Asymptotic Behavior of Greedy Policies

Ralph P. Russo and Nariankadu D. Shyamalkumar

August 2005

Technical Report No. 354

Asymptotic Behavior of Greedy Policies

Ralph P. Russo★**and Nariankadu D. Shyamalkumar**★

Consider i.i.d. samples from two discrete distributions, each with a finite support on the set of positive integers. A match is said to occur between two sampled members if they are from different populations and have the same value. In an earlier work, motivated by a database problem in computer science, we showed that the *greedy* policy for choosing the order of sampling from the two sources maximizes the expected number of matches, uniformly across all steps - hence beating the commonly used *alternating* policy. Here we study the asymptotic performance of both these policies. First, in contrast to the optimality of the greedy, the almost sure limits and weak limits of the number of matches are shown to be the same under both these policies. Second, we show that the difference in the expected number of matches between the greedy and alternating policies grows at the rate of n to a positive limit - unearthing a measure in which the greedy is asymptotically superior to the alternating. Third, study of the weak limit of the difference in the number of matches between the greedy and alternating policies (on the same sampled values, normalized by $n^{1.25}$, yields a scale mixture of normals centered at zero - weakening the impact of the former result.

KEY WORDS: Optimal Strategies, Greedy Policy, Martingales, Coupling

★ Department of Statistics and Actuarial Science, University of Iowa, Iowa City, IA 52242; E-mail: rrusso@stat.uiowa.edu, shyamal-kumar@uiowa.edu; Corresponding Author: Ralph P. Russo

1. INTRODUCTION

Suppose that we toss a coin repeatedly. On each toss we may choose either of two coins; one fair, the other two-headed. Let $R(n)$ denote the number of heads obtained with the fair coin, and $S(n)$ the number obtained with the two-headed, after n tosses. The object is to find an algorithm for choosing the coin to use on the n-th toss so as to maximize $\mathbb{E}(R(n)S(n))$, the expected number of *matches* after n tosses, uniformly in $n >= 1$. If we alternate between the two coins, then $\mathbb{E}(R(n)S(n)) \approx n^2/8$. However, it is shown in Shyamalkumar et al (2005) that this policy is not optimal. In fact, an optimal policy in this case is to toss the fair coin until heads occurs, switch to the two-headed coin for two tosses, and then return to the fair coin to repeat the cycle (Corollary 1 of Shyamalkumar et al (2005)).

The above problem is a simple special case of a more general problem that we investigate in Shyamalkumar et al (2005). That work was motivated by a specific database problem in computer science (refer to it for details). Suppose that observations are made sequentially and without replacement from two sources (populations) **R** and **S** whose members each carry a single positive integer valued label. A match is said to occur between two sampled members if they are drawn from different populations and carry the same label. The goal is to generate matches as quickly as possible. An algorithm that chooses where (**R** or **S**) to obtain the n-th observation, $n >= 1$, is referred to as a *reading policy*. An optimal *reading policy* is one which maximizes the expected number of matches after n observations have been made, uniformly in n.

Two reading policies of interest are the *alternating* policy, which alternately samples from **R**and **S** and the *myopic* (or *greedy*) policy, which chooses a source at each step so as to maximize the expected gain in matches for that step. The alternating policy is easy to implement, and requires no knowledge of **R** or **S**. Any policy with a fixed sampling order (such as **RSSRRSSRR**...) for which the **R** sample size is always within one of the sample size, is considered an *alternating* policy, as such policies all produce the same number of expected matches after k steps (Theorem 1 of [1]). In contrast to the *alternating* policy, the greedy policy requires a complete knowledge of **R**and **S**. It is a short term strategy that optimizes the expected one step gain, with no explicit regard to future (two step and beyond) gains. Note that there may be more than one greedy policy, as the greedy criterion may be ambivalent between **R** and **S** on some steps.

When **R** or **S** is finite, it is shown in Shyamalkumar et al (2005) that an optimal policy need not exist, that the alternating policies are optimal among the restricted class of *non-adaptive* policies (those that ignore the information obtained from the samples), and that any *greedy* policy dominates (and in most cases is strictly better than) any alternating policy. When **R** and **S** are infinite, the problem reduces to i.i.d. sampling from those distributions. In this case it is shown in Shyamalkumar et al (2005) that the alternating policy is again optimal among the *non-adaptives*, and that the *greedy* policy is optimal among all reading policies.

Our goal here is to explore the asymptotics associated with the *alternating* and *greedy* policies in the i.i.d. (infinite populations) case. Of interest is the strong and weak limiting behaviors of $M_A(n)$ and $M_G(n)$ (the numbers of matches formed through the first n steps by the alternating and greedy policies), and also their difference $M_G(n) - M_A(n)$.

In our coin example, it is easy to check that the policy we claimed as optimal is a member of the greedy class of policies. Note that this policy generates a renewal process $\{T_n\}_{n\geq 1}$, with a renewal occurring upon the observance of a tail, and an inter-arrival variable $X \overset{d}{=} 3Y+2$, where $Y \approx$ Geometric(1/2) with$\mathbb{E}(Y) = 2$. Now, using the **CLT** for renewal counting processes (see Ross (1983), page 62) and that $M_G(n) \approx (2/9)(n - N(n))^2$, we have

$$\left(\frac{M_G(n) - n^2/8}{n^{3/2}}\right) \xrightarrow{d} \mathrm{N}\left(0, \frac{1}{32}\right) \tag{1}$$

It is easy to show (via the ordinary **CLT** for i.i.d. variables) that the above result holds when $M_{\mathbf{G}}$ is replaced by $M_{\mathbf{A}}$. Law of the iterated logarithm results for $N(\cdot)$, $M_{\mathbf{G}}(\cdot)$ and $M_{\mathbf{A}}(\cdot)$ are also easily obtainable. Dealing with $\mathbb{E}(M_{\mathbf{G}}(n) - M_{\mathbf{A}}(n))$ requires a lighter touch. Let P denote the transition probability matrix of the Markov chain whose states are **RT** (tail), **RH** (heads with fair coin), **SH** (first heads with 2-headed coin), and **SHH** (second heads with 2-headed coin). Since P is doubly stochastic, the vector of stationary probabilities is uniform. This fact, together with the exact expressions relating $N(n)$ and $M_{\mathbf{G}}(n)$ for the various states (for example, $M_{\mathbf{G}}(n) = (1/9)(n - N(n) + 1)(2n - 2N(n) - 1)$ when the process is in state **SH**), geometric rate of convergence to stationarity and expectation of the residual lifetime or *overshoot* (which, in this example, can be easily calculated without recourse to the key renewal theorem) yield the following result:

$$\left(\frac{1}{n}\right)\mathbb{E}(M_G(n) - M_A(n)) \rightarrow \frac{1}{16}. \tag{2}$$

This is precise statement regarding the asymptotic superiority of the *greedy* policy over the *alternating* policy. In this work we show that results of the type stated above for the coin example hold true in the general case - the final theorem going beyond.

Conventions: All vectors carry a tilde, for example $\tilde{x}$. $\tilde{1}$ denotes a generic vector of ones whose dimension should be clear from context. For a vector $\tilde{x}$, $|\tilde{x}|$ will denote the absolute sum of its components, i.e. its L^1 norm. And for two vectors, $\tilde{x}$ and $\tilde{y}$, with the same dimensions, $\tilde{x} \cdot \tilde{y}$ will denote their inner product. Almost sure convergence will be denoted by $\xrightarrow{a.s.}$, convergence in probability by $\xrightarrow{P}$ and weak convergence (convergence in distribution) by $\xrightarrow{d}$. In the case where the weak limit is a constant we interchangeably use $\xrightarrow{d}$ and $\xrightarrow{P}$. We denote the iterated logarithms by $\log_2$, i.e. $\log_2(n) = \log(\log(n))$.

2. Alternating and Greedy Policies

A record from either source, **R** or **S**, carries one of the possible l (possible infinite) labels $1, \ldots, l$. The probability that a single record from the **R** source (resp., the **S** source) carries the i-th label, is r_i (resp., s_i). The probability vectors $(r_1, \ldots, r_l)$ and $(s_1, \ldots, s_l)$ are denoted by $\tilde{r}$ and $\tilde{s}$, respectively. The inner product of $\tilde{r}$ with $\tilde{s}$ is denoted by μ, i.e. $\mu = \tilde{r} \cdot \tilde{s}$. We shall always assume μ to be positive as otherwise there will be no common label between the two sources. The labels on the n-th records read from the **R** and **S** sources are denoted by $L_{\mathbf{R}}(n)$ and $L_{\mathbf{S}}(n)$. The above implies that $\{L_{\mathbf{R}}(n)\}_{n \geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n \geq 1}$ are sequences of independent and identically distributed random variables with

$$\Pr(L_{\mathbf{R}}(1) = i) = r_i \quad \text{and} \quad \Pr(L_{\mathbf{S}}(1) = i) = s_i, \quad i = 1, \ldots, l \tag{3}$$

Associated with the sequences $\{L_{\mathbf{R}}(n)\}_{n \geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n \geq 1}$ are the discrete time vector counting process $\{\tilde{N}_{\mathbf{R}}(n)\}_{n \geq 1}$ and $\{\tilde{N}_{\mathbf{S}}(n)\}_{n \geq 1}$; the first is defined by

$$\tilde{N}_{\mathbf{R}}(n) = (N_{\mathbf{R}}(n,1), \ldots, N_{\mathbf{R}}(n,l)), \quad \text{with } N_{\mathbf{R}}(n,i) = \sum_{j=1}^{n} I_{\{L_{\mathbf{R}}(j)=i\}}, \quad i = 1 \ldots, l; \; n = \geq 1 \tag{4}$$

and the second is defined analogously.

A reading policy is a zero-one valued stochastic process with the convention that the value 1 denotes a selection from source **R** and the value 0 a selection from source **S**. Hence

$$C(n) = \begin{cases} 1 & \text{if the } n\text{-th selection is from } \mathbf{R}; \\ \\ 0 & \text{if the } n\text{-th selection is from } \mathbf{S}; \end{cases}, \quad n = 1, 2, \ldots \tag{5}$$

Associated with each reading policy are two counting processes $\{R(n)\}_{n \geq 1}$ and $\{S(n)\}_{n \geq 1}$ defined by

$$R(n) := \sum_{j=1}^{n} C(j) \quad \text{and} \quad S(n) := n - R(n), \quad n = 1, 2, \ldots \tag{6}$$

These processes keep track of the number of records read from **R** and **S**, respectively, after a total of n records have been read. Also associated with a reading policy is a non-decreasing process $\{M(n)\}_{n \geq 1}$ which counts the number of join returns, i.e. matches, generated by the first n records. Hence,

$$M(n) = \tilde{N}_{\mathbf{R}}(R(n)) \cdot \tilde{N}_{\mathbf{S}}(S(n)), \quad n = 1, 2, \ldots \tag{7}$$

Observe that all of the processes $\{M(n)\}_{n \geq 1}$, $\{R(n)\}_{n \geq 1}$ and $\{S(n)\}_{n \geq 1}$ depend on the reading policy even though the notation does not make it explicit.

The filtration $\{\mathcal{F}_n\}_{n\geq 1}$ is defined by

$$\mathcal{F}_{n+1} := \mathcal{F}_1 \vee \sigma \langle L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}(R(n)); L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}(S(n)) \rangle, \quad n = 1, 2, \ldots \tag{8}$$

with $\mathcal{F}_1$ being arbitrary. $\mathcal{F}_1$ for example could contain all the information needed for randomization. All reading policies will henceforth be assumed to be adapted to the above filtration - they form the set of all *implementable* reading policies. Note that the filtration itself depends on the reading policy.

We define a alternating policy as one for which

$$R(2n) = n, \quad n = 1, 2, \ldots \tag{9}$$

Any reading policy in C_{NA} which satisfies (9) is called an alternating policy. In words, an alternating policy is one which does not use any information from the records, and under which at any step the numbers of records read from the two sources are within one of each other. There exists a large, if not an infinite, number of alternating policies. For our purposes $C_{\mathbf{A}}$ will denote the canonical alternating policy which strictly alternates between the two sources with the first pick being from **R**. Hence,

$$C_{\mathbf{A}}(n) = n \bmod 2, \quad n = 1, 2, \ldots \tag{10}$$

From the point of view of implementation, it may be more efficient to work with the alternating policy given by

$$C(n) = I_{\{n \bmod 4 < 2\}}, \quad n = 1, 2, \ldots \tag{11}$$

as it, leaving apart the first record, reads two records at a time from the chosen source.

Policies that utilize knowledge of $\tilde{r}$ and $\tilde{s}$, together with the information contained in the records that have been read, in order to optimize the choice for the next step are referred to as *greedy* policies. Towards a more precise definition, we observe that

$$\begin{aligned} \mathbb{E}(M(n+1) - M(n)|\mathcal{F}_{n+1}) = \mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right]|\mathcal{F}_{n+1})C(n+1) \\ + \mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right]|\mathcal{F}_{n+1})\left(1 - C(n+1)\right) \end{aligned} \tag{12}$$

The above implies that any $C(\cdot)$ maximizing the above conditional expectation should satisfy, for $n \geq 1$,

$$C(n+1) = \begin{cases} 1 & \text{if } \mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right]|\mathcal{F}_{n+1}) > \mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right]|\mathcal{F}_{n+1}); \\ \\ 0 & \text{if } \mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right]|\mathcal{F}_{n+1}) < \mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right]|\mathcal{F}_{n+1}); \end{cases} . \tag{13}$$

Note that at every step where

$$\mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right] | \mathcal{F}_{n+1}) = \mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right] | \mathcal{F}_{n+1}), \tag{14}$$

two greedy policies may differ as the greedy criterion is ambivalent.

As $\{L_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{L_{\mathbf{S}}(n)\}_{n\geq 1}$ are sequences of i.i.d. random variables, we have

$$\mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right] | \mathcal{F}_{n+1}) = \tilde{N}_{\mathbf{S}}\left(S(n)\right) \cdot \tilde{r}, \quad n = 1, 2, \ldots \tag{15}$$

and

$$\mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right] | \mathcal{F}_{n+1}) = \tilde{N}_{\mathbf{R}}\left(R(n)\right) \cdot \tilde{s}, \quad n = 1, 2, \ldots \tag{16}$$

For further analysis it is important to realize that $\tilde{N}_{\mathbf{S}}\left(S(n)\right) \cdot \tilde{r}$ and $\tilde{N}_{\mathbf{R}}\left(R(n)\right) \cdot \tilde{s}$ are both sums of i.i.d. observations. To make this explicit we define

$$X_{\mathbf{R}}\left(n\right) := s_{L_{\mathbf{R}}(n)} \quad \text{and} \quad X_{\mathbf{S}}\left(n\right) := r_{L_{\mathbf{S}}(n)}, \quad n = 1, 2, \ldots \tag{17}$$

The two sequences $\{X_{\mathbf{R}}\left(n\right)\}_{n\geq 1}$ and $\{X_{\mathbf{S}}\left(n\right)\}_{n\geq 1}$ are sequences of i.i.d. random variables with common mean μ and variances $\sigma^2_{\mathbf{R}}$ and $\sigma^2_{\mathbf{S}}$, respectively. We shall denote their partial sums by $\Gamma_{\mathbf{R}}\left[\cdot\right]$ and $\Gamma_{\mathbf{S}}\left[\cdot\right]$, i.e.

$$\Gamma_{\mathbf{R}}\left[n\right] = \sum_{j=1}^{n} X_{\mathbf{R}}\left(j\right) \quad \text{and} \quad \Gamma_{\mathbf{S}}\left[n\right] = \sum_{j=1}^{n} X_{\mathbf{S}}\left(j\right), \quad n = 1, 2, \ldots \tag{18}$$

Now, we can write

$$\tilde{N}_{\mathbf{S}}\left(S(n)\right) \cdot \tilde{r} = \Gamma_{\mathbf{S}}\left[S(n)\right] \quad \text{and} \quad \tilde{N}_{\mathbf{R}}\left(R(n)\right) \cdot \tilde{s} = \Gamma_{\mathbf{R}}\left[R(n)\right], \quad n = 1, 2, \ldots \tag{19}$$

Hence (13) simplifies to

$$C(n+1) = \begin{cases} 1 & \text{if } \Gamma_{\mathbf{S}}\left[S(n)\right] > \Gamma_{\mathbf{R}}\left[R(n)\right]; \\ \\ 0 & \text{if } \Gamma_{\mathbf{S}}\left[S(n)\right] < \Gamma_{\mathbf{R}}\left[R(n)\right]; \end{cases} , n = 1, 2, \ldots \tag{20}$$

We observe that the implementation of the greedy algorithm is greatly facilitated by the representation (20).

For our purposes $C_{\mathbf{G}}$ will denote the canonical greedy policy which chooses from the source **R** whenever there is a tie, i.e. whenever $\mathbb{E}(N_{\mathbf{S}}\left[S(n), L_{\mathbf{R}}\left(R(n)+1\right)\right] | \mathcal{F}_{n+1})$ is equal to $\mathbb{E}(N_{\mathbf{R}}\left[R(n), L_{\mathbf{S}}\left(S(n)+1\right)\right] | \mathcal{F}_{n+1})$, and whose first pick is from **R**.

All quantities with a subscript of **G** (resp., **A**) will denote quantities corresponding to the canonical greedy (resp., alternating) policy $C_{\mathbf{G}}$ (resp., $C_{\mathbf{A}}$).

3. Selection Ratio of a Greedy Policy

In this sub-section we discuss the asymptotic behavior of $R_G(n)$ as n increases to infinity. For this purpose an important fact, following from the definition of a greedy algorithm, is that

$$|\Gamma_{\mathbf{R}}[R_G(n)] - \Gamma_{\mathbf{S}}[S_G(n)]| \le \gamma, \quad n = 1, 2, \ldots; \quad \text{where } \gamma := \max_{1\le i\le l} r_i \vee \max_{1\le i\le l} s_i \tag{21}$$

Theorem 1 For the greedy reading policy C_G we have,

$$\frac{(R_G(n) - n/2)}{\sqrt{n}} \xrightarrow{\mathrm{d}} N\left(0, \sigma_{R_G}^2\right) \quad \text{as } n \to \infty; \sigma_{R_G}^2 := \left(\frac{\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2}{8\mu^2}\right) \tag{22}$$

Theorem 2 For the greedy reading policy C_G we have,

$$\limsup_{n\to\infty} \frac{R_G(n) - n/2}{\sqrt{2\sigma_{R_G}^2 n \log_2 n}} \xrightarrow{\mathrm{a.s.}} 1 \quad \text{and} \quad \liminf_{n\to\infty} \frac{R_G(n) - n/2}{\sqrt{2\sigma_{R_G}^2 n \log_2 n}} \xrightarrow{\mathrm{a.s.}} -1 \tag{23}$$

Lemma 1 In the case of the greedy algorithm we have

$$\Gamma_{\mathbf{R}}[\lceil x\rceil] < \Gamma_{\mathbf{S}}[n - \lceil x\rceil] - \gamma \implies R_G(n) > x \implies \Gamma_{\mathbf{R}}[\lfloor x\rfloor] \le \Gamma_{\mathbf{S}}[n - \lfloor x\rfloor] + \gamma, \quad 0 \le x \le n \tag{24}$$

Proof Since $\Gamma_{\mathbf{S}}[\cdot]$ and $\Gamma_{\mathbf{R}}[\cdot]$ are non-decreasing, for $0 \le x \le n$, we have by (21)

$$R_G(n) > x \quad \implies \quad \Gamma_{\mathbf{S}}[\lfloor x\rfloor] \le \Gamma_{\mathbf{R}}[R_G(n)] \le \Gamma_{\mathbf{S}}[S_G(n)] + \gamma \le \Gamma_{\mathbf{S}}[n - \lfloor x\rfloor] + \gamma. \tag{25}$$

The other half follows by observing that for $0 \le x \le n$,

$$R_G(n) \le x \quad \implies \quad \Gamma_{\mathbf{R}}[\lceil x\rceil] \ge \Gamma_{\mathbf{R}}[R_G(n)] \ge \Gamma_{\mathbf{S}}[S_G(n)] - \gamma \ge \Gamma_{\mathbf{S}}[n - \lceil x\rceil] - \gamma. \tag{26}$$

§

Proof of Theorem 1 First, we will argue that

$$Z(n) := \frac{\Gamma_{\mathbf{R}}[k_n] - \Gamma_{\mathbf{S}}[n - k_n]}{\sqrt{n/2}} \xrightarrow{\mathrm{d}} N\left(2^{1.5}\mu x, \sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2\right), \quad \text{as } n \to \infty \tag{27}$$

for any sequence $\{k_n\}_{n\ge 1}$ satisfying

$$\lim_{n\to\infty} \left(\frac{k_n - n/2}{\sqrt{n}}\right) = x \tag{28}$$

To this end note that

$$Z(n) = a_n \underbrace{\left[\frac{\sum_{j=1}^{k_n} X_{\mathbf{R}}(j) - k_n\mu}{k_n^{0.5}}\right]}_{\xrightarrow{\mathrm{d}} N(0,\sigma_{\mathbf{R}}^2)} - b_n \underbrace{\left[\frac{\sum_{j=1}^{n-k_n} X_{\mathbf{S}}(j) - (n-k_n)\mu}{(n-k_n)^{0.5}}\right]}_{\xrightarrow{\mathrm{d}} N(0,\sigma_{\mathbf{S}}^2)} + c_n 2^{1.5}\mu x \tag{29}$$

where the three sequence $\{a_n\}_{n\geq 1}$, $\{b_n\}_{n\geq 1}$ and $\{c_n\}_{n\geq 1}$ all converge to 1. The above observed weak limits are due to the ordinary central limit theorem. By independence of the first two terms in (29) and Slutsky theorem we have (27). Now defining Z_* and Z^* as Z but with the sequence $\{k_n\}_{n\geq 1}$ taken to be $\{\lceil n/2 + x\sqrt{n}\rceil\}_{n\geq 1}$ and $\{\lfloor n/2 + x\sqrt{n}\rfloor\}_{n\geq 1}$, respectively, we have

$$Z_*(n) \xrightarrow{\text{d}} N\left(2^{1.5}\mu x, \sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2\right) \quad \text{and} \quad Z^*(n) \xrightarrow{\text{d}} N\left(2^{1.5}\mu x, \sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2\right), \quad \text{as } n \to \infty \tag{30}$$

Now by Lemma 1 we have

$$\Pr\left(Z_*(n) < \frac{-\gamma}{\sqrt{n/2}}\right) \leq \Pr\left(\frac{(R_{\mathbf{G}}(n) - n/2)}{\sqrt{n}} > x\right) \leq \Pr\left(Z^*(n) < \frac{\gamma}{\sqrt{n/2}}\right), \text{ for large } n \text{ and } x \in \mathbb{R} \tag{31}$$

which combined with (30) completes the proof. §

Proof of Theorem 2 Let $\{k_n\}_{n\geq 1}$ be sequence of non-negative integers and $\{a_n\}_{n\geq 1}$ a sequence of reals such that

$$\frac{k_n - n/2}{\sqrt{2\sigma_{R_{\mathbf{G}}}^2 n \log_2 n}} \overset{n\to\infty}{\longrightarrow} C > 0 \quad \text{and} \quad a_n = \frac{1}{\sqrt{2(\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2)(n-k_n)\log_2(n-k_n)}}, \quad \forall n \geq 1 \tag{32}$$

For such a sequence $\{k_n\}_{n\geq 1}$,

$$\begin{aligned} &a_n\left(\Gamma_{\mathbf{R}}[k_n] - \Gamma_{\mathbf{S}}[n-k_n]\right) \\ &= \underbrace{a_n\left(\Gamma_{\mathbf{R}}[n-k_n] - \Gamma_{\mathbf{S}}[n-k_n]\right)}_{\liminf_{n\to\infty} = -1} + \underbrace{a_n\left(\Gamma_{\mathbf{R}}[k_n] - \Gamma_{\mathbf{R}}[n-k_n] - (2k_n - n)\mu\right)}_{\overset{\text{a.s.}}{\longrightarrow} 0} + \underbrace{b_n}_{\overset{n\to\infty}{\longrightarrow} C} \end{aligned} \tag{33}$$

where the first limit infimum is due to the standard law of iterated logarithm, the second limit due to Theorem 5.1 of Hanson and Russo (1983) on lag sums and the third limit as a consequence of (32). Hence for a sequence $\{k_n\}_{n\geq 1}$ satisfying (32) we have

$$\liminf_{n\to\infty} a_n\left(\Gamma_{\mathbf{R}}[k_n] - \Gamma_{\mathbf{S}}[n-k_n]\right) = C - 1 \tag{34}$$

Using (34) and Lemma 1, with k_n equal to $\left\lceil n/2 + (1-\epsilon)\sqrt{2\sigma_{R_{\mathbf{G}}}^2 n \log_2 n}\right\rceil$ we have

$$\frac{R_{\mathbf{G}}(n) - n/2}{\sqrt{2\sigma_{R_{\mathbf{G}}}^2 n \log_2 n}} > 1 - \epsilon \text{ infinitely often (i.o.) a.s., } \forall \epsilon > 0 \tag{35}$$

Now similarly, working instead with k_n equal to $\left\lfloor n/2 + (1+\epsilon)\sqrt{2\sigma_{R_{\mathbf{G}}}^2 n \log_2 n}\right\rfloor$, we have

$$\frac{R_{\mathrm{G}}(n) - n/2}{\sqrt{2\sigma^2_{R_{\mathrm{G}}} n \log_2 n}} > 1 + \epsilon \text{ only finitely often a.s., } \forall \epsilon > 0 \tag{36}$$

Now (35) and (36) are equivalent to the first statement in (23). A similar argument leads to the other. §

4. Asymptotic Behavior of $M(\cdot)$

Theorem 3

i. For any reading policy satisfying

$$n^{0.25}\left(\frac{R(n)}{n} - 0.5\right) \xrightarrow{\text{a.s.}} 0, \tag{37}$$

we have

$$\sqrt{n}\left[\frac{M(n)}{(n/2)^2} - \mu\right] \xrightarrow{\text{d}} N\left(0, 2(\sigma^2_{\mathrm{R}} + \sigma^2_{\mathrm{S}})\right) \tag{38}$$

ii. For any reading policy satisfying

$$\frac{R(n)}{n} \xrightarrow{\text{a.s.}} \alpha \in (0,1) \quad \text{and} \quad \sqrt{n}\left(\frac{R(n)}{n} - \alpha\right) \xrightarrow{\text{a.s.}} 0 \tag{39}$$

we have

$$\sqrt{n}\left[\frac{M(n)}{\alpha(1-\alpha)n^2} - \mu\right] \xrightarrow{\text{d}} N\left(0, \left[\frac{(1-\alpha)\sigma^2_{\mathrm{R}} + \alpha\sigma^2_{\mathrm{S}}}{\alpha(1-\alpha)}\right]\right) \tag{40}$$

Theorem 3 can be generalized to the case when $\{R(n)\}_{n\geq 1}$ is independent of the labels and is asymptotically normal.

Theorem 4 For any reading policy satisfying

$$\sqrt{\frac{n}{\log_2 n}}\left(\frac{R(n)}{n} - 0.5\right) \xrightarrow{\text{a.s.}} 0.5, \tag{41}$$

we have

$$\limsup_{n\to\infty} \sqrt{\frac{n}{4(\sigma^2_{\mathrm{R}} + \sigma^2_{\mathrm{S}})\log_2 n}}\left[\frac{M(n)}{(n/2)^2} - \mu\right] = 1, \text{ a.s.} \tag{42}$$

and

$$\liminf_{n\to\infty} \sqrt{\frac{n}{4(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)\log_2 n}} \left[\frac{M(n)}{(n/2)^2} - \mu\right] = -1, \text{ a.s.} \tag{43}$$

Corollary 1 For both the canonical greedy and alternating policies we have

$$\sqrt{n}\left[\frac{M(n)}{(n/2)^2} - \mu\right] \xrightarrow{\text{d}} N\left(0, 2(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)\right) \tag{44}$$

Proof The result follows by (10), Theorem 2 and Theorem 3. §

Corollary 2 For both the canonical greedy and alternating policies we have (42) and (43)

Proof The result follows by (10), Theorem 2 and Theorem 4. §

Lemma 2 Consider a reading policy satisfying

$$\sqrt{a_n}\left(\frac{R(n)}{n} - \alpha\right) \xrightarrow{\text{a.s.}} 0, \quad \text{for some } \alpha \in (0,1) \tag{45}$$

where $\{a_n\}_{n\geq 1}$ is a sequence such that

$$\lim a_n = \infty \quad \text{and} \quad a_n = o\left(\frac{n}{\log_2 n}\right) \tag{46}$$

For such a reading policy and sequence $\{a_n\}_{n\geq 1}$ we have,

$$a_n\left[\frac{M(n)}{\alpha(1-\alpha)n^2} - \frac{(1-\alpha)\Gamma_{\mathbf{R}}[R(n)] + \alpha\Gamma_{\mathbf{S}}[S(n)]}{\alpha(1-\alpha)n} + \mu\right] \xrightarrow{\text{a.s.}} 0, \tag{47}$$

Proof First, we observe that

$$\frac{M(n)}{\alpha(1-\alpha)n^2} - \frac{(1-\alpha)\Gamma_{\mathbf{R}}[R(n)] + \alpha\Gamma_{\mathbf{S}}[S(n)]}{\alpha(1-\alpha)n} + \mu = \left(\frac{\tilde{N}_{\mathbf{R}}(R(n))}{\alpha n} - \tilde{r}\right)\cdot\left(\frac{\tilde{N}_{\mathbf{S}}(S(n))}{(1-\alpha)n} - \tilde{s}\right) \tag{48}$$

Now by Cauchy-Schwartz inequality we have

$$\begin{aligned} &\left|a_n\left(\frac{\tilde{N}_{\mathbf{R}}(R(n))}{\alpha n} - \tilde{r}\right)\cdot\left(\frac{\tilde{N}_{\mathbf{S}}(S(n))}{(1-\alpha)n} - \tilde{s}\right)\right| \\ &\qquad \leq \sqrt{\sum_{i=1}^{l}\left[\sqrt{a_n}\left(\frac{N_{\mathbf{R}}[R(n),i]}{\alpha n} - r_i\right)\right]^2}\sqrt{\sum_{i=1}^{l}\left[\sqrt{a_n}\left(\frac{N_{\mathbf{S}}[S(n),i]}{(1-\alpha)n} - s_i\right)\right]^2} \end{aligned} \tag{49}$$

All of the summands in (49) go to zero almost surely due to the standard law of iterated logarithm and (45) by arguments similar to

$$\sqrt{a_n}\left(\frac{N_{\mathbf{R}}[R(n),i]}{\alpha n}-r_i\right)$$

$$=\underbrace{\sqrt{\frac{a_n \log_2 R(n)}{R(n)}}\left(\frac{R(n)}{\alpha n}\right)}_{\xrightarrow{\text{a.s.}}0 \text{ by (45) and (46)}}\underbrace{\sqrt{\frac{R(n)}{\log_2 R(n)}}\left(\frac{N_{\mathbf{R}}[R(n),i]}{R(n)}-r_i\right)}_{\text{bounded by LIL}}+r_i\underbrace{\sqrt{a_n}\left(\frac{R(n)}{\alpha n}-1\right)}_{\xrightarrow{\text{a.s.}}0 \text{ by (45)}}\xrightarrow{\text{a.s.}}0 \quad (50)$$

Hence the proof. §

For the following result we shall need the filtration, say $\{\mathcal{G}_n\}_{n\geq 0}$, defined as

$$\mathcal{G}_n=\mathcal{G}_0\vee\sigma\langle L_{\mathbf{R}}(1),\ldots,L_{\mathbf{R}}(R(n));L_{\mathbf{S}}(1),\ldots,L_{\mathbf{S}}(S(n))\rangle,\quad n=1,2,\ldots \quad (51)$$

with $\mathcal{G}_0$ containing all the information needed for randomization by C.

Lemma 3 For any reading policy such that

$$\frac{R(n)}{n}\xrightarrow{\text{a.s.}}\alpha\in(0,1), \quad (52)$$

we have $\{Y_n\}_{n\geq 1}$ defined by

$$Y_n:=(1-\alpha)\Gamma_{\mathbf{R}}[R(n)]+\alpha\Gamma_{\mathbf{S}}[S(n)]-[\alpha n+(1-2\alpha)R(n)]\mu,\quad n=1,2,\ldots \quad (53)$$

is a $\{\mathcal{G}_n\}_{n\geq 1}$ martingale with bounded increments. Moreover, it satisfies the following:

i. $\{Y_n\}_{n\geq 1}$ suitably normalized is asymptotically normal, that is

$$\frac{Y_n}{\sqrt{\alpha(1-\alpha)n}}\xrightarrow{\text{d}}N\left(0,(1-\alpha)\sigma_{\mathbf{R}}^2+\alpha\sigma_{\mathbf{S}}^2\right) \quad (54)$$

ii. Moreover, we have

$$\limsup_{n\to\infty}\frac{Y_n}{\sqrt{2\alpha(1-\alpha)n\log_2 n}}=\sqrt{(1-\alpha)\sigma_{\mathbf{R}}^2+\alpha\sigma_{\mathbf{S}}^2},\ \text{a.s.} \quad (55)$$

and

$$\liminf_{n\to\infty}\frac{Y_n}{\sqrt{2\alpha(1-\alpha)n\log_2 n}}=-\sqrt{(1-\alpha)\sigma_{\mathbf{R}}^2+\alpha\sigma_{\mathbf{S}}^2},\ \text{a.s.} \quad (56)$$

Proof First note that

$$D_n:=Y_n-Y_{n-1}=\begin{cases}(1-\alpha)\left(X_{\mathbf{R}}(R(n))-\mu\right), & \text{if } C(n)=1;\\ \alpha\left(X_{\mathbf{S}}(S(n))-\mu\right), & \text{if } C(n)=0;\end{cases},\quad n=2,3,\ldots;\ D_1:=Y_1 \quad (57)$$

As $C(n)$ is $\mathcal{G}_{n-1}$ measurable and both $X_{\mathbf{R}}(R(n))$ and $X_{\mathbf{S}}(S(n))$ are independent of $\mathcal{G}_{n-1}$ we have,

$$\mathbb{E}(D_n|\mathcal{G}_{n-1}) = 0 \quad \text{and} \quad \mathbb{E}\left(D_n^2|\mathcal{G}_{n-1}\right) = (1-\alpha)^2\sigma_R^2 C(n) + \alpha^2\sigma_S^2(1-C(n)) \tag{58}$$

Moreover as D_n is bounded and Y_n is $\mathcal{G}_n$ measurable, we have $\{Y_n\}_{n\geq 1}$ is a $\{\mathcal{G}_n\}_{n\geq 1}$ martingale with bounded increments. Now as a consequence of (58) we have

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\left(D_k^2|\mathcal{G}_{k-1}\right) = (1-\alpha)^2\sigma_R^2\frac{R(n)}{n} + \alpha^2\sigma_S^2\frac{S(n)}{n} \xrightarrow{\text{a.s.}} \alpha(1-\alpha)\left((1-\alpha)\sigma_{\mathbf{R}}^2 + \alpha\sigma_{\mathbf{S}}^2\right) \tag{59}$$

which by using the martingale central limit theorem, see for example Theorem 7.4 of Durrett (1991), leads to (54). The law of the iterated logarithm, (55) and (56), follows from the same for martingales - see for example Theorems 1 and 2 of Stout (1970). §

Proof of Theorem 3 First, observe that by Lemma 2,

$$\sqrt{n}\left[\frac{M(n)}{\alpha(1-\alpha)n^2} - \mu\right] \quad \text{and} \quad \sqrt{n}\left[\frac{(1-\alpha)\Gamma_{\mathbf{R}}[R(n)] + \alpha\Gamma_{\mathbf{S}}[S(n)]}{\alpha(1-\alpha)n} - 2\mu\right] \tag{60}$$

have the same weak limit. Now observe that the latter term can be written as

$$\underbrace{\sqrt{n}\left[\frac{Y_n}{\alpha(1-\alpha)n}\right]}_{\xrightarrow{\text{d}} N(0,\cdot) \text{ by Lemma 3}} + \underbrace{\sqrt{n}\left(\frac{1-2\alpha}{\alpha(1-\alpha)}\right)\left[\frac{R(n)}{n} - \alpha\right]}_{\xrightarrow{\text{a.s.}} 0 \text{ by (37) or (39)}} \tag{61}$$

where $\{Y_n\}_{n\geq 1}$ is as defined in (53). The first term on the right is asymptotically normal by Lemma 3 and the second term converges almost surely to zero by the assumptions (37) or (39). Hence by Slutsky theorem we have (38) and (40). §

Proof First, observe that by Lemma 2,

$$\sqrt{\frac{n}{4(\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2)\log_2 n}}\left[\frac{M(n)}{(n/2)^2} - \mu\right] \quad \text{and} \quad \sqrt{\frac{n}{(\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2)\log_2 n}}\left[\frac{\Gamma_{\mathbf{R}}[R(n)] + \Gamma_{\mathbf{S}}[S(n)]}{n} - \mu\right] \tag{62}$$

have the same limiting behavior in the almost sure sense. The latter term is nothing but

$$\frac{Y_n}{\sqrt{0.25n(\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2)\log_2 n}}, \tag{63}$$

where $\{Y_n\}_{n\geq 1}$ is as defined in (53) with α taking the value 0.5. Now (42) and (43) follow from Lemma 3. §

5. Relative Performance of Greedy v/s Alternating

All the asymptotic results for the number of matches so far have been exactly the same for both the greedy and alternating policies. But Corollary 1 of Shyamalkumar et al (2005), showing optimality of the greedy, makes it interesting to study the asymptotic behavior of the difference in the number matches. First, we study the expected difference in the numbers of matches which concludes with Theorem 5 below. Second, we study the weak limit of the difference in the numbers of matches and this is identified in Theorem 6.

In the following we will need the filtration $\{\mathcal{G}_n\}_{n\geq 0}$ defined as

$$\mathcal{G}_n = \mathcal{G}_0 \vee \sigma \langle L_{\mathbf{R}}(1), \ldots, L_{\mathbf{R}}(R_{\mathrm{G}}(n)); L_{\mathbf{S}}(1), \ldots, L_{\mathbf{S}}(S_{\mathrm{G}}(n))\rangle, \quad n = 1, 2, \ldots \tag{64}$$

with $\mathcal{G}_0$ containing all the information needed for randomization by not only C_{G} but also $C_{\mathbf{A}}$. The argument below is based on the sequence of random times $\{T_n\}_{n\geq 1}$ defined as

$$T_n = \inf\left\{k \geq 1 \middle| S_{\mathrm{G}}(k+1) = \left\lceil \frac{n}{2} \right\rceil + 1 \quad \text{or} \quad R_{\mathrm{G}}(k+1) = \left\lceil \frac{n}{2} \right\rceil + 1\right\}, \quad n = 1, 2, \ldots \tag{65}$$

It is easily checked that $\{T_n\}_{n\geq 1}$ is a sequence of $\{\mathcal{G}_n\}_{n\geq 0}$ stopping times. Similar to Lemma 1, it can be shown that for positive x

$$(n - T_n) > x \implies \Gamma_{\mathbf{R}}[\lceil n/2 \rceil] \leq \Gamma_{\mathbf{S}}[\lfloor n/2 \rfloor - \lfloor x \rfloor] + \gamma \text{ or } \Gamma_{\mathbf{S}}[\lceil n/2 \rceil] \leq \Gamma_{\mathbf{R}}[\lfloor n/2 \rfloor - \lfloor x \rfloor] + \gamma \tag{66}$$

and

$$\Gamma_{\mathbf{R}}[\lceil n/2 \rceil] < \Gamma_{\mathbf{S}}[\lfloor n/2 \rfloor - \lceil x \rceil] \text{ or } \Gamma_{\mathbf{S}}[\lceil n/2 \rceil] < \Gamma_{\mathbf{R}}[\lfloor n/2 \rfloor - \lceil x \rceil] \implies (n - T_n) > x \tag{67}$$

Theorem 5 For the canonical greedy and alternating policies we have

$$\lim_{n\to\infty} \left(\frac{\mathbb{E}(M_{\mathrm{G}}(n) - M_{\mathbf{A}}(n))}{n} \right) = \frac{\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2}{8\mu} \tag{68}$$

Theorem 6 For the canonical greedy and alternating policies we have

$$\left(\frac{M_{\mathrm{G}}(n) - M_{\mathbf{A}}(n)}{n\sqrt{n - T_n}} \right) \xrightarrow{\mathrm{d}} N\left(0, \frac{\sigma_{\mathbf{R}}^2 + \sigma_{\mathbf{S}}^2}{8}\right) \text{ as } n \to \infty \tag{69}$$

The following lemma describes the weak limit of T_n.

Lemma 4 For the above defined stopping times $\{T_n\}_{n\geq 1}$ corresponding to the greedy reading policy C_{G} we have,

$$\frac{(n - T_n)}{2\sqrt{n}} \xrightarrow{\mathrm{d}} \left| N\left(0, \sigma_{R_{\mathrm{G}}}^2\right) \right| \quad \text{as } n \to \infty \tag{70}$$

where $\sigma_{R_{\mathrm{G}}}^2$, the asymptotic variance of $R_{\mathrm{G}}(n)$, is defined in (22).

Proof The proof follows along similar lines as Theorem 1 - the only change being using (66) and (67) instead of Lemma 1. §

The following lemma describes the behavior of $\{R_G(n) - R_G(T_n)\}_{n\geq 1}$ relative to $\{n - T_n\}_{n\geq 1}$.

Lemma 5 For the above defined stopping times $\{T_n\}_{n\geq 1}$ corresponding to the greedy reading policy C_G we have,

$$\frac{R_G(n) - R_G(T_n)}{n - T_n} \xrightarrow{P} \frac{1}{2} \quad \text{as } n \to \infty \tag{71}$$

Proof First, we show that

$$\frac{R_G(n) - R_G(T_n)}{\log(n)} \xrightarrow{P} \infty, \quad \text{as } n \to \infty \tag{72}$$

Towards this end, observe that using (21) twice we have

$$|(\Gamma_R[R_G(n)] - \Gamma_R[R_G(T_n)]) - (\Gamma_S[S_G(n)] - \Gamma_S[S_G(T_n)])| \leq 2\gamma \tag{73}$$

which implies that for any positive K

$$R_G(n) - R_G(T_n) < K\log(n) \implies \sum_{i=1}^{n-T_n-K\log(n)} X_S(i + S_G(T_n)) - 2\gamma < \sum_{i=1}^{K\log(n)} X_R(i + R_G(T_n)) \tag{74}$$

The second expression can be rewritten as

$$\left(\sum_{i=1}^{n-T_n-K\log(n)} \frac{X_S(i + S_G(T_n)) - \mu}{\sqrt{n - T_n - K\log(n)}}\right) - \left(\sum_{i=1}^{K\log(n)} \frac{X_R(i + R_G(T_n)) - \mu}{\sqrt{K\log(n)}}\right) < 2\gamma - \mu\left(\sqrt{n - T_n - K\log(n)} - \sqrt{K\log(n)}\right) \tag{75}$$

As $n^{-0.25}(n - T_n)$ goes to infinity in probability (Lemma 4) we have the independent terms on the left converging to normal distributions and the term on the right converging to negative infinity. Hence the probability of the above event converges to zero. Now using (74) we have (72). Combining (72) with Theorem 2 we have

$$R_G(n) \xrightarrow{P} \infty \quad \text{and} \quad \frac{R_G(n) - R_G(T_n)}{\log(R_G(n))} \xrightarrow{P} \infty, \quad \text{as } n \to \infty \tag{76}$$

The above with Theorem 5.1 of Hanson and Russo (1983) on lag sums gives us

$$\left(\frac{\Gamma_R[R_G(n)] - \Gamma_R[R_G(T_n)]}{R_G(n) - R_G(T_n)}\right) \xrightarrow{P} \mu \quad \text{and} \quad \left(\frac{\Gamma_S[S_G(n)] - \Gamma_S[S_G(T_n)]}{S_G(n) - S_G(T_n)}\right) \xrightarrow{P} \mu \tag{77}$$

where the second part follows by symmetry. This with (73) along with (72) gives us

$$\frac{R_{\mathrm{G}}(n) - R_{\mathrm{G}}(T_n)}{S_{\mathrm{G}}(n) - S_{\mathrm{G}}(T_n)} \xrightarrow{\mathrm{P}} 1 \tag{78}$$

which is equivalent to (71). §

Observe that the above implies that

$$\frac{R_{\mathrm{G}}(n) - R_{\mathrm{G}}(T_n)}{\sqrt{n}} = \left(\frac{R_{\mathrm{G}}(n) - R_{\mathrm{G}}(T_n)}{n - T_n}\right)\left(\frac{n - T_n}{\sqrt{n}}\right) \xrightarrow{\mathrm{d}} \left|N\left(0, \sigma_{R_{\mathrm{G}}}^2\right)\right| \quad \text{as } n \to \infty \tag{79}$$

where the convergence in probability of the first term follows from (71) and the weak convergence to the folded normal of the second term follows from (70). In fact, the stronger result

$$\frac{R_{\mathrm{G}}(n) - R_{\mathrm{G}}(T_n) - 0.5(n - T_n)}{\sqrt{n - T_n}} \xrightarrow{\mathrm{d}} N\left(0, \sigma_{R_{\mathrm{G}}}^2\right) \quad \text{as } n \to \infty \tag{80}$$

can be shown using an argument similar to that of Theorem 1 - we refrain from doing so as this result is not needed for our purpose.

The following lemma addresses the following uniform integrability of two sequences that we shall need below.

Lemma 6 For the canonical greedy policy, the following two sequences

$$\left\{\left(\frac{R_{\mathrm{G}}(n) - n/2}{\sqrt{n}}\right)^2\right\}_{n\geq 1} \quad \text{and} \quad \left\{\left(\frac{n - T_n}{\sqrt{n}}\right)^2\right\}_{n\geq 1} \tag{81}$$

are uniformly integrable.

Proof First we show the uniform integrability of the first sequence above. Let $k = \lfloor n/2 + \sqrt{tn} \rfloor$. By Lemma 1 we have

$$\Pr\left(R_{\mathrm{G}}(n) > n/2 + \sqrt{tn}\right) \leq \Pr\left(\Gamma_{\mathrm{R}}[k] \leq \Gamma_{\mathrm{S}}[n-k] + \gamma\right) \tag{82}$$

Observe that

$$\Gamma_{\mathrm{S}}[n-k] - \Gamma_{\mathrm{R}}[k] + (2k-n)\mu = [\Gamma_{\mathrm{S}}[n-k] - \Gamma_{\mathrm{R}}[n-k]] - [\Gamma_{\mathrm{R}}[k] - \Gamma_{\mathrm{R}}[n-k] - (2k-n)\mu] \tag{83}$$

which implies that the term on the left is a sum of k independent zero mean random variables taking values in $[-\gamma, \gamma]$. This along with Hoeffding inequality, for example see Serfling (1980), implies

$$\Pr\left(\Gamma_{\mathrm{R}}[k] \leq \Gamma_{\mathrm{S}}[n-k] + \gamma\right) \leq \exp\left\{-\frac{0.5}{k\gamma^2}[(2k-n)\mu - \gamma]^2\right\} \tag{84}$$

Working with the upper bound above and using the inherent symmetry we get the following simple upper bound

$$\Pr\left(\left[\frac{R_{\mathrm{G}}(n)-n/2}{\sqrt{n}}\right]^2 > t\right) \le 2\exp\left\{-\left(\frac{\mu}{2\gamma}\right)^2\sqrt{t}\right\}, \quad \forall t \ge 1;\ \forall n \ge \left(2+\frac{\gamma}{\mu}\right)^2 \tag{85}$$

And since this bound is integrable and is free of n we have the uniform integrability of the sequence in (81).

For the second sequence the uniform integrability follows from an argument similar to the above - the only change being using (66) and (67) instead of Lemma 1. §

Lemma 7 For the canonical greedy policy and the sequence of stopping times $\{T_n\}_{n\ge1}$ defined in (65) we have

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\left(\left[\tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(T_n))\right]\cdot\left[\tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(T_n))\right]\right) = \frac{\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2}{8\mu} \tag{86}$$

Proof First, we will show that

$$\left[\frac{\tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(T_n))}{R_{\mathrm{G}}(n)-R_{\mathrm{G}}(T_n)}\right]\cdot\left[\frac{\tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(T_n))}{S_{\mathrm{G}}(n)-S_{\mathrm{G}}(T_n)}\right] \xrightarrow{\mathrm{P}} \mu \tag{87}$$

Note that, by symmetry and Slutsky theorem, it suffices to show that for any j in $\{1,2,\ldots,l\}$, we have

$$\left[\frac{N_{\mathbf{R}}[R_{\mathrm{G}}(n),j] - N_{\mathbf{R}}[R_{\mathrm{G}}(T_n),j]}{R_{\mathrm{G}}(n)-R_{\mathrm{G}}(T_n)}\right] \xrightarrow{\mathrm{P}} r_j \tag{88}$$

Now (76) coupled with Theorem 5.1 of Hanson and Russo (1983) on lag sums gives us (88) and hence (87).

Second, we have by Lemma 5 and Slutsky theorem that

$$\left(\frac{R_{\mathrm{G}}(n)-R_{\mathrm{G}}(T_n)}{0.5(n-T_n)}\right)\left(\frac{S_{\mathrm{G}}(n)-S_{\mathrm{G}}(T_n)}{0.5(n-T_n)}\right) \xrightarrow{\mathrm{P}} 1 \tag{89}$$

Combining (87) and (89) with Lemma 4 and using Slutsky theorem, we have

$$\left(\frac{1}{n}\right)\left[\tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{R}}(R_{\mathrm{G}}(T_n))\right]\cdot\left[\tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(n)) - \tilde{N}_{\mathbf{S}}(S_{\mathrm{G}}(T_n))\right] \xrightarrow{\mathrm{d}} \left(\frac{\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2}{8\mu}\right)\chi^2(1) \tag{90}$$

Now observe that the sequence in (87) is a non-negative sequence bounded above by $l\gamma^2$ and that the non-negative sequence in (89) is bounded - by say, 2. These along with the uniform integrability of the sequence $\{n^{-1}(n-T_n)^2\}_{n\ge1}$ provided by Lemma 6 gives us the convergence of the expectation. Hence the proof. §

Proof of Theorem 4 Working on the set $\{R_{\mathrm{G}}(T_n) = \lceil n/2\rceil\}$, we have

$$
\begin{aligned}
M_{\mathbf{G}}(n) - M_{\mathbf{A}}(n) = &\left(\tilde{N}_{\mathbf{R}}(R_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{R}}(\lceil n/2 \rceil)\right) \cdot \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(T_n)) \\
&+ \left(\tilde{N}_{\mathbf{R}}(R_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{R}}(\lceil n/2 \rceil)\right) \cdot \left(\tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(T_n))\right) \\
&- \tilde{N}_{\mathbf{R}}(\lceil n/2 \rceil) \cdot \left(\tilde{N}_{\mathbf{S}}(\lfloor n/2 \rfloor) - \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(n))\right)
\end{aligned} \tag{91}
$$

The first term on the right can be written as

$$
\left(\tilde{N}_{\mathbf{R}}(R_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{R}}(\lceil n/2 \rceil) - [R_{\mathbf{G}}(n) - \lceil n/2 \rceil]\,\tilde{r}\right) \cdot \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(T_n)) + [R_{\mathbf{G}}(n) - \lceil n/2 \rceil]\,\Gamma_{\mathbf{S}}[S_{\mathbf{G}}(T_n)] \tag{92}
$$

where the first term has a conditional expectation of zero on the set $\{R_{\mathbf{G}}(T_n) = \lceil n/2 \rceil\}$ as it is the $(n - T_n)$-th term of a zero martingale. The argument for the latter is similar to that found in Lemma 3. Now the third term on the right of (91) can be written as

$$
\tilde{N}_{\mathbf{R}}(\lceil n/2 \rceil) \cdot \left(\tilde{N}_{\mathbf{S}}(\lfloor n/2 \rfloor) - \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(n)) - [R_{\mathbf{G}}(n) - \lceil n/2 \rceil]\,\tilde{s}\right) + [R_{\mathbf{G}}(n) - \lceil n/2 \rceil]\,\Gamma_{\mathbf{R}}[\lceil n/2 \rceil] \tag{93}
$$

where the first term has a conditional expectation of zero on the set $\{R_{\mathbf{G}}(T_n) = \lceil n/2 \rceil\}$ as it is independent of $\mathcal{G}_n$ ($\supseteq \mathcal{G}_{T_n}$) and conditioned on $\mathcal{G}_n$ has zero mean. Using symmetry along with (92) and (93) leads to

$$
\begin{aligned}
&\frac{1}{n}\left|\mathbb{E}(M_{\mathbf{G}}(n) - M_{\mathbf{A}}(n)) - \mathbb{E}\left(\left[\tilde{N}_{\mathbf{R}}(R_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{R}}(R_{\mathbf{G}}(T_n))\right] \cdot \left[\tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(n)) - \tilde{N}_{\mathbf{S}}(S_{\mathbf{G}}(T_n))\right]\right)\right| \\
&\leq \frac{1}{n}\mathbb{E}(|\Gamma_{\mathbf{R}}[R_{\mathbf{G}}(T_n)] - \Gamma_{\mathbf{R}}[S_{\mathbf{G}}(T_n)]|\,|R_{\mathbf{G}}(n) - \lceil n/2 \rceil|) \\
&\leq \gamma \mathbb{E}\left(\frac{|R_{\mathbf{G}}(n) - \lceil n/2 \rceil|}{n}\right) \to 0
\end{aligned} \tag{94}
$$

where the convergence to zero of the last term follows by Lemma 6 and Theorem 1. The theorem follows now by using Lemma 7. §

For our final theorem we will need a uniform central limit theorem for a class of policies which can be described as *greedy with offsets*. This is the content of the next lemma; below we describe some needed notations. Let $\mathbf{G}_\delta$, for $\delta \in [-\gamma, \gamma]$, be a policy such that

$$
C_{\mathbf{G}_\delta}(n+1) = \begin{cases} 1 & \text{if } \Gamma_{\mathbf{S}}[S(n)] > \Gamma_{\mathbf{R}}[R(n)] + \delta; \\ 0 & \text{if } \Gamma_{\mathbf{S}}[S(n)] < \Gamma_{\mathbf{R}}[R(n)] + \delta; \end{cases} \quad, n = 1, 2, \ldots \tag{95}
$$

Let $\{X^*_{\mathbf{R}}(n)\}_{n\geq 1}$ and $\{X^*_{\mathbf{S}}(n)\}_{n\geq 1}$ denote two auxiliary sequences of i.i.d. random variables with $X^*_{\mathbf{R}} \stackrel{d}{=} X_{\mathbf{R}}$ and $X^*_{\mathbf{S}} \stackrel{d}{=} X_{\mathbf{S}}$. Also let $\Gamma^*_{\mathbf{R}}(\cdot)$ and $\Gamma^*_{\mathbf{S}}(\cdot)$ denote their respective partial sums. Now we define the sequence of random variables $\{Y^\delta_n\}_{n\geq 1}$ and $\{Z^\delta_n\}_{n\geq 1}$, for $\delta \in [-\gamma, \gamma]$, as

$$Y_n^\delta := \left(\frac{1}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)n}}\right)\left[\Gamma_{\mathbf{R}}\left[R_{\mathbf{G}_\delta}(n)\right]-\Gamma_{\mathbf{S}}^*\left(R_{\mathbf{G}_\delta}(n)\right)\right], \quad n \geq 1 \tag{96}$$

and

$$Z_n^\delta := \left(\frac{1}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)n}}\right)\left[\Gamma_{\mathbf{S}}\left[S_{\mathbf{G}_\delta}(n)\right]-\Gamma_{\mathbf{R}}^*\left(S_{\mathbf{G}_\delta}(n)\right)\right], \quad n \geq 1 \tag{97}$$

Lemma 8 There exists a $K > 0$ such that

$$\sup_{t\in\mathbb{R}}\left|\Pr\left(Y_n^\delta \leq t\right)-\Phi(t)\right| \leq K n^{\frac{-1}{4}}\log(n) \tag{98}$$

and

$$\sup_{t\in\mathbb{R}}\left|\Pr\left(Z_n^\delta \leq t\right)-\Phi(t)\right| \leq K n^{\frac{-1}{4}}\log(n) \tag{99}$$

Proof It suffices to show (98) as (99) follows by symmetry. Towards this, we need $\{\mathcal{G}_m\}_{m\geq 0}$, a filtration, defined for $m \geq 1$ as

$$\mathcal{G}_m = \mathcal{G}_0 \vee \sigma\left\langle L_{\mathbf{R}}(1),\ldots,L_{\mathbf{R}}\left(R_{\mathbf{G}_\delta}(m)\right);L_{\mathbf{S}}(1),\ldots,L_{\mathbf{S}}\left(S_{\mathbf{G}_\delta}(m)\right);X_{\mathbf{R}}^*(1),\ldots,X_{\mathbf{R}}^*\left(R_{\mathbf{G}_\delta}(m)\right)\right\rangle \tag{100}$$

with $\mathcal{G}_0$ containing all the information needed for randomization by $C_{\mathbf{G}_\delta}$. Also, we define, for a fixed $n \geq 1$,

$$D_m := \begin{cases} \frac{X_{\mathbf{R}}\left(R_{\mathbf{G}_\delta}(m)\right)-X_{\mathbf{S}}^*\left(R_{\mathbf{G}_\delta}(m)\right)}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)n}}, & \text{if } C_{\mathbf{G}_\delta}(m)=1; \\ 0, & \text{if } C_{\mathbf{G}_\delta}(m)=0; \end{cases}, \quad m = 2,3,\ldots,n;\ D_1 := \frac{Y_1}{\sqrt{n}} \tag{101}$$

By construction,

$$\sum_{i=1}^{n} D_i = Y_n^\delta \quad \text{and} \quad \max_{i\leq n} D_i \leq n^{-1/2}\left(\frac{\gamma}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)}}\right). \tag{102}$$

As $C_{\mathbf{G}_\delta}(m)$ is $\mathcal{G}_{m-1}$ measurable and both $X_{\mathbf{R}}\left(R_{\mathbf{G}_\delta}(m)\right)$ and $X_{\mathbf{S}}^*\left(R_{\mathbf{G}_\delta}(m)\right)$ are independent of $\mathcal{G}_{m-1}$ we have,

$$\mathbb{E}(D_m|\mathcal{G}_{m-1}) = 0 \quad \text{and} \quad \mathbb{E}\left(D_m^2|\mathcal{G}_{m-1}\right) = \left(\frac{2}{n}\right)C_{\mathbf{G}_\delta}(m), \quad m = 1,2,\ldots,n \tag{103}$$

Hence, as D_m^δ is $\mathcal{G}_m$ measurable, $\{\sum_{i=1}^m D_i\}_{1\le m\le n}$ is a martingale. Now, as a consequence of (103) we have

$$V_n^2 := \sum_{i=1}^n \mathbb{E}\left(D_i^2|\mathcal{G}_{i-1}\right) = \left(\frac{2}{n}\right)R_{\mathbf{G}_\delta}(n). \tag{104}$$

This implies that

$$\Pr\left(\left|V_n^2 - 1\right| > n^{-1/2}(\log(n))^2\right) = \Pr\left(\left|\frac{R_{\mathbf{G}_\delta}(n) - n/2}{\sqrt{n}}\right| > \left(\frac{1}{2}\right)(\log(n))^2\right) \tag{105}$$

By an argument similar to that in the proof of Lemma 6 we get, analogous to (85),

$$\Pr\left(\left[\frac{R_{\mathbf{G}_\delta}(n) - n/2}{\sqrt{n}}\right]^2 > t\right) \le 2\exp\left\{-\left(\frac{\mu}{4\gamma}\right)^2\sqrt{t}\right\}, \quad \forall t \ge 1;\ \forall n \ge 4\left(1+\frac{\gamma}{\mu}\right)^2 \tag{106}$$

Combining (105) and (85) we get

$$\Pr\left(\left|V_n^2 - 1\right| > n^{-1/2}(\log(n))^2\right) \le \exp\left(\frac{32\gamma^2}{\mu^2}\right)\left(\frac{1}{n}\right), \quad \forall n \ge 1 \tag{107}$$

Using the inequalities in (102) and (107) we have (98), for some K free of δ, as an application of Theorem 3.7 (along with the remark (ii) following it) of Hall and Heyde (1980). Hence the proof. §

For the following lemma, we define the sequence of random variables $\{G_n\}_{n\ge1}$ as

$$2G_n := \begin{cases} (\Gamma_{\mathbf{R}}[R_{\mathbf{G}}(n)] - \Gamma_{\mathbf{R}}[\lceil n/2\rceil]) - (\Gamma_{\mathbf{S}}[\lfloor n/2\rfloor] - \Gamma_{\mathbf{S}}[S_{\mathbf{G}}(n)]) & \text{on } A_n; \\ \\ (\Gamma_{\mathbf{S}}[S_{\mathbf{G}}(n)] - \Gamma_{\mathbf{S}}[\lceil n/2\rceil]) - (\Gamma_{\mathbf{R}}[\lfloor n/2\rfloor] - \Gamma_{\mathbf{R}}[R_{\mathbf{G}}(n)]) & \text{on } A_n^c; \end{cases} \tag{108}$$

where the sequence of events $\{A_n\}_{n\ge1}$ is defined by $A_n := \{R_{\mathbf{G}}(T_n) = \lceil n/2\rceil\}$.

Lemma 9

$$\left(\frac{M_{\mathbf{G}}(n) - M_{\mathbf{A}}(n)}{n\sqrt{n-T_n}}\right) - \frac{G_n}{\sqrt{n-T_n}} \xrightarrow{\mathrm{d}} 0 \text{ as } n \to \infty \tag{109}$$

Proof We start with a decomposition analogous to (91),

$$
\begin{aligned}
\frac{M_{\mathrm{G}}(n)-M_{\mathrm{A}}(n)}{n}-G_n = &\left(\frac{1}{n}\right)\left(\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{R}}\left(\lceil n/2\rceil\right)\right)\cdot\left(\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(T_n)\right)\right)\\
&+\mathrm{I}_{A_n}\left[\left(\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)\right)\cdot\left(\frac{\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(T_n)\right)}{n}-0.5\tilde{s}\right)\right.\\
&\left.-\left(\tilde{N}_{\mathrm{S}}\left(\lfloor n/2\rfloor\right)-\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(n)\right)\right)\cdot\left(\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)}{n}-0.5\tilde{r}\right)\right]\\
&+\mathrm{I}_{A_n^c}\left[\left(\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(T_n)\right)\right)\cdot\left(\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)}{n}-0.5\tilde{r}\right)\right.\\
&\left.-\left(\tilde{N}_{\mathrm{R}}\left(\lfloor n/2\rfloor\right)-\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(n)\right)\right)\cdot\left(\frac{\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(T_n)\right)}{n}-0.5\tilde{s}\right)\right]
\end{aligned}
\tag{110}
$$

We will now proceed to show that each term on the right of (110) divided by $\sqrt{n-T_n}$ converges in probability to zero. For the first term the result follows from (90) and Lemma 4. The second and third terms on the right of (110) are similar (by symmetry) and hence it suffices to show that the second term divided by $\sqrt{n-T_n}$ converges in probability to zero. First,

$$
\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)}{\sqrt{n-T_n}\,n^{3/8}}=\underbrace{\left(\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(n)\right)-\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)}{n-T_n}\right)}_{\text{bounded}}\underbrace{\left[\frac{\sqrt{n-T_n}}{n^{3/8}}\right]}_{\xrightarrow{\mathrm{P}}0\text{ (Lemma 4)}}\xrightarrow{\mathrm{P}}0. \tag{111}
$$

And, by a similar argument,

$$
\left(\frac{\tilde{N}_{\mathrm{S}}\left(\lfloor n/2\rfloor\right)-\tilde{N}_{\mathrm{S}}\left(S_{\mathrm{G}}(n)\right)}{\sqrt{n-T_n}\,n^{3/8}}\right)\xrightarrow{\mathrm{P}}0. \tag{112}
$$

Second , we observe that

$$
\frac{n/2-R_{\mathrm{G}}(T_n)}{n^{5/8}}=\underbrace{\left(\frac{n/2-R_{\mathrm{G}}(T_n)}{\max(n-T_n,1)}\right)}_{\text{bounded}}\underbrace{\left[\frac{\max(n-T_n,1)}{n^{5/8}}\right]}_{\xrightarrow{\mathrm{P}}0\text{ (Lemma 4)}}\xrightarrow{\mathrm{P}}0. \tag{113}
$$

Third,

$$
\begin{aligned}
&n^{-5/8}\left(\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)}{n}-0.5\tilde{r}\right)\\
&=\underbrace{\left[\frac{R_{\mathrm{G}}(T_n)}{n}\right]^{5/8}}_{\text{bounded by 1}}\underbrace{\left(\frac{\tilde{N}_{\mathrm{R}}\left(R_{\mathrm{G}}(T_n)\right)-R_{\mathrm{G}}(T_n)\,\tilde{r}}{R_{\mathrm{G}}(T_n)^{5/8}}\right)}_{\xrightarrow{\mathrm{P}}0\text{ by (88)}}-\underbrace{\left(\frac{n/2-R_{\mathrm{G}}(T_n)}{n^{5/8}}\right)}_{\xrightarrow{\mathrm{P}}0\text{ by (113)}}\xrightarrow{\mathrm{P}}0.
\end{aligned}
\tag{114}
$$

The convergence in probability to zero of the second term divided by $\sqrt{n-T_n}$ now follows by appealing to symmetry and repeated application of (111), (112) and (114). Hence the proof. §

Proof of Theorem 5 In view of Lemma 9, it is sufficient to show that

$$\frac{G_n}{\sqrt{n-T_n}} \xrightarrow{\mathrm{d}} N\left(0, \frac{\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2}{8}\right) \text{ as } n \to \infty \tag{115}$$

Towards this, we first observe that, for $u \in \mathbb{R}$, on A_n

$$\Pr\left(\frac{(\Gamma_{\mathbf{R}}\left[R_{\mathrm{G}}(n)\right]-\Gamma_{\mathbf{R}}\left[\lceil n/2\rceil\right])-(\Gamma_{\mathbf{S}}\left[\lfloor n/2\rfloor\right]-\Gamma_{\mathbf{S}}\left[S_{\mathrm{G}}(n)\right])}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)(n-T_n)}} \le u \middle| \mathcal{G}_{T_n}\right) = \Pr\left(Y_{n-T_n}^{\Delta_n} \le u\right) \tag{116}$$

and on A_n^c,

$$\Pr\left(\frac{(\Gamma_{\mathbf{S}}\left[S_{\mathrm{G}}(n)\right]-\Gamma_{\mathbf{S}}\left[\lceil n/2\rceil\right])-(\Gamma_{\mathbf{R}}\left[\lfloor n/2\rfloor\right]-\Gamma_{\mathbf{R}}\left[R_{\mathrm{G}}(n)\right])}{\sqrt{0.5(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)(n-T_n)}} \le u \middle| \mathcal{G}_{T_n}\right) = \Pr\left(Z_{n-T_n}^{\Delta_n} \le u\right) \tag{117}$$

where $\Delta_n := \Gamma_{\mathbf{S}}\left[S_{\mathrm{G}}(T_n)\right]-\Gamma_{\mathbf{R}}\left[R_{\mathrm{G}}(n)\right]$. This, along with Lemma 8, leads to

$$\begin{aligned} &\left|\Pr\left(\frac{G_n}{\sqrt{0.125(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)(n-T_n)}} \le u\right)-\Phi(u)\right| \\ &\qquad = \left|\int_{A_n}\Pr\left(Y_{n-T_n}^{\Delta_n} \le u\right)\mathrm{dP}+\int_{A_n^c}\Pr\left(Z_{n-T_n}^{\Delta_n} \le u\right)\mathrm{dP}-\Phi(u)\right| \\ &\qquad \le K\,\mathbb{E}\left(\max\left[1,(n-T_n)^{\frac{-1}{4}}\log(n-T_n)\right]\right) \to 0, \text{ as } n \to \infty. \end{aligned} \tag{118}$$

Hence the proof. §

Corollary 3 For the canonical greedy and alternating policies we have

$$\left(\frac{M_{\mathrm{G}}(n)-M_{\mathrm{A}}(n)}{n^{1.25}}\right) \xrightarrow{\mathrm{d}} \mathrm{F} \text{ as } n \to \infty \tag{119}$$

where F is a scale mixture of normals centered at zero given by

$$\mathrm{F} = \int N(0,\sigma^2)\mathrm{dG}(\sigma^2), \quad \text{where } \mathrm{G} := \left|N\left(0, \frac{(\sigma_{\mathbf{R}}^2+\sigma_{\mathbf{S}}^2)^3}{8\mu^2}\right)\right| \tag{120}$$

Proof The proof follows by using Theorem 6, Lemma 4 and the asymptotic independence between the two terms on the right side of the equation below.

$$\left(\frac{M_{\mathrm{G}}(n) - M_{\mathrm{A}}(n)}{n^{1.25}}\right) = \left(\frac{M_{\mathrm{G}}(n) - M_{\mathrm{A}}(n)}{n\sqrt{n - T_n}}\right)\left(\sqrt{\frac{n - T_n}{\sqrt{n}}}\right) \tag{121}$$

§

References

Durrett, R. 1991. *Probability: Theory and Examples.* First Edition. Wadsworth & Brooks/Cole, USA.

Hall, P. and Heyde, C. C. 1980. *Martingale Limit Theory and its Application.* First Edition. Academic Press, USA.

Hanson, D. L. and Russo, R. P. 1983. 'Some Results on Increments of the Wiener Process with Applications to Lag Sums of I.I.D. Random Variables'. *Annals of Probability* **11**, No. 3, 609-623.

Lindvall, T. 1992. *Lectures on the Coupling Method.* New York:John Wiley & Sons.

Ross, T. 1983. *Stochastic Processes.* New York:John Wiley & Sons.

Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics.* John Wiley & Sons, USA.

Shyamalkumar, N. D., Russo, R. P. and Lawrence, R. 2005. 'Optimal Policies to Obtain the Most Join Results'. *Technical Report No.* **353**, Department of Statistics and Actuarial Science, The University of Iowa.

Stout, W. F. 1970. 'A Martingale Analogue of Kolmogorov's Law of the Iterated Logarithm'. *Z. Wahrsch. Verw. Gebiete* **15**, 279-290.