ON SEMI-RESTRICTED ROCK, PAPER, SCISSORS

SVANTE JANSON

Abstract. Spiro, Surya and Zeng (Electron. J. Combin. 2023) recently studied a semi-restricted variant of the well-known game Rock, Paper, Scissors; in this variant the game is played for $3n$ rounds, but one of the two players is restricted and has to use each of the three moves exactly n times. They show that the optimal strategy for the restricted player is the greedy strategy, and show that it results in an expected score for the unrestricted player $\Theta(\sqrt{n})$; they conjecture, based on numerical evidence, that the expectation is $\approx 1.46\sqrt{n}$. We analyse the result of the strategy further and show that the average is $\sim c\sqrt{n}$ with $c = 3\sqrt{3}/2\sqrt{\pi} \doteq 1.466$, verifying the conjecture.

The proof is based on considering the case when both players play greedily, which leads to the same expectation as optimal play; for this case we also find the asymptotic distribution of the score, and compute its variance.

1. Introduction

A semi-restricted variant of the well-known game Rock, Paper, Scissors (RPS) was recently studied by Spiro, Surya and Zeng [6]. In the standard version of RPS, two players simultaneously select one of the three choices *rock*, *paper*, *scissors*, where *paper* beats *rock*, *scissors* beats *paper*, and *rock* beats *scissors*; if both select the same, the result is a draw. The game is symmetric, so there is obviously no advantage to any of the players. It is easy to see that the optimal strategy for both players is to choose randomly, with equal probability for each choice (see further Section 2.2).

In the semi-restricted variant in [6], two players R (restricted) and N (normal) agree to play $3n$ rounds of RPS for some integer n, but R is restricted to choose *rock*, *paper*, and *scissors* exactly n times each, while N plays without restriction. Clearly, the restriction is a disadvantage for R. (In particular, N will always win the last round, since R then has only one choice, and N knows which one.) How large is this disadvantage? More precisely, let S_n be the final score of N, defined as the number of rounds won by N minus the number lost. We assume (as [6]) that the objective of both players is the expectation $\mathbb{E}\, S_n$, which N wants as high as possible, while R wants the opposite. Semi-restricted RPS is a two-player zero-sum game, and thus by the theory of von Neumann [5], each player has an optimal randomized strategy, see further e.g. [4, Chapter 2]. We use S_n^{op} to denote the final score when both players use their optimal strategies. (This is a random variable, since the strategies are randomized.)

The main result of [6] is that the unique optimal strategy for R is to play greedily, i.e., as if each round were the last; see further Section 2.2. (This is far from obvious, and rather surprising.) It is also shown in [6] that with optimal strategies, the expected gain $\mathbb{E}\, S_n^{\mathrm{op}} = \Theta(\sqrt{n})$, and it is asked [6, Question 21] whether $\mathbb{E}\, S_n^{\mathrm{op}} \sim c\sqrt{n}$

Date: 22 February, 2024; revised 26 March 2024.
Supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

for some constant $c > 0$ as $n \to \infty$; [6] says further that numerical calculations for $n \leqslant 100$ suggest that this might hold with $c \approx 1.46$.

The main purpose of the present note is to verify this conjecture, and to identify the constant.

Theorem 1.1. *For semi-restricted RPS played over $3n$ rounds, the expected score for N with optimal plays for both players is, as $n \to \infty$,*

$$\mathbb{E}\, S_n^{\mathrm{op}} \sim \sqrt{\frac{27n}{4\pi}} = \frac{3\sqrt{3}}{2\sqrt{\pi}}\sqrt{n}. \tag{1.1}$$

The constant $3\sqrt{3}/(2\sqrt{\pi}) \doteq 1.4658$, which verifies also the numerical conjecture in [6].

The optimal strategy for R is thus the greedy strategy. Given that R uses this strategy, there are many strategies for N that give the optimal expectation $\mathbb{E}\, S_n^{\mathrm{op}}$. One of them is the greedy strategy for N, but as pointed out to me by Sam Spiro [personal communication], the greedy strategy is *not* the optimal strategy for N; see Section 2.3. We let S_n^{gr} denote S_n when *both* players play with their greedy strategies. This is also a random variable, and as just said, we have

$$\mathbb{E}\, S_n^{\mathrm{gr}} = \mathbb{E}\, S_n^{\mathrm{op}}. \tag{1.2}$$

The random variable S_n^{gr} can be analysed asymptotically using standard tools from probability theory. This is done in Sections 3 and 4 and yields the asymptotics of $\mathbb{E}\, S_n^{\mathrm{gr}}$; Theorem 1.1 then follows by (1.2). Moreover, our analysis also yields the asymptotic distribution of S_n^{gr}, see Theorem 4.1.

In Section 5 we give some partial results on the asymptotic distribution of S_n if R uses the optimal (greedy) strategy and N uses a rather arbitrary strategy, including the case S_n^{op} when both play optimally. We leave as an open problem whether S_n^{op} and S_n^{gr} have the same asymptotic distribution.

In Section 6, we discuss the probability that the disadvantaged player R nevertheless wins the game; we compute it for the case that both players play greedily, but leave the case of optimal play for the objective of maximizing the probability of winning as an open problem.

Acknowledgement. I am grateful to Sam Spiro for pointing out a serious error in a previous version, and for showing me Example 2.1. I also thank an anonymous referee for helpful comments.

2. Preliminaries

2.1. **Notation.** The three choices *rock*, *paper*, *scissors* will be numbered $1, 2, 3$; thus $i + 1$ beats $i \pmod 3$.

The random variable $S(t)$ is the score of N after round $t = 1, \dots, 3n$, i.e., the number of rounds won by N so far minus the number of rounds won by R. As in the introduction, $S_n := S(3n)$ is the score at the end of the game. (Except for S_n, we do not show n explicitly in the notation, although $S(t)$ and many variables introduced below depend on n.)

If X_n is a sequence of random variables, and a_n a sequence of (positive) numbers, we write $X_n = O_{\mathrm{p}}(a_n)$ if the family $\{X_n/a_n\}$ is bounded in probability (also called *tight*), i.e., if for every $\varepsilon > 0$ there exists C such that $\mathbb{P}(|X_n| > Ca_n) < \varepsilon$ for all n. Furthermore, we write $X_n = O_{L^p}(a_n)$ (where $p > 0$ is a parameter) if the family $\{X_n/a_n\}$ is bounded in L^p, i.e., $\sup_n \mathbb{E}\, |X_n/a_n|^p < \infty$.

$N(0,\sigma^2)$ denotes the normal distribution with mean 0 and variance $\sigma^2 \geqslant 0$. More generally, if Σ is a symmetric positive semidefinite $d \times d$ matrix, then $N(0,\Sigma)$ is the normal distribution with mean 0 and covariance matrix Σ; this is a distribution of a random vector in $\mathbb{R}^d$.

The basis vectors in $\mathbb{R}^3$ are denoted $\mathbf{e}_1 := (1,0,0)$, $\mathbf{e}_2 := (0,1,0)$, $\mathbf{e}_3 := (0,0,1)$.

We use C_p, C_p', C_p'' for some constants that depend on the parameter p.

Unspecified limits are as $n \to \infty$.

2.2. **The greedy strategy.** Recall that in any two-person zero-sum game, each player has an optimal strategy which in general is randomized; the different alternatives are selected with some probabilities chosen such that they maximize the minimum over all strategies of the opponent of the expected gain; see [5] and e.g. [4].

As said above, it was shown by Spiro, Surya and Zeng [6] that in semi-restricted RPS, the best strategy of R is to play greedily, i.e., to analyse each round separately and use the optimal strategy for the expected score in that round. (This is far from obvious, since the best play in one specific round may be punished by lower expected score in later rounds; nevertheless, [6] shows that the expected later gains by any alternative strategy are offset by the immediate expected loss.) This optimal strategy for a single round is easy to find (as was done in [6]):

(i) If R still has all three choices available, then the optimal strategy is (obviously, by symmetry), to choose one of them randomly, with probability 1/3 each. And the best strategy for N is the same. (This game was one of the examples in the original paper by von Neumann [5].) The outcome for N is -1, 0, or $+1$ with probability 1/3 each.

(ii) If R has only two choices available, say 1 (*rock*) and 2 (*paper*), then the game is described by the matrix in Figure 1. N should never play 1 (which in this case can lose but never win). A simple calculation shows [6] that the best strategy for R is to play 1 with probability 1/3 and 2 with probability 2/3; similarly N plays 2 with probability 2/3 and 3 with probability 1/3. The expected gain for N is 1/3.

	rock	*paper*	*scissors*
rock	0	1	−1
paper	−1	0	1

FIGURE 1. Score matrix for N when R is restricted to $\{\mathit{rock}, \mathit{paper}\}$; rows show the move by R; columns the move by N.

(iii) If R has only one choice, then R has to play that, and N obviously plays the next choice (mod 3) and is sure to win. Gain for N is 1.

2.3. **Strategies for N.** Suppose that R plays optimally, i.e., greedily. Then R plays each time with a random move that depends only on the available moves, and thus on the history of the moves made by R. However, these moves are not affected by the moves made by N. Hence, the moves made by R will be the same regardless of the strategy chosen by N. It thus follows from the discussion above of the greedy strategy that the expected gain for N will be the same for any strategy of N that does not do anything stupid (here and in the sequel meaning making a move that cannot win); for example, as long as R is able to make all three moves, the expected gain of each round is 0 for any strategy of N. In particular, the expected gain for N when both players use their optimal strategies is the same as when both play greedily, which shows (1.2). Nevertheless, the greedy strategy is not optimal for N, since it may be worse if R chooses a different strategy as shown by the following simple example.

Example 2.1. (Sam Spiro, personal communication.) Suppose that N plays with the greedy strategy described above. If R chooses to play (deterministically) $1, 2, 3, 1, 2, 3, \dots$ for all $3n$ rounds, then for all but the last two rounds, the greedy strategy by N makes him play randomly, with probability $1/3$ for each choice, and therefore the expected gain is 0 for each round. Hence the total gain $\mathbb{E}\, S_n$ will in this case be only $1/3 + 1 = 4/3$ (from the last two rounds), while we know from von Neumann's theorem [5] that N has some strategy guaranteeing an expected gain of at least $\mathbb{E}\, S_n^{\mathrm{op}}$ against every strategy of R. (Note that $\mathbb{E}\, S_n^{\mathrm{op}} > 4/3$ at least for large n by Theorem 1.1. In fact, it is can easily be seen from (3.3) below that the inequality holds for every $n \geqslant 2$.) $\triangle$

It seems likely that the optimal strategy of N is very complicated. See further Section 5.1.

3. Analysis for the greedy strategies

In this section we assume that R uses the greedy strategy, which is known to be optimal. For simplicity, we assume here that also N uses the greedy strategy. In fact, most of the analysis is valid for almost any strategy by N; we discuss the few but important differences in Section 5.

Let $N_{t,i}$ be the number of times that R plays i during rounds $1, \dots, t$. The vector $\mathbf{N}_t = (N_{t,i})_{i=1}^3$ then evolves as a random walk which changes character each time some $N_{t,i}$ hits n and R thus cannot choose i in the future. We let T_j, $j = 1, 2, 3$, be the first time that R has used up j of the three choices; in particular, $T_3 := 3n$, when the game ends.

Since R uses the greedy strategy described above, $\mathbf{N}_t$ evolves as follows, for $t = 0, \dots, n$, starting at $\mathbf{N}_0 = (0, 0, 0)$:

I. A random walk $\mathbf{N}_0, \dots, \mathbf{N}_{T_1}$ with increments that are independent and uniformly chosen from $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, until
$$T_1 := \inf\bigl\{t : N_{t,i} = n \text{ for some } i \in \{1, 2, 3\}\bigr\}. \tag{3.1}$$

II. A random walk $\mathbf{N}_{T_1}, \dots, \mathbf{N}_{T_2}$ with increments chosen independently and randomly from the remaining two choices by the strategy above; for example, if $N_{t,1}$ hits n first, so $N_{T_1,1} = n > N_{T_1,2}, N_{T_1,3}$, then the increments are chosen as $\mathbf{e}_2$ and $\mathbf{e}_3$ with probabilities $1/3$ and $2/3$. This goes on until
$$T_2 := \inf\bigl\{t : N_{t,i} = n \text{ for at least two } i \in \{1, 2, 3\}\bigr\}. \tag{3.2}$$

III. A deterministic walk $\mathbf{N}_{T_2}, \dots, \mathbf{N}_{T_3}$ where all increments are $\mathbf{e}_i$ for the only i that still has $N_{t,i} < n$.

The expected gain for N is 0 for each step in phase I, $1/3$ for each step in phase II, and 1 for each step in phase III, so the expected score for N is
$$\mathbb{E}\, S_n = \mathbb{E}\, S(3n) = \mathbb{E}\, [(T_2 - T_1)/3 + T_3 - T_2]. \tag{3.3}$$

We will analyse this more carefully below and also both bound and asymptotically describe the random fluctuations. We do this by analysing the constrained random walk $\mathbf{N}_t$ and the stopping times T_1 and T_2 in some detail. A central role in the analysis is played by the (somewhat arbitrary) non-random time

$$T_0 := 3n - 3\lceil n^{2/3} \rceil. \tag{3.4}$$

3.1. **Phase I: until T_1.** Let $(\boldsymbol{\xi}_t)_{t=1}^{\infty}$ be an i.i.d. sequence of random vectors with the distribution $\mathbb{P}(\boldsymbol{\xi}_t = \mathbf{e}_i) = 1/3$ for $i = 1, 2, 3$. We may assume that $\mathbf{N}_t - \mathbf{N}_{t-1} = \boldsymbol{\xi}_t$ for $1 \leqslant t \leqslant T_1$. Let

$$\mathbf{N}'_t = (N'_{t,i})_{i=1}^3 := \sum_{u=1}^{t} \boldsymbol{\xi}_u, \qquad t \geqslant 0; \tag{3.5}$$

thus $\mathbf{N}'_t = \mathbf{N}_t$ for $t \leqslant T_1$. (We may interpret $\boldsymbol{\xi}_t$ and $\mathbf{N}'_t$ as how R would have played if the restriction had not existed.) In particular, for $t \leqslant T_1$ we have $N'_{t,i} = N_{t,i} \leqslant n$ for all i, and for $t \geqslant T_1$ we have $\max_i N'_{t,i} \geqslant \max_i N_{T_1,i} = n$; thus T_1 is also the time that $\max_i N'_{t,i}$ hits n.

At time T_0, the central limit theorem shows that

$$N'_{T_0,i} = \tfrac{1}{3} T_0 + O_{\mathrm{p}}(n^{1/2}) = n - n^{2/3} + O_{\mathrm{p}}\bigl(n^{1/2}\bigr). \tag{3.6}$$

This is less that n for each i w.h.p. (with high probability, i.e., with probability $1 - o(1)$ as $n \to \infty$), and thus w.h.p. $T_1 > T_0$. More precisely, the Chernoff inequality (e.g. in the version in [2, Remark 2.5]) yields

$$\mathbb{P}(T_1 \leqslant T_0) \leqslant \sum_{i=1}^{3} \mathbb{P}(N'_{T_0,i} \geqslant n) = 3\,\mathbb{P}\bigl(N'_{T_0,1} - \tfrac{1}{3} T_0 \geqslant \lceil n^{2/3} \rceil\bigr) \leqslant e^{-2n^{4/3}/T_0} \leqslant e^{-n^{1/3}}. \tag{3.7}$$

Hence, this probability decreases faster than any polynomial, which means that we can ignore the event $T_1 \leqslant T_0$ also when calculating moments below (since the random variables we consider all are deterministically $O(n)$).

Similarly, concentrating on the time after T_0, define

$$M' := \max_{i=1,2,3} \max_{T_0 \leqslant t \leqslant 3n} \bigl| N'_{t,i} - N'_{T_0,i} - \tfrac{1}{3}(t - T_0) \bigr|. \tag{3.8}$$

By classical results on moment convergence in the central limit theorem together with Doob's inequality (since $N'_{t,i} - N'_{T_0,i} - \frac{1}{3}(t - T_0)$ is a martingale), see for example [1, Theorem 7.5.1, Corollary 3.8.2, and Theorem 10.9.4], we have, for any $p > 1$

$$\begin{aligned}
\mathbb{E}\,(M')^p &\leqslant \sum_{i=1}^{3} \mathbb{E} \max_{T_0 \leqslant t \leqslant 3n} \bigl| N'_{t,i} - N'_{T_0,i} - \tfrac{1}{3}(t - T_0) \bigr|^p \\
&\leqslant C_p \sum_{i=1}^{3} \mathbb{E} \bigl| N'_{3n,i} - N'_{T_0,i} - \tfrac{1}{3}(3n - T_0) \bigr|^p \\
&\leqslant C'_p (3n - T_0)^{p/2} \leqslant C''_p n^{p/3}.
\end{aligned} \tag{3.9}$$

Consequently,

$$M' = O_{L^p}\bigl(n^{1/3}\bigr) \tag{3.10}$$

for every $p < \infty$. (The case $p \leqslant 1$ follows from the case $p > 1$ by Lyapounov's inequality [1, Theorem 3.2.5].)

We introduce some further notation. Let, for $i = 1, 2, 3$,

$$X_i := N'_{T_0,i} - \mathbb{E}\, N'_{T_0,i} = N'_{T_0,i} - \tfrac{1}{3} T_0. \tag{3.11}$$

(If we ignore the minor technical difference between $N_{t,i}$ and $N'_{t,i}$, these measure thus the deviation from the expectation at time T_0 of the choices made by R.) Note for later use that

$$X_1 + X_2 + X_3 = \sum_{i=1}^{3} N'_{T_0,i} - T_0 = 0. \tag{3.12}$$

Furthermore, let

$$X_{\max} := \max_{i=1,2,3} X_i. \tag{3.13}$$

(As we will see in detail below, this largest deviation will give us a good estimate of the time T_1 when R runs out of one choice.)

Condition on the event $T_1 > T_0$, which has probability $1-o(1)$. Then, $N'_{T_0} = N_{T_0}$. Moreover, we may take $t = T_1$ in (3.8) and obtain, using (3.11),

$$N_{T_1,i} = N'_{T_1,i} = N'_{T_0,i} + \tfrac{1}{3}(T_1 - T_0) + O(M') = X_i + \tfrac{1}{3}T_1 + O(M'). \tag{3.14}$$

Hence, recalling the definitions of T_1 and $X_{\max}$,

$$n = \max_i N_{T_1,i} = \max_i X_i + \tfrac{1}{3}T_1 + O(M') = X_{\max} + \tfrac{1}{3}T_1 + O(M'). \tag{3.15}$$

Consequently,

$$T_1 = 3n - 3X_{\max} + O(M') \tag{3.16}$$

and thus, using (3.10),

$$T_1 = 3n - 3X_{\max} + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.17}$$

This was derived conditioned on $T_1 > T_0$, but by (3.7) and the comment after it, (3.17) holds also unconditionally.

Furthermore, for every $i \in \{1,2,3\}$, by (3.14) and (3.10),

$$N_{T_1,i} - \tfrac{1}{3}T_1 = X_i + O_{L^p}\bigl(n^{1/3}\bigr) \tag{3.18}$$

and thus by (3.17)

$$n - N_{T_1,i} = n - \tfrac{1}{3}T_1 - X_i + O_{L^p}\bigl(n^{1/3}\bigr) = X_{\max} - X_i + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.19}$$

Thus, at time T_1, when R runs out of one of the three choices, she has approximatively $X_{\max} - X_i$ left of each other choice i.

To find the score in Phase I, consider first the score at T_0, and condition again on $T_1 > T_0$. Then in each round up to T_0, R plays normally and thus R and N win with probability 1/3 each, and draw otherwise; thus $\Delta S(t) := S(t) - S(t-1) \in \{\pm1, 0\}$ with probability 1/3 each. Consequently, the central limit theorem shows that, since $\mathbb{E}\,\Delta S(t) = 0$ and $\operatorname{Var}\Delta S(t) = 2/3$, and $T_0 \sim 3n$,

$$\frac{S(T_0)}{n^{1/2}} \xrightarrow{\mathrm{d}} N(0,2), \qquad \text{as } n \to \infty, \tag{3.20}$$

together with all moments. Moreover, since also N is assumed to use the optimal strategy, which for these t means uniformly randomly, the score in each round is independent of the choices made by R, and thus of the vectors $\mathbf{N}_t$. Consequently, $S(T_0)$ is independent of (X_1, X_2, X_3). We conditioned here on $T_1 > T_0$, but in the unlikely event $T_1 \leqslant T_0$, we may modify $S(T_0)$ (similarly as we defined $\mathbf{N}'$ above) and define a sum $S'(T_0)$ that is independent of (X_1, X_2, X_3) and satisfies $S'(T_0) = S(T_0)$ whenever $T_1 > T_0$, and thus, by (3.7), (rather coarsely)

$$S(T_0) = S'(T_0) + O_{L^p}\big(n^{1/3}\big). \tag{3.21}$$

For $T_0 < t \leqslant T_1$, we still have the same distribution of $\Delta S(t)$, and by the same argument as in (3.9), if we condition on $T_1 > T_0$, then

$$S(T_1) - S(T_0) = O_{L^p}\big(n^{1/3}\big). \tag{3.22}$$

By (3.7), this holds also unconditionally.

3.2. **Phase II: T_1 to T_2.** Since the entire game is symmetric under cyclic permutations of the three choices *rock*, *paper*, *scissors*, we may for the next phase assume that R first uses up all n *rock*, i.e., that $N_{T_1,1} = 0$. Note, however, that the game is not symmetric under odd permutations, so having made this assumption, choices 2 (*paper*) and 3 (*scissors*) play different roles, since 3 beats 2.

By the discussion of the greedy strategy in Section 2.2, for $t \in [T_1, T_2)$, R should play randomly and choose 2 or 3 with probabilities 1/3 and 2/3. We argue as in the preceding subsection (and therefore omit some details); we now let $(\boldsymbol{\eta}_t)_1^\infty$ be an i.i.d. sequence of random vectors with $\mathbb{P}(\boldsymbol{\eta}_t = \mathbf{e}_i) = p_i$ for $i = 1, 2, 3$, with $(p_1, p_2, p_3) = (0, \frac{1}{3}, \frac{2}{3})$, and we assume as we may that $\mathbf{N}_t - \mathbf{N}_{t-1} = \boldsymbol{\eta}_t$ for $T_1 < t \leqslant T_2$. Let

$$\mathbf{N}''_t = (N''_{t,i})_{i=1}^3 := \mathbf{N}_{T_1} + \sum_{u=T_1+1}^{t} \boldsymbol{\eta}_u, \qquad t \geqslant T_1. \tag{3.23}$$

Then $\mathbf{N}''_t = \mathbf{N}_t$ for $T_1 \leqslant t \leqslant T_2$. Let

$$M'' := \max_{i=1,2,3} \max_{T_1 \leqslant t \leqslant 3n} \big|N''_{t,i} - N''_{T_1,i} - p_i(t - T_1)\big|. \tag{3.24}$$

If we again condition on $T_1 > T_0$, we obtain, by conditioning on T_1 and arguing as in (3.9) and using $3n - T_1 < 3n - T_0 = O(n^{2/3})$,

$$M'' = O_{L^p}\big(n^{1/3}\big). \tag{3.25}$$

By (3.7) again, this holds also unconditionally. We obtain from (3.24) and (3.25), taking $t = T_2$, for every i,

$$N_{T_2,i} = N''_{T_2,i} = N_{T_1,i} + p_i(T_2 - T_1) + O_{L^p}\big(n^{1/3}\big) \tag{3.26}$$

and thus, by (3.19),

$$\begin{aligned} n - N_{T_2,i} &= n - N_{T_1,i} - p_i(T_2 - T_1) + O_{L^p}\big(n^{1/3}\big) \\ &= X_{\max} - X_i - p_i(T_2 - T_1) + O_{L^p}\big(n^{1/3}\big). \end{aligned} \tag{3.27}$$

We have assumed $N_{T_1,1} = n$, and then T_2 is the first t such that $N_{t,2} = n$ or $N_{t,3} = n$. In particular, (3.27) implies

$$0 = \min_{i=2,3}\big(n - N_{T_2,i}\big) = \min_{i=2,3}\big(X_{\max} - X_i - p_i(T_2 - T_1)\big) + O_{L^p}\big(n^{1/3}\big). \tag{3.28}$$

Consequently,

$$\min_{i=2,3}\big(X_{\max} - X_i - p_i(T_2 - T_1)\big) = O_{L^p}\big(n^{1/3}\big). \tag{3.29}$$

It follows that also

$$\min_{i=2,3} p_i^{-1}\big(X_{\max} - X_i - p_i(T_2 - T_1)\big) = O_{L^p}\big(n^{1/3}\big), \tag{3.30}$$

which can be written

$$\min_{i=2,3} p_i^{-1}\bigl(X_{\max} - X_i\bigr) - (T_2 - T_1) = O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.31}$$

Thus

$$T_2 - T_1 = \min_{i=2,3} \frac{X_{\max} - X_i}{p_i} + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.32}$$

We repeat that this holds assuming that choice 1 is the first to be used up by R.

In this phase, the gain $\Delta S(t)$ of N has expectation 1/3 in each round (and its absolute value is bounded by 1, so all moments are bounded); moreover, the gains in different rounds are i.i.d. Hence, similarly to (3.9) again, the central limit theorem with moment convergence together with Doob's inequality yields

$$S(T_2) - S(T_1) = \tfrac{1}{3}(T_2 - T_1) + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.33}$$

3.3. **Phase III: T_2 to T_3.** This phase is deterministic, and not very fun to play (at least not for R): R has only one choice, and N wins every round. The total gain for N in this phase are thus, using (3.17) and recalling that $T_3 = 3n$,

$$\begin{aligned} S(T_3) - S(T_2) &= T_3 - T_2 = T_3 - T_1 - (T_2 - T_1) \\ &= 3X_{\max} - (T_2 - T_1) + O_{L^p}\bigl(n^{1/3}\bigr). \end{aligned} \tag{3.34}$$

3.4. **Collecting the gains.** By (3.22), (3.33), and (3.34), the final score of N is

$$S_n = S(T_3) = S(T_0) + 3X_{\max} - \tfrac{2}{3}(T_2 - T_1) + O_{L^p}\bigl(n^{1/3}\bigr), \tag{3.35}$$

where furthermore $T_2 - T_1$ is given by (3.32) when choice 1 (*rock*) is the first to be used up by R. We develop (3.35) as follows.

Lemma 3.1. *We have*

$$S_n - S(T_0) = \max\bigl\{X_1 + 2X_2, X_2 + 2X_3, X_3 + 2X_1\bigr\} + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.36}$$

Proof. We may again, by symmetry, suppose that R first uses up 1. Typically, this is the case when $X_{\max} = X_1$, but it is possible that X_1 is not the maximum. (Then $N_{t,1}$ is not the largest at $t = T_0$, but $N_{t,1}$ overtakes the other two components and hits n first.) In any case, $N_{T_2,1} = N_{T_1,1} = n$, and thus (3.27) yields, recalling $p_1 = 0$,

$$0 = n - N_{T_2,1} = X_{\max} - X_1 + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.37}$$

Hence,

$$X_{\max} = X_1 + O_{L^p}\bigl(n^{1/3}\bigr). \tag{3.38}$$

We obtain from (3.35), (3.32) and (3.38), recalling $p_2 = \frac{1}{3}$ and $p_3 = \frac{2}{3}$,

$$\begin{aligned} S_n - S(T_0) &= 3X_{\max} - \tfrac{2}{3}(T_2 - T_1) + O_{L^p}\bigl(n^{1/3}\bigr) \\ &= 3X_{\max} - \min\bigl\{2(X_{\max} - X_2), (X_{\max} - X_3)\bigr\} + O_{L^p}\bigl(n^{1/3}\bigr) \\ &= \max\bigl\{X_{\max} + 2X_2, 2X_{\max} + X_3\bigr\} + O_{L^p}\bigl(n^{1/3}\bigr) \\ &= \max\bigl\{X_1 + 2X_2, 2X_1 + X_3\bigr\} + O_{L^p}\bigl(n^{1/3}\bigr). \end{aligned} \tag{3.39}$$

Furthermore, (3.12) implies that $X_{\max} \geqslant 0$ and that, using also (3.38),

$$\begin{aligned} 2X_1 + X_3 &= 3X_1 + X_2 + 2X_3 = 3X_{\max} + X_2 + 2X_3 + O_{L^p}\bigl(n^{1/3}\bigr) \\ &\geqslant X_2 + 2X_3 + O_{L^p}\bigl(n^{1/3}\bigr). \end{aligned} \tag{3.40}$$

Hence (3.39) yields (3.36) in the case when R first uses up choice 1. By symmetry (3.36) holds in general. □

We may now summarize the analysis in the following limit result.

Theorem 3.2. *As $n \to \infty$, we have convergence in distribution, together with all moments,*

$$n^{-1/2} S_n \xrightarrow{\mathrm{d}} \mathcal{S}^{\mathrm{gr}} := W + \max\bigl\{V_1 + 2V_2, V_2 + 2V_3, V_3 + 2V_1\bigr\}, \tag{3.41}$$

where W, V_1, V_2, V_3 are jointly normal with W independent of (V_1, V_2, V_3) and

$$W \in N(0, 2), \tag{3.42}$$

$$(V_1, V_2, V_3) \in N\left(0, \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\right). \tag{3.43}$$

Proof. The random vectors $\boldsymbol{\xi}_t$ in (3.5) are i.i.d. with $\mathbb{E}\,\boldsymbol{\xi}_t = 0$ and covariance matrix (regarding $\boldsymbol{\xi}_t$ as a column vector)

$$\mathrm{Var}(\boldsymbol{\xi}_t) := \mathbb{E}\,\boldsymbol{\xi}_t^{\mathsf{tr}} \boldsymbol{\xi}_t - (\mathbb{E}\,\boldsymbol{\xi}_t^{\mathsf{tr}})(\mathbb{E}\,\boldsymbol{\xi}_t) = \Sigma := \begin{pmatrix} \frac{2}{9} & -\frac{1}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{2}{9} \end{pmatrix}. \tag{3.44}$$

Since $T_0 \sim 3n$ by (3.4), the central limit theorem yields, recalling (3.11),

$$n^{-1/2}(X_1, X_2, X_3) \xrightarrow{\mathrm{d}} (V_1, V_2, V_3) \in N(0, 3\Sigma), \tag{3.45}$$

which agrees with (3.43). Similarly, as noted in (3.20), $n^{-1/2} S(T_0) \xrightarrow{\mathrm{d}} W$. Furthermore, by (3.21) we may here replace $S(T_0)$ be the approximation $S'(T_0)$ which, as noted above, is independent of (X_1, X_2, X_3). Hence,

$$n^{-1/2}\bigl(S(T_0), X_1, X_2, X_3\bigr) \xrightarrow{\mathrm{d}} (W, V_1, V_2, V_3), \tag{3.46}$$

and thus (3.36) and the continuous mapping theorem yield (3.41). Moreover, all moments converge in the central limit theorems (3.45) and (3.20) [1, Theorem 7.5.1], and it follows (e.g. using uniform integrability) that all moments converge also in (3.46) and (3.41). □

In the following section, we give more convenient expressions for the limit $\mathcal{S}^{\mathrm{gr}}$.

4. The distribution of the limit for greedy strategies

We give several alternative descriptions of the asymptotic distribution found in Theorem 3.2; using them we then prove Theorem 1.1. See also Section 6 for another use of these descriptions.

Theorem 4.1. *The limit $\mathcal{S}^{\mathrm{gr}}$ in Theorem 3.2 can be described by any of the following equivalent formulas:*

(i) *We have*

$$\mathcal{S}^{\mathrm{gr}} = W + \max\bigl\{Z_1, Z_2, Z_3\bigr\} \tag{4.1}$$

where W, Z_1, Z_2, Z_3 are jointly normal with W independent of (Z_1, Z_2, Z_3) and

$$W \in N(0, 2), \tag{4.2}$$

$$(Z_1, Z_2, Z_3) \in N\left(0, \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\right). \tag{4.3}$$

(ii) *We have*

$$\mathcal{S}^{\mathrm{gr}} = W' + \sqrt{3}\max\{Z_1', Z_2', Z_3'\} \tag{4.4}$$

where W', Z_1', Z_2', Z_3' are independent standard normal $N(0,1)$.

(iii) *We have*

$$\mathcal{S}^{\mathrm{gr}} = \max\{Z_1'', Z_2'', Z_3''\} \tag{4.5}$$

where Z_1'', Z_2'', Z_3'' are jointly normal with

$$(Z_1'', Z_2'', Z_3'') \in N\left(0, \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}\right). \tag{4.6}$$

(iv) *We have*

$$\mathcal{S}^{\mathrm{gr}} = W + R\cos\Theta, \tag{4.7}$$

where W, R, Θ are independent with $W \in N(0,2)$ as in (4.2)*, R has a Rayleigh distribution with density $\frac{1}{2}re^{-r^2/4}$, $r > 0$, and Θ has a uniform distribution $U(0, \pi/3)$.*

We will use the notation

$$Z_{\max} := \max\{Z_1, Z_2, Z_3\}. \tag{4.8}$$

Note also that (4.3) implies that $Z_1 + Z_2 + Z_3$ has variance 0, and thus the normal variables Z_1, Z_2, Z_3 in (4.1) satisfy $Z_1 + Z_2 + Z_3 = 0$ almost surely; thus (Z_1, Z_2, Z_3) lives in a 2-dimensional space.

Proof of Theorem 4.1. (i): Define

$$Z_1 := V_1 + 2V_2, \qquad Z_2 := V_2 + 2V_3, \qquad Z_3 = V_3 + 2V_1. \tag{4.9}$$

Then (3.41) shows that (4.1) holds, and a simple calculation shows that (Z_1, Z_2, Z_3) has the distribution (4.3).

(iii): Define $Z_i'' := W + Z_i$, $i = 1, 2, 3$. Then (4.1) yields (4.5), and (4.2)–(4.3) yield (4.6).

(ii): We may write $W = W' + \widetilde{W}$, where $W', \widetilde{W} \in N(0,1)$, and W' and $\widetilde{W}$ are independent of each other and of (Z_1, Z_2, Z_3). Define $Z_i' := (\widetilde{W} + Z_i)/\sqrt{3}$, $i = 1, 2, 3$. Then (4.1) yields (4.4), and it follows from (4.3) that the covariance matrix of (Z_1', Z_2', Z_3') is the identity matrix; thus the jointly normal variables W', Z_1', Z_2', Z_3' are independent $N(0,1)$.

(iv): As said above, $Z_1 + Z_2 + Z_3 = 0$ almost surely, so (Z_1, Z_2, Z_3) has really a 2-dimensional normal distribution. In fact, if $\boldsymbol{\zeta} = (\zeta_1, \zeta_2)$ is a centered normal distribution in $\mathbb{R}^2$ with $\operatorname{Var}\zeta_1 = \operatorname{Var}\zeta_2 = 2$ and $\operatorname{Cov}(\zeta_1, \zeta_2) = 0$, then we can construct (Z_1, Z_2, Z_3) with the desired distribution (4.3) by

$$Z_i := \mathbf{f}_i \cdot \boldsymbol{\zeta}, \tag{4.10}$$

where $\mathbf{f}_1 := (1, 0)$, $\mathbf{f}_2 := (-\frac{1}{2}, \frac{\sqrt{3}}{2})$, $\mathbf{f}_3 := (-\frac{1}{2}, -\frac{\sqrt{3}}{2})$. We define $R := |\boldsymbol{\zeta}|$ and $\Theta := \arg(\zeta_1 + \mathrm{i}\zeta_2) \in [-\pi, \pi)$; thus

$$\boldsymbol{\zeta} = (R\cos\Theta, R\sin\Theta), \tag{4.11}$$

and it follows from (4.10) by simple calculations (which are made even simpler by identifying $\mathbb{R}^2$ and $\mathbb{C}$ and regarding $\boldsymbol{\zeta}$ as a complex random variable) that

$$Z_1 = R\cos\Theta, \qquad Z_2 = R\cos(\Theta - 2\pi/3), \qquad Z_3 = R\cos(\Theta + 2\pi/3). \tag{4.12}$$

The normal distribution of $\boldsymbol{\zeta}$ is rotationally symmetric, and thus, as is well-known, R and Θ are independent, with Θ uniformly distributed on $[-\pi, \pi)$; furthermore, R has the Rayleigh distribution stated in the theorem. To find the distribution of $Z_{\max} := \max\{Z_1, Z_2, Z_3\}$, we may by symmetry condition on $Z_{\max} = Z_1$, which by (4.12) is equivalent to $\Theta \in [-\pi/3, \pi/3]$, and since $\cos\Theta$ is an even function, we may further restrict to $\Theta \in [0, \pi/3]$. Then $Z_{\max} = Z_1 = R\cos\Theta$, and thus (4.7) follows from (4.1) □

Proof of Theorem 1.1. By the moment convergence in Theorem 3.2, it suffices to find $\mathbb{E}\,\mathcal{S}^{\mathrm{gr}}$. For this we use (4.7). We have $\mathbb{E}\,W = 0$, and by simple calculations

$$\mathbb{E}\,R = \int_0^\infty \tfrac{1}{2} r^2 e^{-r^2/4}\,\mathrm{d}r = \sqrt{\pi}, \tag{4.13}$$

$$\mathbb{E}\,\cos\Theta = \frac{3}{\pi}\int_0^{\pi/3} \cos\vartheta\,\mathrm{d}\vartheta = \frac{3\sqrt{3}}{2\pi}. \tag{4.14}$$

Hence, by the independence,

$$\mathbb{E}\,\mathcal{S}^{\mathrm{gr}} = \mathbb{E}\,R \cdot \mathbb{E}\,\cos\Theta = \frac{3\sqrt{3}}{2\sqrt{\pi}} \doteq 1.4658075. \tag{4.15}$$

□

Remark 4.2. Alternatively, we can use (4.4) and conclude

$$\mathbb{E}\,\mathcal{S}^{\mathrm{gr}} = \sqrt{3}\,\mathbb{E}\,\max\bigl\{Z_1', Z_2', Z_3'\bigr\}, \tag{4.16}$$

where the right-hand side contains the expectation of the maximum of three i.i.d. standard normal variables which is known to be $3/(2\sqrt{\pi})$ [3]. △

Higher moments of $\mathcal{S}^{\mathrm{gr}}$ can be computed in the same way. For example, we have

$$\mathbb{E}\,(\mathcal{S}^{\mathrm{gr}})^2 = \mathbb{E}\,W^2 + \mathbb{E}\,R^2\,\mathbb{E}\,\cos^2\Theta = 2 + 4\Bigl(\frac{1}{2} + \frac{3\sqrt{3}}{8\pi}\Bigr) = 4 + \frac{3\sqrt{3}}{2\pi} \doteq 4.82699 \tag{4.17}$$

and hence

$$\operatorname{Var}\mathcal{S}^{\mathrm{gr}} = 4 - \frac{27 - 6\sqrt{3}}{4\pi} \doteq 2.67840. \tag{4.18}$$

Hence, we have

$$\operatorname{Var} S_n^{\mathrm{gr}} \sim \Bigl(4 - \frac{27 - 6\sqrt{3}}{4\pi}\Bigr) n. \tag{4.19}$$

5. Analysis when N does not play greedily

Assume as above that R uses the optimal strategy, i.e., the greedy strategy. In Section 3 we assumed that N uses the greedy strategy. More generally, suppose now that N uses any strategy that does not do anything stupid (a move that cannot win when R has only one or two choices). (This includes both the unknown optimal strategy for N, and the greedy strategy, but also many others.) Then, as noted in Section 2.3, the expected gain for N is still 1/3 in each round where R has two choices left, and 1 in each round where R has only one choice. Hence, (3.3) still holds. Moreover, the strategy of R is not affected by the moves made by N, and thus the random walk $\mathbf{N}_0, \dots, \mathbf{N}_{3n}$ and the variables $T_1, T_2, X_1, X_2, X_3, X_{\max}$ (and others) are the same as in Section 3. In particular, (3.45) still holds.

For the score of N, recall first that in Phase I, when R still has three choices, R plays each with the same probability. It follows that regardless of the strategy of N, the outcome $\Delta S(t)$ of each round has the same distribution as discussed in Section 3, i.e., 1, 0, or -1 with probability 1/3 each; moreover, this is independent of the previous history, so the outcomes of different rounds in this phase are independent. Consequently, $S(T_1)$ has the same distribution as for the greedy strategy, and so has $S(T_0)$ if we condition on $T_1 > T_0$. It follows that (3.19) still holds, and so do (3.20)–(3.21). However, there is one important difference from the case of the greedy strategy in Section 3: there the score $S(T_0)$ is independent of (X_1, X_2, X_3) (again conditioned on $T_1 > T_0$). This is no longer true in general, since the strategy of N may cause dependencies. We give a simple example showing that this actually may happen in Example 5.3.

In Phase II, R has two choices, and uses the greedy strategy described in Section 2.2(ii). We have assumed that the strategy of N is not stupid, and that leaves two choices for N. Both give an expected gain $\mathbb{E}\,\Delta S(t) = 1/3$, but the distributions are different. The precise distribution of $S(T_2) - S(T_1)$ may therefore depend on the strategy of N, but if we define $M_i := S(T_1 + i) - S(T_1) - \frac{1}{3}i$, then the sequence $(M_{i\wedge(T_2-T_1)})_{i\geqslant 1}$ (where we stop at $T_1 + i = T_2$) is, for any non-stupid strategy of N, a martingale with uniformly bounded increments, and Doob's inequality shows that (3.33) holds.

In Phase III, N has only one choice that is not stupid, so the strategy is the same is in Section 3, and (3.34) still holds.

It follows that (3.35) holds, and thus Lemma 3.1 holds, by the same proof as above. This leads to the following result.

Theorem 5.1. *Suppose that R uses the optimal (i.e., greedy) strategy, and that N uses any non-stupid strategy. (For example, his optimal strategy.) If we decompose*

$$n^{-1/2}S_n = n^{-1/2}S(T_0) + n^{-1/2}(S_n - S(T_0)), \tag{5.1}$$

then the two terms individually converge in distribution to the limits W and $Z_{\max}$ in (4.1)*; however, in general the two terms are dependent, so Theorems 3.2 and Theorem 4.1 do not hold.*

Note that it does not follow from Theorem 5.1 that $n^{-1/2}S_n$ converges in distribution. By general principles, the convergence in distribution implies that each of the sequences $n^{-1/2}S(T_0)$ and $n^{-1/2}(S_n - S(T_0)$ is tight, and thus so is their sum $n^{-1/2}S_n$; this implies that there are subsequences that converge in distribution, but it is conceivable that different subsequences have different limits. (This can easily happen if the strategy explicitly depends on, for example, whether n is even or odd, but it is not expected for "natural" strategies.)

Remark 5.2. In general, any (subsequential) limit in distribution $\mathcal{S}$ can be written as $W + Z_{\max}$ with W and $Z_{\max}$ as in Theorem 4.1, but possibly dependent. It follows from Minkowski's inequality and calculations as in (4.17)–(4.18) that, with the notation $[a \pm b] := [a - b, a + b]$,

$$(\operatorname{Var}\mathcal{S})^{1/2} \in \bigl[(\operatorname{Var}W)^{1/2} \pm (\operatorname{Var}Z_{\max})^{1/2}\bigr] = \Bigl[\sqrt{2} \pm \sqrt{2 - \frac{27 - 6\sqrt{3}}{4\pi}}\Bigr] \tag{5.2}$$

and thus, numerically, $(\operatorname{Var}\mathcal{S})^{1/2} \in [0.590\ldots, 2.237\ldots]$ and thus

$$\operatorname{Var}\mathcal{S} \in [0.348\ldots, 5.008\ldots]. \tag{5.3}$$

Since we have moment convergence by the same arguments as before, it follows that, for any non-stupid strategy for N, $\liminf n^{-1} \operatorname{Var} S_n$ and $\limsup n^{-1} \operatorname{Var} S_n$ lie in the interval (5.3). Furthermore, (5.3) shows that $\operatorname{Var}\mathcal{S} > 0$, so the limit distribution is non-degenerate. △

We give next a simple example showing that there are strategies for N for which $n^{-1/2}S_n$ has a limit in distribution that is different from $\mathcal{S}^{\mathrm{gr}}$; we then discuss briefly the optimal strategy.

Example 5.3. Let the strategy of N be to always play *rock* as long as R has three choices, and then switch to the greedy strategy for the endgame. (This is obviously a risky strategy if R would guess it, but we assume that R is a mathematician and knows that the greedy strategy is proven to be optimal, and therefore sticks to it.) We do not claim that this is a clever strategy, but it is not stupid in the sense above; thus the results above hold for it. Moreover, in Phase I, N wins when R plays *scissors*, and loses when R plays *paper*; hence $S(t) = N_{t,3} - N_{t,2}$ for all $t \leqslant T_1$. Consequently, assuming $T_1 > T_0$, we have

$$S(T_0) = N_{T_0,3} - N_{T_0,2} = X_3 - X_2. \tag{5.4}$$

It follows that (3.41) still holds, with (V_1, V_2, V_3) and (Z_1, Z_2, Z_3) as before and

$$W = V_3 - V_2 = -V_1 - 2V_2 = -Z_1. \tag{5.5}$$

Hence, instead of (3.41) and (4.1) we find

$$n^{-1/2}S_n \xrightarrow{\mathrm{d}} \mathcal{S} := W + Z_{\max} = -Z_1 + Z_{\max} = \max\{0, Z_2 - Z_1, Z_3 - Z_1\}. \tag{5.6}$$

Note that (3.42)–(3.43) and (4.2)–(4.3) still hold, but W and $Z_{\max}$ are no longer independent. To see that the dependence really matters and leads to a different limit distribution $\mathcal{S}$ than for the greedy strategy, we compute, using symmetry and the representation in Theorem 4.1(iv),

$$\begin{aligned}
\mathbb{E}\left[W^2 Z_{\max}\right] &= \mathbb{E}\left[Z_1^2 Z_{\max}\right] = \tfrac{1}{3}\,\mathbb{E}\left[Z_{\max}\sum_1^3 Z_i^2\right] \\
&= \tfrac{1}{3}\,\mathbb{E}\left[(R\cos\Theta)R^2\right] = \tfrac{1}{3}\,\mathbb{E}\,R^3\ \mathbb{E}\,\cos\Theta \\
&> \tfrac{1}{3}\,\mathbb{E}\,R^2 \cdot \mathbb{E}\,R\,\mathbb{E}\,\cos\Theta = \mathbb{E}\,W^2\,\mathbb{E}\,Z_{\max}.
\end{aligned} \tag{5.7}$$

($\mathbb{E}\,R^3 > \mathbb{E}\,R^2\,\mathbb{E}\,R$ follows from Lyapounov's inequality, or because a calculation yields $\mathbb{E}\,R = \sqrt{\pi}$, $\mathbb{E}\,R^2 = 4$, $\mathbb{E}\,R^3 = 6\sqrt{\pi}$.) Similarly,

$$\mathbb{E}\left[W Z_{\max}^2\right] = -\,\mathbb{E}\left[Z_1 Z_{\max}^2\right] = -\tfrac{1}{3}\,\mathbb{E}\left[Z_{\max}^2 \sum_1^3 Z_i\right] = 0 = \mathbb{E}\,W\,\mathbb{E}\,Z_{\max}^2. \tag{5.8}$$

It follows that if $W' \sim N(0,2)$ is independent of $Z_{\max}$, then

$$\mathbb{E}\,\mathcal{S}^3 = \mathbb{E}\,(W + Z_{\max})^3 > \mathbb{E}\,(W' + Z_{\max})^3 = \mathbb{E}\,(\mathcal{S}^{\mathrm{gr}})^3. \tag{5.9}$$

Hence the limit distribution $\mathcal{S}$ differs from $\mathcal{S}^{\mathrm{gr}}$ for the greedy distribution. △

5.1. **On the optimal strategy for N.** Consider now the unknown optimal strategy for N. Theorem 4.1 leads to an obvious conjecture:

Conjecture 5.4. *If both players play optimally, then*

$$n^{-1/2}S_n^{\mathrm{op}} \xrightarrow{\mathrm{d}} \mathcal{S}^{\mathrm{op}} = W + Z_{\max}, \tag{5.10}$$

where W and $Z_{\max} := \max\{Z_1, Z_2, Z_3\}$ each are as in Theorem 4.1, but they now may be dependent.

Note that if this holds, then (5.2)–(5.3) hold for $\mathcal{S}^{\mathrm{op}}$.

The optimal strategy for N has to punish strategies for R like the one in Example 2.1. Intuitively, it therefore seems likely that if R plays greedily, then the optimal strategy of N will punish R in games where the times T_1 and T_2 in our analysis in Section 3 are unusually large (and conversely reward R when they are small; remember that the expectation is the same as if N plays greedily). It therefore seems likely that if both players play optimally, there is a negative correlation between the two terms in (5.1). However, even if this is correct, it is possible that the dependency vanishes asymptotically so that we have the same limit $\mathcal{S}^{\mathrm{gr}}$ as in Theorem 4.1. We have no guess, and leave this as a problem.

Problem 5.5. If both players play optimally, does $n^{-1/2}S_n$ have the same asymptotic distribution $\mathcal{S}^{\mathrm{gr}}$ as in Theorem 4.1 for greedy play? If not, is there an asymptotic distribution $\mathcal{S}^{\mathrm{op}}$ (as conjectured above), and what is it?

6. The probability of winning for greedy play

Finally, we return to the case of both players using their greedy strategies and note that we may also calculate the asymptotic probability that R wins the game, in spite of her restriction, i.e., that the final score $S_n < 0$. (Recall that S_n is the score for N.)

Theorem 6.1. *If both players use their greedy strategies, then the probability that R wins has as $n \to \infty$ the limit*

$$\mathbb{P}(S_n < 0) \to \frac{3\arccos(1/4) - \pi}{4\pi} = \frac{\arccos(11/16)}{4\pi} \doteq 0.064677. \tag{6.1}$$

Proof. By Theorem 4.1, we have $\mathbb{P}(S_n < 0) \to \mathbb{P}(\mathcal{S}^{\mathrm{gr}} < 0)$ (since $\mathcal{S}^{\mathrm{gr}}$ has a continuous distribution, e.g. by (4.1)). We compute this probability using Theorem 4.1(iii). By (4.5), we have

$$\mathcal{S}^{\mathrm{gr}} < 0 \iff Z_i'' < 0 \ \forall i. \tag{6.2}$$

We may, similarly to (4.10), construct Z_i'' as

$$Z_i'' := \widehat{\mathbf{f}}_i \cdot \widehat{\boldsymbol{\zeta}}, \tag{6.3}$$

where $\widehat{\boldsymbol{\zeta}}$ is a standard normal distribution in $\mathbb{R}^3$, and $\widehat{\mathbf{f}}_1, \widehat{\mathbf{f}}_2, \widehat{\mathbf{f}}_3$ are three vectors in $\mathbb{R}^3$ such that

$$\widehat{\mathbf{f}}_i \cdot \widehat{\mathbf{f}}_j = \begin{cases} 4, & i = j, \\ -1, & i \neq j. \end{cases} \tag{6.4}$$

By (6.3), the condition (6.2) means that $\widehat{\boldsymbol{\zeta}}$ lies in the intersection of three open half-spaces H_1, H_2, H_3, which are bounded by hyperplanes orthogonal to $\widehat{\mathbf{f}}_1$, $\widehat{\mathbf{f}}_2$ and $\widehat{\mathbf{f}}_3$. The angle between any two of these vectors is, by (6.4), $\alpha := \arccos(-1/4)$. Hence, the interior angle between any of the two hyperplanes is $\beta := \pi - \alpha = \arccos(1/4)$, and thus the intersection of the unit sphere and $H_1 \cap H_2 \cap H_3$ is a spherical triangle Δ with all three angles β. Consequently, the area $|\Delta|$ of Δ is $3\beta - \pi$. The distribution of $\widehat{\boldsymbol{\zeta}}$ is rotationally symmetric, and thus we may project $\widehat{\boldsymbol{\zeta}}$ onto the unit sphere, and find, recalling that the area of the sphere is 4π,

$$\mathbb{P}(Z_{\max} < 0) = \mathbb{P}\bigl(\widehat{\boldsymbol{\zeta}} \in H_1 \cap H_2 \cap H_3\bigr) = \frac{|\Delta|}{4\pi} = \frac{3\beta - \pi}{4\pi} = \frac{3\arccos(1/4) - \pi}{4\pi}. \tag{6.5}$$

Finally, note that

$$\cos(3\beta - \pi) = -4\cos^3\beta + 3\cos\beta = -4\bigl(\tfrac14\bigr)^3 + 3 \cdot \tfrac14 = \tfrac{11}{16}. \tag{6.6}$$

□

Theorem 6.1 assumes that the players use their greedy strategies; we know that this is optimal for R, and yields the same expectation for N as his optimal strategy, if their objectives are to maximize the expected gain; if they instead want to maximize the probability of winning (but do not care about how much they win or lose), the optimal strategies are presumably different (see Example 6.3), and most likely much more complex; hence we do not know whether (6.1) holds or not in that case.

Problem 6.2. Suppose that both players want to maxime $\mathbb{P}(\text{win}) - \mathbb{P}(\text{lose})$. What is (asymptotically) the probability that R wins?

It is possible that the asymptotic answer is the same as in Theorem 6.1, although the probabilities for finite n are different. (See Example 6.3.) It might seem likely that a strategy that gives one of the players a significantly lower expected score will also give a lower probability that this score is positive. However, Example 6.4 shows that strategies with the same expectation still might give different distributions of the score and therefore different probabilities of winning, so it seems that there is no simple solution to Problem 6.2.

Example 6.3. Here is simple example showing that the greedy strategy is not the optimal strategy for R if the objective is to win, as in Problem 6.2. Let $n = 2$, and suppose that in the first four rounds, R has (by chance) chosen *rock*, *paper*, *scissors*, *scissors*, and that R won two of these while two were draws. Thus the score (for N) $S(4) = -2$. Hence, N cannot win, but since he will win the last round, the game will be a draw if he wins round 5. Therefore, in round 5, the objective for R is to minimize the probability of losing (but a draw is as good as a win). In this round R plays the game in Figure 1; if she wants to minimize the probability of losing this round the best strategy is to play *rock* or *paper* with equal probabilities, and not with the probabilities in Section 2.2 that minimize the expected loss. (The example can be extended to any $n \geqslant 2$ by assuming that R has played the three choices $n-2$ times each in the first $3(n-2)$ rounds, and that each of these rounds was a draw; the play then continues as above.) △

Example 6.4. Suppose that R uses the greedy strategy above, but that N uses the strategy in Example 5.3. As seen in Example 5.3, then

$$n^{-1/2} S_n \xrightarrow{\mathrm{d}} \mathcal{S} := -Z_1 + Z_{\max} = \max\bigl\{0, Z_2 - Z_1, Z_3 - Z_1\bigr\}. \tag{6.7}$$

Thus $\mathcal{S} \geqslant 0$ with a point mass $\mathbb{P}(\mathcal{S} = 0) = 1/3$ (by symmetry). In this case, we cannot immediately find the limit of $\mathbb{P}(S_n < 0)$, but if the strategy is perturbed a little, and N plays normally for the first $\varepsilon_n n$ rounds with $\varepsilon_n \to 0$ very slowly, it can be seen that $\mathbb{P}(S_n < 0) \to \frac{1}{2}\,\mathbb{P}(\mathcal{S} = 0) = 1/6$.

In this case, the new strategy for N is worse for him; it gives the same expected score but a lower probability that the score is positive (given that R plays greedily). However, it suggests that there also might be other strategies that instead increase the probability that N wins. $\triangle$

References

[1] Allan Gut. *Probability: A Graduate Course*, 2nd ed. Springer, New York, 2013.

[2] Svante Janson, Tomasz Łuczak & Andrzej Ruciński. *Random Graphs.* Wiley, New York, 2000.

[3] Howard L. Jones. Exact lower moments of order statistics in small samples from a normal distribution. *Ann. Math. Statistics* **19** (1948), 270–273.

[4] Anna R. Karlin & Yuval Peres. *Game Theory, Alive.* American Mathematical Society, Providence, RI, 2017.

[5] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Math. Ann.* **100** (1928), no. 1, 295–320.

[6] Sam Spiro, Erlang Surya & Ji Zeng. Semi-restricted rock, paper, scissors. *Electron. J. Combin.* **30** (2023), no. 4, Paper No. 4.32, 29 pp.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Email address: `svante.janson@math.uu.se`