mathcal{B})$ for any $t' > t$, then so is the case for t. We have already shown that, when t is the time of the last negotiation step, the proposer's expected benefit is the expected benefit from the exploration carried out, and the expected benefit of the responder is zero. Therefore the sum of utilities is $EV(t,v,\mathcal{B})$. For any other t, we sum $U_1(t,v,\mathcal{B})$ and $U_2(t,v,\mathcal{B})$ according to Equations 13 and 14, obtaining:

$$\begin{aligned} U_1(t,v,\mathcal{B}) + U_2(t,v,\mathcal{B}) = \delta^{j+k} \int_{y=-\infty}^{\infty} & (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y)dy \\ + \delta^{j+k} \int_{y=-\infty}^{\infty} & \Big(U_1(t+j+k, \max(v,y), \mathcal{B} - b_i) \\ + U_2(t+j+k, & \max(v,y), \mathcal{B} - b_i) \Big) f_i(y)dy - \delta^{j+k} c_i \end{aligned} \quad (19)$$

Since, according to the induction hypothesis, $U_1(t+j+k, \max(v,y), \mathcal{B} - b_i) + U_2(t+j+k, \max(v,y), \mathcal{B} - b_i) = EV(t+j+k, \max(v,y), \mathcal{B} - b_i)$, and $\delta^{j+k} \int_{y=-\infty}^{\infty} (\max(v,y) - \frac{v}{\delta^{j+k}}) f_i(y)dy - \delta^{j+k} c_i$ is the immediate benefit from the current exploration, the sum $U_1(t,v,\mathcal{B}) + U_2(t,v,\mathcal{B})$ as calculated in (19) is in fact $EV(t,v,\mathcal{B})$.

Having established the equality $U_1(t,v,\mathcal{B}) + U_2(t,v,\mathcal{B}) = EV(t,v,\mathcal{B})$, and given that the responder's utility $U_{\bar{p}}(t,v,\mathcal{B})$ that the proposer needs to guarantee is

$\delta^j U_{\bar{p}}(t+j, v, \mathcal{B})$ (according to (15)), the proposer's expected utility is $U_p(t, v, \mathcal{B}) = EV(t, v, \mathcal{B}) - \delta^j U_{\bar{p}}(t+j, v, \mathcal{B})$. Since the value of $U_{\bar{p}}(t+j, v, \mathcal{B})$ does not depend on b_i (since it is the case when an offer is rejected and the values v and $\mathcal{B}$ do not change), the choice of b_i affects only the value of $EV(t, v, \mathcal{B})$. Therefore the proposer will choose the opportunity b_i that maximizes $EV(t, v, \mathcal{B})$. Since according to the inductive assumption future explorations will follow the optimal exploration strategy, the opportunity b_i that maximizes $EV(t, v, \mathcal{B})$ is necessarily the one that maximizes the stand-alone problem $(V, \mathcal{B}, T-t+1, j+k)$. Therefore the proposer will always follow the optimal exploration strategy as stated in the theorem.

Finally, we note that since $U_1(t, v, \mathcal{B}) + U_2(t, v, \mathcal{B}) = EV(t, v, \mathcal{B})$, the utility of the responder if it will choose to reject the proposal is at most $EV(t, v, \mathcal{B}) > 0$ (as when rejecting, less periods remain for exploration and also benefits are further discounted). Therefore, as long as $EV(t, v, \mathcal{B}) > 0$, the proposer will have an incentive to resume the negotiation. The case where $EV(t, v, \mathcal{B}) < 0$ precludes a positive utility for the proposer and thus it will opt out. Similarly, the optimal choice for the corresponding stand-alone problem in this case is to terminate the exploration.

The arguments for the uniqueness of the equilibrium are the same as in the proof of Theorem 1. ■

4.4 Comparing the Non-Interleaved and Interleaved Protocols

While the same set of opportunities is available to and can be potentially explored by the agents in both of the interleaved and non-interleaved protocols, and the deadline is the same, the resulting exploration in each case is different. This is evident from Theorems 1 and 2, as the exploration takes place according to the optimal solution to two different stand-alone underlying problems. In fact, despite the similar deadline, when using the interleaved protocol the agents are more constrained by the number of opportunities they can potentially explore, as they have less "effective" exploration time (compared to the non-interleaved case). Another difference between the protocols is in the ability to share the risk between the agents. For example in the interleaved protocol, Agt_1 will offer Agt_2 a payment before each exploration is executed, and this payment always exceeds the cost of the exploration. Hence, Agt_2 can never walk away with negative utility. In the non-interleaved model, insuring that Agt_2 always walks away with a non negative utility may require signing a complex contract.

From the computational aspect, the interleaved negotiation protocol is more complex than the non-interleaved negotiation protocol. In the non-interleaved negotiation the computational complexity of calculating the solution is linear in the number of negotiation steps, since the negotiation terminates once an offer is received. Computing the offers of the interleaved protocol using the naive backward induction approach requires an exponential number of steps. The reason for this is that regardless of whether an offer is accepted or not, the negotiation may continue. The additional negotiations from that point on, influence the offer made by the proposer (as it will want to offer a side-payment with which the responder is indifferent between accepting and rejecting) hence it is solved by the recursion prior to analyzing the current step, for all

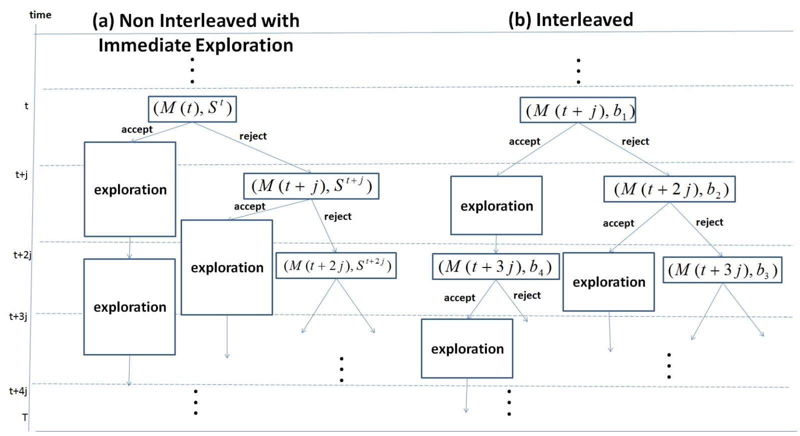


Fig. 5: Equivalence between the interleaved and non-interleaved variants for the case of $j_{accept} = 0$ and $k = j_{reject} > 0$.

possible future outcomes of the exploration. In order to deal with the computational complexity we introduce Theorem 3, which avoids the naive induction.

Theorem 3 *The agents' expected utilities, $U_1(t, v, \mathcal{B})$ and $U_2(t, v, \mathcal{B})$, resulting from negotiating according to the interleaved negotiation protocol given a setting $(v, \delta, \mathcal{B}, T-t+1, j, k)$ are equal to $U_1(t)$ and $U_2(t)$ resulting from a non-interleaved negotiation with immediate exploration setting $(v, \delta, \mathcal{B}, T-t+1, j_{accept} = 0, j_{reject} = j, j+k)$ (see Section 4.2).*

Proof In both negotiation protocols the resulting exploration follows the optimal exploration strategy for the stand-alone problem $(v, \delta, \mathcal{B}, T - t + 1, j + k)$, according to Theorem 1 and Theorem 2. Therefore, the expected sum $U_1(t, v, \mathcal{B}) + U_2(t, v, \mathcal{B})$ in both cases equals $EV(t, v, \mathcal{B})$. This enables an inductive proof for the theorem. In the final negotiation iteration $T - (2j + k) + 1 < t \leq T - (j + k) + 1$ (i.e., when no further exploration can be negotiated and executed), the expected utility of the responder is zero in both protocols, and the expected utility of the proposer in both cases is the expected benefit of the corresponding stand-alone problem. Now assume that the agents' expected utilities when reaching any state $(t' > t, v' \geq v, \mathcal{B}' \subseteq \mathcal{B})$ are identical in both negotiation protocols. We need to prove that the agents' expected utilities when in state $(t, v, \mathcal{B})$ are also identical. The expected utility of the proposer in the interleaved protocol is $U_p(t, v, \mathcal{B}) = EV(t, v, \mathcal{B}) - \delta^j U_{\bar{p}}(t+j, v, \mathcal{B})$ and in the non-interleaved protocol with immediate exploration $U_p(t) = EV(t) - \delta^j U_{\bar{p}}(t + j)$ where $v_0 = v$ and $B = \mathcal{B}$. For the responder, the expected utility is $\delta^j U_{\bar{p}}(t + j, v, \mathcal{B})$ and $\delta^j U_{\bar{p}}(t + j)$ for the interleaved and non-interleaved protocols respectively. Since the values of $EV(t, v, \mathcal{B}) = EV(t)$ and $U_{\bar{p}}(t + j, v, \mathcal{B}) = U_{\bar{p}}(t + j)$ are identical in both protocols (the first due to following the same exploration strategy as proven above and the second due to the induction assumption), so is the utility in both protocols, i.e., $U_p(t, v, \mathcal{B}) = U_p(t)$. The reason for using the non-interleaved negotiation with immediate exploration rather than the regular non-interleaved protocol is that in the first, the resulting exploration follows the optimal exploration strategy for the stand-alone problem $(v, \delta, \mathcal{B}, T-t+1, j+k)$, whereas in the second it is the optimal exploration strategy for the stand-alone problem $(v, \delta, \mathcal{B}, T-t-j+1, j+k)$. ■

The equivalence is demonstrated in Figure 5 for the case where $j_{reject} = k > 0$. In the non-interleaved protocol with immediate exploration, the proposing agent offers, at any time t, an exploration strategy for a payment $M(t)$ and in the interleaved protocol the proposing agent offers, at any time t, the exploration of an opportunity b_i for a payment $M(t+j)$. In the non-interleaved protocol with immediate exploration, if an offer is accepted then the exploration is carried out in the same time period, however if the offer is rejected the negotiation resumes in the next period. When the responding agent always accepts the offer, the number of explorations in both protocols is identical (and equals $\lfloor \frac{T}{j+k} \rfloor$). From the figure we can see that in both protocols the negotiation flow when proposals are accepted is the same and is fully aligned, making the agents' utilities identical.

Using Theorem 3 the agents can replace the interleaved negotiation with a simpler, non-interleaved one that yields the same outcomes for both agents. Computing the outcomes using the non-interleaved variant requires computing the optimal exploration sequence for $\lfloor \frac{T}{j+k} \rfloor$ stand-alone exploration problems of the form $(t, v_0, \mathcal{B})$. Therefore, instead of computing the outcomes using the interleaved protocol with exponential complexity, we can use the non-interleaved protocol and decrease the complexity to linear in the negotiation horizon T.

In many settings, as discussed in Subsection 4.1, the optimal solution for the exploration problem is immediate and thus the complexity of solving the negotiation problem, when using the non-interleaved variant, becomes linear in the number of allowed negotiation steps. Furthermore, Theorem 3 can also be used as an efficient means for computing the payments that need to be made by Agt_1 even if the interleaved protocol is preferred for the reasons discussed in the introduction: given state $(t, v, \mathcal{B})$, the proposer will offer to explore the next opportunity according to the optimal exploration strategy for the corresponding stand-alone problem. The payment $M(t+j, v, \mathcal{B})$ will be determined as the difference: $\delta^j U_{\bar{p}}(t+j) - \int_{x=v}^{\infty} U_{\bar{p}}(t+j+k) f_i(x)dx$, where $U_{\bar{p}}(t+j)$ and $U_{\bar{p}}(t+j+k)$ are calculated as the solution to the two appropriate instances of the non-interleaved negotiation problem with $(v, \mathcal{B})$ and $(\max(x, v), \mathcal{B} - b_i)$ respectively.

Before continuing to the next section, we map some cases where the interleaved protocol dominates its corresponding non-interleaved variant and vice-versa. As part of this analysis we refer both to individual utilities and social welfare, where the latter in our case is considered to be $U_1 + U_2$. For exposition purposes we assume in all the examples bellow, WLOG, that Agt_1 is the first proposer in the negotiation.

Proposition 2 *The non-interleaved protocol will always produce a greater or equal social welfare, in comparison to the equivalent interleaved protocol.*

Proof The proof is straightforward. We have already established the fact that in both protocols the exploration will take place according to the optimal strategy for the corresponding stand-alone problem (for the non-interleaved protocol see Theorem 1 and for the interleaved negotiation protocol see Theorem 2). The resulting stand-alone problems in both protocols consider the same set of opportunities, however while the number of allowed explorations potentially enabled in the non-interleaved protocol is $\lfloor \frac{T-j}{k} \rfloor$, in the interleaved protocol it is $\lfloor \frac{T}{j+k} \rfloor$. Therefore, the value of $EV(t)$ for the non-interleaved protocol is necessarily greater than the value of $EV(t, v, B)$ for the interleaved one, when using the optimal strategy in both cases. Since we have already established that the social welfare (i.e, the utilities sum) is captured by the expected net-benefit of the exploration that actually takes place, the social welfare in the case of the non-interleaved protocol is necessarily greater than in the case of the interleaved one. ■

An increase in the social welfare, though, does not guarantee that both agents will increase their utility individually. Still, from Proposition 2 we conclude that a setting in which both agents prefer to negotiate according to the interleaved negotiation protocol cannot exist. This is simply because if both benefit, then the total should be greater when using the interleaved protocol with that setting, contradicting Proposition 2. Interestingly, settings in which both agents will prefer the non-interleaved protocol over the interleaved protocol do exist, as proven in the following proposition.

Proposition 3 *There are settings for which both agents can increase their individual utility by preferring to negotiate according to the non-interleaved protocol, rather than according to the interleaved protocol.*

Proof We demonstrate that both agents may prefer the non-interleaved protocol using a simplified example. Consider a case where there are $n \geq 22$ homogeneous opportunities such that each opportunity is associated with uniform distribution between 0 and 1. There is no discounting ($\delta = 1$) and the exploration cost is $c = 0$. The negotiation horizon is $T = 32$, each negotiation iteration takes $j = 10$ periods and each exploration iteration takes $k = 1$ periods. We note that for the uniform distribution function the expected maximum of a k-sized sample is given by $k/(k+1)$ [79]. Furthermore, since $c = 0$ and $\delta = 1$, the optimal strategy for the stand-alone problem is always to explore as many opportunities as possible, subject to the negotiation horizon constraint.

We begin with the interleaved protocol. In this case we can solve using Theorem 3, i.e., using the non-interleaved protocol with immediate exploration (where exploration takes $k' = 11$ periods, $j_{reject} = 10$ and $j_{accept} = 0$). When reaching the last negotiation step, i.e., at time $t = 21$, Agt_1 is the proposer and it offers Agt_2 nothing in exchange for one exploration, hence $U_1(21) = \frac{1}{2}$ and $U_2(21) = 0$. Consequently, at time $t = 11$ Agt_2 will offer the exploration of two opportunities in exchange for a payment of $M = \frac{2}{3} - \frac{1}{2}$, in order to guarantee that Agt_1 obtains a utility equal to $U_1(21) = \frac{1}{2}$ (as $\delta = 1$). Hence $U_1(11) = \frac{1}{2}$ and $U_2(11) = \frac{2}{3} - \frac{1}{2}$. Finally, at time $t = 1$ there are only two explorations that can be executed if Agt_1's proposal is accepted, hence the utilities are similar to those at time $t = 11$, i.e., $U_1(1) = \frac{1}{2}$ $U_2(1) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$ (again, since $\delta = 1$).

Next, we move to the non-interleaved case. If the non-interleaved protocol is used, then when the last negotiation step is reached, i.e., at time $t = 21$, Agt_1 is the proposer, offering Agt_2 nothing in exchange for exploring two opportunities, hence $U_1(21) = \frac{2}{3}$ and $U_2(21) = 0$. Consequently, at time $t = 11$, Agt_2 will offer the exploration of 12 opportunities in exchange for a payment $M = \frac{12}{13} - \frac{2}{3}$, in order to guarantee that Agt_1 will obtain a utility of $U_1(21) = \frac{2}{3}$. Hence $U_1(11) = \frac{2}{3}$ and $U_2(21) = \frac{12}{13} - \frac{2}{3}$. Finally at time $t = 1$ Agt_1 will offer Agt_2 the exploration of 22 opportunities in exchange for $M = \frac{22}{23} - \frac{12}{13} + \frac{2}{3}$. The agents' utilities are thus: $U_1(1) = \frac{22}{23} - \frac{12}{13} + \frac{2}{3} = 0.7$ and $U_2(1) = \frac{12}{13} - \frac{2}{3} = 0.256$, i.e., both agents' expected utilities in the non-interleaved case are greater than their expected utilities when using the interleaved protocol (which are $U_1(1) = \frac{1}{2}$ $U_2(1) = \frac{1}{6}$ as calculated above). ■

Despite the superiority of the non-interleaved protocol over the interleaved one from the social-welfare point of view, it is possible that one of the two agents will prefer the latter protocol over the non-interleaved one. As an example of such a case, consider the same setting used for the proof of Proposition 3, except that $T = 31$ (instead of $T = 32$ in the original setting). Once again we use the non-interleaved protocol with immediate exploration for solving the interleaved protocol, using the same transformation as in the proof for Proposition 3. This time, however, at time $t = 11$ Agt_2 is the proposer and there is only one exploration that can be offered as at time $t = 21$ (because $T = 31$), therefore $U_1(11) = \frac{1}{2}$ and $U_2(11) = 0$. When $t = 1$, two opportunities can be explored given the new negotiation horizon constraint, and since Agt_1 is the proposer, and since the payment to Agt_2 needs to guarantee a zero utility, we obtain: $U_1(1) = \frac{2}{3}$ and $U_2(1) = 0$.

The computations for the non-interleaved protocol are also similar to those given in the proof for Proposition 3 except that there is one less opportunity that can be explored in the proposals made. Therefore: $U_1(21) = \frac{1}{2}$ and $U_2(21) = 0$. Consequently, at time $t = 11$ Agt_2 will offer the exploration of 11 opportunities in exchange for a payment $\frac{11}{12} - \frac{1}{2}$, and the corresponding utilities are: $U_1(11) = \frac{1}{2}$ and $U_2(1) = \frac{11}{12} - \frac{1}{2}$. Finally, at time $t = 1$ Agt_1 will offer Agt_2 the exploration of 21 opportunities in exchange for $\frac{21}{22} - \frac{11}{12} + \frac{1}{2}$. Agt_2's utility is thus $U_2(1) = \frac{11}{12} - \frac{1}{2} > 0$, i.e., Agt_2 would prefer to negotiate according to the interleaved protocol.

Interestingly, when using the interleaved protocol in the last example, Agt_1 received all the benefit of the process while the expected benefit of Agt_2 was zero. Theorem 4 defines a class of settings, for which the solution will necessarily be of that form.

Theorem 4 *In the interleaved protocol when there is no discounting ($\delta = 1$), then the agent to propose last, Agt_p, will gain the entire expected benefit from the exploration and the other agent will gain zero, if: (a) for every time period $i \cdot (j + k) + 1$ ($i \geq 0$), where Agt_p is the proposer, the following holds: $\lfloor \frac{T-i(j+k)-1-j-2wj}{j+k} \rfloor = \lfloor \frac{T-i(j+k)-1-2j-2wj}{j+k} \rfloor$ for any w such that $T - i(j + k) - 1 - 2j - 2wj > 0$; and (b) for every time period $i \cdot (j + k) + 1$ ($i \geq 0$), where the other agent is the proposer, the following holds: $\lfloor \frac{T-i(j+k)-1-2wj}{j+k} \rfloor = \lfloor \frac{T-i(j+k)-1-j-2wj}{j+k} \rfloor$ for any w such that $T - i(j + k) - 1 - j - 2wj > 0$.*

Proof Assume that the agents are in period $i \cdot (j + k) + 1$ and it is Agt_p's turn to make a proposal. We prove that if the condition specified in (a) holds, then the other agent's expected utility, $U_{\bar{p}}(i \cdot (j + k) + 1)$, is necessarily zero. The proof is inductive and shows that under the above conditions, if for some $w > 0$ Agt_p is the proposer at time $i \cdot (j+k)+2wj+1$ and $U_{\bar{p}}(i \cdot (j+k)+2wj+1) = 0$, then necessarily $U_{\bar{p}}(i \cdot (j+k)+2j+2wj+1) = 0$. For the highest w value for which $i \cdot (j+k)-k-j+2wj+1 > 0$, denoted w', we know from $\lfloor \frac{T-i(j+k)-1-2wj}{j+k} \rfloor = \lfloor \frac{T-i(j+k)-1-j-2wj}{j+k} \rfloor$ that Agt_p is necessarily the last that will have a chance to make a proposal that will result in some exploration (as the condition $\lfloor \frac{T-i(j+k)-1-2wj}{j+k} \rfloor = \lfloor \frac{T-i(j+k)-1-j-2wj}{j+k} \rfloor$ in this case suggests that if $Agt_{\bar{p}}$ can still make a proposal that will enable exploration then so can Agt_p). Therefore, $U_{\bar{p}}(i \cdot (j+k) - k - j + 2w'j + 1) = 0$. Assume that for some $w > 0$, Agt_p is the proposer at time $i \cdot (j+k) + 2wj + 1$ and $U_{\bar{p}}(i \cdot (j+k) + 2wj + 1) = 0$. We now show that $U_{\bar{p}}(i \cdot (j+k) + 2wj - 2j + 1) = 0$. This is straightforward, because regardless of the payment suggested by $Agt_{\bar{p}}$ at time $i \cdot (j+k)+2wj-j+1$, Agt_p can reject the proposal and offer $Agt_{\bar{p}}$ its utility if further rejecting, i.e., $U_{\bar{p}}(i \cdot (j+k) + 2wj + 1) = 0$, and gain the entire expected benefit. The number of opportunities that can be explored at times $i \cdot (j+k) + 2wj - j + 1$ and $i \cdot (j+k) - 2wj - 2j$ is equal according to the condition given in (a) as no loss is incurred due to the rejection (as $\delta = 1$). Therefore, at time $i \cdot (j+k) + 1$ (which is equivalent to $w = 0$), $U_{\bar{p}}(i \cdot (j+k) + 1) = 0$.

Now assume that the agents are in period $i \cdot (j+k) + 1$ and it is $Agt_{\bar{p}}$'s turn to make a proposal. The same logic used for proving (a) can be applied to show that if Agt_p rejects the offer then, under the conditions given in (b) the expected utility of $Agt_{\bar{p}}$ is zero and the same amount of opportunities can still be explored (and no loss is incurred due to discounting).

Having established the above, the proof is completed since proposals will be made only at times $i \cdot (j+k)$ (according to Theorem 2, as an offer is made according to the optimal solution to the stand-alone problem, always accepted, and an exploration takes place after each proposal), thus $U_{\bar{p}}(i \cdot (j+k) + 1) = 0$ for any $i > 0$. ■

The above theorem can best be understood with the specific example where $j = k$ (which complies with the theorem's condition). Here, the last agent to propose ends up with the entire revenue when $\delta = 1$, because only when it is that agent's turn to make a proposal the number of opportunities that can potentially be explored changes. Therefore, the other agent cannot benefit from rejecting the proposal since, once it becomes the proposer, the first agent can reject it during his turn and become the proposer once again without sacrificing the overall benefit that can be achieved by the exploration process.

Finally, we show that for a specific interleaved protocol variant the results obtained are exactly the same as with its non-interleaved equivalent (i.e., without changing the problem setting as in Theorem 3). This case is when, upon acceptance of a proposal in both the interleaved and non-interleaved protocols, the exploration process starts immediately (i.e., a non-interleaved and interleaved protocols with immediate exploration). Similar to its non-interleaved variant, the interleaved protocol with immediate exploration uses $j_{accept} = 0$ and $j_{reject} \geq 0$, and its applicability derives from the same justifications given above for this variant (see Section 4.2). Figure 6 describes both protocols when used in their immediate exploration form for the case where $j = k$. The figure illustrates that whenever an offer is accepted the exploration executes in the same time period in both protocols.

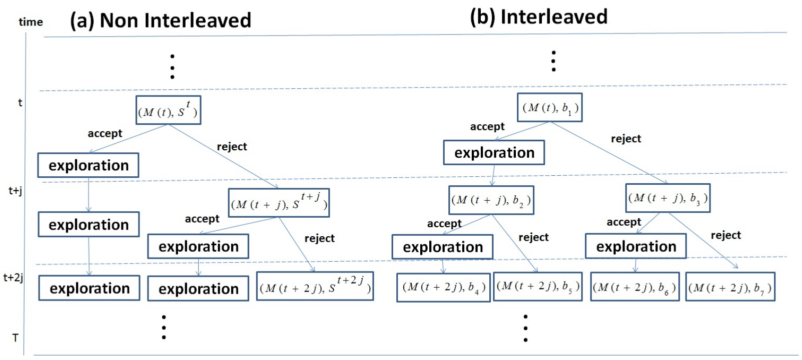


Fig. 6: Interleaved and non-interleaved protocols with immediate exploration.

Proposition 4 *The interleaved and non-interleaved protocol variants with immediate exploration always result in the same solution (exact same utilities to the two players).*

Proof Following the same steps made in the proof of Theorem 2 (with the only change being that the acceptance does not require waiting j time periods), we obtain that in the interleaved negotiation protocol with immediate exploration the agent will propose according to the solution to the stand-alone problem $(v, \delta, \mathcal{B}, T - t + 1, k)$. Given the latter result, we can use a proof identical to the one given for Theorem 3 (with the only change being considering the stand-alone problem $(v, \delta, \mathcal{B}, T - t + 1, k)$ instead of $(v, \delta, \mathcal{B}, T - t + 1, j + k)$) in order to show that the agents' utilities using both protocols are identical. ■

One important implication of Proposition 4 is that both agents will be indifferent to the protocol to be used whenever the exploration is immediate, i.e., when acceptance does not induce further delays. This means that the agents can be interchanged between the two protocols, based on external preferences. For example, if Agt_1 wants to guarantee that Agt_2 actually follows the optimal exploration strategy, then it can request to use the interleaved protocol, without inflicting any change to the utility of the two. Similarly, in cases where the negotiation itself results with some additional overhead, the agents can switch to the non-interleaved protocol, reducing the necessary overhead while guaranteeing the same utilities as in the more complex protocol.

Finally, we note that none of the above proofs relied on the equivalence between the discounting of payments and the values associated with the different opportunities. Therefore a model in which payments and opportunity values are discounted differently, or a model where each agent is assigned a different discounting factor are also supported, by multiplying by the appropriate discounting factor whenever applicable.

5 Controlling the Negotiation Protocol

In this section we illustrate the differences in the resulting individual utilities and in social welfare when using the interleaved and non-interleaved protocols. The two protocols are also compared with the legacy negotiation protocol in which the divided surplus is fixed throughout the negotiation (i.e., when $k = 0$, hence in each step all available opportunities can potentially be explored) as discussed in Section 4.2 (and also see [68]). The numerical illustrations given in this section also enable a thorough discussion of the choice of controlling the negotiation setting parameters for improving one's individual expected utility. In particular we focus on the choice of the negotiation horizon and the intensity of the exploration. Understanding the influence of allowing any of the agents control over these negotiation parameters is important, as in many scenarios the agents are not equal in their power over the negotiation.

Since results of this section are mostly of an existential nature, the illustrations use, in large, a synthetic, simplistic environment where opportunities are homogeneous. Therefore, unless stated otherwise, the opportunities share the same exploration cost and distribution of values (uniform between 0 to 1 in this case), and $j = k = 1$. Similarly, for the legacy negotiation protocol we assume $j = 1$. We also assume in the illustrations that Agt_1 is the first to propose (though the same phenomena occur if Agt_2 proposes first).

5.1 Affecting Parameters

The parameters that characterize our archetypal model are the negotiation horizon, the discounting factor and the set of exploration opportunities available (characterized by their distribution of values and exploration cost). For the legacy negotiation protocol in which the divided surplus is fixed throughout the negotiation, the influence of these parameters (individually, given that the other two are fixed) over individual utilities is given as follows:

- Negotiation horizon - the expected-utility-maximizing negotiation horizon for the agent that is set to be the first to issue a proposal is $T = 1$. For the other agent it is $T = 2$. This result is straightforward from Equation 9.[13]
- Discounting of gains and costs - the expected-utility-maximizing discounting factor for the agent that is set to be the last to issue a proposal is $\delta = 1$. This derives directly from Equation 9 as the term by which the exploration's expected benefit is multiplied becomes equal to 1 whenever the agent is the last proposer (i.e., whenever the negotiation horizon is even if the player is the first proposer and odd otherwise). Therefore, by using $\delta = 1$ the agent that proposes last guarantees obtaining the entire expected benefit from the exploration.[14] If the agent is not

[13] One specific case worth mentioning within this context is when $\delta = 1$. Here, the agent that is set to be the first to issue a proposal will prefer any odd negotiation horizon and the other agent will prefer any even horizon.

[14] Notice that while that EV is fixed for the legacy negotiation protocol, even if the exploration that takes place was affected by δ this would not change the result, because the exploration's expected benefit is also maximized for $\delta = 1$.

the last one to issue a proposal, then the expected-utility-maximizing discounting factor is some $0 < \delta^* < 1$ that maximizes the agent's expected utility according to the fixed surplus variant of Equation 9. For example, if $T = 5$ and the agent is not the last to issue a proposal, then $\delta^* = 0.73$ is the discounting factor for which $\Sigma_{i=2}^{T-t+1}(-1)^i\delta^i EV$ is maximized.

– Opportunities - in general, the more favorable the stand-alone exploration problem is, the greater the individual expected utility of both agents, regardless of the values the other problem parameters obtain. This is because the agents divide a fixed sum EV among themselves, and the portion that each of them receives (according to Equation 9) does not depend on EV. For example, if the exploration cost of any of the opportunities decreases, then both agents' individual utilities will increase.

In the following paragraphs we demonstrate that the above does not generally hold in our interleaved and non-interleaved protocols. We begin with the exploration cost as a representative of a more or less favorable exploration stand-alone problem. Consider a case where there are two opportunities b_1 and b_2, where b_1 is associated with a value 100 with probability 0.5 and zero otherwise, and b_2 is associated with a value 50 with probability 0.5 and 70 otherwise. The exploration costs are $c_1 = c_2 = 2$ and the other setting parameters are $k = j = 1$ and $\delta = 1$. The optimal exploration strategy according to Section 4.1 when both opportunities can be explored is to explore opportunity b_1 first, and if the value found is zero to continue with the exploration of b_2 (and terminate exploration otherwise). The expected benefit of this exploration process is $-2 + 100 * 0.5 + 0.5 * (0.5 * 70 + 0.5 * 50 - 2) = 77$. When only one opportunity can be explored, the optimal exploration strategy is to explore b_2, yielding 58. Consequently, when negotiating according to the non-interleaved protocol with $T = 3$, the expected utilities of the different agents are: $U_1 = 77 - 58 = 19$ and $U_2 = 58$. If the cost of exploring b_2 now becomes $c_2 = 0$, then the expected benefit from the stand-alone exploration becomes 78 when the exploration of two opportunities is enabled and 60 when only one opportunity can be explored. Therefore, $U_1 = 78 - 60 = 18$ and $U_2 = 60$. The example for the interleaved protocol is similar to the one illustrated above, however uses $T = 6$, $j = 2$ and $k = 1$. In this case Agt_2 gains the utility of exploring b_2 and Agt_1 gains the utility from adding the exploration of b_1 to the process. Therefore, once again, when $c_1 = c_2 = 2$ we obtain $U_1 = 19$ and $U_2 = 58$ and when $c_2 = 0$ we obtain $U_1 = 18$ and $U_2 = 60$. Therefore, in contrast to the general rule that applies to the change in the exploration cost for the legacy protocol, in our case the improvement in the underlying stand-alone exploration problem may result in an increase in one agent's expected utility and a decrease in another's.

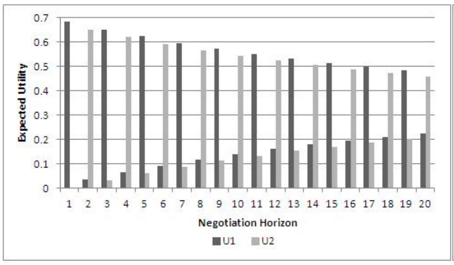
As for the discounting factor, the property according to which by using $\delta = 1$ the agent that proposes last guarantees obtaining the entire expected benefit from the exploration, can be proven to hold only for a specific case, as shown in Theorem 4. For other cases, the expected benefit is divided between the agents with some proportion, even when $\delta = 1$. For example, consider a setting with two opportunities to explore, b_1 and b_2, both associated with a value 1 with probability 0.5 and zero otherwise, and the other setting parameters are: $j = 1$, $k = 2$, $T = 6$ and $c = 0$.

In this example Agt_2, when using the interleaved protocol, is the last to propose if following a sequence of rejections from the start. When $\delta = 1$ we obtain $U_1 = \frac{1}{4}$ and $U_2 = \frac{1}{2}$, thus Agt_2 does not obtain the entire expected benefit. Furthermore, using a similar example we can demonstrate that even when the non-interleaved protocol is used, the agent that is the last to propose if following a sequence of rejections from the start does not take over the entire benefit when $\delta = 1$. The example uses $T = 3$, $j = k = 1$, two opportunities associated with value 1 with probability 0.5 and 0 otherwise and $c = 0$. In this case we obtain once again $U_1 = \frac{1}{4}$ and $U_2 = \frac{1}{2}$.

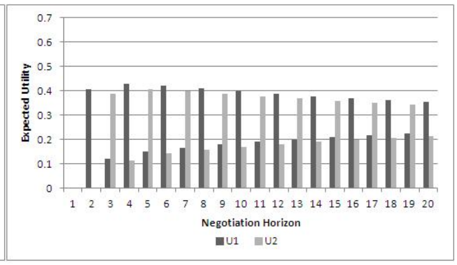
5.2 The Choice of Controlling the Negotiation Horizon

Among the three parameters specified above, the negotiation horizon is of the greatest importance, since unlike the other two it is in most cases the only one that the agents can actually influence. Taking the recruiting through a headhunter problem as an example, the discounting of gains and costs is usually exogenous as it derives from global policies and market liquidity. Similarly, the opportunities are given and their values cannot be influenced by the agents in any way, and also the exploration costs (e.g., the cost of a single interview) are fixed or depend on the interviewer's salary. The negotiation horizon, on the other hand, is of a different nature in the sense that is commonly controlled by one or both of the agents. For example, the headhunter doing the exploration may choose to shorten the negotiation horizon by committing to another recruiting project that is supposed to begin in a few weeks, thus limiting the amount of time allowed in the negotiation and exploration for the current project. Similarly, the company benefiting from the interviews may initially allocate a shorter time for the recruiting part, justifying it by project schedule constraints, once again limiting the amount of time allowed for the negotiation and exploration.

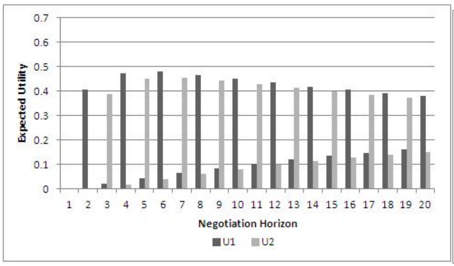
We demonstrate the effect of the negotiation horizon T over the expected utility of the two agents in the non-interleaved and interleaved protocols as well as in the corresponding legacy negotiation protocol that was defined above. The agents' utilities as a function of the negotiation horizon (the horizontal axis) in the different protocols are depicted in Figures 7a-7c for the case where the exploration cost is $c = 0.05$ and the discounting factor is $\delta = 0.95$. The number of opportunities in this example is 20, therefore the number of opportunities that can potentially be explored in the legacy protocol is 20. In the non-interleaved and interleaved protocols the number of opportunities that can potentially be explored depends on T: in the non-interleaved protocol $T-1$ opportunities can potentially be explored (i.e., at least one period is used for negotiating) and in the interleaved protocol $\lfloor \frac{T}{2} \rfloor$ opportunities can potentially be explored. From Figure 7 we observe that, given the option to set the negotiation horizon, in the non-interleaved protocol Agt_1 would have chosen the negotiation horizon to be $T = 4$ whereas Agt_2 would have chosen $T = 5$, while in the interleaved protocol Agt_1 would have chosen $T = 6$ and Agt_2 would have chosen $T = 7$. This contrasts the results known for the legacy negotiation protocol where the agents negotiate over the division of the expected benefit from an optimal exploration process [68]. As specified at the beginning of the section, the agent that proposes first will prefer $T = 1$ and the one that is second will prefer $T = 2$. Fur-



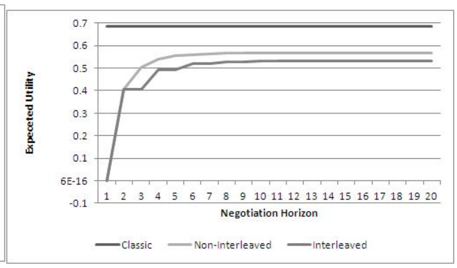
(a) Classic protocol



(b) Non-interleaved protocol



(c) Interleaved protocol



(d) Social welfare comparison

Fig. 7: Agents' utilities and joint utility (social welfare), when $c = 0.05$ and $\delta = 0.95$.

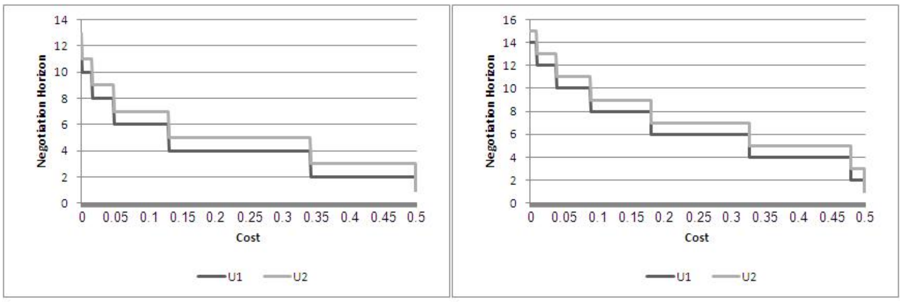
thermore, in constraining the number of periods to be odd, the agent to propose first prefers $T = 1$ but the agent to propose second prefers $T \to \infty$. When the number of periods is constrained to be even, the agent to propose second prefers $T = 2$ and the agent to propose first prefers $T \to \infty$. This is illustrated in Figure 7a where the number of available opportunities is set to 20.

The explanation of the difference between the interleaved and non-interleaved negotiation protocols and the legacy negotiation protocol is that in the first two there are factors that influence the agents' utilities that conflict with each other. On one hand, the increase in the negotiation horizon reflects an increase in the number of opportunities that can potentially be explored, and consequently an increase in the sum of the agents' expected utilities (social welfare). On the other hand, when the number of opportunities that can potentially be explored increases, the agents split the expected joint utility more equally (similar to the case of the legacy negotiation protocol [70]). In particular the advantage in being the last proposer, in terms of ending up with a higher expected utility, becomes less apparent as the number of available exploration periods increases. Nevertheless, the positive effect of the increase in the negotiation horizon on the sum of the agents' utilities is relevant only for the interleaved and non-interleaved protocols, whereas its effect over the division of the sum is relevant for all three protocols. Therefore, in the non-interleaved and interleaved protocols the agents need to find the balance between increasing the negotiation horizon and potentially enabling the exploration of more opportunities, and the effect it has on their individual utilities. Each agent will thus attempt to set the negotiation horizon such that its expected utility is maximized (e.g., Agt_1 prefers $T = 4$ in the non-interleaved protocol and $T = 6$ in the interleaved one).

Due to the importance of the negotiation horizon as discussed above, we propose additional numerical illustrations, demonstrating the effect of the two other parameters, δ and c, on the choice of the negotiation horizon. As discussed at the beginning of this section, the agents' preference of the negotiation horizon in the legacy negotiation protocol is not affected by these two other parameters. This, however, does not hold in the negotiation protocols analyzed in this paper, as demonstrated below. Figure 8 depicts the utility-maximizing negotiation horizon for each agent as a function of the exploration cost c for $\delta = 0.99$ in the non-interleaved and interleaved protocols (graphs 8a and 8b, respectively). The maximum negotiation horizon allowed is $T = 20$ thus the agents are limited to choosing their preferred negotiation horizon in the interval $1 - 20$. From the figure we observe that as the cost of exploration increases, the benefit in further exploration decreases and thus the agents prefer to use a shorter negotiation horizon. Figure 9 depicts a similar analysis, however this time, as a function of the discounting factor δ where $c = 0.05$ and the maximal allowed negotiation horizon of $T = 20$ in the non-interleaved and interleaved protocols (graphs 9a and 9b, respectively). Here, as the discounting factor increases, the overall sum divided between the two agents increases and the agents prefer a longer negotiation horizon.

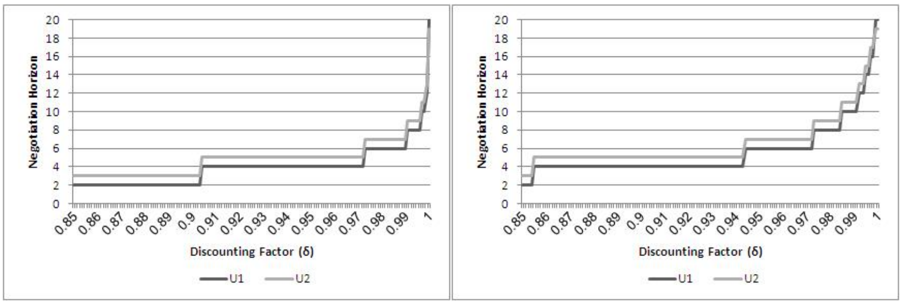
Another observation made based on Figures 8 and 9 is that both agents change their preferred negotiation horizon for the same cost (or discounting factor) values, and for any given cost (or discounting factor) the difference between their preferred negotiation horizon is j (in this example $j = 1$). The intuition for this is as follows: Agt_1 necessarily sets its expected-benefit-maximizing negotiation horizon such that it is the last proposer (as there is no benefit for Agt_1 from the additional exploration carried out when Agt_2 is the last proposer). By increasing the negotiation horizon that benefits Agt_1 th most by j, Agt_2 becomes the last proposer, while the number of explorations that can be carried out does not change. Therefore, Agt_2 prefers a negotiation horizon similar to Agt_1's plus j. It is notable that while this observation is true for the example given in Figures 8 and 9, it is not generally true, as by increasing the negotiation horizon by j the agent also loses due to the discounting of gains.

The choice of which protocol is more beneficial from each agent's point of view is setting-dependent. This is illustrated in Figure 10 which compares the agents' expected utilities as a function of the negotiation horizon (ranging between 1 and 20), for the case where $c = 0.05$ and $\delta = 0.95$, in the non-interleaved and interleaved protocols. Figure 10a depicts Agt_1's expected utility and Figure 10b depicts Agt_2's expected utility. When there is only one period (i.e., no further exploration can be executed), both utilities are zero and the agents are indifferent between the two negotiation protocols. When the negotiation horizon is greater than 1 we observe that the choice of the negotiation protocol is setting-dependent: when there is an even number of periods, Agt_1 prefers to negotiate according to the interleaved protocol, and when the number of time periods is odd Agt_1 prefers the non-interleaved protocol. Agt_2's preferences will be the opposite of Agt_1's, i.e., with an even number of periods it will prefer the non-interleaved protocol and with odd numbers it will prefer the interleaved one. While the preference of the non-interleaved method is intuitive (as the overall so-



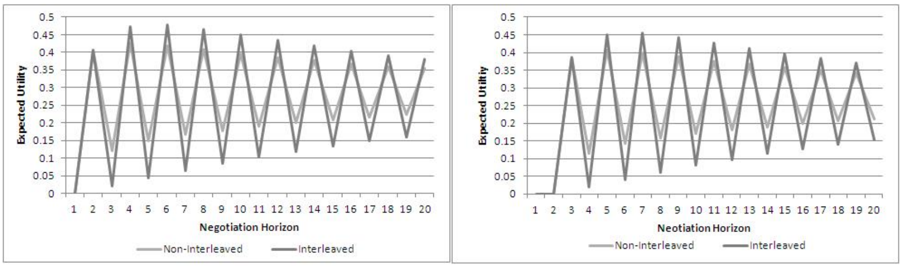
(a) Non-interleaved protocol (b) Interleaved protocol

Fig. 8: The effect of the exploration cost over the preferred negotiation horizon in the non-interleaved and interleaved protocols, when $\delta = 0.99$.



(a) Non-interleaved protocol (b) Interleaved protocol

Fig. 9: The effect of the discounting factor over the preferred negotiation horizon in the non-interleaved and interleaved protocols, when $c = 0.05$.

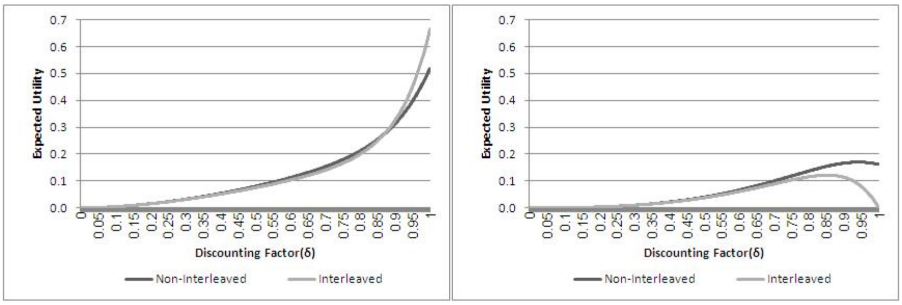


(a) Agt_1's expected utilities (b) Agt_2's expected utilities

Fig. 10: Agents' expected utilities for different negotiation horizons in the non-interleaved and interleaved protocols, when $c = 0.05$ and $\delta = 0.95$.

cial welfare is greater with that method according to Proposition 2), the preference of the interleaved protocol may seem counter-intuitive, though such a case was already demonstrated in Section 4.4. In our case, the preference of the negotiation protocol is explained mainly by the choice of which of the agents will be the last to propose in a sequence of rejections starting from any negotiation step. Similar to the logic of Theorem 4, for relatively high values of δ, the agent that proposes last will get hold of a substantial portion of the overall expected benefit (social welfare). Therefore, despite the potential increase in the social welfare due to the increase in the number of exploration periods allowed in the non-interleaved protocol, when weighing in the portion that the agent obtains out of this sum, the last-proposing agent finds it beneficial to prefer the interleaved protocol. Similarly, for the last-responding agent there are two factors which support its preference of the non-interleaved protocol in our case. First, the overall social welfare increases when the non-interleaved protocol is used. Second, the portion that the last-proposing agent obtains out of the overall sum is lower with the non-interleaved protocol, and consequently the portion of the last-responding agent is greater with that protocol. Therefore, the last-proposing agent necessarily gains more when using the non-interleaved protocol in our case.

The choice of which protocol to be used depends not only on the negotiation horizon but also on the other parameters. We illustrate this for the discounting factor parameter using Figure 11. The figure depicts the agents' expected utility as a function of the discounting factor, when $T = 10$ and $c = 0.05$. Figure 11a depicts Agt_1's expected utility and Figure 11b depicts Agt_2's expected utility. From the figure we observe that Agt_1 chooses the non-interleaved protocol for $\delta < 0.871$ and the interleaved protocol otherwise. Agt_2 in this example prefers the non-interleaved protocol for any value that δ obtains. The explanation is as follows: in the interleaved protocol the increase in δ results in an increase in the portion that the last-proposing agent (Agt_1 in our case, as $T = 10$) obtains out of the overall sum (based on the logic of Theorem 4, as discussed above). For low δ values, the portion obtained out of the total with the interleaved protocol is low, and since the overall social welfare with the interleaved protocol is necessarily lower, the non-interleaved protocol is preferred. As δ increases, the benefit from the increase in the portion obtained by the last-proposing agent outweighs the difference in the overall social welfare, hence once the interleaved protocol becomes the preferred protocol, it remains the preferred protocol also for greater δ values (hence only one intersection point between the two curves). As for Agt_2, the preference of the non-interleaved protocol is explained as follows: when δ is relatively low, the (negative) effect of the additional discounting of gains, due to the negotiation that takes place between explorations in the interleaved protocol, over the total divided utility is substantial. The divided total is thus substantially greater with the non-interleaved protocol. In this case the benefit from the increased overall divided sum when using the non-interleaved protocol outweighs any potential benefit the agent may have in terms of the relative portion it receives for low δ values (i.e., the opposite case for the logic of Theorem 4) out of the total with the interleaved protocol. As δ increases Agt_1's portion out of the total obtained with the interleaved protocol increases, as explained above. Therefore, since Agt_2's share in the interleaved protocol keeps decreasing and, on the other hand, the total that can be divided in the non-interleaved protocol increases, Agt_2 prefers the latter



(a) Agt_1's expected utilities (b) Agt_2's expected utilities

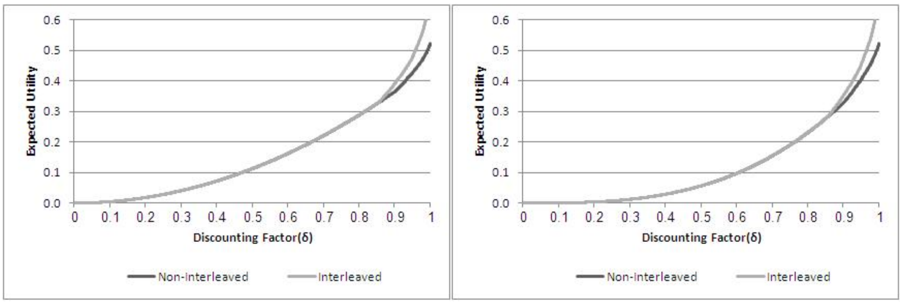
Fig. 11: The effect of the discounting factor over the agents' expected utilities (and consequently the social welfare) when $T = 10$ and $c = 0.05$.

protocol. The same (reversed) phenomena occurs when the negotiation horizon is odd (i.e., when Agt_2 becomes the last-proposing agent).

Interestingly, in the above example, if one of the agents can control the negotiation horizon it will always prefer the interleaved protocol. This is illustrated in Figure 12, which depicts the maximum value that each agent could have obtained if it had exclusive control over the negotiation horizon, under the $T \leq 20$ constraint, as a function of δ. As observed from the figure, the choice of the interleaved protocol weakly dominates the non-interleaved alternative, regardless of the agent's role. Furthermore, as the discounting factor increases, the interleaved protocol becomes even more favorable for the agent. This is because in the interleaved protocol the last proposer ends up with most of the revenue and the other agent gains only the discounted differences as described in Theorem 4. Therefore, the agent can choose the negotiation horizon such that it will be the last proposer. The fact that both curves in each graph are the same for small δ values is explained as follows: for low δ values, the preferred negotiation horizon is 2 and 3 for Agt_1 and Agt_2, respectively, for the same cost $c = 0.05$, hence the expected sum of utilities is the same (based on a single exploration). Since the last-proposing agent gains the entire sum in this case, it is indifferent between the two protocols.

Naturally, the agent's ability to control the negotiation horizon affects individual utilities and the social welfare. Therefore, exploring bounds for the effect of such choices is of high importance. We illustrate the extraction of such bounds for the interleaved protocol. Proposition 5 proves that the amount by which an agent can improve its utility if given the option to choose the negotiation horizon compared to its utility when the social-welfare-maximizing negotiation horizon is used, is unbounded. Let $SW(H)$ denote the social welfare when the negotiation horizon is H and let H_{max} be the social-welfare-maximizing negotiation horizon. Let $U_p(H)$ denote the expected utility of an agent if the negotiation horizon is H and let H_p be the agent's utility-maximizing negotiation horizon.

Proposition 5 *The ratio $U_2(H_2)/U_2(H_{max})$ is unbounded.*



(a) Agt_1's expected utilities (b) Agt_2's expected utilities

Fig. 12: The preference of a negotiation protocol when one of the agents receives exclusive control over the negotiation horizon (ranging from $1 - 20$), for different values of δ and $c = 0.05$.

Proof Consider a case where the maximum negotiation horizon is $T = 3$ and $j = k = 1$. There is one opportunity with a uniform distribution of values in the range $0 - 1$, the cost is $c = 0$ and the discounting factor is $\delta = 0.9$. In this case Agt_2 prefers $H_2 = 3$ in order to make the last proposal, in which case its expected utility is $U_2(H_2) = 0.9^3 * 0.5$ and Agt_1's utility is zero. The social-welfare-maximizing horizon is, however, $H_{max} = 2$, in which case Agt_1 gains the utility of the single exploration executed while $U_2(H_{max}) = 0$. Therefore, the ratio between $U_2(3)/U_2(2)$ is unbounded. ■

As far as the social welfare is concerned, the decrease in this value due to allowing one of the agents to choose the negotiation horizon in a way that maximizes its own utility (compared to the social-welfare-maximizing negotiation horizon) is bounded, as proven in Proposition 6.

Proposition 6 *The ratio $SW(H_{max})/SW(H_p)$ is bounded by $\delta^j \frac{1}{2} SW(H_{max})$.*

Proof Consider the negotiation setting $(v_0, \delta, B, T, k, j)$ and an agent Agt_p. If the agent's expected utility $U_p(H_{max})$ is greater than $\frac{1}{2}SW(H_{max})$, then the bound given in the proposition necessarily holds. Otherwise, if $U_p(H_{max}) < \frac{1}{2}SW(H_{max})$ then the agent can choose to use a negotiation horizon $H_{max} + j$ and offer a proposal that will necessarily be rejected at time $t = 1$ (if it is the first to propose in the negotiation) or reject the proposal made by the other agent at time $t = 1$ (if it is not the first to propose in the negotiation). Once the first proposal is rejected, the remaining negotiation with H_{max} periods will result in an overall sum $\delta^j SW(H_{max})$, and since the agents "switched places" compared to the original setting, Agt_p's expected utility is $\delta^j(SW(H_{max}) - U_p(H_{max}))$ and since $U_p(H_{max}) < \frac{1}{2}SW(H_{max})$, then $\delta^j(SW(H_{max}) - U_p(H_{max})) > \frac{1}{2}\delta^j SW(H_{max})$. Therefore, Agt_p will end up with at least $\frac{1}{2}\delta^j SW(H_{max})$ and consequently the social welfare will be greater than $\frac{1}{2}\delta^j SW(H_{max})$. ■

5.3 The Choice of Controlling the Exploration Intensity

The control of the exploration intensity is a prevalent theme in costly exploration literature [12,23,82]. Executing some of the explorations in parallel is beneficial whenever the agent is limited by an negotiation horizon. It enables the exploration of more opportunities within the time constraints. The downside of parallel exploration is that it is wasteful in the sense that exploration of some of the opportunities is not actually required (i.e., those explored in the exploration stage in which the exploration terminates).

In this section we demonstrate that the agents can exploit the intensity of the exploration in order to increase their expected utility. As with the case of controlling the negotiation horizon, we demonstrate the effect of the exploration intensity choice over the agents' individual utilities and the social welfare and prove some related bounds. We assume that the agent that controls the negotiation intensity sets the number of opportunities to explore in parallel q before the negotiation starts, and this rule remains the same throughout the negotiation. In this case the agents negotiate in each step over the payment that Agt_1 should pay Agt_2 for exploring q opportunities at a time (i.e., in parallel) with an exploration cost equal to the sum of the cost incurred by exploring the q opportunities. The proposing agent can either choose to offer a payment for exploring exactly q opportunities in parallel or choose to terminate the negotiation.

The analysis of the parallel negotiation protocol uses backward induction similar to that used with the non-interleaved and interleaved negotiation protocols [78]. The number of opportunities to explore in parallel becomes an additional parameter describing the underlying corresponding stand-alone exploration problem $(v_0, \delta, B, T, k, q)$ (see Section 4.1).The solution to the parallel stand-alone exploration problem is setting-dependent, similar to the case of pure sequential exploration, as discussed in Subsection 4.1. In a case where the deadline enables the agent to potentially explore all of the opportunities in $\mathcal{B}$ or when the opportunities are homogeneous, the optimal solution is based on reservation values and can be extracted with a polynomial complexity [54,53]. The appropriate modification of Equation 1, from which the corresponding reservation values can be extracted, is:

$$\delta^k c_i = \delta^k \int_{x=r_i}^{\infty} (x - r_i) q f_i(x) F_i(x)^{q-1} dx - (1 - \delta^k) r_i. \qquad (20)$$

where $F_i(x)$ is the commutative distribution function (c.f.d) of $f_i(x)$. In this case, once again, if one of the agents is given control over the number of opportunities that can be explored over each negotiation step, it will choose the number q that maximizes its own expected utility, which may result in a decrease in the expected joint utility.

We now explore the ratio between an agent's utility with and without such control, and the ratio between the social welfare with and without such control. This lets us bound the effect of the extra control for every set of opportunities.

Suppose that there are several identical opportunities, and that Agt_2 is capable of exploring exactly q opportunities in parallel each time (by paying the cost of all of them). Let $SW(q)$ denote the social welfare when q opportunities are explored at each step (until the agents decide to terminate the exploration). Let q_{max} be the number of opportunities which maximizes $SW(q)$. Meaning that q_{max} is the number of explorations used in parallel in the optimal solution to the stand-alone problem. Consider the setting in which Agt_1 is the first proposer that chooses a number q in the beginning of the negotiation (though the analysis is equally true when Agt_2 is the proposer), and then at each negotiation step each of the agents in its turn offers a price for exploring q opportunities at a time. Note that Agt_1 will choose q in order to maximize its utility from the entire procedure, taking into account both the negotiation and the exploration. Let $U_1(q)$ denote Agt_1's utility if q opportunities are explored at each step, and let q_1 be the value which maximizes Agt_1's utility.

We now show that Agt_1 can greatly improve its utility using its control over the negotiation.

Proposition 7 *The ratio $U_1(q_1)/U_1(q_{max})$ is unbounded.*

Proof Consider an interleaved negotiation protocol where the negotiation horizon is $T = 6$, $j = 2$ and $k = 1$. Assume that the set of opportunities is comprised of three identical opportunities that are associated with value 1 with probability 0.5 and 0 otherwise, the exploration cost is zero and there is no discounting ($\delta = 1$). In this case q can be either 1 or 3 (as these are the only dividors of the three opportunities). When $q = 3$, Agt_2 will gain the benefit from three explorations in the last negotiation step, i.e., $U_2(3) = \frac{7}{8}$. In the previous negotiation iteration Agt_1 will offer $M = \frac{7}{8}$ hence $U_1(3) = 0$. When $q = 1$, Agt_2's utility from a single exploration in the last negotiation step is $U_2(1) = \frac{1}{2}$. Therefore, the payment that Agt_1 pays in the previous step is $M = \frac{1}{2}$ and Agt_1's utility in the first step is $U_1(1) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$. Therefore, $q_{max} = 3$, $q_1 = 1$ and the ratio between $U_1(3)/U_1(1)$ is unbounded because Agt_1's utility increases from zero to $\frac{1}{4}$. ■

Next we show that Agt_1 can greatly decrease the social welfare while using its control over the negotiation.

Proposition 8 *The ratio $SW(q_{max})/SW(q_1)$ is unbounded.*

Proof Consider an interleaved negotiation protocol where the negotiation horizon is $T = 22$, negotiation steps take $j = 10$ and exploration steps take $k = 1$. Assume that the set of opportunities available is comprised of $\frac{10}{\epsilon}$ identical opportunities, for some small constant ϵ, such that each opportunity is associated with value 1 with probability $\frac{\epsilon}{10}$ and zero otherwise. Also, assume that there is no exploration cost ($c = 0$) and no discounting ($\delta = 1$). Assume that Agt_1 can control the number of opportunities that can be explored in parallel in each step of the negotiation, in a way that q obtains either the number of opportunities available or one only. In this case $q_{max} = \frac{10}{\epsilon}$ because as the number of opportunities increases the expected benefit increases as well (i.e., the greater the chance to get a value of 1 at the end). The social welfare in this case is $SW(\frac{10}{\epsilon}) = 1 - (1 - \frac{\epsilon}{10})^{\frac{10}{\epsilon}} = 1 - \frac{1}{e}$. However, when $q = 1$ only two opportunities will eventually be explored and the social welfare becomes $SW(1) = 1 - (1 - \frac{\epsilon}{10})^2 \leq \frac{2\epsilon}{10}$. The ratio $SW(\frac{10}{\epsilon})/SW(1)$ is unbounded, as we can continue to decrease ϵ indefinitely, and in order to complete the proof we need to show that Agt_1 prefers $q_1 = 1$. When $q = \frac{10}{\epsilon}$, in the last negotiation step Agt_2 will gain the marginal benefit from the exploration: $U_2(\frac{10}{\epsilon}) = 1 - (1 - \frac{\epsilon}{10})^{\frac{10}{\epsilon}} = 1 - \frac{1}{e}$. Therefore, in the first negotiation iteration Agt_1 will offer a payment, $M = 1 - \frac{1}{e}$, hence $U_1(\frac{10}{\epsilon}) = 0$. When $q_1 = 1$, in the last negotiation step Agt_2's utility is $U_2(1) = \frac{\epsilon}{10}$. Therefore, the payment that Agt_1 pays in the previous step is $M = \frac{\epsilon}{10}$ and Agt_1's utility in the first step is $U_1(1) = 1 - (1 - \frac{\epsilon}{10})^2 - \frac{\epsilon}{10} = \frac{\epsilon}{10} - (\frac{\epsilon}{10})^2 > 0$. Therefore, $U_1(1) > 0 = U_1(\frac{10}{\epsilon})$ and Agt_1 prefers $q_1 = 1$. ■

We now consider a setting with heterogeneous opportunities in which Agt_1 chooses exactly which opportunities will be negotiated over at each step. Agt_1 can still greatly improve its own utility in this settings, as when choosing how many opportunities will be negotiated in each step as shown in Proposition 7. Surprisingly, Proposition 8 does not apply in this setting and the decrease in the social welfare is bounded. Let $SW_{OPT}(B, \delta)$ denote the maximal social welfare of the exploration problem, when both agents cooperate during the negotiation. Let $SW_{1choose}(B, \delta)$ denote the social welfare when Agt_1 chooses which opportunities will be negotiated at each step in order to maximize its own utility. To generalize the result concerning the bound in this case, we divide the parameter to j_{accept} and j_{reject} as used in Section 4.2.

Proposition 9 *The welfare loss when one of the agents gets to decide the exploration of which opportunities will be negotiated at each step is bounded,* $SW_{1choose}(B, \delta) \geq SW_{OPT}(B, \delta^{1+\frac{j_{reject}}{j_{accept}+k}})$.

Proof Consider an interleaved negotiation setting $(v_0, \delta, B, T, k, j_{accept}, j_{reject})$ with an exploitation value v_0, opportunities $B = \{b_1, \ldots, b_n\}$, discounting factor δ, where accepting an offer takes j_{accept} periods, rejecting an offer takes j_{reject} and exploration step takes k periods. We show a strategy which guarantees Agt_1 a utility of $SW_{OPT}(B, \delta^{1+\frac{j_{reject}}{j_{accept}+k}})$. As the social welfare of the protocol is at least the utility of Agt_1, this proves the proposition. Let S be an optimal strategy for the original exploration problem if the discounting factor is $\delta^{1+\frac{j_{reject}}{j_{accept}+k}}$. When it is Agt_1's turn to make an offer, the negotiation will be over the set of q opportunities which S would utilize at that step, and when Agt_2 makes the offers, the negotiations will be over an empty set of opportunities. Therefore, the agents will explore q opportunities in each $j_{accept} + j_{reject} + k$ time periods ($k + j_{accept}$ time periods associated with exploration proposed by Agt_1 and reject time periods associated with Agt_2's rejected proposal). It is easy to see that if Agt_1 uses this strategy, then Agt_2 gains no utility at all, and Agt_1 gains a utility of $SW_{OPT}(B, \delta^{1+\frac{j_{reject}}{j_{accept}+k}})$, as required. ■

6 Conclusions and Future Work

As is often common in real life, both negotiation and the execution of the negotiated underlying exploration are time consuming processes. Therefore binding the two of them to the same deadline enables a more realistic problem modeling, hence improving the applicability of such models. Such binding, however, increases the model complexity, whenever the two interleave, as each influences the other: the extent of the negotiation affects the amount of time that can be allocated to exploration, and the results of the exploration affect the proposals made in consequent negotiation rounds. As discussed and demonstrated throughout the paper, the use of such a negotiation scheme results both in analysis and solutions that are different from those obtained with legacy negotiation protocols that completely separate between the two.

The paper focuses on the full information model, which lets us distill the essence of combining negotiation and exploration. However, we see this as a stepping stone to analyzing models when there is a prior, and players have varying knowledge about the state of the system. Naturally, any Bayesian setting will have to include our results as a special case.

The paper studies two fundamental negotiation protocols of the above type, one where the exploration is interleaved with the negotiation itself and the other where the exploration details are negotiated first and then the negotiation executes within the pre-defined time constraints. As discussed in the introduction, each of the two protocols has its unique advantages compared to the other. In the non-interleaved protocol more time can potentially be allocated for the exploration, hence increasing the social welfare. In the interleaved protocol the agents do not decide beforehand how the exploration will be performed and how much the beneficiary agent will pay the exploring agent, but rather negotiate it as the exploration progresses. The payment is thus better correlated with the actual exploration that takes place and the risk taken by the exploring agent in running into an extensive non-beneficial exploration is minimized. Moreover, the interleaved negotiation protocol does not bind the agents to a comprehensive exploration process, potentially enabling the agents to opt out even between exploration executions. It also enables the agent interested in the negotiation verifying an agreed execution of the exploration and is more intuitive for people, as it does not require agreeing on the entire exploration strategy in advance.

The analysis given in the paper shows that both of the agents' strategies, their expected utilities and the social welfare are different whenever each of the two negotiation protocols are used. Therefore, being able to choose between the two protocols (which is facilitated through the analysis given) is important, as the agents may benefit (individually or jointly) from such a choice. Furthermore, enabling any of the agents partial control over the value of some of the negotiation parameters, such as the negotiation horizon, as is often the case in real-life, may substantially affect the results. The implications are not limited only to individual utility but also concern the social welfare — an agent may choose to deviate to a sub-optimal (social-welfare-wise) negotiation if this increases its own expected utility. Some of the implications are counter-intuitive. For example, with the non-interleaved protocol, a greater social welfare is always obtained, however, for some settings, when using the interleaved protocol, if the discounting factor is relatively high (close to 1), one of the agents may end up taking over almost the entire social welfare, thus preferring that protocol over the first. Moreover, we find that for some environments, if given the option to choose the negotiation protocol and the negotiation horizon, agents always prefer the interleaved protocol (as they can decide on a value such that they will be the last proposer) while if the negotiation horizon is fixed and the agents can only choose the protocol to be used, then their decision is setting-dependent. The degree of control that the agent obtains has a substantial effect on the degradation in the social welfare. For example, if the agent is allowed to choose a fixed number of opportunities that will be explored at every step, there can be a sharp (potentially unbounded) drop in social welfare, whereas if the agent can a priori chose a different number for every step, the drop is bounded.

The inherent advantages of the interleaved protocol, as surveyed above, as well as its dominance for one of the agents in some settings (as illustrated in the former section) suggest that it is often likely to be preferred in some real-life. Yet, the protocol is also associated with a substantially increased computational complexity. Theorem 3 reduces the complexity of the interleaved protocol to the same order of magnitude as solving the stand-alone exploration problem, thus enabling both to determine the exploration that will be offered and compute the payments at each step.

Finally, we emphasize that the proof given in Theorem 3 can be generalized to other negotiation settings (not necessarily exploration-based), such as the general case of task allocation and planning. Here, whenever considering a multi-task setting, it is also common to find the task execution and negotiation bound to the same negotiation deadline. The agents may prefer to negotiate and execute the tasks agreed upon as two separate processes (non-interleaved) or one at a time (interleaved). The analysis of such domains is, of course, beyond the scope of the current paper and is thus left to future research. Other interesting ideas for future research that can benefit from the analysis given in the paper include: (a) identifying some guidelines for preferring any of the negotiation protocols without having to calculate the resulting utilities (e.g., characterize environments where one of the protocols is generally better than the other); (b) analyzing the negotiation enabling proposing the (sequential) exploration of a partial set of opportunities each time; and (c) allowing several negotiation steps to take place before each exploration or set of explorations.

References

1. L. Amgoud, Y. Dimopoulos, and P. Moraitis. A unified and general framework for argumentation-based negotiation. In *proceedings of the 6th international joint conference on Autonomous agents and multiagent systems (AAMAS-07)*, pages 158:1–158:8, 2007.
2. L. Amgoud and S. Vesic. A formal analysis of the role of argumentation in negotiation dialogues. *Journal of Logic and Computation*, 22(5):957–978, 2012.
3. B. An, N. Gatti, and V. Lesser. Bilateral bargaining with one-sided uncertain reserve prices. *Autonomous Agents and Multi-Agent Systems*, 26(3):420–455, 2013.
4. A. Atakan. Assortative matching with explicit search costs. *Econometrica*, 74(3):667–680, 2006.
5. J. Atlas and K. Decker. Coordination for uncertain outcomes using distributed neighbor exchange. In *proceedings of the Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS-10)*, pages 1047–1054, 2010.
6. A. Azaria, Y. Aumann, and S. Kraus. Automated strategies for determining rewards for human work. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.
7. A. Azaria, Y. Aumann, and S. Kraus. Automated agents for reward determination for human work in crowdsourcing applications. *Autonomous Agents and Multi-Agent Systems*, 28(6):934–955, 2014.
8. T. Baarslag, K. Fujita, E. H. Gerding, K. V. Hindriks, T. Ito, N. R. Jennings, C. M. Jonker, S. Kraus, R. Lin, V. Robu, and C. R. Williams. Evaluating practical negotiating agents: Results and analysis of the 2011 international competition. *Artif. Intell.*, 198:73–103, 2013.
9. T. Baarslag, K. Hindriks, C. M. Jonker, S. Kraus, and R. Lin. The first automated negotiating agents competition (anac 2010). In T. Ito, M. Zhang, V. Robu, S. Fatima, and T. Matsuo, editors, *New*

Trends in Agent-based Complex Automated Negotiations, volume 383 of *Studies in Computational Intelligence*, pages 113–135. Springer-Verlag, Springer-Verlag, 2012.
10. M. Bac and H. Raff. Issue-by-issue negotiations: the role of information and time preference. *Games and Economic Behavior*, 13(1):125–134, 1996.
11. L. Barbulescu, Z. Rubinstein, S. Smith, and T. Zimmerman. Distributed coordination of mobile agent teams: the advantage of planning ahead. In *proceedings of the Ninth International Conference on Autonomous Agents and Multiagent Systems (AAMAS-10)*, pages 1331–1338, 2010.
12. J. Benhabib and C. Bull. Job search: The choice of intensity. *Journal of Political Economy*, 91(5):747–764, 1983.
13. R. Buttner. A classification structure for automated negotiations. In *proceedings of the IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*, WI-IATW '06, pages 523–530, 2006.
14. Y. Chevaleyre, U. Endriss, and N. Maudet. Simple negotiation schemes for agents with simple preferences: Sufficiency, necessity and maximality. *Autonomous Agents and Multi-Agent Systems*, 20(2):234–259, 2010.
15. M. Chhabra, S. Das, and D. Sarne. Expert-mediated sequential search. *European Journal of Operational Research*, 234(3):861–873, 2014.
16. P. Dunne, M. Wooldridge, and M. Laurence. The complexity of contract negotiation. *Artificial Intelligence*, 164(1-2):23–46, 2005.
17. E. Ephrati and J. Rosenschein. A framework for the interleaving of execution and planning for dynamic tasks by multiple agents. In *From Reaction to Cognition*, pages 139–153. Springer, 1995.
18. P. Faratin, C. Sierra, and N. Jennings. Using similarity criteria to make issue trade-offs in automated negotiations. *Artificial Intelligence*, 142:205–237, 2002.
19. S. Fatima, M. Wooldridge, and N. Jennings. An agenda-based framework for multi-issue negotiation. *Artificial Intelligence*, 152(1):1–45, 2004.
20. S. Fatima, M. Wooldridge, and N. Jennings. Multi-issue negotiation with deadlines. *Journal of Artificial Intelligence Research*, 27(1):381–417, 2006.
21. T. S. Ferguson. Who solved the secretary problem? *Statistical Science*, 4(3):282–289, 1989.
22. C. Fershtman. The importance of the agenda in bargaining. *Games and Economic Behavior*, 2(3):224–238, 1990.
23. S. Gal, M. Landsberger, and B. Levykson. A compound strategy for search in the labor market. *International Economic Review*, 22(3):597–608, 1981.
24. N. Gatti. Extending the alternating-offers protocol in the presence of competition: models and theoretical analysis. *Annals of Mathematics and Artificial Intelligence*, 55:189–236, 2009.
25. B. Grosz and S. Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.
26. M. Hadad, S. Kraus, and I. B.-A. Hartman. Cooperative planning with time constraints. *Annals of Mathematics and Artificial Intelligence*, forthcoming, 2013.
27. G. Haim, Y. Gal, M. Gelfand, and S. Kraus. A cultural sensitive agent for human-computer negotiation. In *proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, AAMAS-12, pages 451–458, 2012.
28. J. Harsanyi and R. Selten. A generalized nash solution for two-person bargaining games with incomplete information. *Management Science*, 18(5):80–106, 1972.
29. N. Hazon, Y. Aumann, S. Kraus, and D. Sarne. Physical search problems with probabilistic knowledge. *Artificial Intelligence*, 196:26–52, 2013.
30. N. Jennings, P. Faratin, A. Lomuscio, S. Parsons, C. Sierra, and M. Wooldridge. Automated negotiation: prospects, methods and challenges. *International Journal of Group Decision and Negotiation*, 10(2):199–215, 2001.
31. N. Jennings, P. Faratin, A. Lomuscio, S. Parsons, M. Wooldridge, and C. Sierra. Automated negotiation: prospects, methods and challenges. *Group Decision and Negotiation*, 10(2):199–215, 2001.
32. J. Kephart and A. Greenwald. Shopbot economics. *Journal of Autonomous Agents and Multi-Agent Systems*, 5(3):255–287, 2002.
33. J. Kim and A. Segev. A framework for dynamic ebusiness negotiation processes. In *E-Commerce, 2003. CEC 2003. IEEE International Conference on*, pages 84–91. IEEE, 2003.
34. S. Krasteva and H. Yildirim. On the role of confidentiality and deadlines in bilateral negotiations. *Games and Economic Behavior*, 2012.
35. S. Kraus. Negotiation and cooperation in multi-agent environments. *Artificial Intelligence*, 94:79–97, 1997.

36. S. Kraus. *Strategic Negotiation in Multiagent Environments*. MIT Press, Cambridge, USA, 2001.
37. S. Kraus and D. Lehmann. Designing and building a negotiating automated agent. *Computational Intelligence*, 11(1):132–171, 1995.
38. S. Kraus and J. Wilkenfeld. The function of time in cooperative negotiations. In *proceedings of the ninth National conference on Artificial intelligence*, AAAI'91, pages 179–184, 1991.
39. S. Kraus and J. Wilkenfeld. Negotiations over time in a multi agent environment: Preliminary report. In *proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91)*, pages 56–61, 1991.
40. S. Kraus, J. Wilkenfeld, and G. Zlotkin. Multiagent negotiation under time constraints. *Artificial Intelligence*, 75(2):297–345, 1995.
41. K. Larson and T. Sandholm. Bargaining with limited computation: Deliberation equilibrium. *Artificial Intelligence*, 132(2):183 – 217, 2001.
42. K. Larson and T. Sandholm. An alternating offers bargaining model for computationally limited agents. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1*, AAMAS '02, pages 135–142, New York, NY, USA, 2002. ACM.
43. K. Larson and T. Sandholm. An alternating offers bargaining model for computationally limited agents. In *proceedings of the first international joint conference on Autonomous agents and multiagent systems (AAMAS-02)*, pages 135–142, 2002.
44. R. Lin and S. Chou. Mediating a bilateral multi-issue negotiation. *Electronic Commerce Research and Applications*, 3(2):126–138, 2004.
45. R. Lin, Y. Gal, S. Kraus, and Y. Mazliah. Training with automated agents improves people's behavior in negotiation and coordination tasks. *Decision Support Systems*, 60(0):1 – 9, 2014.
46. R. Lin and S. Kraus. Can automated agents proficiently negotiate with humans? *Communications of the ACM*, 53(1):78–88, 2010.
47. R. Lin, S. Kraus, J. Wilkenfeld, and J. Barry. An automated agent for bilateral negotiation with bounssded rational agents with incomplete information. pages 270–274, 2006.
48. S. Lippman and J. McCall. The economics of job search: A survey. *Economic Inquiry*, 14(3):155–189, 1976.
49. A. Lomuscio, M. Wooldridge, and N. Jennings. A classification scheme for negotiation in electronic commerce. In *Proceedings of Agent Mediated Electronic Commerce*, pages 19–33, 2001.
50. A. Ma and M. Manove. Bargaining with deadlines and imperfect player control. *Econometrica*, 61(6):1313–39, 1993.
51. J. McMillan and M. Rothschild. Search. In *Handbook of Game Theory with Economic Applications*, pages 905–927. 1994.
52. D. Miller and E. Gat. proceedings of exploiting known topologies to navigate with low-computation sensing. In *Fibers' 91*, pages 425–435. International Society for Optics and Photonics, 1991.
53. P. Morgan. Search and optimal sample sizes. *The Review of Economic Studies*, 50(4):659, 1983.
54. P. Morgan and R. Manning. Optimal search. *Econometrica*, pages 923–944, 1985.
55. D. Musliner, E. Durfee, and K. Shin. World modeling for the dynamic construction of real-time control plans. *Artificial Intelligence*, 74(1):83–127, 1995.
56. R. Myerson. Two-person bargaining problems with incomplete information. *Econometrica*, 52(2):461–487, 1984.
57. Y. Nahum, D. Sarne, S. Das, and O. Shehory. Two-sided search with experts. *Autonomous Agents and Multi-Agent Systems*, 29(3):364–401, 2015.
58. J. Nash. The Bargaining Problem. *Econometrica*, 18:155–162, 1950.
59. N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic game theory*, volume 1. Cambridge University Press Cambridge, 2007.
60. M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1994.
61. M. Paolucci, O. Shehory, and K. Sycara. Interleaving planning and execution in a multiagent team planning environment. *Electronic Transaction Artificial Intelligence*, 4:23–43, 2000.
62. H. Raïffa. *The art and science of negotiation*. Belknap Press of Harvard University Press, 1982.
63. I. Rochlin, Y. Aumann, D. Sarne, and L. Golosman. Efficiency and fairness in team search with self-interested agents. In *International conference on Autonomous Agents and Multi-Agent Systems, AAMAS'14*, pages 365–372, 2014.
64. I. Rochlin, D. Sarne, and M. Mash. Joint search with self-interested agents and the failure of cooperation enhancers. *Artificial Intelligence*, 214:45–65, 2014.
65. A. Rosenfeld, I. Zuckerman, E. Segal-Halevi, O. Drein, and S. Kraus. NegoChat-A: a chat-based negotiation agent with bounded rationality. *Autonomous Agents and Multi-Agent Systems*, pages 1–22, 2015.

66. J. Rosenschein and G. Zlotkin. *Rules of encounter: designing conventions for automated negotiation among computers*. the MIT Press, 1994.
67. M. Rothschild. Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, 82(4):689–711, 1974.
68. A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109, 1982.
69. A. Rubinstein. A bargaining model with incomplete information about time preferences. *Econometrica*, 53(5):1151–72, 1985.
70. T. Sandholm and N. Vulkan. Bargaining with deadlines. In *proceedings of the National Conference on Artificial Intelligence (AAAI-99)*, pages 44–51, 1999.
71. D. Sarne and S. Kraus. Managing parallel inquiries in agents two-sided search. *Artificial Intelligence*, 172(4-5):541–569, 2008.
72. D. Sarne, E. Manisterski, and S. Kraus. Multi-goal economic search using dynamic search structures. *Autonomous Agents and Multi-Agent Systems*, 21(2):204–236, 2010.
73. R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.
74. C. Sierra, P. Faratin, and N. Jennings. A service-oriented negotiation model between autonomous agents. In *proceedings of the 8th European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW)*, pages 17–35, 1997.
75. C. Sierra, N. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In *proceedings of the 4th International Workshop on Agent Theories, Architectures, and Languages (ATAL)*, pages 177–192, 1997.
76. R. Simmons. An architecture for coordinating planning, sensing, and action. In *proceedings of DARPA Workshop on Innovative Approaches to Planning, Scheduling and Control*, pages 292–297, 1990.
77. L. Smith. Frictional matching models. *Annual Reviews in Economics*, 3(1):319–338, 2011.
78. I. Stahl. *Bargaining theory*. Economics Research Institute, Stockholm School of Economics, Stockholm, 1972.
79. G. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961.
80. T. Tan and S. Werlang. A guide to knowledge and games. In *proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 163–177. Morgan Kaufmann Publishers Inc., 1988.
81. X. Tong, W. Zhang, and H. Huang. Agent negotiation on resources with nonlinear utility functions. *Computer Science and Information Systems*, pages 1697–1720, 2012.
82. G. van den Berg. Nonstationarity in job search theory. *Review of Economic Studies*, 57:255–277, 1990.
83. T. Wagner, J. Phelps, V. Guralnik, and R. VanRiper. An application view of Coordinators: Coordination managers for first responders. In *proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, pages 908–915, 2004.
84. M. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–54, 1979.
85. M. Wooldridge. *Reasoning about rational agents*. the MIT Press, 2000.
86. R. Zheng, N. Chakraborty, T. Dai, K. Sycara, and M. Lewis. Automated bilateral multiple-issue negotiation with no information about opponent. In *proceedings of the 46th Hawaii International Conference on System Sciences (HICSS)*, pages 520–527, 2013.
87. G. Zlotkin and J. Rosenschein. Mechanism design for automated negotiation, and its application to task oriented domains. *Artificial Intelligence*, 86(2):195–244, 1996.