# Modeling Risk-Averse Behavior in $k$-Lookahead Search

Matthew Eichhorn, Shawn Ong, Zhen Zhang

## 1 Introduction

In deterministic sequential games, there is an implicit trade-off between payoff and computational efficiency. A player with complete information and unlimited computational resources could fully calculate the game tree and use Zermelo's backward induction algorithm to choose the strategy that maximizes their payoff. On the other hand, a player who chooses to myopically best-respond on each of their turns performs very little computation, at the expense for no guarantee in their eventual payoff. In their recent paper [1], Mirrokni, Thain, and Vetta introduce $k$-lookahead search, a strategy selection mechanism that captures the essence of this trade-off. In $k$-lookahead search, a player computes the next $k$ levels of the game tree from the current node, estimates the value at the leaf nodes, and performs backward induction in order to select an optimal strategy. In this way, $k$-lookahead search naturally captures both myopic best-response ($k = 1$) and Zermelo's algorithm ($k = \infty$), as well as a spectrum of intermediary approaches.

One assumption that is made in the analysis of Mirrokni et al. is that each player is risk-neutral; that is, players assume that each player acts to maximize their (expected) payout each turn. This assumption is realized during the exploration of the search tree, wherein all branches which do not maximize the payout to the acting player are pruned. While this assumption greatly simplifies the analysis, it is not necessarily accurate in all scenarios. Notably, it does nothing to account for irrational decisions, or malicious players who seek not to maximize their own payout but rather to wreak havoc and decrease the payout to others. A natural response for these behaviors is for a player to adopt of a risk-averse viewpoint, in which they are willing to sacrifice some of their payout in order to minimize their risk later in the game. In this way, a player seeks to maximize their worst case outcome, thus minimizing the effects of an irrational or spiteful player. Through the remainder of this paper, we focus on examining the results of risk-averse players using $k$-lookahead search.

### 1.1 Definitions and Notations

The strategic game $G$ is defined as a tuple $G(\mathcal{P}, \mathcal{S}, \{v_i : i \in \mathcal{P}\})$. $\mathcal{P}$ is the set of $n$ players $\{1, \ldots, n\}$. The strategy space $\mathcal{S}$ is the Cartesian product $(S_1 \times S_2 \times \cdots \times S_n)$, where each $S_i$ is the set of strategies available for player $p_i$. $v_i : \mathcal{S} \to \mathbb{R}$ is the value function for player $i$ given a strategy profile from $\mathcal{S}$. We assume the set of players $\mathcal{P}$ and the value functions are independent of time, i.e. they do not change throughout the game. Thus, a node in the game tree is defined by the strategy vector $s \in \mathcal{S}$.

A player $i \in \mathcal{P}$ using lookahead search to move from $s \in \mathcal{S}$ in the game tree does so based on the best child node. To be precise, suppose $i$ is about to move from $s$ and $T_i$ is the local search tree. First $i$ assigns a valuation $\Pi_j(l)$ for each player $j$ (including $i$) and each leaf node $l$ in $T_i$. Then for each remaining vertex $v$ in $T_i$, player $i$ evaluates its value for all players, assigning values $\Pi_j(v)$ (intuitively, what $i$ believes $j$'s value to be at $v$). We define values recursively for player $i$ (who is about to move at state $v$) as:

$$\Pi_i(v) = \max_{v' \in \mathcal{C}(v)} [r_{i,v} + \Pi_i(v')]$$

And at other states, $\Pi_i(v) = r_{i,v} + \Pi_i(v')$ where $v'$ is the state after the next player moves (for leaves, we have instead $\Pi_i(l) = r_{i,l}$). For risk-averse player $i$, assume any non-moving player $j \neq i$ places a value of $\Pi_j(v) = -\Pi_i(v)$ on node $v$. Intuitively, $j$ maximizing their value corresponds to choosing the states which have the worst value for $i$. From this, $i$ is able to compute lookahead payoff $\Pi_i(s')$ for all children of the root of $T_i$ and move accordingly. This allows us to define a *lookahead improving move* $s'_i$ for player $i$ at state $s \in \mathcal{S}$ to be one where $\Pi_i(s) \leq \Pi_i(s'_i, s_{-i})$, and a *lookahead best response* for $i$ at state $s$ to be $s_i$ such that $\forall s'_i \in \mathcal{S}_i : \Pi_i(s_i, s_{-i}) \leq \Pi_i(s'_i, s_{-i})$.

It is important to distinguish what $j$ would do in $i$'s opinion and what $j$ actually would do. Suppose player $i$ moved from $s$ to $s'$ so that player $j$ is about to move at node $s'$, and $j$ applies the same lookahead procedure with a new search tree $T_j$. Player $i$ may not be able to fully simulate $j$'s computation so that even if both players are aware of $j$'s mechanism, they may compute different moves. For example, if both $i$ and $j$ use $k$-lookahead search, then $i$ can only perform a $k-1$ lookahead search when considering $j$'s move, while $j$ can do a $k$-lookahead search when it is actually her move.

### 1.1.1 Path model vs. Leaf model

For games such as chess where only the final state matters, we have $r_{i,s} = 0$ for all $i$ and $s$. There is no intermediate gain associated with any node before the game ends, so the value of a node is entirely determined by the values of its child nodes. We call models associated with such games the *leaf model* since players are playing to reach a desired leaf in the local search tree (intuitively, the values of the leaves should correspond to estimates of the value of their descendant leaves on the game tree). On the other hand, for games such as business investment, each business invested has its short-term payoff as well as long-term payoff. In these games different paths leading to the same node may have different values. Players strive to pick the path with highest value for themselves. These models are called the *path model*.

### 1.1.2 Equilibrium vs. Dynamics

Two ways of studying the performance of lookahead search are via equilibria and via dynamics. *A lookahead equilibrium*, strategy profile where each player plays a lookahead best response. The price of anarchy is defined as the ratio of the worst lookahead equilibrium and the global optimal social welfare. *A lookahead dynamic* is a walk on the state graph $\mathcal{G}$, where nodes are set of states and there is an edge from $s$ to $s'$ with label $i$ if $s'$ is the result of $i$ switching to a lookahead best response at $s$. The price of anarchy for lookahead dynamics is defined as the worst ratio between expected social welfare on a random walk polynomial in $\mathcal{G}$ and the global optimal social welfare. It may be more enlightening to study lookahead dynamics, as equilibria are not guaranteed to exist; and if they do, similar techniques for proving bounds on dynamics often apply to equilibria as well.

#### 1.1.3 Order of Moves

Lookahead can also depend on the order in which players move, as this affects the shape of the game tree. When there are more than two players, we must distinguish between the *worst-case* and *average-case* models; in the former, the player assumes that the sequence of players chosen to move is adversarial. In the latter case, the player assumes that at each round, a player is selected to move uniformly at random (this may lead to consecutive moves by the same player). We focus our attention here on the average-case model.

### 1.2 Paper Layout

In Section 2, we discuss the Cournot model for duopolistic competition. We categorize optimal risk-averse lookahead play, and consider the long term behavior of competition between risk-averse firms, as well as between a risk-averse and a risk-neutral firm. In Section 3, we discuss atomic selfish routing. We begin by proving a price of anarchy bound on a risk-averse 2-lookahead equilibrium. Finally, we establish a constant bound on the price of anarchy for lookahead dynamics with risk-averse players. In Section 4 we discuss limitations of our work and possible extensions.

## 2 Cournot Competition

We consider the competition of firms in a duopolostic market under the Cournot model [2]. Under this model, we assume that the price of the item is a function of the market saturation. To be more concrete, let $p$ be the production quantity of Company 1 and $q$ be the production quantity of Company 2. Assume that it costs each company $c$ to produce one item[1], and that $d$ represents the market demand for the item. Then, the price for each item is set to $d - p - q$, so that the utility to the companies can be represented as

$$u_1(p, q) = p(d - p - q - c) \qquad\qquad u_2(p, q) = q(d - p - q - c),$$

the profit to each company at these production levels. Without loss of generality, we can assume $d - c = 1$, so that

$$u_1(p, q) = p(1 - p - q) \qquad\qquad u_2(p, q) = q(1 - p - q).$$

By setting $q = \frac{1}{3}$, we see that company 1 can maximize their profits by setting $p = \frac{1}{3}$, so by symmetry $(\frac{1}{3}, \frac{1}{3})$ is a Nash equilibrium, the unique equilibrium in pure strategies. We measure social welfare in this game as $SW(p, q) = p + q$, the fraction of the demand which is satisfied. The social welfare at this equilibrium is $\frac{2}{3}$, and each company has profit $\frac{1}{9}$.

[1]Economics of scale make it unreasonable to assume that the marginal cost of production is fixed. However, this assumption was made in Cournot's original model, so we adopt this view here.

## 2.1 Applying a Risk-Averse $k$-Lookahead Strategy

In order to apply lookahead strategies to Cournot Competition, we must make it into a sequential game. To do this, we assume that the companies take turns adjusting their prices, realizing their profit between each adjustment. A risk-neutral, or utility maximizing, company will seek to maximize their overall profits summed over all rounds (this is an example of the path model) and will assume that the other company is doing the same. A risk-averse company will seek also seek to maximize their overall profits, but will assume that the other company is trying to maliciously run them out of business. Since a company can, theoretically, raise their production arbitrarily high, making the price of the good negative, there is no per se "worst-case" outcome for the risk-averse company to compute. Therefore, we make the additional assumption that neither company will ever act in a way which causes them to lose money in a round.[2] Thus, the risk-averse company will assume that the other company will always respond in a way that makes total production 1 (and thus makes profits 0). The resulting strategy is discussed in the following lemma.

**Lemma 1.** *Assume that Company 1 is employing a risk-averse (as described above) $k$-lookahead search. Then, if Company 2 currently has Production $q$, Company 1 will set their production to $p = \frac{j}{j+1}(1-q)$ where $j = \lceil \frac{k}{2} \rceil$.*

*Proof.* We argue by induction on $k$ that this is the optimal strategy according to $k$-lookahead and that Company 1 will anticipate earning $\frac{j}{2(j+1)}(1-q)^2$ during the lookahead period. Note that it is sufficient to argue for $k$ odd because in every even round (where Company 2 adjusts their production) Company 1 will anticipate making no profit. As a base case, the myopic best response of Company 1 is the maximize their immediate profit. By applying univariate calculus, we see that

$$0 = \frac{\partial}{\partial p}\left[p - p^2 - pq\right] = 1 - 2p - q \implies p = \frac{1}{2}(1-q),$$

and Company 1 anticipates earning a profit of

$$\frac{1-q}{2}\left(1 - \frac{1-q}{2} - q\right) = \left(\frac{1-q}{2}\right)^2.$$

Now, consider a $k$ lookahead strategy. If Company 1 responds by producing $p$, then in the next round Company 2 would produce $1-p$. Applying the inductive hypothesis, Company 1's best $(k-2)$-lookahead strategy from this point would earn them $\frac{(j-1)}{2j}p^2$ over the next $k-2$ rounds. Therefore, Company 1 anticipates earning

$$p(1-p-q) + \frac{(j-1)}{2j}p^2$$

[2] This assumption may not be entirely accurate. There are numerous examples of *predatory pricing*, instances where companies lower prices beyond their profitable capacity in order to force competition out of business, in modern economics. For more information about predatory pricing and mechanisms to alleviate it, see [3].

over the $k$ rounds. Again, by univariate calculus, we find that

$$0 = \frac{\partial}{\partial p}\left[p(1-p-q) + \frac{(j-1)}{2j}p^2\right] = 1 - 2p - q + \frac{j-1}{j}p \implies p = \frac{j}{j+1}(1-q),$$

and Company 1 anticipates earning

$$\begin{aligned}
&\frac{j}{j+1}(1-q)\left(1 - \frac{j}{j+1}(1-q) - q\right) + \frac{(j-1)j}{2(j+1)^2}(1-q)^2 \\
&= \frac{j}{(j+1)^2}(1-q)^2 + \frac{(j-1)j}{2(j+1)^2}(1-q)^2 \\
&= \frac{j}{2(j+1)}(1-q)^2,
\end{aligned}$$

completing the induction. □

## 2.2 Two Risk-Averse Companies

We can use this lemma to determine the behavior of companies utilizing this risk-averse lookahead strategy. First, suppose that both of the companies are using this strategy with the same lookahead window $k$. Then, their productions will each quickly converge to the fixed point of the map $x \mapsto \frac{j}{j+1}(1-x)$. Doing the algebra, we have

$$(j+1)x = j(1-x) \implies jx + x = j - xj \implies x = \frac{j}{2j+1}.$$

Note that when $k = 1$ (so $j = 1$), we recover the myopic equilibrium $(\frac{1}{3}, \frac{1}{3})$, and as $k \to \infty$, the production by each company approaches $\frac{1}{2}$. Therefore, the risk aversion of the companies leads to an overall increase in social welfare (that is, an increase in production) at the expense of company profitability. Intuitively, when one company acts under the assumption that the other is going to saturate the market, it is in their best interest to ramp up production now to reap as much in short-term profits as they can.

Next, we consider the case where both companies are risk-averse, but they have different lookahead windows. Without loss of generality, we let $k$ denote Company 1's lookahead and $k'$ denote Company 2's lookahead with $k > k'$ (correspondingly, we define $j, j'$ as above). Then, the production levels $p$ and $q$ will converge to the solution of the system

$$p = \frac{j}{j+1}(1-q) \qquad\qquad q = \frac{j'}{j'+1}(1-p),$$

which can be computed to be

$$p = \frac{j}{j+j'+1} \qquad\qquad q = \frac{j'}{j+j'+1}.$$

Note that when $j = j'$, this captures the fixed point solution from above. Moreover, since $j > j'$, $p > q$, so Company 1 will reap a greater profit than Company 2. This confirms the intuition that a risk-averse company with greater capacity for market prediction should realize greater profits.

## 2.3 Risk-Averse versus Risk-Neutral Companies

As a final application of the Cournot model, we consider a risk-neutral company competing with a risk-averse company. Specifically, we focus on the case where the companies have the same lookahead window $k$ and determine which mindset results in greater profitability. We utilize the recurrence in Lemma 2.1 from [1] to determine the optimal lookahead strategy for the risk-neutral company and the above lemma to reason about the risk-averse company.

In the case where $k = 2$, Company 1 (the risk-averse company) best responds to Company 2 (the risk-neutral company) producing $q$ units by producing $p = \frac{1}{2}(1-q)$ units. Company 2 best responds to Company 1 producing $p$ units by producing $q \approx 0.30508 - 0.48305p$ units. A sequence of these responses converges to the production levels $(p, q) \approx (0.43854, 0.12293)$. For any $k \geq 3$, we have

$$p \geq \frac{2}{3}(1-q) \qquad\qquad q \leq 0.298 - 0.478p,$$

for which a sequence of best responses will eventually drive $q \to 0$ and $p \to \frac{j}{j+1}$. Therefore, for any $k \geq 2$, a risk-averse lookahead strategy results in greater profitability than a risk-neutral strategy. Moreover, having even one risk-averse company ensures a greater production level (and thus greater social welfare) than 2 myopic companies and also ensures a greater production level than 2 risk-neutral companies for $k \geq 5$.

# 3 Atomic Selfish Routing

We now examine the atomic selfish routing game. This game consists of a directed graph $G = (V, E)$, and $n$ players, each of whom wishes to route 1 unit of flow from $s_i$ to $t_i$ along a path $P_i$. The paths together define a flow $f$; we assume that each edge has a linear latency $\lambda_e(f_e) = a_e f_e + b_e$ with $a_e, b_e \geq 0$. The total latency of a flow $f$ (i.e. the social cost) is then denoted $l(f) = \sum_{e \in E} \lambda_e(f_e) f_e$; likewise, denote a player's latency as $l_i(f) = \sum_{e \in P_i} \lambda_e(f_e)$. Again, to apply lookahead, we turn this into a sequential game by allowing players to update their choices over time. Like the original paper, we will consider this game under 2-lookahead in the leaf model. We use an analogous definition of risk-aversion here: players want to minimize their own latency, under the assumption that all other players in the game are acting to maximize it.

## 3.1 Price of Anarchy for Lookahead Equilibrium

We present a proof based on largely based on that of Theorem 4.1 in [1]. Note that this proof itself is based on work in Awerbuch, Azar, Epstein [4].

**Theorem 2.** *In the average-case 2-lookahead model with risk-averse players, the price of anarchy for a lookahead equilibrium is at most* 4.

*Proof.* Let $f$ be a flow in lookahead equilibrium and let $f^*$ be an optimal flow. Let $P_j$ denote the path which $j$ takes in $f$ and $P_j^*$ the same for $f^*$. Let $J(e)$ be the set of players using edge

$e$ in $f$, and $J^*(e)$ the equivalent for $f^*$. Note that at a 2-lookahead equilibrium, $j$ prefers playing $P_j$ to $P_j^*$, i.e. it must be that the worst outcome from a move after choosing $P_j$ is at least as good as the worst from a move after $P_j^*$. In particular, the former has latency no less than that obtained when choosing $P_j$ if the next player remains at their current position. The worst possible outcome from choosing $P_j^*$ is no worse than if another player who originally shared no edges with $P_j^*$ moves entirely onto the path. This gives:

$$\sum_{e \in P_j^*} a_e(f_e + 2) + b_e \geq \sum_{e \in P_j} a_e f_e + b_e$$

Now sum over all players:

$$\sum_j \sum_{e \in P_j^*} a_e(f_e + 2) + b_e \geq \sum_j \sum_{e \in P_j} a_e f_e + b_e = \sum_{e \in E} (a_e f_e + b_e) f_e = \sum_{e \in E} \lambda_e(f_e) f_e$$

We reverse this inequality and continue:

$$\begin{aligned} \sum_{e \in E} \lambda_e(f_e) f_e &\leq \sum_j \sum_{e \in P_j^*} a_e(f_e + 2) + b_e \\ &= \sum_{e \in E} (a_e(f_e + 2) + b_e) f_e^* \\ &\leq \sum_{e \in E} a_e f_e f_e^* + (2a_e + b_e) f_e^* \end{aligned}$$

Now note that since $f_e^* \in \mathbb{Z}_{\geq 0}$, we have that $f_e^{*2} \geq f_e^*$.

$$\begin{aligned} &\leq \sum_{e \in E} a_e f_e f_e^* + 2a_e f_e^{*2} + b_e f_e^* \\ &\leq \sum_{e \in E} a_e f_e f_e^* + 2 \sum_{e \in E} \lambda_e(f_e^*) f_e^* \end{aligned}$$

Applying the Cauchy-Schwartz inequality to the first term:

$$\begin{aligned} &\leq \sqrt{\sum_{e \in E} a_e f_e^2} \sqrt{\sum_{e \in E} a_e f_e^{*2}} + 2 \sum_{e \in E} \lambda_e(f_e^*) f_e^* \\ &\leq \sqrt{\sum_{e \in E} \lambda_e(f_e) f_e} \sqrt{\sum_{e \in E} \lambda_e(f_e^*) f_e^*} + 2 \sum_{e \in E} \lambda_e(f_e^*) f_e^* \end{aligned}$$

Take $\rho = \sqrt{\frac{\sum_e \lambda_e(f_e) f_e}{\sum_e \lambda_e(f_e^*) f_e^*}}$; note that $\rho^2$ is the PoA if $f$ is the worst lookahead equilibrium. Dividing the resulting inequality by $\sum_{e \in E} \lambda_e(f_e^*) f_e^*$ gives $\rho^2 \leq \rho + 2$. Thus, $\rho \leq 2$ and the price of anarchy is at most 4. □

In particular, we may compare this to the price of anarchy bound of $(1 + \sqrt{5})^2 \approx 10.472$ for risk-neutral players in [1]. Since the optimal social cost remains the same, we have a tighter bound on social cost at lookahead equilibrium for risk-averse players.

## 3.2 Price of Anarchy for Lookahead Dynamics

We now show that the price of anarchy for lookahead dynamics is also bounded by a constant when players are risk-averse. The proof once again follows similarly to that in [1] with a number of lemmas, but we may use risk aversion to improve the bounds. This proof also draws heavily on work in Goemans, Mirrokni, Vetta [5].

**Lemma 3.** *If player $i$ makes a lookahead-improving move from path $P_i$ to $P_i'$ which changes the flow from $f$ to $f_i'$, then $l_i(f_i') \leq 2l_i(f)$*

*Proof.* Note that $l_i(f_i')$ is at most $i$'s worst-case latency after a move from $f_i'$. But this is at most the worst-case latency if $i$ were to remain on $P_i$, which in turn is bounded by:

$$\sum_{e \in P_i} a_e(f_e + 1) + b_e \leq \sum_{e \in P_i} 2a_e f_e + b_e \leq 2l_i(f)$$

Here, the first term assumes that the next player to move did not share any edges with $P_i$ and moved entirely onto the path in the next turn.. Thus, we have $l_i(f_i') \leq 2l_i(f)$. □

**Lemma 4.** *Let $f$ be the current flow and choose at random a player to make a lookahead best response. The resulting flow $f'$ satisfies $\mathbb{E}[l(f') \mid f] \leq (1 + \frac{3}{n})l(f)$.*

*Proof.* Let $f_i'$ denote the flow obtained by player $i$ making a lookahead best response. We first appeal to Lemma 4.2 of [5], which states that if $i$ changes its path from $P_i$ to $P_i'$ to give a new flow $f_i'$, then $l(f_i') \leq l(f) + 2l_i(f_i') - l_i(f)$ (when delays are linear). It follows from Lemma 3 that $l(f_i') \leq l(f) + 3l_i(f)$. Thus,

$$\begin{aligned}
\mathbb{E}[l(f') \mid f] &= \frac{1}{n}\sum_i l(f_i') \\
&\leq \frac{1}{n}\sum_i (l(f) + 3l_i(f)) \\
&= l(f) + \frac{1}{n}\sum_i 3l_i(f) \\
&= l(f) + \frac{3}{n}l(f),
\end{aligned}$$

□

We prove now our final lemma:

**Lemma 5.** *Let $f$ be the current flow. Choose a player at random to make a lookahead best response, resulting in flow $f'$. Then we have either $\mathbb{E}[l(f') \mid f] \leq (1 - \frac{1}{2n})l(f)$ or $l(f) < (2 + 2\sqrt{3})\mathsf{OPT}$.*

*Proof.* Let $f_i'$ denote the flow which occurs when $i$ moves to a lookahead best response. We consider the following cases:

*Case 1:* $4\sum_i l_i(f_i') \le l(f)$.

$$\begin{aligned}
\mathbb{E}\left[l(f') \mid f\right] &= \frac{1}{n}\sum_i l(f_i') \\
&\le \frac{1}{n}\sum_i (l(f) + 2l_i(f_i') - l_i(f)) && \text{(Lemma 4.2 of [5])} \\
&\le \frac{1}{n}\sum_i (l(f) + \frac{1}{2}l_i(f) - l_i(f)) \\
&= (1 - \frac{1}{2n})l(f)
\end{aligned}$$

*Case 2:* $4\sum_i l_i(f_i') > l(f)$.

Let $f^*$ denote an optimal flow and let $P_i^*$ denote player $i$'s path in this flow. Also let $J^*(e)$ be the set of players on edge $e$ in $f^*$. Since $P_i'$ is a lookahead best response, we must have $\sum_{e\in P_i^*} a_e(f_e+2) + b_e \ge l_i(f_i')$ (this is by the same reasoning as in Theorem 2). Thus,

$$\begin{aligned}
l(f) &< 4\sum_i \sum_{e\in P_i^*} a_e(f_e+2) + b_e \\
&= 4\sum_{e\in E}\sum_{i\in J^*(e)} a_e(f_e+2) + b_e \\
&= 4\sum_{e\in E}(a_e(f_e+2) + b_e)f^*(e) \\
&\le 4\sum_{e\in E} a_e f_e f_e^* + 8\sum_{e\in E}(a_e f_e^* + b_e f_e^*)
\end{aligned}$$

Applying Cauchy-Schwarz to the first term and noting again that ${f_e^*}^2 \ge f_e^*$ for the second:

$$\begin{aligned}
&\le 4\sqrt{\sum_{e\in E} a_e f_e^2}\sqrt{\sum_{e\in E} a_e {f_e^*}^2} + 8\sum_{e\in E}(a_e f_e^* + b_e)f_e^* \\
&\le 4\sqrt{l(f)l(f^*)} + 8l(f^*)
\end{aligned}$$

If we take $\rho = \sqrt{\frac{l(f)}{l(f^*)}}$, we may rewrite this as $\rho^2 < 4\rho + 8$. This gives a bound of $l(f) < (2+2\sqrt{3})\mathsf{OPT}$. □

**Theorem 6.** *In the average-case 2-lookahead leaf model, the price of anarchy for lookahead dynamics by risk-averse players is bounded by a constant. Starting from a flow with latency $C$, the expected latency after $O(n\log\frac{C}{\mathsf{OPT}})$ random lookahead best responses is at most* 27.5 OPT.

*Proof.* With lemmas 4 and 5 in hand, the rest of the proof follows as in Theorem 4.5 of [5]. We omit the details of the proof as they are identical, up to the values of the constants in our lemmas. We find instead that after $j \ge n\log\frac{1}{\epsilon}\log\frac{C}{\mathsf{OPT}}$ steps, the expected social cost is at most $(27.32+\epsilon)\mathsf{OPT}$. □

# 4 Discussion

We note some limitations and possible extensions of our work. One fairly heavy restriction is that throughout our work, we assumed that the mechanisms of all players were common knowledge. It may be interesting to consider the effect of uncertainty in players' decisions, as well as situations with more heterogeneous players, as most of our examples dealt exclusively with risk-averse players.

There are also more nuanced approaches to model risk-averse players. For example, our model of a risk-averse player is one who assumes that all other players are acting in their worst interest, with little regard to their own utility (e.g. in Cournot competition, the adversarial company is expected to move so that they end up with 0 profit). It may be that a more realistic model would have other players acting against them, but not in such a way so as to cause themselves more harm than they would the target. Additionally, one could implement a model where a risk-averse player places slightly more faith in other players to not play antagonistically, and place valuations on the game tree such that nodes with more high-value children are valued more highly than those with fewer.

For atomic selfish routing, we note that similar analysis may be applied if we were to consider the path model instead of the leaf model. Additionally, Theorem 2 would hold with the same proof were we to consider the worst-case move order rather than average case. However, the proof for Theorem 6 would fail if we were to examine the worst-case model rather than average case; in particular, Lemma 4 and Case 1 of Lemma 5 rely heavily on this assumption. It is not clear whether there exists a weakening of these lemmas which hold for the worst-case move order model, yet are still strong enough to prove an analogue of Theorem 6.

# References


[1] V. S. Mirrokni, N. Thain, and A. Vetta, "A theoretical examination of practical game playing: Lookahead search," in *SAGT*, pp. 251–262, 2012.

[2] A. Cournot, *Recherces sur les Principes Mathématiques de la Théorie des Richesse*. 1838.

[3] D. Besanko, U. Doraszelski, and Y. Kryukov, "The economics of predation: What drives pricing when there is learning-by-doing?," *American Economic Review*, vol. 104, pp. 868–97, March 2014.

[4] B. Awerbuch, Y. Azar, and A. Epstein, "The price of routing unsplittable flow," in *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005* (H. N. Gabow and R. Fagin, eds.), pp. 57–66, ACM, 2005.

[5] M. Goemans, V. Mirrokni, and A. Vetta, "Sink equilibria and convergence," vol. 2005, pp. 142– 151, 11 2005.