A Markov Chain Analysis of Blackjack Strategy

Michael Wakin and Christopher Rozell

Rice University

I. INTRODUCTION

The game of blackjack naturally lends itself to analysis using Markov chains. Constructing a state space where each state represents the value of a hand, for example, a sequence of draws can be viewed as a random walk with transition probabilities dictated by the (unseen) cards remaining in the deck. If we assume an infinite deck (equivalently, that the probability of the next card does not depend on previously dealt cards), then the process is first-order Markov: the probability distribution on the next state depends only on the value of the current hand. Such simplifications are somewhat artificial, of course, but they allow us to ask a series of interesting questions that may someday lead to better playing strategies. As evidenced by a trip to Harrah's Lake Charles Casino during the writing of this paper, intelligent play can be both entertaining and profitable.

In this paper, we explore several methods for blackjack analysis using Markov chains. First, we develop a collection of Markov chains used to model the play of a single hand, and we use these chains to compute the player's expected advantage when playing according to a basic strategy. Next, we analyze a simple card-counting technique called the Complete Point-Count System, introduced by Harvey Dubner in 1963 and discussed in Edward Thorp's famous book, *Beat the Dealer* [1]. This system relies on tracking the state of the deck using a High-Low Index (HLI); we construct a Markov chain that models the evolution of the HLI throughout multiple

Email: {*wakin,crozell*}@*rice.edu*

rounds of play. By computing the equilibrium distribution on this chain, we estimate how much time the player may spend in favorable states, and we evaluate the expected advantage for a card-counting player.

This paper is organized as follows. Section II explains the basic rules of blackjack. Section III develops a collection of Markov chains for modeling the play of a single hand, and explains how these chains can be used to compute the player's advantage. Section IV introduces the Complete Point-Count System and describes the construction of a Markov chain using the HLI state space; Section V explores the problem of finding the equilibrium for this chain. In Section VI, we present our analysis of the Complete Point-Count System. Finally, we conclude in Section VII.

II. Blackjack Rules

We describe in this section a basic collection of blackjack rules. We assume these rules for the analysis that follows. Many variations on these rules exist [2]; in most cases these variations can easily be considered in a similar analysis.

A. Object of the game

The player's objective is to obtain a total that exceeds the dealer's hand, without exceeding 21.

B. The cards

The value of a hand is computed as the total of all cards held. Face cards each count as 10 toward the total; we refer to any such card as a 10. Aces may count as either 1 or 11, whichever yields a larger total that is less than 21. A hand with an 11-valued Ace is called *soft*. A hand with a 1-valued Ace, or with no Aces, is called *hard*. A total of 21 on the first two cards is called a *blackjack* or *natural*. Note that a natural must consist of a 10 and an Ace.

For our analysis, the number D of 52-card decks in play will be specified when relevant. The value $D = 6$ is typical in today's casinos.

A box called the *shoe* holds the cards to be dealt. When a certain number of cards have been played (roughly $3/4$ of a shoe), the entire shoe is reshuffled when the next hand is complete. Reshuffling is discussed in greater detail in Section IV.

C. The deal

The player places a bet B at the beginning of each hand. To begin, the player and dealer each receive two cards. Both of the player's cards are face up. One of the dealer's cards is face up; one is face down.

1) Insurance: If the dealer shows an Ace, the player has the option of placing a side bet called *insurance*. A player taking insurance bets the amount $\frac{B}{2}$ that the dealer holds a natural. If the dealer does hold a natural, the player's insurance bet is returned with a profit of B. If the dealer does not hold a natural, the player loses his insurance bet.

The insurance bet has no impact on the remaining aspects of play.

2) Blackjack (Natural): If the dealer holds a natural and the player does not, the player loses his bet. If the dealer and player both hold a natural, the hand is over with no money exchanged. If the player holds a natural, but the dealer does not, his original bet B is returned with a profit of $\frac{3}{2}B$.

If neither the dealer nor the player holds a natural, the player proceeds according to the options described below. When he finishes his turn, the dealer proceeds according to a fixed strategy, drawing until her total exceeds 16.

3) Hitting and standing: If his total is less than 21, the player may *hit*, requesting another card. The player may choose to continue hitting until his total exceeds 21, in which case he *busts* and loses his bet. This is an advantage for the house, who wins even if the dealer subsequently busts. The player may also elect to *stand* at any time, drawing no additional cards and passing play to the dealer.

At the conclusion of the hand, the player wins if his total exceeds the dealer's total. In this case, his bet is returned with a profit B. If the dealer holds a higher total than the player, the player loses his bet. In the case of a tie, called a *push*, no money changes hands.

4) Doubling down: When holding his first two cards, the player may elect to *double down*, increasing his original bet to $2B$ and drawing a single additional card before passing play to the dealer.

5) Splitting pairs: If the player's first two cards have the same value, the player may elect to split the pair. In this case, the two cards are divided into two distinct hands (the player places an additional bet B to cover the new hand), and play proceeds as normal with two small exceptions. First, a total of 21 after a split is never counted as a natural. Second, a player who splits a pair of Aces is allowed only a single card drawn to each Ace. Otherwise, the player is allowed to double a bet after splitting, or even split again if he receives another pair. At the moment, we place no limit on the number of times a player may split during a hand.

III. Analyzing a Basic Strategy

In this section, we use Markov chains to model the play of a single hand. First, we construct a Markov chain for the dealer's hand, and we use the transition matrix to compute the probability distribution among the possible dealer outcomes. Next, we analyze a playing strategy for the player (known as the "Basic Strategy" [1]–[3]); in this system, the player makes firm decisions that depend strictly on the content of his own hand and the value of the dealer's up card. Again, we use a Markov chain to determine the probability distribution among the possible player outcomes; we then use this distribution to compute the player's expected profit on a hand.

For this section we assume a uniform, infinite shoe – that is, we assume for each draw the probability d_i of drawing card i is as follows:

$$d_A = d_2 = d_3 = \cdots = d_9 = 1/13, \tag{1}$$

$$d_{10} = 4/13. \tag{2}$$

A. The dealer's hand

The dealer plays according to a fixed strategy, hitting on all hands 16 or below, and standing on all hands 17 or above. To model the play of the dealer's hand, we construct a Markov chain (Ψ_d, Z_d). The state space Ψ_d contains the following elements:

- $\{first_i : i \in \{2, \ldots, 11\}\}$: the dealer holds a single card, valued i. All other states assume the dealer holds more than one card.
- $\{hard_i : i \in \{4, \ldots, 17\}\}$: the dealer holds a hard total of i.
- $\{soft_i : i \in \{12, \ldots, 17\}\}$: the dealer holds a soft total of i.
- $\{stand_i : i \in \{17, \ldots, 21\}\}$: the dealer stands with a total of i.
- bj: the dealer holds a natural.
- $bust$: the dealer busts.

TABLE I

Distribution of dealer outcomes: before the first card is dealt, and given that an Ace is dealt first.

Outcome	*Probability, before start*	*Probability, given Ace*
$stand_{17}$	0.1451	0.1308
$stand_{18}$	0.1395	0.1308
$stand_{19}$	0.1335	0.1308
$stand_{20}$	0.1803	0.1308
$stand_{21}$	0.0727	0.0539
bj	0.0473	0.3077
$bust$	0.2816	0.1153

In total, we obtain a state space with $|\Psi_d| = 37$. The dealer's play corresponds to a random walk on this state space, with initial state corresponding to the dealer's first card, and with each transition corresponding to the draw of a new card. Transition probabilities are dictated by the shoe distribution d_i.

For the situations where the dealer must stand (i.e. when her total is above 16), we specify that each state transitions to itself with probability 1.[1] The states $stand_{17}, \ \ldots, \ stand_{21}, \ bj,$ and $bust$ then become *absorbing states*. Because the dealer's total increases with each transition (except possibly once when a soft hand transitions to hard), it is clear that within $n = 17$ transitions, any random walk will necessarily reach one of these absorbing states.

To compute a probability distribution among the dealer's possible outcomes, we need only to find the distribution among the absorbing states. This is accomplished by constructing the transition matrix Z_d, and observing $(Z_d^{17})_{i,j}$. Given that the dealer shows initial card γ, for example, we may the compute her possible outcomes using $(Z_d^{17})_{first_\gamma,\cdot}$. Averaging over all possible initial cards γ (and weighting by the probability d_γ that she starts with card γ), we may compute her overall distribution on the absorbing states. As an example, Table I lists the distribution of the dealer's outcomes, before the first card is dealt, as well as the distribution of the dealer's outcomes, given that she starts with an Ace face up.

[1]Note that states $hard_{17}$ and $soft_{17}$ first transition to $stand_{17}$ with probability 1, however – these are provided to accommodate possible rules variations.

Notice that our assumption regarding the infinite shoe is key for this analysis. Otherwise, the cards played would impact the distribution of the remaining cards. Our assumption about a uniform distribution is not necessary, however. In Section IV, we perform a similar analysis for nonuniform, infinite shoe distributions.

B. The player's hand

The player's decisions under Basic Strategy depend only on the cards held in his hand and the single card shown by the dealer. Basic Strategy is known as the optimal technique which maximizes the player's expected return without considering the cards dealt in previous hands.[2] Under Basic Strategy, the player sometimes will elect to double his bet, or to split a pair, but never to take insurance. Full details of the strategy are provided in [1]–[3].

We again use Markov chains to compute the distribution of the player's outcomes under the Basic Strategy. Because the player's strategy depends on the dealer's card, we use 10 different Markov chains, one for each card $\gamma \in \{2, \ldots, 11\}$ that the dealer may be showing. These chains will be analyzed together in the next section.

For each Markov chain (Ψ_γ, Z_γ) , we use a state space containing the following elements:

- $\{first_i : i \in \{2, \ldots, 11\}\}$: the player holds a single card, valued i, and will automatically be dealt another.
- $\{twoHard_i : i \in \{4, \ldots, 21\}\}$: the player holds two different cards for a hard total of i and may hit, stand, or double.
- $\{twoSoft_i : i \in \{12, \ldots, 21\}\}$: the player holds two different cards for a soft total of i and may hit, stand, or double.
- $\{pair_i : i \in \{2, \ldots, 11\}\}$: the player holds two cards, each of value i, and may hit, stand, double, or split.
- $\{hard_i : i \in \{5, \ldots, 21\}\}$: the player holds more than two cards for a hard total of i and may hit or stand.
- $\{soft_i : i \in \{13, \ldots, 21\}\}$: the player holds more than two cards for a soft total of i and may hit or stand.

[2] This system is described as optimal among "total-based" systems, those which do not depend on the particular cards composing the player's hand.

- $\{stand_i : i \in \{4, \ldots, 21\}\}$: the player stands with his original bet and a total of i.
- $\{doubStand_i : i \in \{6, \ldots, 21\}\}$: the player stands with a doubled bet and a total of i.
- $\{split_i : i \in \{2, \ldots, 11\}\}$: the player splits a pair, each card valued i (modeled as an absorbing state).
- bj: the player holds a natural.
- $bust$: the player busts with his original bet.
- $doubBust$: the player busts with a doubled bet.

Note that different states with the same total often indicate that different options are available to the player. In total, we obtain a state space with $|\Psi_\gamma| = 121$.

Analysis on this Markov chain is similar to the dealer's chain described above. The Basic Strategy dictates a particular move by the player (hit, stand, double, or split) for each of the states. Transition probabilities then depend on the moves of the Basic Strategy, as well as the distribution d_i of cards in the shoe. Because the player's total increases with each transition (except possibly once when a soft hand transitions to hard), it is clear that within $n = 21$ transitions, any random walk will necessarily reach one of the absorbing states. The primary difference in analysis comes from the possibility of a split hand.

We include a series of absorbing states $\{split_i\}$ for the event where the player elects to split a pair of cards i. Intuitively, we imagine that the player then begins two new hands, each in the state $first_i$. To model the play of one of these hands, we create another Markov chain (similar to the player's chain described above), but we construct this chain using the particular rules for a hand that follows a split (see Section II-C.5). Because multiple splits are allowed, this chain also includes the absorbing state $split_i$.

C. Computing the player's advantage

Assume the dealer shows card γ face up. As described in Section III-A, we may use $(Z_d^{17})_{first_\gamma,\cdot}$ to determine the distribution u_γ on her absorbing states $U \subset \Psi_d$. Similarly, we may use Z_γ^{21} to determine the distribution v_γ on the player's absorbing states $V \subset \Psi_\gamma$. Note that these outcomes are independent, given γ, so the probability that any pair of player/dealer outcomes occurs can be computed from the product of the distributions.

If the player never elected to split, then each of the player's absorbing states would correspond to the end of the player's turn. Using any combination of the dealer's absorbing state i and the player's absorbing state j, we could refer to the rules in Section II and determine the exact profit (or gain) $g(i,j)$ to the player. Averaging this gain over all possible combinations of the player's and dealer's absorbing states (and weighing each combination by its probability), we could compute precisely the player's expected gain on a single hand as follows

$$G = \sum_{\gamma=2}^{11} d_\gamma \sum_{i \in U} \sum_{j \in V} u_\gamma(i) v_\gamma(j) g(i,j). \tag{3}$$

The situation is more complicated, however, if the player elects to split. Suppose $j \in V$ is one of the player's split-absorbing states. We need to compute the player's expected profit $g_s(i,j)$, given that he starts a single post-split hand. This can be accomplished as usual using the post-split Markov chain described in Section III-B, *except* that there is some probability that he will elect to split again. In that case, the player plays two more post-split hands, each with expected gain $g_s(i,j)$. This recursion allows us to compute $g_s(i,j)$ precisely. Letting x be the probability of splitting again, and letting y be the payoff given that he does *not* split again, we have:

$$g_s = (1-x)y + 2xg_s = \frac{(1-x)y}{1-2x}. \tag{4}$$

As long as $x < 0.5$, we may use this formula to compute the expected gain $g(i,j) = 2g_s(i,j)$, and combining with (3), we can compute the player's overall expected gain G.

For the Basic Strategy table published in [2], with the player placing a unity bet $B = \$1$, we compute $G = -0.0052$. This corresponds to an average house advantage of 0.52%, or an expected *loss* for the player of 0.52 cents per hand.

IV. THE COMPLETE POINT-COUNT SYSTEM

The Complete Point-Count System [1] is based on a few simple observations:

- The player may vary his bets and playing strategy at will, while the dealer must play a fixed strategy.
- When the shoe has relatively many high cards remaining, there is a playing strategy that gives the player an expected advantage over the house. The advantage occurs because the dealer must hit on totals 12-16, even when there are disproportionately many tens left in the shoe.
- When the shoe has relatively few high cards remaining, the house has a small advantage over the player, regardless of his strategy.

These observations are fundamental to most card-counting strategies and are also the basic reasons why card-counting is not permitted by casinos. Because the player can place bets in such a way to minimize losses during unfavorable times, card-counting can give the player an overall expected gain over the house.

The Complete Point-Count System is one method for tracking the relative numbers of high cards remaining in the shoe. We assume now that the shoe contains a finite number D of decks, so that the cards played throughout a round have an impact on the distribution of cards remaining. In this section, we explain how Markov chains can be used to analyze such a scheme.

A. Details of the card-counting system

In this system, all cards in the deck are classified as *low* (2 through 6), *medium* (7 through 9), and *high* (10 and Ace). Each 52-card deck thus contains 20 low cards, 12 medium cards, and 20 high cards.

As the round progresses, the player keeps track of the cards played. For convenience, we assume that he keeps track of an ordered triple (L, M, H), representing the number of low, medium and high cards that have been played. This triple is sufficient to compute the number of cards remaining in the shoe, $R = 52D - (L + M + H)$.

The player uses the ordered triple to compute a *high-low index* (HLI):

$$HLI = 100 \cdot \frac{L - H}{R}. \tag{5}$$

The HLI gives an estimate for the favorability of the shoe: when positive, the player generally has an advantage and should bet high; when negative, the player generally has a disadvantage and should bet low. Thorp suggests a specific betting strategy according to the HLI [1]:

$$B = \begin{cases} b & \text{if} \quad -100 \leq HLI \leq 2 \\ \left\lfloor \frac{HLI}{2} \right\rfloor b & \text{if} \quad 2 < HLI \leq 10 \\ 5b & \text{if} \quad 10 < HLI \leq 100. \end{cases} \tag{6}$$

where b is the player's fundamental unit bet. For the simplicity of this paper, we assume $b = \$1$. It is important also to note that, although the player's advantage is presumably high when $HLI > 10$, Thorp recommends an upper limit on the bets for practical reasons. If a casino suspects that a player is counting cards, they will often remove that player from the game.

The player's optimal playing strategy (hit, stand, double, or split) changes as a function of HLI. Thorp gives a series of tables to be used by the player [1]. To be precise, the player's decisions depend on HLI, the dealer's face card, and the state of the player's own hand. For simplicity, we assume the player fixes his strategy at the beginning of each hand – that is, he does not track changes in HLI until the hand is complete.

Finally, Thorp recommends taking the insurance bet when $HLI > 8$.

B. Computing the player's expected gain

We would like to compute the expected gain per hand of a player using the Complete Point-Count System. We cannot immediately apply the techniques of Section III, however, because the player's strategy is a function of HLI.

Suppose, however, that we are given an ordered triple (L, M, H) that the player observes prior to beginning a hand. From the triple we may compute HLI and obtain his betting and playing strategies. In order to able to apply the techniques of Section III, we must make two key assumptions. First, we assume that the pdf for the next card drawn is *uniform* over each category: low, medium, and high. To be precise, we assume

$$d_2 = \cdots = d_6 = \left(\tfrac{1}{5}\right) \tfrac{20D-L}{R} \tag{7}$$

$$d_7 = d_8 = d_9 = \left(\tfrac{1}{3}\right) \tfrac{12D-M}{R} \tag{8}$$

$$d_{10} = \left(\tfrac{4}{5}\right) \tfrac{20D-H}{R} \tag{9}$$

$$d_A = \left(\tfrac{1}{5}\right) \tfrac{20D-H}{R}. \tag{10}$$

Second, we assume that the shoe pdf *does not change* during the play of the next hand. This is a kind of "locally infinite" assumption, enough to permit the techniques of Section III. With these two assumptions, we are able to compute the player's expected gain $G_{(L,M,H)}$ on that hand.[3] Accounting for the expected gain of an insurance bet is also simple, given these assumptions, as the probability of winning an insurance bet is precisely d_{10}.

If we were able to determine the overall probability $\pi_{(L,M,H)}$ that the player begins a hand with triple (L, M, H), we would be able to compute his overall expected gain by averaging over

[3] When the HLI is very high, we may violate the assumption of (4) that $x < 0.5$. To avoid the danger of a negative probability, we make a slight change and use $g_s = (1 - x)y$, which is similar to limiting the player to a single split.

all triples:

$$G = \sum_{(L,M,H)} \pi_{(L,M,H)} G_{(L,M,H)}. \tag{11}$$

We turn once again to Markov chains to find the probabilities $\pi_{(L,M,H)}$.

C. Markov chain framework for shoe analysis

In the Complete Point-Count Strategy, the state of the shoe after n cards have been played is determined by the proportion of high, medium and low cards present in the first n cards. To calculate the state of the shoe after $n+1$ cards have been played, it is enough to know the $(n+1)^{th}$ card and the state of the shoe at time n. The finite memory of the system makes it perfect for a Markov chain analysis framework. We will study the state of a changing shoe in isolation from the analysis of a playing strategy. You can imagine that we sit with a shoe of cards containing D decks and turn over one card at a time while we watch how the state of the remaining cards change. Once we know the equilibrium properties of a shoe as you draw cards from it, we can incorporate that information into an analysis of playing strategies.

In the Complete Point-Count Strategy, the only information about a card that matters is whether it belongs to the low, medium or high category. Consider a Markov chain (Σ, P) where each state of Σ is an ordered triple (L, M, H), representing the number of low, medium and high cards that have been played. This (assuming knowledge of the shoe size) is clearly enough to determine the current HLI, as well as the probability distribution on the category of the next card. As mentioned earlier, in D decks of cards, there are $N = 52D$ total cards, distributed as $12D$ medium cards and $20D$ each of high and low cards. The total number of states in the chain is therefore given by

$$|\Sigma| = (20D+1)(12D+1)(20D+1) = 4800D^3 + 880D^2 + N + 1.$$

Clearly $|\Sigma|$ grows as N^3. Table II shows the number of states for some example shoe sizes.

For now, each state will have only three potential transitions out. From the state representing (L, M, H) the chain can transition to $(L+1, M, H), (L, M+1, H)$ and $(L, M, H+1)$ with probabilities equal to the probability the next card drawn is a low, medium or high card, respectively. To be more explicit, if the current state is (L, M, H), the transition matrix for that row is given by

TABLE II

Summary of the number of cards and $|\Sigma|$ for some common shoe sizes discussed in this report.

D	N	$\|\Sigma\|$	$N \times N$	$\|\Sigma\| \times \|\Sigma\|$
1/4	13	144	169	20,736
1/2	26	847	676	717,409
1	52	5733	2704	3.29×10^7
2	104	42,025	10,816	1.7×10^9
4	208	321,489	43,264	1.03×10^{11}
6	312	1,068,793	97,344	1.14×10^{12}

$$\tilde{P}_{(L,M,H)(a,b,c)} = \begin{cases} \frac{(20D-L)}{R} & \text{if} \quad (a,b,c) = (L+1, M, H) \\ \frac{(12D-M)}{R} & \text{if} \quad (a,b,c) = (L, M+1, H) \\ \frac{(20D-H)}{R} & \text{if} \quad (a,b,c) = (L, M, H+1) \\ 0 & \quad \text{otherwise} \end{cases} \tag{12}$$

Note that some of these probabilities could be zero, but these are the only three possible transitions in one step. For simplicity right now, assume that the last state $(20D, 12D, 20D)$ transitions to the first state $(0, 0, 0)$ with probability one.

The current simple chain is unrealistic because it plays through the entire shoe before reshuffling back to the beginning. In a typical casino situation, a dealer will normally play through most, but not all, of a shoe before reshuffling. To eliminate most of the advantage from counting cards, the casino could reshuffle after every hand. However, this desire opposes the casino's desire to play as many hands as quickly as possible to maximize the payout from their advantage over most players. In reality, a dealer will cut $\sim$75% into the shoe and play up to this point before reshuffling.

To model the typical reshuffle point as closely centered around $\sim$75% of the shoe, we made the reshuffle point a (normalized) Laplacian random variable centered around $.75(N)$, with support over the second half of the shoe. The variance is scaled with the size of the shoe in order to keep a constant proportion of the shoe in a region with a high probability of a reshuffle. Precisely, the probability of the reshuffle point being after the n^{th} card is played is given by

$$\mathrm{Prob}\left[\text{reshuffle} = n\right] = \begin{cases} \frac{C}{\sqrt{2\sigma^2}} \exp\left\{\frac{-|n-.75(N)|}{\sqrt{\sigma^2/2}}\right\} & \text{if} \quad n \geq \lceil N/2 \rceil \\ 0 & \text{otherwise} \end{cases}$$

where $\sigma^2 = N/10$, and C is a normalizing constant to make the distribution sum to one,

$$C^{-1} = \sum_{n=\lceil N/2 \rceil}^{N} \mathrm{Prob}\left[\text{reshuffle} = n\right].$$

To translate this into the Markov chain, every state will now be allowed four possible transitions. Three possible transitions were described earlier, resulting from the drawing of a low, medium or high card. The fourth possible transition is the possibility of “reshuffling”, or transitioning back to the $(0, 0, 0)$ state.

Let $I_n \subset \Sigma$ be the set of all states such that n cards have been played, $I_n = \{(L, M, H) \in \Sigma : L + M + H = n\}$. To calculate the probability ρ_n of a reshuffle from a state $(L, M, H) \in I_n$, we must calculate the probability that the reshuffle point is n conditioned on the assumption that the reshuffle point is *at least* n,

$$P_{(L,M,H)(0,0,0)} = \rho_n = \mathrm{Prob}\left[\text{reshuffle} = n | \text{reshuffle} \geq n\right] = \frac{\mathrm{Prob}\left[\text{reshuffle} = n\right]}{\sum_{m=n}^{N} \mathrm{Prob}\left[\text{reshuffle} = m\right]}. \quad (13)$$

The probability distribution on the reshuffle location as well as the reshuffle transition probabilities are shown in Figure 1. The rest of the transition matrix is filled in with the reweighted values from the chain described in (12):

$$P_{(L,M,H)(a,b,c)} = (1 - \rho_n)\tilde{P}_{(L,M,H)(a,b,c)}, \quad \forall\ (a, b, c) \neq (0, 0, 0). \quad (14)$$

Before we can make any claims about the equilibrium of this chain, it is necessary to examine the properties of the transition matrix P. By inspection, it is clear that the chain is irreducible. From any state, it is possible to reach any other state with some non-zero probability. It is also clear that this chain aperiodic. To see this, we note that starting at state $(0, 0, 0)$, two possible return times are $\lceil N/2 \rceil$ and $\lceil N/2 \rceil + 1$. Since $\gcd(\lceil N/2 \rceil, \lceil N/2 \rceil + 1) = 1$, the period for $(0, 0, 0)$ is one, and the state is aperiodic. Because the chain is irreducible, all of the elements have the same period and (Σ, P) itself is aperiodic. The combination of irreducibility and aperiodicity give us that the chain converges to a unique equilibrium, $\lim_{n\to\infty} P^n = \pi$, where $\pi = \pi P$. Finally, it is also clear by inspection that the chain is *not* π-reversible. If $i, j \in \Sigma$ and $P_{i,j} > 0$, then we know from the properties of the chain that $P_{j,i} = 0$. Once a card has been played, it cannot be taken back.

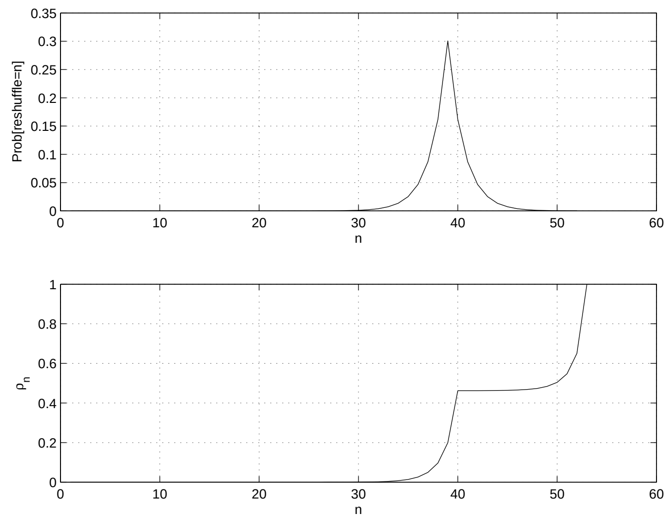


Fig. 1. *Reshuffle point PDF and state reshuffle prob. for 1 deck.*

V. DECK EQUILIBRIUM

A. Analytic calculations of shoe equilibrium

To evaluate the effectiveness over the long run of a playing strategy that depends on the state of the shoe, we must be able to determine the relative proportion of time the shoe is in favorable or unfavorable states. In other words, we must be able to calculate or estimate the equilibrium distribution of the shoe Markov chain described in section IV-C. Furthermore, in order for the results to be most applicable to real game situations, we must be able to analyze multiple deck games (at least two decks, and preferably four or six). Equations (13) and (14) give an explicit expression for the transition matrix P. Knowing P, the unique equilibrium can be solved analytically as

$$\pi = (1, 1, \ldots, 1)(I - P + E)^{-1}, \tag{15}$$

where I and E are the $|\Sigma| \times |\Sigma|$ identity and ones matrices, respectively [4]. Referring back to Table II, we see that even for one deck shoe, P would have 33 million entries. To store P as a matrix of 8-byte, double-precision floating point numbers, it would require approximate 263MB of memory. To analyze a two deck shoe, P would require approximately 13.6GB of memory. Aside from the issue of storing P in memory, one would also need to create the I and E matrices, and then invert $(I - P + E)$. Clearly, this is a situation in which we have perfect knowledge of local transitions, but it is impossible to deal with P as a whole. In practice, using MATLAB on a Pentium III PC with 512MB of memory, we can calculate π through direct matrix inversion for 1/4, 1/2 and 3/4 deck shoes, but not for anything larger.

If we cannot use the direct inversion of equation (15) to analytically determine π, we could turn to simulation methods. The ergodic theorem tells us that if we let the walk run, it will asymptotically converge to the equilibrium distribution. Once a walk is sufficiently well-mixed (within some error tolerance), we could stop it and take the final state to be *one* sample from π. Alternately, we could use a technique such as "coupling from the past" to draw samples exactly from π. However, a histogram estimator over an alphabet with $|\Sigma|$ entries requires many samples. To get estimates that match (with reasonable probability) the true distribution with moderate error, we calculated that we would need on the order of 10^7 samples in the $D = 1$ case and 10^9 samples in the $D = 4$ case [5]. Considering the convergence bounds available to us (discussed in section V-B), estimating π through simulation could be very computationally intensive.

Looking more carefully at the Markov chain we have constructed, there is a great deal of structure. The form of the chain is more clear in a graphical representation. Imagine that we arranged all of the states so that states representing the same number of cards played (belonging to the same set I_n) are in the same column, and each column represents one more card played than in the previous column (depicted graphically in Figure 2). Note that when a card is played, a state in I_n can only move to a state in I_{n+1}. Only the states in I_n for $n \geq \lceil N/2 \rceil$ are capable of causing a reshuffle (transitioning back to the $(0, 0, 0)$ state), and each state in I_n reshuffles with the same probability (ρ_n). Columns closer to the midpoint of the shoe contain more states, and the first and last columns taper down to one state in each ($|I_0| = |I_N| < |I_1| = |I_{N-1}| < |I_2| \ldots$). The fan out to many states followed by a taper back down to one state happens because there are many valid ways to make valid triples representing $\lceil N/2 \rceil$ cards played, but the $(0, 0, 0)$ and $(20D, 12D, 20D)$ states are the only ways to have played 0 and N cards, respectively.

Starting at state $(0, 0, 0)$ a walk will take one particular path from left to right, moving one column with each step and always taking exactly $\lceil N/2 \rceil$ steps to reach the midpoint. After the midpoint, the states can also reshuffle at each step with a probability ρ_n that *only* depends on the the current column and not on the path taken up to that point or even the exact current state within the column. Essentially, this structure allows us to separate the calculation of π into

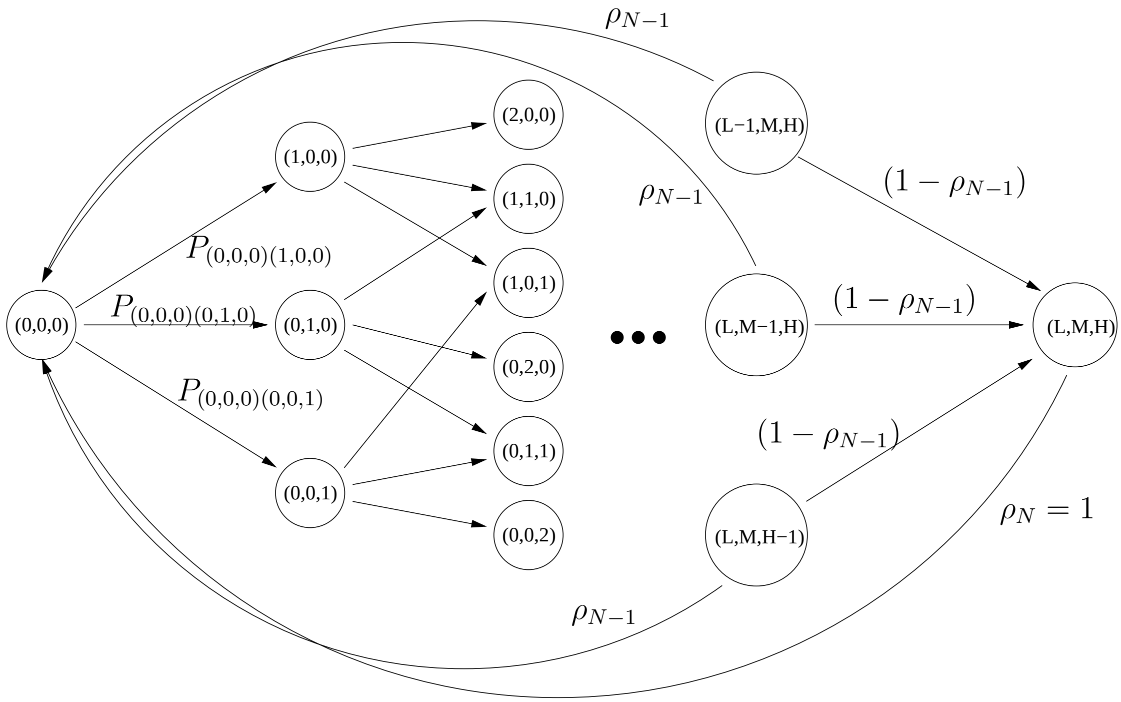


Fig. 2. *Graphical depiction of full state space* Σ.

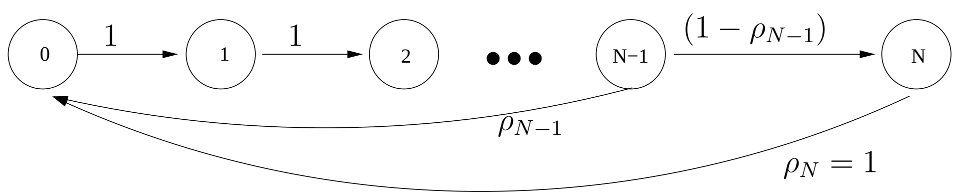


Fig. 3. *Graphical depiction of reduced column state space* Γ.

two components: how much relative time is spent in each column, and within each individual column, what proportion of time is spent in each state. To investigate the relative time spent in each column, we create a reduced chain (Γ, Q) where each column in Figure 2 is represented by one state in Γ (depicted graphically in Figure 3). The transition matrix is given by

$$Q_{n,m} = \begin{cases} (1-\rho_n) & \text{if} \quad m = n+1 \\ \rho_n & \text{if} \quad m = 0 \\ 0 & \quad\text{otherwise} \end{cases} \tag{16}$$

It is clear that the chain (Γ, Q) is also irreducible, aperiodic and not reversible. From this, we know that the chain does converge to a unique equilibrium μ, representing the relative proportion

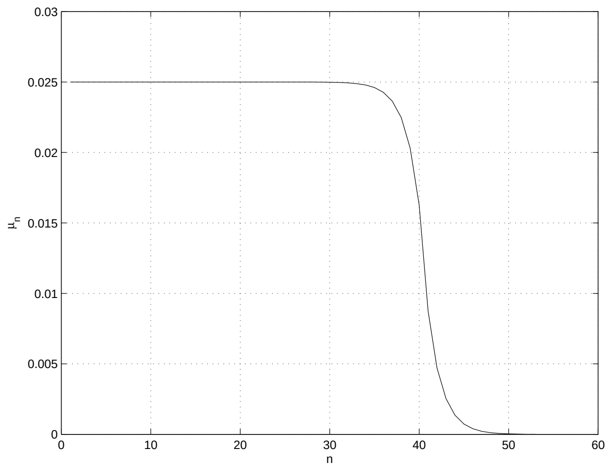


Fig. 4. μ *for* $D = 1$.

of time that the original chain (Σ, P) spent in each column of Figure 2. This is stated more precisely as

$$\mu_n = \sum_{(L,M,H)\in I_n} \pi_{(L,M,H)}. \tag{17}$$

Figure 4 shows μ for $D = 1$. Importantly, the dimension of the reduced column-space chain is much smaller than the original chain, with $|\Gamma| = N$ and $|\Sigma| = O(N^3)$. Even in the case when $D = 6$, the direct inversion calculation of

$$\mu = (1, 1, \ldots, 1)(I - Q + E)^{-1}$$

is easily done in minutes. It is also important to note here that because $|I_0| = 1$, $\pi_0 = \mu_0$.

The structure of the Markov chain allows us to compute π once we have μ. Using the relation $\pi = P\pi$, we observe that

$$\pi_{(L,M,H)} = \sum_{k\in\Sigma} \pi_k P_{k,(L,M,H)}. \tag{18}$$

Suppose $(L, M, H) \in I_n$ with $n > 0$. The only states that transition to (L, M, H) are contained in I_{n-1}, and so we have

$$\pi_{(L,M,H)} = \sum_{k\in I_{n-1}} \pi_k P_{k,(L,M,H)}. \tag{19}$$

Because we know that $\pi_0 = \mu_0$, we are able to compute π_1, followed by π_2, and so on. With knowledge only of π_0, we are able to completely determine the equilibrium π. The technique

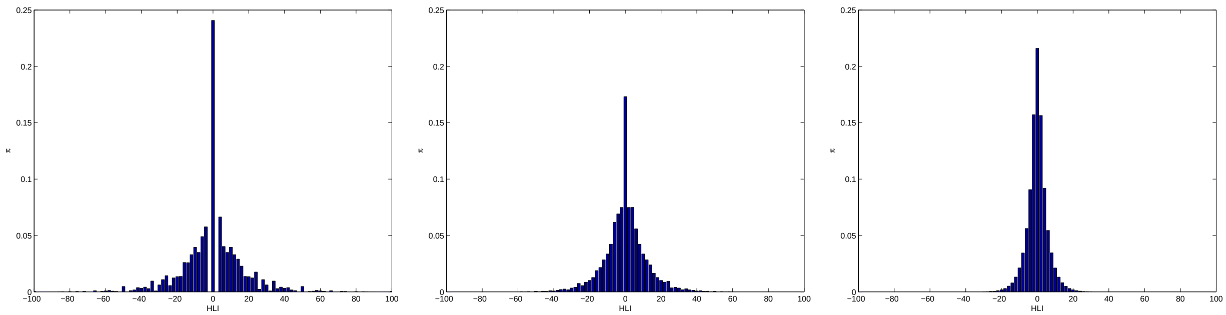


Fig. 5. *Equilibria for 1/2, 1 and 4 deck shoes.*

described here takes advantage of the rich structure in the chain to analytically calculate π exactly using only local knowledge of P, and the inversion of a $|N| \times |N|$ matrix. The technique gives numerically identical results (accurate up to the precision of the computer) to equilibrium calculated through the direct inversion in (15) for 1/4 and 1/2 deck shoes. The algorithm can calculate the equilibrium when $D = 6$ in under an hour, making it *much* more efficient and accurate than estimation through simulation. Equilibria calculated through this method for $D = \{1/2, 1, 4\}$ are shown in Figure 5. Because the equilibrium would be difficult to plot in the three dimensional state space Σ, states with the same HLI are combined and the equilibria are plotted vs. HLI.

B. Convergence bounds and actual mixing rates

Even though we have a method for analytically calculating π exactly for the cases of interest to us, we are still interested in investigating the mixing rate of the chain (Σ, P). If we analyze the player advantage according to the equilibrium distribution on the state of the deck, it is important to know how long a player would have to wait in order for the equilibrium assumption to be reasonably valid. For an irreducible, aperiodic chain with an K such that $P^K > 0$, we have two known bounds on the convergence speed:

$$|P_{i,j}^n - \pi_j| \leq (1 - 2\epsilon)^{(n/K)-1} \tag{20}$$

$$|P_{i,j}^n - \pi_j| \leq (1 - |\Sigma|\epsilon)^{(n/K)-1}, \tag{21}$$

where $\epsilon = \min_{i,j}\left(P_{i,j}^K\right)$. It should be noted that calculating ϵ will be very difficult when P is so large that we are unable to construct or manipulate it. Because (Σ, P) is not reversible, eigenvalue techniques for bounding convergence speed cannot be applied. Coupling techniques could be applied to (Σ, P) and indeed constructing a suitable coupling is not difficult. However, we were unable to find a tractable bound for the expected coupling time so we were unable to bound the convergence speed using coupling.

For $D = 1/2$ (and no larger) we can construct and exponentiate P. Also calculating π as described in section V-A, we can compute the mixing rate exactly and compare it to the bounds in (21). To get $P^K > 0$, we need $K \geq 41$. However, with $K = 41$, ϵ is *very* small. The bounds in (21) hold for any K such that $P^K > 0$, and by choosing $K = 100$ we can get a significant improvement in the bounds. However, it is clear that these bounds are not at all tight. According to the bounds in equation (21), to guarantee that $|P^n_{i,j} - \pi_j| < .1$, we need to wait for $n \geq 10^{12}$ cards to be played. Direct calculation shows us that $|P^n_{i,j} - \pi_j| < .01$ for $n \approx 400$ cards! In order to compute a tighter bound for $|P^n_{i,j} - \pi_j|$, we focus once again on the column structure of the state space.

Suppose $n > N + 1$, and let $i, j \in \Sigma$ be ordered triples. We define c_i to be the column index of triple i; that is, $i \in I_{c_i}$. We wish to investigate the behavior $|P^n_{i,j} - \pi_j|$. Due to our reshuffle scheme, a path from i to j in n steps must involve a reshuffle after precisely $n - c_j$ steps. Therefore we have

$$P^n_{i,j} = P^{n-c_j}_{i,0} P^{c_j}_{0,j}. \tag{22}$$

Note that the first term is dependent only on the reshuffle probabilities, so we have

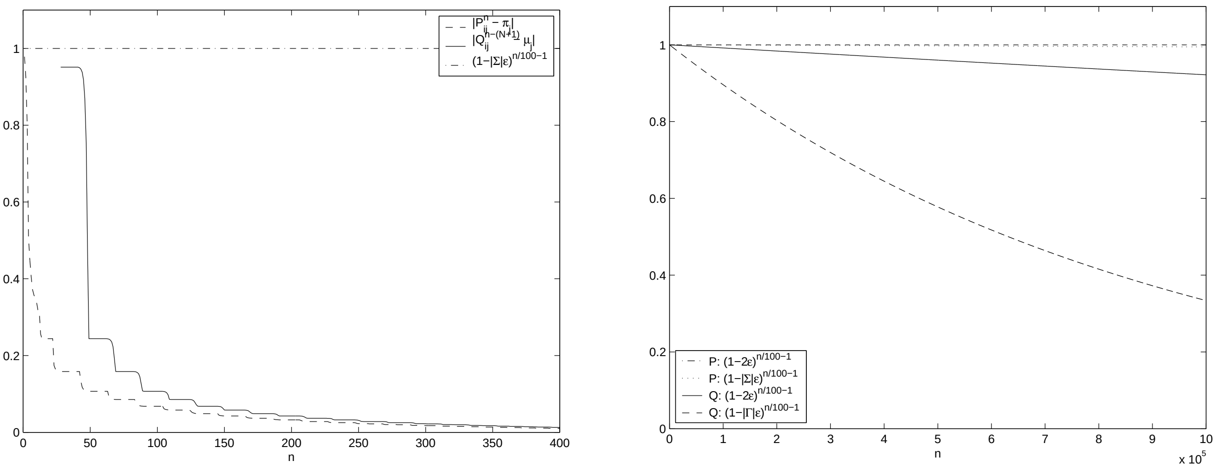
$$P^{n-c_j}_{i,0} = Q^{n-c_j}_{c_i,0}. \tag{23}$$

Also, it follows from recursively applying (19) that

$$\pi_j = \pi_0 P^{c_j}_{0,j}. \tag{24}$$

Therefore we have

$$\begin{aligned}
\left|P^n_{i,j} - \pi_j\right| &= \left|Q^{n-c_j}_{c_i,0} P^{c_j}_{0,j} - \pi_0 P^{c_j}_{0,j}\right| && (25)\\
&= P^{c_j}_{0,j} \left|Q^{n-c_j}_{c_i,0} - \pi_0\right| && (26)\\
&\leq \left|Q^{n-c_j}_{c_i,0} - \mu_0\right|. && (27)
\end{aligned}$$



(a) (b)

Fig. 6. *Good and bad convergence bounds for* $D = 1/2$.

We see that a convergence bound for (Σ, P) is closely related to a convergence bound for (Γ, Q). Continuing, we have

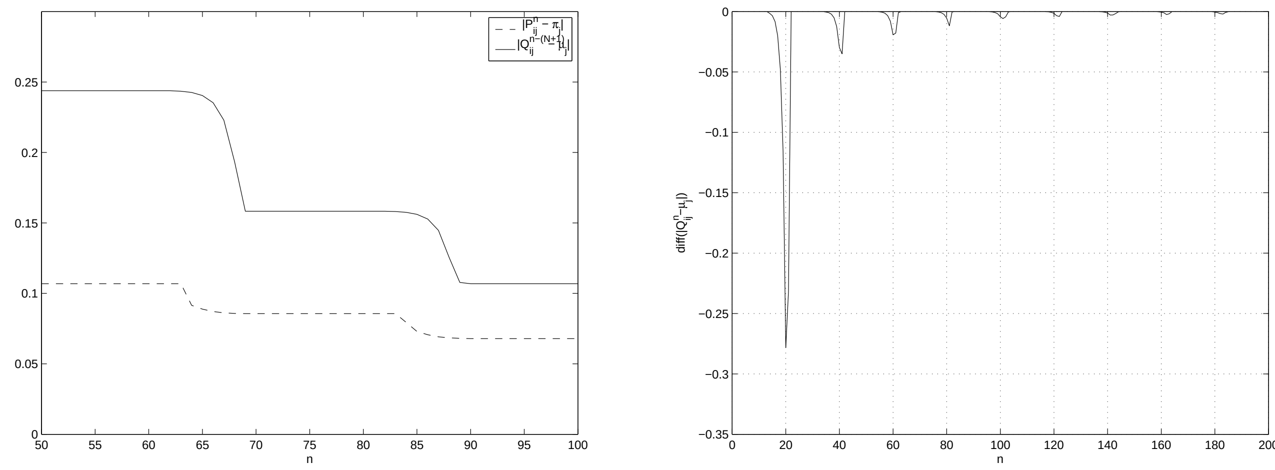
$$\left|P_{i,j}^{n} - \pi_j\right| \leq \max_{k,l}\left|Q_{k,l}^{n-c_j} - \mu_l\right| \tag{28}$$

$$\leq \max_{k,l}\left|Q_{k,l}^{n-(N+1)} - \mu_l\right| \tag{29}$$

where the last step follows because we observe that the convergence is nonincreasing. The bound achieved in (29) (under the assumption of (29) not increasing with n) is significant because though we cannot exponentiate P, we can exponentiate Q (for any reasonable number of decks).

Figure 6(a) plots $\max_{k,l}\left|Q_{k,l}^{n-(N+1)} - \mu_l\right|$ as a bound for $|P_{i,j}^n - \pi_j|$. On this plot, we also show the actual deviation of P^n from π. Our bound using $Q^{n-(N+1)}$ is rather close. As a stark comparison, we show on the same plot our best bound that results from (21) with $K = 100$. This bound is approximately equal to 1 for all interesting values of n. By exploiting the column structure, we have improved our bound on running time by approximately 10 orders of magnitude!

We briefly mention a few interesting facts about the convergence of the chain (Γ, Q). For the sake of completeness, we plot in Figure 6(b) the bounds for $|Q_{i,j}^n - \mu_j|$ that result from the analysis of (20) and (21). These bounds converge slightly faster than the corresponding bounds



(a) (b)

Fig. 7. *Interesting bound results, $D = 1/2$.*

for $|P^n_{i,j} - \pi_j|$. We can, of course, compute Q^n directly, and we observe that it still converges much more quickly than the bounds indicate. Figure 7(a) shows a close-up zoom of Figure 6(a). It is quite interesting to note that the convergence of these Markov chains occurs in rather abrupt stages. Figure 7(b) plots the (discrete) derivative of the error $\max_{i,j} |Q^n_{i,j} - \mu_j|$. We observe that these changes are nonpositive (this was necessary to obtain (29)), and also that the changes are periodic with period of roughly $n = 20$ cards, or 75% of the size of the shoe. Computing $|Q^n_{i,j} - \mu_j|$ for other values of D, we see that the periodicity is always approximately $\frac{3}{4} \cdot 52D$. Though we have no precise explanation here, we believe that the behavior is intimately tied to the reshuffling strategy of our model. The constant segments in Figure 7(a) have a width that corresponds to $\lceil N/2 \rceil$ and Figure 7(b) has a period that is the expected reshuffle time, $.75N$. The constant segments in Figure 7(a) suggest to us that the chain we have developed is almost periodic, and the only real mixing occurs because of a reshuffle. The reader interested in observing connections between the mixing behavior seen in Figure 7(a) and the reshuffling scheme is referred back to Figure 1, upside-down and held backwards up to the light.

Finally, because we cannot compute P^n directly for the case $D = 4$, we plot in Figure 8 the bounds that arise from (20), (21), and (29). Notice here that the bound from (29) is roughly 12 orders of magnitude better than the bounds that would be available to us from the general results on Markov Chains.

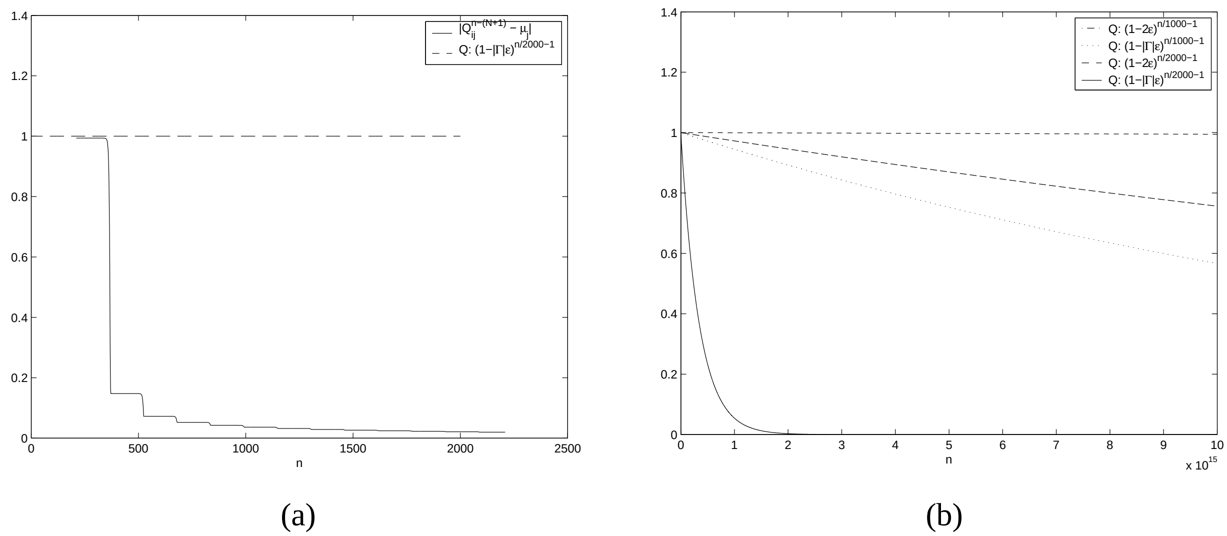


Fig. 8. *Good and bad convergence bounds for $D = 4$.*

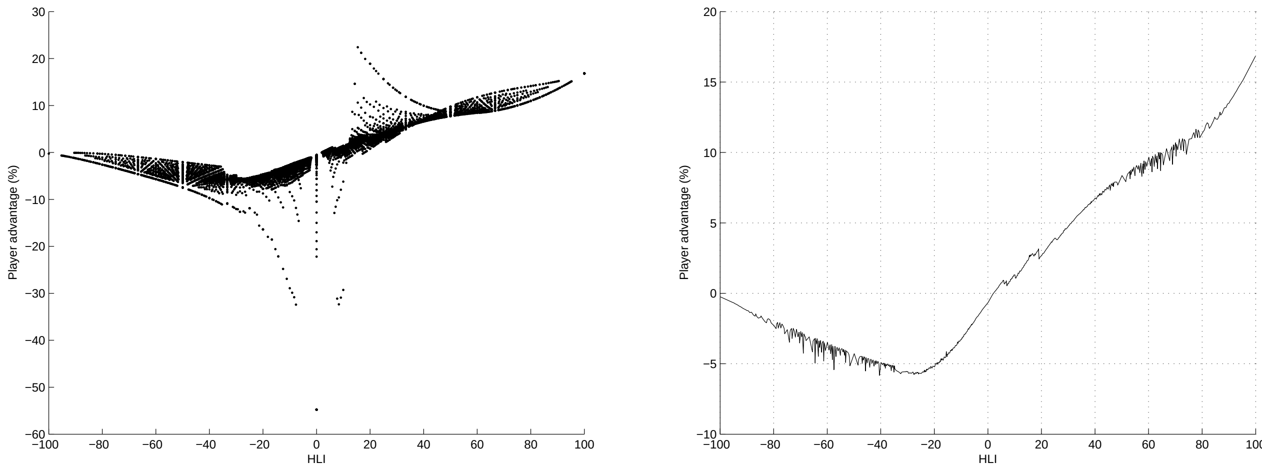
VI. Analysis

We present in this section our analysis of the Complete Point-Count System. Because the betting is not constant in this system, it is important now to distinguish between the player's *advantage* and the player's expected *gain*. The player's *expected gain* G is defined as the dollar amount he expects to profit from a single hand when betting according to (6). The player's advantage A is the percent of his bet he expects to win:

$$A = 100 \cdot \frac{G}{B}. \tag{30}$$

A. Player advantage vs. HLI

For a game with $D = 1$ deck, we use the method described in Section IV-B to compute the player's advantage for each possible triple (L, M, H). Figure 9(a) plots the player's advantage against the corresponding HLI for each possible triple in the 1-deck game (assuming for the moment that the player does not take insurance). It is interesting to note that a single value of HLI may correspond to several possible deck states; these states may, in turn, correspond to widely varying advantages for the player. Some of these triples may be highly unlikely, however. To get a better feel for the relation between HLI and player advantage, we use the analytic method of Section V-A to compute the equilibrium distribution of the states. Figure 9(b) shows a plot



(a) (b)

Fig. 9. *(a) Player advantages generally increase with HLI, with some outliers ($D = 1$). (b) Average player advantage given state HLI.*

where we use the relative equilibrium time to average all triples that correspond to the same HLI value.

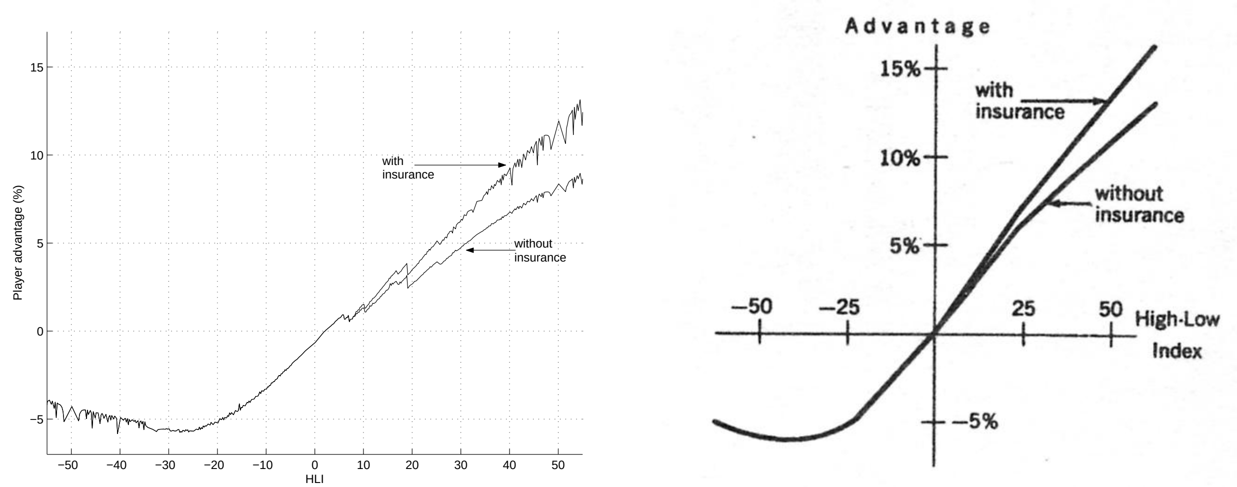
As expected, the player's advantage generally increases with HLI, and the player is at a disadvantage when HLI is negative. Surprisingly, though, as the HLI approaches -100, the player's disadvantage diminishes.[4] Figure 10 focuses on a more typical range of HLI values, and includes the player's advantage when playing with insurance. For comparison purposes, we include the corresponding plot that appears in Thorp's description of the Complete Point-Count System [1].

B. Expected gains

Figure 11 shows the average amount of time the player expects to play with different advantages (we assume $D = 1$ and that the player plays with insurance). The player spends a considerable amount of time in states with a disadvantage. In fact, if the player placed a unity bet on every hand, he would play at a disadvantage of 0.64%.

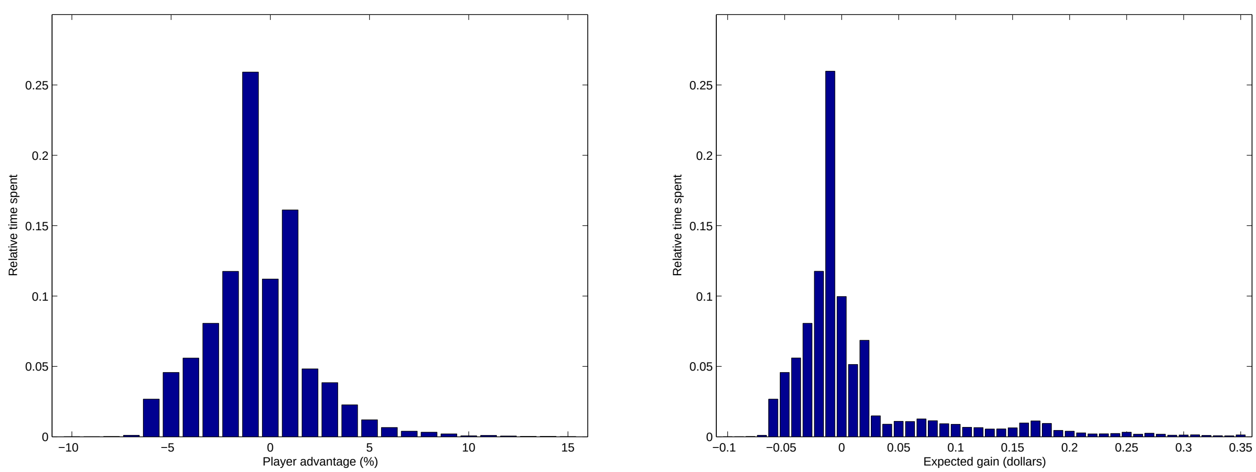
Adjusting the bets is key to the player's hope for a positive expected gain. Figure 11 also shows the average amount of time the player expects to play with different expected gains. The

[4]This is also mentioned in [1], although Thorp claims a significant advantage for the player as $HLI \rightarrow -100$.



(a) (b)

Fig. 10. *Player advantage vs. HLI. (a) Our analysis. (b) Thorp's result [1].*



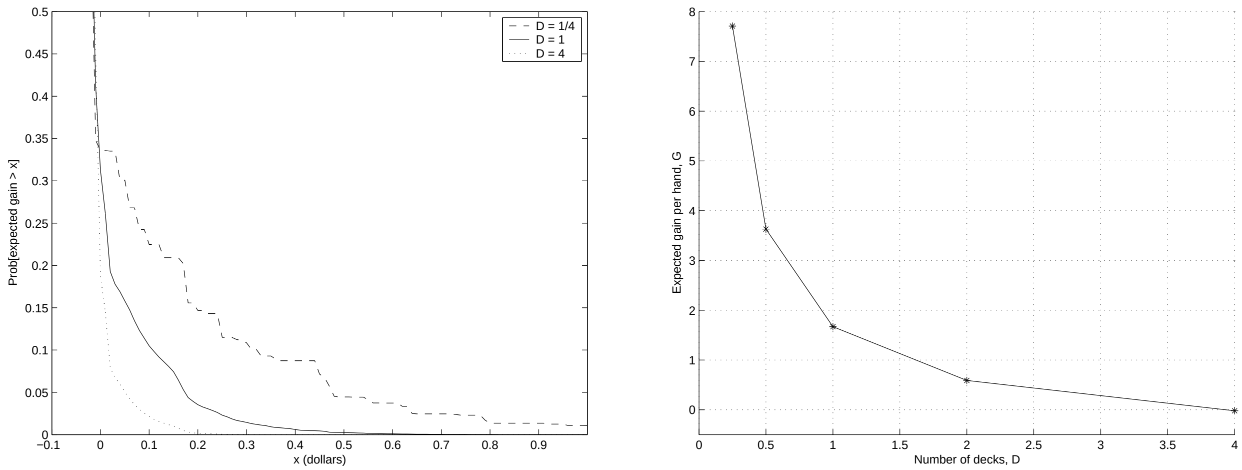
(a) (b)

Fig. 11. *(a) Relative time spent with different player advantages ($D = 1$). (b) Relative time spent with different expected gains.*

larger bets that are placed when the shoe is favorable (according to (6)) allow the player to win more money in those states. We compute, in fact, that the player plays with an expected gain of $G = 0.0167$, or 1.67 cents per hand.

C. Dependency on deck size

Not surprisingly, the number of decks in play has a direct impact on the player's expected gain. We notice, however, that plots such as Figure 9 change little as the number of decks



(a) (b)

Fig. 12. *(a) Time in favorable states depends on number of decks. (b) Expected player gain (per hand) using Complete Point-Count System.*

changes, indicating that HLI is a universally good indicator of the player's advantage. As we observed in Figure 5, however, the relative amount of time spent in each HLI depends strongly on the number of decks in play. As we saw in the previous section, much of the player's gain comes from placing large bets when HLI is large. With more decks in play, he is less likely to encounter these extreme situations. Figure 12 illustrates this dependency and plots the player's overall expected gain, as the number of decks changes.

VII. CONCLUSIONS

Blackjack is a non-trivial game, and any precise, analytic statements about a player's odds when using a particular strategy would be overwhelming (if not impossible) to calculate without the framework of a Markov chain analysis. We have seen in this work that under a very few mild simplifying assumptions, the long-term advantage and expected winnings of a playing strategy can be completely determined using the equilibrium analysis of a combination of Markov chains. Though blackjack is only a casino game, our exercise illustrates the power of Markov chains in analyzing complicated, real-world problems.

Our basic strategy analysis is a simple application of a Markov chain based on the deterministic choices made by the player. We observe that the infinite shoe assumption is critical to this analysis because it allows us a tractable method for calculating the distributions on the absorbing states.

In our analysis of the Complete Point-Count system, we deal with the situation where the shoe is finite. Our "locally infinite" assumption, however, allows us to compute the player's approximate advantage for each state. By computing the average time the player spends in each state, we are able to compute his overall advantage.

Our Markov chain to model the High-Low Index contains a large number of states, and only through our knowledge of its column structure can we perform a precise analysis. This paper truly highlights the importance of exploiting the known structure of the specific Markov chain under analysis. Perhaps the most powerful example we provide are the bounds for convergence to equilibrium. Using our column-structure analysis, we obtain bounds that are immensely more useful than the bounds applicable to general Markov chains. Though taking advantage of the structure can be a big win in achieving better bounds, it is sometimes difficult (or impossible) to do so. This is illustrated by the inapplicability of an eigenanalysis (due to the non-reversibility) and the intractability of a coupling analysis.

The mixing bounds we obtain using a column analysis show that the player can expect the shoe to approach equilibrium within a reasonable amount of time. The 1000 or so cards the player needs to observe is quite a few, but is also easily achievable in a few hours. In practical terms, the expected gain provided by the Complete Point-Count System is subtle (compared to, say, the Basic Strategy). It does allow the player, however, to be on the lookout for the occasional highly favorable deck. It is in these situations, when the player increases his bet, that the card-counter's game can truly be both entertaining and profitable.

References

[1] E. O. Thorp, *Beat the Dealer: A Winning Strategy for the Game of Twenty-One*. New York: Vintage, 1966.

[2] S. Wong, *Professional Blackjack*. La Jolla, CA, USA: Pi Yee Press, 1994.

[3] P. A. Griffen, *The Theory of Blackjack: The Compleat Card Counter's Guide to the Casino Game of 21*. Las Vegas, NV, USA: Huntington Press, 1996.

[4] J. R. Norris, *Markov Chains*. Cambridge University Press, 1997.

[5] T. Cover and J. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 1991.