Combinatorics, Probability and Computing (2022), **31**, pp. 1–20
doi:10.1017/S0963548321000134

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Card guessing with partial feedback

Persi Diaconis[1,†], Ron Graham[2] , Xiaoyu He[3,‡] and Sam Spiro[4,*,ᘔ]

[1]Department of Mathematics and Statistics, Stanford University, Stanford, CA 94305, USA; [2]Department of Computer Science and Engineering, UCSD, La Jolla, CA 92093, USA; [3]Deparment of Mathematics, Stanford University, Stanford, CA 94305, USA and [4]Department of Mathematics, UCSD, La Jolla, CA 92093, USA
*Corresponding author. Email: sspiro@ucsd.edu

(Received 19 October 2019; revised 23 March 2021; accepted 25 April 2021; first published online 15 June 2021)

Abstract

Consider the following experiment: a deck with m copies of n different card types is randomly shuffled, and a guesser attempts to guess the cards sequentially as they are drawn. Each time a guess is made, some amount of 'feedback' is given. For example, one could tell the guesser the true identity of the card they just guessed (the complete feedback model) or they could be told nothing at all (the no feedback model). In this paper we explore a partial feedback model, where upon guessing a card, the guesser is only told whether or not their guess was correct. We show in this setting that, uniformly in n, at most $m + O(m^{3/4} \log m)$ cards can be guessed correctly in expectation. This resolves a question of Diaconis and Graham from 1981, where even the $m = 2$ case was open.

Keywords: Card Guessing, Shuffling, Games

2020 MSC Codes: 60C05

1. Introduction

Let $\mathfrak{S}_{m,n}$ be the set of words π over the alphabet $[n] := \{1, 2, \ldots, n\}$ where each character in $[n]$ appears exactly m times in π. We think of π as some way to shuffle a deck of cards which has m suits and n card types. For example, a standard deck of 52 cards has $n = 13$ values (Ace, Two, …, King), each appearing $m = 4$ times. We find it helpful to think that m is for multiplicity and n is for number of values. We refer to the elements of $\mathfrak{S}_{m,n}$ as permutations, even though for $m > 1$ this is technically not the case. If X is a finite set, we write $x \sim X$ to indicate that x is chosen uniformly at random from X.

Consider the following experiment: a deck with m copies of n different card types is randomly shuffled according to some $\boldsymbol{\pi} \sim \mathfrak{S}_{m,n}$, and a guesser attempts to guess each card as it is drawn, and the drawn card is discarded after the guess is made (i.e. this is sampling without replacement). Each time a guess is made, some amount of 'feedback' is given. For example, one could tell the guesser the true identity of the card they just guessed (the complete feedback model) or they could be told nothing at all (the no feedback model). This can also be viewed as a one player game where

[†]Research supported by NSF Grant DMS-1954042.
[‡]Research supported by NSF Graduate Research Fellowship Grant No. DGE-1656518.
[ᘔ]Research supported by NSF Graduate Research Fellowship Grant No. DGE-1650112.

the guesser tries to either maximise or minimise the number of times their guesses are correct, and we will often refer to these models as games.

These sorts of models were considered by Blackwell and Hodges [2] and Efron [10] in relation to clinical trials. Here, aiming for a fixed number of subjects, say 100, in a medical trial comprising say 4 treatments, a deck of 100 cards with 25 labelled with each treatment is prepared. Subjects are assigned to treatments as they come into the clinic, sequentially, using the next card (which is then discarded). Hospital staff has the option of ruling subjects ineligible. If the staff has strong opinions about the efficacy of treatments and observes which treatments have been given out, they may guess what the next treatment is and bias the experiment by ruling a sickly subject ineligible. It is clearly of interest to be able to evaluate the expected potential bias.

Card guessing is also a mainstay of classical experiments to test 'Extra Sensory Perception' (ESP). The most common experiment utilises a deck of 25 cards where there are 5 copies of 5 different types of cards (so $m = n = 5$ in our language) where the subjects iteratively try and guess the identity of the next card, and experimenters routinely give various kinds of feedback to enhance 'learning'. Diaconis [6] and Diaconis and Graham [7] give a review of these problems.

In the no feedback model every strategy guesses m cards correctly in expectation. The distribution of correct guesses depends on the guessing strategy: if the guesser always guesses the same card type then the variance is 0, and it can be shown that the variance is largest if the guesser uses a permutation of the mn values, see [7].

The complete feedback model is more complicated, but optimal strategies were determined in [7]. Given a strategy $\mathcal{G}$ for the guesser, let $C(\mathcal{G}, \pi)$ denote the number of correct guesses the guesser gets in the complete feedback model if they use strategy $\mathcal{G}$ and the deck is shuffled according to π. Let $\mathcal{C}^{+}_{m,n} = \max_{\mathcal{G}} \mathbb{E}[C(\mathcal{G}, \boldsymbol{\pi})]$, where $\boldsymbol{\pi} \sim \mathfrak{S}_{m,n}$ and the maximum ranges over all possible strategies $\mathcal{G}$. Similarly define $\mathcal{C}^{-}_{m,n} = \min_{\mathcal{G}} \mathbb{E}[C(\mathcal{G}, \boldsymbol{\pi})]$. The following is proven in [7].

Theorem 1.1. [7]. *If $\mathcal{G}^{+}$ (respectively $\mathcal{G}^{-}$) is the strategy where one guesses a most likely (respectively least likely) card at each step, then $\mathcal{C}^{\pm}_{m,n} = \mathbb{E}[C(\mathcal{G}^{\pm}, \boldsymbol{\pi})]$. Moreover,*

$$\mathcal{C}^{\pm}_{m,n} = m \pm M_n \sqrt{m} + o_n(\sqrt{m}),$$

where $M_n = \Theta(\sqrt{\log n})$ is the expected maximum value of n independent standard normal variables.

One can also consider variants of these models where $\boldsymbol{\pi}$ is chosen according to some non-uniform distribution. For $m = 1$ the case when π is obtained from a single riffle shuffle is studied by Ciucu [5] (no feedback), and Liu [12] (complete feedback). Analysis under repeated 'top to random shuffles' is done by Pehlivan [13]. We emphasise that for our results, we only consider the uniform distribution $\boldsymbol{\pi} \sim \mathfrak{S}_{m,n}$.

The main focus in this paper is on a feedback model called the partial feedback model, which returns an intermediate amount of information to the guesser. After each guess, the guesser is only told whether their guess was correct or not (and thus not the identity of the card if they were incorrect). This feedback protocol was recommended when conducting ESP trials and is a natural notion of bias if card guessing experiments are performed with experimenter and subject in the same room. Given a strategy $\mathcal{G}$ for this game, let $P(\mathcal{G}, \pi)$ denote the number of cards the guesser guesses correctly using strategy $\mathcal{G}$ if the deck is shuffled according to π, and define $\mathcal{P}^{+}_{m,n} = \max_{\mathcal{G}} \mathbb{E}[P(\mathcal{G}, \boldsymbol{\pi})]$ and $\mathcal{P}^{-}_{m,n} = \min_{\mathcal{G}} \mathbb{E}[P(\mathcal{G}, \boldsymbol{\pi})]$ for the maximum and minimum expected number of correct guesses possible, respectively.

The partial feedback model is significantly more difficult to analyse than the other two models, and relatively little is known about it. This is in large part due to the fact that we do not understand the optimal strategy in this game, and in particular it is not the case that the strategy $\mathcal{G}^{+}$ of guessing a maximum likelihood card satisfies $\mathbb{E}[P(\mathcal{G}^{+}, \boldsymbol{\pi})] = \mathcal{P}^{+}_{m,n}$ for $m \geq 2$, see [7]. We note that bounds for the strategy $\mathcal{G}^{+}$ in this model were studied recently by Gural, Simper, and So [11].

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Define $\mathcal{N}^{\pm}_{m,n}$ for the no feedback model analogous to how $\mathcal{C}^{\pm}_{m,n}$ and $\mathcal{P}^{\pm}_{m,n}$ were defined, and note that $\mathcal{N}^{\pm}_{m,n} = m$. One can easily show that $\mathcal{N}^{+}_{m,n} \le \mathcal{P}^{+}_{m,n} \le \mathcal{C}^{+}_{m,n}$ for all m and n, with the reverse inequalities holding for $-$ instead of $+$. In particular, by Theorem 1.1 and the fact that $\mathcal{N}^{\pm}_{m,n} = m$, we obtain $\mathcal{P}^{\pm}_{m,n} = (1 + o(1))m$ as m goes to infinity, for any fixed n. Motivated by this, our focus for this paper will be in bounding $\mathcal{P}^{\pm}_{m,n}$ when m is fixed and n is large. As a point of comparison, we first establish the value of $\mathcal{C}^{\pm}_{m,n}$ in this regime. Here and throughout we let log denote the natural logarithm.

Theorem 1.2. *For m fixed and $n \to \infty$, we have*

$$\mathcal{C}^{+}_{m,n} = (1 + o(1))H_m \log n,$$

where $H_m = \sum_{i=1}^{m} j^{-1}$ is the mth harmonic number, and

$$\mathcal{C}^{-}_{m,n} = \Theta(n^{-1/m}).$$

The case $m = 1$ of Theorem 1.2 was proved in [7], and we give its simple proof to provide some intuition. Assuming we always guess a card that is in the deck, the chance of getting the first guess correct is $1/n$, then the second is $1/(n-1)$, and so on. Thus the expected value is exactly $1 + 1/2 + \cdots + 1/n = \log n + O(1)$ as claimed.

Theorem 1.2 shows that for any fixed m, in expectation the guesser can achieve arbitrarily many or few correct guesses as n grows in the complete feedback model. In sharp contrast, we show that the guesser cannot obtain arbitrarily many correct guesses in the partial feedback model.

Theorem 1.3. *If n is sufficiently large in terms of m, we have*

$$\mathcal{P}^{+}_{m,n} = m + O(m^{3/4} \log^{1/4} m).$$

This resolves a 40-year-old problem of Diaconis and Graham [7], which was open even for $m = 2$ (i.e. a deck with composition $\{1, 1, 2, 2, \ldots, n, n\}$). In particular, this shows that the information from the partial feedback model is not enough for the guesser to correctly guess asymptotically more cards compared to when they are given no feedback at all. We suspect that the error term in Theorem 1.3 can be improved to $m^{1/2+o(1)}$, which would be best possible; see the discussion in Section 4.

We conclude this introduction with some brief remarks about the related literature. In the partial feedback model, the enumeration of the number of permutations consistent with a given sequence of guesses can be reduced to the evaluation of certain permanents, see Chung, Diaconis, Graham, and Mallows [4] and Diaconis, Graham, and Holmes [8]. These papers contain applications to the partial feedback model, as well as a fascinating 'persistence conjecture': whenever the guesser guesses a card type i incorrectly, it is optimal for them to continue to guess i in the next step.

Throughout, we focused on evaluating the expected number of correct guesses. The distribution of the number of correct guesses is treated in [7], see also Proschan [14]. A variety of other feedback mechanisms have also been explored, such as less feedback if the guesser is doing well, and telling the guesser that their guess is 'high' or 'low', see Samaniego and Utts [15].

Our evaluation for these models gives one point for each card guessed correctly. It is also natural to consider weighted scores: a correct guess early on might be weighted more heavily than a correct guess towards the end since more information is available to the guesser later on. This is known as skill scoring and is discussed in [7] and Briggs and Ruppert [3].

Organisation. This paper is organised as follows. In Section 2, we prove Theorem 1.2 by analysing the number of correct guesses made by the maximum (or minimum) likelihood guessing strategy, which is guaranteed to be optimal by Theorem 1.1. In Section 3, we prove our main result Theorem 1.3 that one cannot do much better than randomly guessing in the partial feedback

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

model. The key ingredient is Lemma 3.1, which shows that in 'typical' game states, no matter what card the guesser guesses the probability of guessing correctly is at most $(1 + o(1))n^{-1}$. Finally, in Section 4 we make some concluding remarks, highlighting some of the many open problems left in this area.

2. The complete feedback model

In this section we prove Theorem 1.2. Throughout this section we treat m as a fixed value, and hence the implicit constants in our asymptotic notation are allowed to depend on m.

Lemma 2.1. *For $1 \le j \le m$, let T_j be the smallest value t such that $\boldsymbol{\pi}_t$ is the jth occurrence of some card type. Equivalently, it is the largest t such that $\{\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_{t-1}\}$ contains no card type with multiplicity at least j. If $t = \gamma n^{1-1/j}$ (here $\gamma > 0$ may depend on n), we have*

$$\mathbb{P}[T_j > t] = 1 - O(\gamma^j),$$
$$\mathbb{P}[T_j > t] = O(\gamma^{-j}).$$

We postpone this proof for the moment and show how this implies the result.

Proof of Theorem 1.2. We start with the proof of the bounds on $\mathcal{C}^-_{m,n}$ (the lowest expected number of correct guesses possible with complete feedback), and recall from Theorem 1.1 that this equals $\mathbb{E}[C(\mathcal{G}^-, \boldsymbol{\pi})]$ where $\mathcal{G}^-$ is the strategy of guessing a least likely card at each stage. Let Y_t denote the indicator function for successfully guessing $\boldsymbol{\pi}_t$ and let J_t denote the largest multiplicity of a card type appearing in $\{\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_{t-1}\}$. Because there are $mn - t + 1$ total cards in the deck when one guesses $\boldsymbol{\pi}_t$, we have

$$\mathbb{P}[Y_t = 1] = \sum_{j=0}^{m} \mathbb{P}[Y_t = 1 | J_t = j]\mathbb{P}[J_t = j] = \sum_{j=0}^{m} \frac{m-j}{mn-t+1}\mathbb{P}[J_t = j].$$

Using $1 \le m - j \le m$ for $j < m$, we find

$$\frac{\mathbb{P}[J_t < m]}{mn-t+1} \le \mathbb{P}[Y_t = 1] \le \frac{m\mathbb{P}[J_t < m]}{mn-t+1}.$$

Note that $J_t < m$ if and only if $\{\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_{t-1}\}$ contains no card with multiplicity m, which happens if and only if $T_m \ge t$ (since T_m is the largest t for which this occurs). Further, $T_m < t$ whenever $t > (m-1)n + 1$ (when there are only $n - 1$ cards left in the deck one of the cards must have appeared m times), and we always have $t \ge 1 \ge 0$. In total we find that

$$(mn)^{-1}\mathbb{P}[T_m \ge t] \le \mathbb{P}[Y_t = 1] \le mn^{-1}\mathbb{P}[T_m \ge t].$$

Because $C(\mathcal{G}^-, \boldsymbol{\pi}) = \sum Y_t$, we conclude by linearity of expectation that

$$(mn)^{-1} \sum \mathbb{P}[T_m \ge t] \le \mathcal{C}^-_{m,n} \le mn^{-1} \sum \mathbb{P}[T_m \ge t],$$

so it will be enough to show $\sum \mathbb{P}[T_m \ge t] = \Theta(n^{1-1/m})$.

For $m = 1$ we have $T_m = 1$ and the result is immediate, so assume $m \ge 2$. By Lemma 2.1, there exists a sufficiently small constant $c > 0$ such that $\mathbb{P}[T_m \ge cn^{1-1/m}] \ge \frac{1}{2}$. Using this and the fact that $\mathbb{P}[T_m \ge t] \ge \mathbb{P}[T_m \ge t + 1]$,

$$\sum \mathbb{P}[T_m \ge t] \ge cn^{1-1/m}\mathbb{P}[T_m \ge cn^{1-1/m}] \ge \frac{1}{2}cn^{1-1/m},$$

proving the lower bound.

For the upper bound, by Lemma 2.1, for all $\gamma \ge 1$,

$$\mathbb{P}[T_m \ge \gamma n^{1-1/m}] \le O(\gamma^{-m}) \le O(\gamma^{-2}).$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Using $\mathbb{P}[T_m \geq t] \leq 1$ for $t \leq n^{1-1/m}$ and the above inequality, we find

$$\sum \mathbb{P}[T_m \geq t] \leq n^{1-1/m} + \sum_{p=1}^{\infty} \sum_{t=pn^{1-1/m}}^{(p+1)n^{1-1/m}-1} O(p^{-2})$$

$$= n^{1-1/m} + n^{1-1/m} \sum_{p=1}^{\infty} O(p^{-2}) = O(n^{1-1/m}),$$

giving the desired result.

We now turn to $\mathcal{C}^+_{m,n}$. Let X_t be the indicator function of the event that we guess $\boldsymbol{\pi}_t$ correctly using strategy $\mathcal{G}^+$. Define J'_t to be the largest multiplicity of a card type in $\{\boldsymbol{\pi}_{mn}, \ldots, \boldsymbol{\pi}_{mn-t+2}\}$ and T'_j the largest value t such that $\{\boldsymbol{\pi}_{mn}, \ldots, \boldsymbol{\pi}_{mn-t+2}\}$ contains no card type with multiplicity at least j. Similar to before we find that $J'_t \geq j$ if and only if $T'_j < t$ and that

$$\mathbb{P}[X_{mn-t+1} = 1] = \sum_{j=1}^{m} \frac{j}{t-1} \mathbb{P}[J'_t = j] = \frac{1}{t-1} \sum_{j=1}^{m} \mathbb{P}[J'_t \geq j] = \frac{1}{t-1} \sum_{j=1}^{m} \mathbb{P}[T_j < t],$$

where this last step used that T_j and T'_j have the same distribution. Using this, $\sum_{k=1}^{N} k^{-1} = \log N + O(1)$, and $\log m = O(1)$, we find

$$\begin{aligned}
\mathcal{C}^+_{m,n} &= \sum_t \mathbb{E}[X_{mn-t+1}] = \sum_t \sum_j \frac{\mathbb{P}[T_j < t]}{t-1} = \sum_j \sum_t \sum_{s<t} \frac{\mathbb{P}[T_j = s]}{t-1} \\
&= \sum_j \sum_s \mathbb{P}[T_j = s] \sum_{t>s} \frac{1}{t-1} \\
&= \sum_j \sum_s \mathbb{P}[T_j = s](\log mn - \log s + O(1)) \\
&= m \log n + O(1) - \sum_j \mathbb{E}[\log T_j].
\end{aligned}$$

Thus to get the desired result it will be enough to show that $\mathbb{E}[\log T_j] = (1 - j^{-1} + o(1)) \log n$ for all j. From now on we fix some $1 \leq j \leq m$. Using summation by parts, we find for any $0 < \varepsilon < (2m)^{-1}$ that

$$\begin{aligned}
\mathbb{E}[\log T_j] &= \sum_{t=2}^{mn} \log t \, \mathbb{P}[T_j = t] = \sum_{t=2}^{mn} \log t (\mathbb{P}[T_j > t-1] - \mathbb{P}[T_j > t]) \\
&= \sum_{t=1}^{mn} \mathbb{P}[T_j > t](\log(t+1) - \log t) = \sum_{t=1}^{mn} \mathbb{P}[T_j > t] \log(1 + t^{-1}) \\
&= O(1) + \sum_{t=1}^{mn} \mathbb{P}[T_j > t] t^{-1} \qquad (1) \\
&= O(1) + \sum_{t \leq n^{1-j^{-1}-\varepsilon}} (1 - o(1)) t^{-1} + \sum_{t > n^{1-j^{-1}-\varepsilon}} \mathbb{P}[T_j > t] t^{-1}, \qquad (2)
\end{aligned}$$

where we used $\log(1 + t^{-1}) = t^{-1} + O(t^{-2})$ to get (1) and Lemma 2.1 with $\gamma \leq n^{-\varepsilon}$ to get (2). By ignoring the second sum in (2), we see that $\mathbb{E}[\log T_j] \geq (1 - j^{-1} - \varepsilon) \log n + o(\log n)$. To get an upper bound, we use (1), the bound $\mathbb{P}[T_j > t] \leq 1$, and Lemma 2.1 with $\gamma \geq n^{\varepsilon}$ to get

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

$$\mathbb{E}[\log T_j] \le O(1) + \sum_{t \le n^{1-j^{-1}+\varepsilon}} t^{-1} + \sum_{t \ge n^{1-j^{-1}+\varepsilon}} o(t^{-1}) = O(1) + (1 - j^{-1} + \varepsilon)\log n + o(\log n).$$

By taking ε to be arbitrarily small, we find that $\mathbb{E}[\log T_j] = (1 - j^{-1} + o(1))\log n$ for all j, giving the desired result. □

It remains to prove Lemma 2.1.

Proof of Lemma 2.1. The first bound is trivial if t is of order n, so assume $t = o(n)$. Let $F_j(i)$ with $i \in [n]$ be the event that $\{\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_t\}$ contains at least j copies of i, and let $F_j = \bigcup_i F_j(i)$. Observe that $T_j \ge t+1$ if and only if F_j does not occur, so it will be enough to show that $\mathbb{P}[F_j] = O(\gamma^j)$ for $t = \gamma n^{1-1/j}$. Indeed, by a simple counting argument we find

$$\mathbb{P}[F_j(i)] = \sum_{j' \ge j} \binom{t}{j'} \binom{mn-t}{m-j'} \frac{(mn-m)!}{(m!)^{n-1}} \cdot \frac{(m!)^n}{(mn)!} = \sum_{j' \ge j} O(t^{j'} n^{-j'}) = O(t^j n^{-j}) = O(\gamma^j n^{-1}),$$

where this second to last step used $tn = o(1)$ when taking the sum. Taking the union bound over all $i \in [n]$ gives the first result.

For a tuple $x = (x_1, \ldots, x_j)$ with $1 \le x_1 < \cdots < x_j \le t$, let $A(x)$ be the Bernoulli variable which is 1 if $\boldsymbol{\pi}_{x_p}$ is the same value for all p, and let $S = \sum_x A(x)$. Observe that $T_j \ge t+1$ if and only if $S = 0$, i.e. if no set of j indices all have the same card type. Thus it will be enough to show $\mathbb{P}[S=0] = O(\gamma^{-j})$, which we do by using Chebyshev's inequality. To this end, let $p_x = \mathbb{P}[A(x) = 1]$ and $p_{x,y} = \mathbb{P}[A(x) = A(y) = 1]$ for $x \ne y$. If μ and σ^2 denote the mean and variance of $S = \sum A(x)$, then by linearity of expectation we get

$$\sigma^2 = \mathbb{E}[S^2] - \mathbb{E}[S]^2 = \sum_x p_x + 2\sum_{x<y} p_{x,y} - \sum_x p_x^2 - 2\sum_{x<y} p_x p_y \le \mu + 2\sum_{x<y} (p_{x,y} - p_x p_y). \quad (3)$$

To compute μ, note that for all x we have

$$p_x = \frac{m-1}{mn-1} \cdots \frac{m-j+1}{mn-j+1} = \frac{(m-1)!(mn-j)!}{(m-j)!(mn-1)!} = \Theta(n^{1-j}), \quad (4)$$

and as there are $\binom{t}{j}$ options for x, we have

$$\mu = \Theta(t^j n^{1-j}) = \Theta(\gamma^j). \quad (5)$$

To bound the rest of σ^2, fix some tuple x and let $V_k = \{y : |\{x_1, \ldots, x_j\} \cap \{y_1, \ldots, y_j\}| = k\}$ for $0 \le k \le j-1$. By symmetry, we see that

$$2\sum_{x<y} (p_{x,y} - p_x p_y) = \binom{t}{j} \sum_{k=0}^{j-1} \sum_{y \in V_k} (p_{x,y} - p_x p_y). \quad (6)$$

Thus it will be enough to bound the inner sum for each k. First consider the case $k > 0$. In this case $p_{x,y}$ is the probability that some given $2j-k$ positions of $\boldsymbol{\pi}$ take on the same value. This is 0 if $2j - k > m$, and otherwise by the same reasoning as above

$$p_{x,y} = \frac{(m-1)!(mn-2j+k)!}{(m-2j+k)!(mn-1)!} = O(n^{1-2j+k}).$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Note that $|V_k| = O(t^{j-k})$, so in total this part of the sum is at most $O(t^{j-k}n^{-2j+k+1})$. Because $t = O(n)$, this quantity is maximised (in order of magnitude) when k is as large as possible, so we have

$$\sum_{k=1}^{j-1}\sum_{y\in V_k}(p_{x,y}-p_xp_y) = O(tn^{-j}) = O(\gamma n^{-j+1-1/j}) = O(\gamma n^{-j+1-1/m}),$$

where this last step used $j \le m$.

It remains to deal with the case $k = 0$. For $y \in V_0$, let $p'_{x,y}$ be the probability that $A(x) = A(y) = 1$ and $\boldsymbol{\pi}_{x_q} = \boldsymbol{\pi}_{y_{q'}}$ for all q, q' and $p''_{x,y}$ the probability that $A(x) = A(y) = 1$ and $\boldsymbol{\pi}_{x_q} \neq \boldsymbol{\pi}_{y_{q'}}$ for any q, q'. Observe that $p_{x,y} = p'_{x,y} + p''_{x,y}$. By the same reasoning as above, we find that $p'_{x,y} = 0$ if $2j > m$ and otherwise it is $\frac{(m-1)!(mn-2j)!}{(m-2j)!(mn-1)!} = O(t^jn^{-2j+1})$. From this and the same reasoning as before, we get

$$\sum_{y\in V_0} p'_{x,y} = O(k^{2j}n^{-2j+1}) = O(tn^{-j+1}) = O(\gamma n^{-j+1-1/m}).$$

It remains to bound $\sum_{y\in V_0} p''_{x,y} - p_xp_y$, and here we will need to be somewhat careful. By first conditioning on the event $A(x) = 1$, we see that

$$\begin{aligned} p''_{x,y} &= p_x \cdot \frac{mn-m}{mn-j}\cdot\frac{m-1}{mn-j-1}\cdots\frac{m-j+1}{mn-2j+1} \le p_x\cdot 1\cdot\frac{(m-1)!(mn-2j)!}{(m-j)!(mn-j)!} \\ &\le p_x\frac{(m-1)!(mn-2j)!}{(m-j)!(mn)!}(mn)^j. \end{aligned}$$

By (4) we have

$$p_xp_y \ge p_x\frac{(m-1)!(mn-2j)!}{(m-j)!(mn)!}(mn-2j)^j,$$

and using $|V_0| = O(t^j)$ and (4) we find

$$\begin{aligned} \sum_{y\in V_0}(p''_{x,y}-p_xp_y) &= O(t^j)\cdot p_x\cdot\frac{(m-1)!(mn-2j)!}{(m-j)!(mn)!}\cdot((mn)^j-(mn-2j)^j) \\ &= O(t^j\cdot n^{-j+1}\cdot n^{-2j}\cdot n^{j-1}) = O(t^jn^{-2j}) = O(\gamma^jn^{-j-1}) = O(n^{-j}) = O(n^{-j+1-1/m}), \end{aligned}$$

where this second to last step used that $\gamma n^{1-1/j} \le mn$ implies $\gamma^j = O(n)$.

In total then by (3), (6), and (5), we have

$$\sigma^2 \le \mu + O(t^j\cdot n^{-j+1-1/m}) = \mu + O(\gamma^jn^{-1/m}) = \mu + o(\mu).$$

In particular, for n sufficiently large we have $\sigma^2 \le 4\mu$ (and the asymptotic bound of the lemma is trivial otherwise). Thus by Chebyshev's inequality, we find

$$\mathbb{P}[S = 0] \le \mathbb{P}[|S-\mu| \ge \mu^{1/2}\sigma/2] \le 4\mu^{-1} = O(\gamma^{-j}),$$

giving the desired result. □

3. The partial feedback model

3.1 Definitions and outline

Throughout this section we fix a guessing strategy $\mathcal{G}$ and a suitable $\varepsilon = \varepsilon(m) > 0$, which will be on the order of $m^{-1/4}\log^{1/4} m$. Our goal is to prove for large enough n that $\mathbb{E}[P(\mathcal{G}, \boldsymbol{\pi})] \le (1 + O(\varepsilon))m$. In this section, we simply refer to the partial feedback model as 'the game.'

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

A *history* $h = (g, y)$ of a completed game is a pair of vectors: the $[n]$-valued vector g of all mn guesses made throughout the game, and the boolean vector y of feedback received, so that $y_t = 1$ if and only if the t-th card in the deck has value g_t. A *history at time* t, denoted as h_t, is a truncation of some complete history h to the first t values in each vector, representing all the information available to the guesser after they make the t-th guess.

We let H denote a sample of the history of the game given the fixed strategy $\mathcal{G}$ and that the deck is shuffled according to a uniform random $\boldsymbol{\pi} \sim \mathfrak{S}_{m,n}$. Similarly H_t denotes a sample of the history of the game at time t.

Given a history $h = (g, y)$, we write $Y(h) := \|y\|$ for the total number of correct guesses, where here and throughout $\|v\| := \sum |v_i|$ is the ℓ^1 norm. Define $a_i(h) := |\{t{:}g_t = i\}|$ to be the number of times card type i has been guessed, and $m_i(h) := m - |\{t{:}g_t = i \text{ and } y_t = 1\}|$ to be the number of copies of card i left to be found in the deck. For a partial history h_t, the values $Y(h_t)$, $a_i(h_t)$, and $m_i(h_t)$ are defined in the same way.

We are ready to outline the proof. The first and most important step is to prove the following 'pointwise' lemma, which roughly shows that for all typical histories h_{t-1}, the probability that the t-th guess is correct is at most $(1 + o(1))n^{-1}$.

Lemma 3.1. *For any history h_{t-1} of the game up to time $t-1$ and any $i \in [n]$,*

$$\mathbb{P}[\boldsymbol{\pi}_t = i | H_{t-1} = h_{t-1}] \leq \frac{m_i(h_{t-1})}{mn - a_i(h_{t-1}) - Y(h_{t-1})}.$$

Note that the fraction on the right-hand side is a natural estimate for $\mathbb{P}[\boldsymbol{\pi}_t = i | H_{t-1} = h_{t-1}]$: the numerator is exactly the number of copies of i in the deck that have yet to be found, and the denominator is approximately the total number of positions amongst $[mn]$ at which such a copy could lie (this may not be exact because $a_i(h_{t-1})$ and $Y(h_{t-1})$ can count the same position twice). We use a simple bijective argument to prove Lemma 3.1 in Section 3.2.

The second step of the proof is to show that the term $Y(H_{t-1})$ in Lemma 3.1 is negligible with high probability, which is done by the following lemma.

Lemma 3.2. *For any $0 < \lambda \leq 1/6$, $n \geq 1152\lambda^{-2}$, and any fixed strategy $\mathcal{G}$,*

$$\mathbb{P}[P(\mathcal{G}, \boldsymbol{\pi}) > \lambda mn] \leq 2e^{-mn}.$$

This bound is proved in Section 3.3 using Lemma 3.1 and Chernoff bounds. Combining Lemmas 3.1 and 3.2, and since $Y(h_{t-1}) \leq Y(h)$, we see that with high probability for any $\varepsilon > 0$,

$$\mathbb{P}[\boldsymbol{\pi}_t = i] \leq \frac{m_i(H_{t-1})}{(1-\varepsilon)mn - a_i(H_{t-1})}.$$

We now break guesses into three types, based on how many times a given card i has already been guessed. We say that guess t, say with $g_t = i$, is *subcritical* if $a_i(H_{t-1}) < \varepsilon mn$, *critical* if $\varepsilon mn \leq a_i(H_{t-1}) < (1-\varepsilon)mn$, and *supercritical* if $a_i(H_{t-1}) \geq (1-\varepsilon)mn$. Note that if even a single supercritical guess is made, then almost all guesses must have been of that same card type, which makes the situation easy to analyse.

By adaptively renumbering the cards during the game if necessary, we may assume without loss of generality that if there are k card types for which critical guesses are made, then they are exactly the first k cards $1, \ldots, k$. For any given history h, let $b_0(h)$ be the number of subcritical guesses made, let $b_i(h)$, $1 \leq i \leq k$ be the number of critical guesses made with $g_t = i$, and let $b_\infty(h)$ be the number of supercritical guesses made. Define $Y_0(h)$, $Y_i(h)$, and $Y_\infty(h)$ to be the number of correct guesses made in each regime.

We finish the proof by showing with high probability that each of the $Y_i(H)$ values are not much larger than their means. The subcritical guesses $Y_0(H)$ are handled in Section 3.4, the critical guesses $Y_i(H)$ in Section 3.5, and the supercritical regime is simple enough to not merit its own subsection. The proof is then completed in Section 3.6.

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Throughout the proof we will often omit floors and ceilings for ease of presentation. For an event E we let $\overline{E}$ denote its complement. For real valued random variables X and Y, we write $X \succeq Y$ if X stochastically dominates Y, i.e. if for all $x \in \mathbb{R}$, $\mathbb{P}[X \geq x] \geq \mathbb{P}[Y \geq x]$. We also recall a standard variant of the Chernoff bound, a proof of which can be found in [1].

Lemma 3.3. [1]. *Let $B(N, p)$ be a binomial random variable with N trials and probability of success p. Then for all $\lambda > 0$,*

$$\mathbb{P}[B(N,p) > (1+\lambda)pN] \leq e^{-\frac{1}{2}\lambda^2 pN}.$$

3.2 The pointwise lemma

In this section we show Lemma 3.1, which is equivalent to an upper bound on the number of $\pi \in \mathfrak{S}_{m,n}$ for which at each position up through t, either π_t is specified or a single value is disallowed for π_t. We reduce to the following setup.

Definition 3.4. Let $\mathbf{m} = (m_1, \ldots, m_n)$ and $\mathbf{a} = (a_1, \ldots, a_n)$ be vectors of non-negative integers satisfying $\|\mathbf{a}\| < \|\mathbf{m}\|$. An $\mathbf{m}$-*permutation* is a word of length $\|\mathbf{m}\|$ over alphabet $[n]$ where i appears exactly m_i times. An $(\mathbf{m}, \mathbf{a})$-*permutation* π is an $\mathbf{m}$-permutation where the first a_1 terms are not 1, the next a_2 terms are not 2, and so on, so that exactly a_i terms in π are forbidden from taking value i.

It is significant that $\|a\| < \|m\|$ strictly in the definition of $(\mathbf{m}, \mathbf{a})$-permutations, guaranteeing that no restrictions are made on the value of the last term. Given a history h_{t-1} up to time $t-1$, we let $\mathbf{m}$ be the vector $(m_1(h_{t-1}), \ldots, m_n(h_{t-1}))$, and $\mathbf{a}$ be the vector $(a_1(h_{t-1}), \ldots, a_n(h_{t-1}))$. We claim that the following bound on $(\mathbf{m}, \mathbf{a})$-permutations implies Lemma 3.1.

Lemma 3.5. *If $f_i(\mathbf{m}, \mathbf{a})$ is the fraction of all $(\mathbf{m}, \mathbf{a})$-permutations for which the last term is i, then*

$$f_i(\mathbf{m}, \mathbf{a}) \leq \frac{m_i}{\|\mathbf{m}\| - a_i}.$$

Indeed, by definition $f_i(\mathbf{m}, \mathbf{a})$ is the probability that the last card in $\boldsymbol{\pi}$ is exactly i given the current history h_{t-1}. But all positions past the first $t-1$ are indistinguishable, so $f_i(\mathbf{m}, \mathbf{a})$ is also the probability that the next card (at index t) is i. Thus it suffices to prove Lemma 3.5.

Proof of Lemma 3.5. It suffices to show the lemma for $i = 1$. First, we make a technical reduction to the case $a_1 = 0$ for convenience. Let $\tilde{\pi}$ be any sequence of a_1 cards in which 1 does not appear and i appears at most m_i times for all $i > 1$. Define an $(\mathbf{m}, \mathbf{a}, \tilde{\pi})$-*permutation* to be an $(\mathbf{m}, \mathbf{a})$-permutation where the first a_1 terms agree with $\tilde{\pi}$.

Define $f_i(\mathbf{m}, \mathbf{a}, \tilde{\pi})$ to be the fraction of $(\mathbf{m}, \mathbf{a}, \tilde{\pi})$-permutations, which have last term i. Since $f_1(\mathbf{m}, \mathbf{a})$ is some convex combination of the values $f_1(\mathbf{m}, \mathbf{a}, \tilde{\pi})$, it suffices to show that for every specific choice of $\tilde{\pi}$,

$$f_1(\mathbf{m}, \mathbf{a}, \tilde{\pi}) \leq \frac{m_1}{\|\mathbf{m}\| - a_1}. \tag{7}$$

Let $\mathbf{m}'$ be the vector of card counts remaining when the cards in $\tilde{\pi}$ are taken out, and let $\mathbf{a}' = (0, a_2, a_3, \ldots, a_n)$, so that an $(\mathbf{m}, \mathbf{a}, \tilde{\pi})$-permutation is just $\tilde{\pi}$ concatenated with an $(\mathbf{m}', \mathbf{a}')$-permutation π'. Since $m_1' \leq m_1$, it suffices to show

$$f_1(\mathbf{m}', \mathbf{a}') \leq \frac{m_1'}{\|\mathbf{m}'\|},$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

which is just the case $a_1 = 0$ in the original lemma statement. Thus, it remains to show that if $a_1 = 0$, we have

$$f_1(\mathbf{m}, \mathbf{a}) \leq \frac{m_1}{\|\mathbf{m}\|}. \tag{8}$$

In fact, we will prove that for any i,

$$\frac{f_1(\mathbf{m}, \mathbf{a})}{f_i(\mathbf{m}, \mathbf{a})} \leq \frac{m_1}{m_i}. \tag{9}$$

The case $i = 1$ is trivial, so we just need to prove this for $i > 1$, and without loss of generality we can assume $i = 2$. We divide the $(\mathbf{m}, \mathbf{a})$-permutations π which end in either 1 or 2 into classes as follows. For each π that ends in either 1 or 2, consider all positions past the first a_2, which contain either a 1 or a 2. Let $S(\pi)$ denote the set of π' obtained by cyclically shifting the 1's and 2's in these positions within π, fixing all other values. Note that with this we never move a 1 into a forbidden position (as $a_1 = 0$) nor a 2 into a forbidden position (as we only shift past the first a_2 positions). It follows that every $\pi' \in S(\pi)$ is a $(\mathbf{m}, \mathbf{a})$-permutation ending in 1 or 2.

Note that the total number of 2's past the first a_2 positions is exactly m_2, since every 2 appears past the first a_2, while the total number of 1's past the first a_2 positions is at most m_1, since there are exactly m_1 1's in total. Thus, we see that the fraction of $\pi' \in S(\pi)$ which end in 1 is at most $\frac{m_1}{m_1+m_2}$ for every π. As the $S(\pi)$ partition all possible $(\mathbf{m}, \mathbf{a})$-permutations π which end in either 1 or 2, (9) follows for $i = 2$.

Finally, to derive (8) it suffices to write (9) as

$$\frac{m_i}{m_1} f_1(\mathbf{m}, \mathbf{a}) \leq f_i(\mathbf{m}, \mathbf{a})$$

and sum over i, noting that $\sum_i f_i(\mathbf{m}, \mathbf{a}) = 1$ since every $(\mathbf{m}, \mathbf{a})$-permutation must end in some i. □

3.3 Weak bound on $\mathcal{P}^+_{m,n}$

The next step is to show that the $Y(h_{t-1})$ term in Lemma 3.1 is negligible with high probability. Since $Y(h_{t-1})$ is bounded by just $Y(h)$, the total number of cards guessed correctly, it suffices to show a weak upper bound on the total number of correct guesses in the form of Lemma 3.2. To do this we first show the following.

Lemma 3.6. *Let $B_1, \ldots, B_k$ be (not necessarily independent) Bernoulli random variables with $\mathbb{P}[B_t = 1 | \sum_{s<t} B_s = x] \leq p$ for all t and x. Then $\sum_{t=1}^k B_t$ is stochastically dominated by a binomial random variable $B(k, p)$.*

This lemma can be proved by induction, with the induction step using the following observation.

Lemma 3.7. *Let X, X', Y, Y' be integer-valued random variables such that X' and Y' are $\{0, 1\}$-valued, $X \geq Y$, and for all $x \in \mathbb{Z}$, $(X'|X = x) \geq (Y'|Y = x)$. Then,*

$$X + X' \geq Y + Y'.$$

Proof. We would like to show that for any $y \in \mathbb{Z}$, $\mathbb{P}[X + X' \geq y] \geq \mathbb{P}[Y + Y' \geq y]$. But clearly

$$\begin{aligned}
\mathbb{P}[X + X' \geq y] &= \mathbb{P}[X \geq y] + \mathbb{P}[(X = y - 1) \wedge (X' = 1)] \\
&= \mathbb{P}[X \geq y] + \mathbb{P}[X = y - 1]\mathbb{P}[X' = 1 | X = y - 1] \\
&\geq \mathbb{P}[X \geq y] + \mathbb{P}[X = y - 1]\mathbb{P}[Y' = 1 | Y = y - 1]
\end{aligned}$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

$$\begin{aligned} &\geq \mathbb{P}[Y \geq y] + \mathbb{P}[Y = y-1]\mathbb{P}[Y' = 1 | Y = y-1] \\ &= \mathbb{P}[Y + Y' \geq y]. \end{aligned} \tag{10}$$

Here only (10) is worth explaining. Since $X \geq Y$ we have $\mathbb{P}[X \geq y] \geq \mathbb{P}[Y \geq y]$ and $\mathbb{P}[X \geq y] + \mathbb{P}[X = y-1] \geq \mathbb{P}[Y \geq y] + \mathbb{P}[Y = y-1]$, so by taking convex combinations of these two inequalities, we have for any $t \in [0,1]$, $\mathbb{P}[X \geq y] + t\mathbb{P}[X = y-1] \geq \mathbb{P}[Y \geq y] + t\mathbb{P}[Y = y-1]$ as well. Taking $t = \mathbb{P}[Y' = 1 | Y = y-1]$ completes the proof. □

Lemma 3.6 follows by iterating Lemma 3.7 with $X = \sum_{s<t} B_s$, $X' = B_t$, Y a binomial random variable $B(t-1, p)$, and Y' a Bernoulli random variable with probability p.

Proof of Lemma 3.2. We first show that few correct guesses are made in the first third of the game, i.e. when $t \leq mn/3$. In this case we apply Lemma 3.1 to find that for any $i \in [n]$,

$$\mathbb{P}[\boldsymbol{\pi}_t = i | H_{t-1} = h_{t-1}] \leq \frac{m_i(h_{t-1})}{mn - a_i(h_{t-1}) - Y(h_{t-1})} \leq \frac{m}{mn - mn/3 - mn/3} = \frac{3}{n},$$

since up to this point there have been at most $mn/3$ correct guesses and each i has been guessed at most $mn/3$ times. It follows that for $t \leq mn/3$, conditional on any h_{t-1}, the probability that the t-th guess is correct is at most $3/n$. In particular the t-th guess is correct with probability at most $3/n$ regardless of the value of $Y(H_{t-1})$, so by Lemma 3.6 the number of correct guesses in the first third of the game $Y(H_{mn/3})$ is stochastically dominated by a binomial random variable $B(mn/3, 3/n)$. Applying Lemma 3.3 gives for all $\delta > 0$,

$$\mathbb{P}[Y(H_{mn/3}) > (1+\delta)m] \leq \mathbb{P}[B(mn/3, 3/n) > (1+\delta)m] \leq e^{-\delta^2 m/2}.$$

Taking $\delta = \lambda n/4 - 1 \geq \lambda n/6$, we find

$$\mathbb{P}[Y(H_{mn/3}) > \lambda mn/4] \leq e^{-\lambda^2 mn^2/72}. \tag{11}$$

Let T be the set of i such that $a_i(h_t) < mn/4$ for all t, and note that there are at most four card types not in T (since only mn total guesses are made). Let E be the event that $Y(H_{mn/3}) \leq \lambda mn/4$, and observe that conditional on E we have $Y(H_t) \leq (2/3 + \lambda/4)mn$ for all t since at most $2mn/3$ correct guesses can be made in the last $2/3$ of the game. Thus by Lemma 3.5 and the above observations, we have for $i \in T$, all $t > mn/3$, and any possible history h_{t-1} for which E occurs,

$$\mathbb{P}[\boldsymbol{\pi}_t = i | H_{t-1} = h_{t-1}] \leq \frac{m}{mn - mn/4 - (2/3 + \lambda/4)mn} \leq \frac{24}{n}, \tag{12}$$

where we used $\lambda \leq 1/6$.

Let $Y'(H)$ denote the total number of correct guesses of card types $i \in T$ and let $Y''(H)$ denote the total number of correct guesses involving $i \notin T$. Observe that

$$Y(H) = Y'(H) + Y''(H) \leq Y'(H) + 4m \leq Y'(H) + \lambda mn/2,$$

where this last step used $n \geq 8\lambda^{-1}$ (which is implicit in our hypothesis of the lemma). By (12) we see that conditional on E, $Y'(H) - Y(H_{mn/3})$ is stochastically dominated by a binomial random variable $B(2mn/3, 24/n)$. Thus

$$\begin{aligned} \mathbb{P}[Y(H) > \lambda mn] &\leq \mathbb{P}[Y'(H) > \lambda mn/2] \leq \mathbb{P}[Y'(H) - Y(H_{mn/3}) > \lambda mn/4 | E] + \mathbb{P}[\overline{E}] \\ &\leq \mathbb{P}[B(2mn/3, 24/n) > \lambda mn/4] + \mathbb{P}[\overline{E}] \leq e^{-\lambda^2 mn^2/1152} + e^{-\lambda^2 mn^2/72}, \end{aligned}$$

where the last inequality used the Chernoff bound with $\delta = \lambda n/64 - 1 \geq \lambda n/96$ and (11). Using $n \geq 1152\lambda^{-2}$ gives the desired result. □

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

3.4 Concentration of subcritical guesses

In this section we handle the subcritical guesses. If X_t denotes the indicator variable that the t-th subcritical guess is correct, then intuitively the X_t variables are dominated by Bernoulli random variables with parameter $p = \frac{1}{(1-2\varepsilon)n}$, so the total number of correct subcritical guesses is dominated by a binomial distribution $B(b_0(H), p)$, where we recall that $b_0(H)$ is the number of subcritical guesses in history H.

We would like to say that this binomial distribution is close to its expectation with high probability. It is not enough, however, to prove this for a fixed binomial distribution. The main technical issue is that the number of trials $b_0(H)$ can be chosen adaptively by the guesser. For example, they can use a strategy where they repeatedly make subcritical guesses until they have guessed an above average number of cards correctly. This is essentially equivalent to the guesser simulating a summation of Bernoulli random variables $\sum_{t=1}^{mn} B_t$, and then choosing some number of trials $b \le mn$ such that the number of correct subcritical guesses is $\sum_{t=1}^{b} B_t$. We thus wish to show that for B_t a sequence of independent Bernoulli variables, $\sum_{t=1}^{b} B_t$ is not much larger than its expectation for all large b. Thus, no matter how the guesser chooses b they can never do much better than pb.

A weak upper bound for this probability comes from applying the Chernoff bound to all $b \le mn$ and then using a union bound. Unfortunately when p is very small this upper bound is not effective. A more careful application of the union bound gives the following technical result, where we think of the Z_k's in its statement as centred binomial random variables.

Lemma 3.8. *Let $0 \le p \le 1$, $c, c' > 0$, and let $0 \equiv Z_0, Z_1, Z_2, \ldots$ be random variables such that $Z_k - Z_{k-1} \ge -p$ for all k, and such that for all integers $0 \le k' < k$ and all $\lambda > 0$,*

$$\mathbb{P}[Z_k - Z_{k'} > \lambda p(k-k')] \le c' e^{-c\lambda^2 p(k-k')}.$$

Then for all $\lambda > 0$ and integers $k_1 \ge k_0 \ge 2\lambda^{-1}$, we have

$$\mathbb{P}[\exists k \in [k_0, k_1], Z_k > \lambda pk] \le \frac{8c'k_1}{\lambda k_0} e^{-\frac{1}{256}c\lambda^3 pk_0}.$$

Proof. Define $\ell = \frac{1}{2}\lambda k_0 \ge 1$. The idea of the proof is to take a union bound over the events $Z_{\ell a} - Z_{\ell(a-1)} > \lambda p\ell$ for all integers $a \le \frac{k_1}{\ell}$, which will turn out to be strong enough to conclude the stated result. To be precise, let $0 = x_0 < x_1 < \cdots < x_r = k_1$ be any sequence of integers such that $\frac{1}{2}\ell \le x_a - x_{a-1} \le \ell$ for all $a > 0$, and note that the number of terms in this sequence satisfies

$$r \le \lceil 2k_1/\ell \rceil \le \frac{8k_1}{\lambda k_0}. \tag{13}$$

Let E be the event that $Z_{x_b} > \frac{1}{8}\lambda pb\ell$ for some b. Observe that $Z_{x_b} = \sum_{a=1}^{b} Z_{x_a} - Z_{x_{a-1}}$, so $Z_{x_b} > \frac{1}{8}\lambda pb\ell$ implies that some $a \le b$ has

$$Z_{x_a} - Z_{x_{a-1}} > \frac{1}{8}\lambda p\ell \ge \frac{1}{8}\lambda p(x_a - x_{a-1}).$$

Thus by the union bound, the hypothesis of the lemma, the fact that $x_a - x_{a-1} \ge \frac{1}{2}\ell$, and inequality (13), we have

$$\mathbb{P}[E] \le \sum_{a=1}^{r} \mathbb{P}[Z_{x_a} - Z_{x_{a-1}} > \frac{1}{8}\lambda p(x_a - x_{a-1})] \le r \cdot c' e^{-\frac{1}{128}c\lambda^2 p\ell} \le \frac{8c'k_1}{\lambda k_0} e^{-\frac{1}{256}c\lambda^3 pk_0}.$$

We claim that if $Z_k > \lambda pk$ for some $k \in [k_0, k_1]$, then E occurs. Indeed, suppose such a k exists and let b be the smallest integer such that $k \le x_b$, which in particular implies $x_b - k \le \ell$. We also have $b \ge 2$ because $k_0 \le k \le b\ell$ and $k_0/\ell = 2\lambda^{-1} \ge 2$, and thus

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

$$k \geq \frac{1}{2}(b-1)\ell \geq \frac{1}{4}b\ell. \tag{14}$$

Note that $Z_{x_b} - Z_k \geq -\ell p$ because $Z_k - Z_{k-1} \geq -p$ for all k. Using this, $\ell = \frac{1}{2}\lambda k_0 \leq \frac{1}{2}\lambda k$, and inequality (14), we have

$$Z_{x_b} > \lambda pk - \ell p \geq \frac{1}{2}\lambda pk \geq \frac{1}{8}pb\ell,$$

so E occurs. Thus,

$$\mathbb{P}[\exists k \in [k_0, k_1], Z_k > \lambda pk] \leq \mathbb{P}[E] \leq \frac{8c'k_1}{\lambda k_0} e^{-\frac{1}{256}c\lambda^3 pk_0}$$

as desired. □

Using Lemma 3.8, we can show that subcritical guesses are well behaved.

Lemma 3.9. *If $\varepsilon \leq \frac{1}{8}$ and n is sufficiently large in terms of ε, m, then*

$$\mathbb{P}\left[Y_0(H) > (1+4\varepsilon)\frac{b_0(H)}{n}\right] \leq c'\varepsilon^{-2}e^{-c\varepsilon^4 m},$$

for some absolute constants $c, c' > 0$.

Proof. Given $t \leq b_0(H)$, let t' be the smallest positive integer for which $b_0(H_{t'}) = t$, so that t' is the time of the t-th subcritical guess (note that t' is itself a random variable), and let $X_t \coloneqq Y_0(H_{t'}) - Y_0(H_{t'-1})$. In other words, X_t is the indicator of the t-th subcritical guess. Let E be the event that $Y(H) > \varepsilon mn$, and define E_t to be the event that $Y(H_{t'}) > \varepsilon mn$. Observe that $\overline{E}$ implies that no E_t occurs.

Note that $Y_0(H) = \sum_{t=1}^{b_0(H)} X_t$. We modify $Y_0(H)$ to ignore the events E_t as follows. Define $X'_t = X_t$ if E_{t-1} does not occur and $X'_t = 0$ otherwise, and let $Y'_0 = \sum_{t=1}^{b_0(H)} X'_t$. With $L \coloneqq (1+4\varepsilon)\frac{b_0(H)}{n}$, we find

$$\begin{aligned}\mathbb{P}[Y_0(H) > L] &\leq \mathbb{P}[(Y_0(H) > L) \wedge \overline{E}] + \mathbb{P}[E] = \mathbb{P}[(Y'_0 > L) \wedge \overline{E}] + \mathbb{P}[E] \\ &\leq \mathbb{P}[Y'_0 > L] + \mathbb{P}[E],\end{aligned}$$

By Lemma 3.2 we know $\mathbb{P}[E] \leq 2e^{-mn}$, so for n sufficiently large the contribution of $\mathbb{P}[E]$ is negligible. It remains to upper bound the probability that Y'_0 is large. Note that $X'_t = 1$ if and only if the next term $\boldsymbol{\pi}_{t'}$ is exactly the next guess i, the total number $a_i(H_{t'-1})$ of times i is guessed is at most εmn, and the total number $Y(H_{t'-1})$ of correct guesses up to this point is also at most εmn. We now have by Lemma 3.1 that

$$\mathbb{P}[X'_t = 1 | X'_1, \ldots, X'_{t-1}] \leq \frac{m_i(H_{t'-1})}{mn - a_i(H_{t'-1}) - Y(H_{t'-1})} \leq \frac{m}{(1-2\varepsilon)mn} = \frac{1}{(1-2\varepsilon)n} =: p.$$

Define $B_1, B_2, \ldots, B_{mn}$ to be independent Bernoulli random variables with $\mathbb{P}[B_t = 1] = p$ and define $Z_k = \sum_{t=1}^{k} B_t - pk$. By the above inequality, we see that given any history $h_{t'-1}$ up to the t'-th guess, X'_t is stochastically dominated by B_t, and hence Z_k stochastically dominates $\sum_{t=1}^{k} X'_t - pk$. Observe that

$$\sum_{t=1}^{b_0(H)} X'_t - pb_0(H) > \frac{\varepsilon b_0(H)}{(1-2\varepsilon)n} \iff Y'_0 > \frac{(1+\varepsilon)b_0(H)}{(1-2\varepsilon)n} \implies Y'_0 > \frac{(1+4\varepsilon)b_0(H)}{n} = L,$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

where the last step used $\varepsilon \le \frac{1}{8}$. Because $Z_{b_0(H)}$ stochastically dominates the above sum, we have

$$\mathbb{P}[Y_0' > L] \le \mathbb{P}\left[\sum_{t=1}^{b_0(H)} X_t' - pb_0(H) > \frac{\varepsilon b_0(H)}{(1-2\varepsilon)n}\right] \le \mathbb{P}\left[Z_{b_0(H)} > \frac{\varepsilon b_0(H)}{(1-2\varepsilon)n}\right]$$
$$\le \mathbb{P}\left[\exists k \in [\varepsilon mn, mn],\ Z_k > \frac{\varepsilon k}{(1-2\varepsilon)n}\right],$$

Where this last step used that the number of subcritical guesses $b_0(H)$ must always be at least εmn and at most mn.

Because Z_k is a centred binomial distribution, $\mathbb{P}[Z_k - Z_{k'} > \lambda p(k-k')] \le e^{-\frac{1}{2}\lambda^2 p(k-k')}$ for $k' < k$ by Lemma 3.3, and also $Z_k - Z_{k-1} \ge -p$ for all k by construction. If n is sufficiently large we have $\varepsilon mn \ge 2\varepsilon^{-1}$, so we can apply Lemma 3.8 to the above inequality with $c' = \frac{1}{2}$ and $c = 1$ to conclude

$$\mathbb{P}[Y_0' > L] \le \frac{8}{(1-2\varepsilon)\varepsilon^2} e^{-\frac{1}{512(1-2\varepsilon)}\varepsilon^4 m} \le 16\varepsilon^{-2} e^{-\frac{1}{512}\varepsilon^4 m},$$

with this last step using $\varepsilon \le \frac{1}{4}$. □

3.5 Concentration of critical guesses

In the subcritical region we were able to bound the number of correct guesses by a binomial random variable. For the critical region, we compare the number of correct guesses with a hypergeometric random variable. We recall that a random variable $S \sim \mathrm{Hyp}(N, m, b)$ has a hypergeometric distribution (with parameters N, m, b) if for all integer $1 \le k \le m$ we have

$$\mathbb{P}[S = k] = \binom{b}{k}\binom{N-b}{m-k}\binom{N}{k}^{-1}. \tag{15}$$

Equivalently one can define this by uniformly shuffling a deck of N cards with m of these cards being 'good', and then letting S be the number of good cards one sees in the first b draws from the deck. From this viewpoint, if we let R_t denote the indicator variable which is 1 if the tth draw is a good card, we see that $S = \sum_{t=1}^{b} R_t$ and that the R_t can be defined by

$$\mathbb{P}[R_t = 1 | R_1, \ldots, R_{t-1}] = \frac{m - (R_1 + \cdots + R_{t-1})}{N - t + 1}. \tag{16}$$

We can use the following lemma to bound random variables by hypergeometric random variables.

Lemma 3.10. *Suppose $P_1, \ldots, P_k$ and $R_1, \ldots, R_k$ are $\{0,1\}$-random variables satisfying*

$$\mathbb{P}[P_t = 1 | P_1, \ldots, P_{t-1}] \le \frac{m - (P_1 + \cdots + P_{t-1})}{N - t + 1}$$
$$\mathbb{P}[R_t = 1 | R_1, \ldots, R_{t-1}] = \frac{m - (R_1 + \cdots + R_{t-1})}{N - t + 1}.$$

Then $R_1 + \cdots + R_k \succeq P_1 + \cdots + P_k$.

The proof of Lemma 3.10 follows from induction and applying Lemma 3.7 with $X = R_1 + \cdots + R_{t-1}$, $X' = R_t$, $Y = P_1 + \cdots + P_{t-1}$, and $Y' = P_t$. The last thing we need is to use Lemma 3.8 in this hypergeometric setting.

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Lemma 3.11. *Let $N \geq m^2 + m$, define the indicator random variables $R_1, R_2, \ldots, R_N$ as in (16), and let $S_b := \sum_{t=1}^{b} R_t$ for all b. Then for all b and $\lambda > 0$,*

$$\mathbb{P}\left[S_b > \frac{(1+\lambda)bm}{N}\right] \leq 3e^{-\frac{\lambda^2 bm}{2N}}.$$

Further, for all $\lambda > 0$ and integers b_0, b_1 satisfying $2\lambda^{-1} \leq b_0 \leq b_1 \leq N$, we have

$$\mathbb{P}\left[\exists b \in [b_0, b_1], S_b > \frac{(1+\lambda)bm}{N}\right] \leq \frac{24b_1}{\lambda b_0} e^{-\frac{\lambda^3 b_0 m}{512N}}.$$

Proof. Observe that $S_b \sim \mathrm{Hyp}(N, m, b)$. Thus if $q := b/N$, we have by (15) that

$$\begin{aligned}
\mathbb{P}[S_b = k] &= \binom{qN}{k}\binom{(1-q)N}{m-k}\binom{N}{m}^{-1} \leq \frac{(qN)^k}{k!}\frac{((1-q)N)^{m-k}}{(m-k)!}\frac{m!}{(N-m)^m} \\
&= \binom{m}{k} q^k (1-q)^{m-k}\left(1 + \frac{m}{N-m}\right)^m \leq \binom{m}{k} q^k (1-q)^{m-k} e^{m^2/(N-m)} \\
&\leq \binom{m}{k} q^k (1-q)^{m-k} \cdot 3,
\end{aligned}$$

where this last step used $N - m \geq m^2$.

We thus see that $\mathbb{P}[S_b > (1+\lambda)qm] \cdot 3^{-1}$ is at most the probability that a binomial distribution with m trials and probability q of success has at least $(1+\lambda)qm$ successes, which is at most $e^{-\lambda^2 qm/2}$ by Lemma 3.3. This gives the first result.

For the second result, define $p = m/N$ and let $Z_b := S_b - pb$. Note that $S_b - S_{b'} \sim \mathrm{Hyp}(N, m, b - b')$ for $b > b'$ (since this is just a sum of $b - b'$ of the R_t variables), so the first result implies

$$\mathbb{P}[Z_b - Z_{b'} > \lambda p(b - b')] = \mathbb{P}[S_b - S_{b'} > (1+\lambda)p(b - b')] \leq 3e^{-\frac{1}{2}\lambda^2 pb}.$$

We can thus apply Lemma 3.8 to the Z_b variables with $c' = 3$ and $c = -\frac{1}{2}$ to conclude the result. □

Using this we can prove the following.

Lemma 3.12. *For $i \geq 1$ finite, $\varepsilon \leq \frac{1}{4}$, and n sufficiently large in terms of ε, m, we have*

$$\mathbb{P}\left[Y_i(H) > (1+4\varepsilon)\frac{b_i(H)}{n} + \varepsilon^2 m\right] \leq c'\varepsilon^{-2} e^{-c\varepsilon^4 m}$$

for some constants $c, c' > 0$ depending only on ε.

Proof. Fix i positive and finite, and let $X_t := Y_i(H_{t'}) - Y_i(H_{t'-1})$ where t' is the smallest positive integer for which $b_i(H_{t'}) = t$ (note that t' is itself a random variable). In other words, X_t is the indicator of the t-th critical guess of i. Define $X'_t = X_t$ if $Y(H) \leq \varepsilon mn$ and define $X'_t = 0$ otherwise.

Let R_t be random variables as in Lemma 3.10 with $N = (1-2\varepsilon)mn$, and define $S_b = \sum_{i=1}^{b} R_i$ for all $1 \leq b \leq (1-2\varepsilon)mn$. By applying Lemma 3.5 (and noting that i was guessed εmn times before its critical guesses started), we see

$$\mathbb{P}[X'_t = 1 | X'_1, \ldots, X'_{t-1}] \leq \frac{m - (X'_1 + \cdots + X'_{t-1})}{(1-2\varepsilon)mn - t + 1}.$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

Thus we can apply Lemma 3.10 with X'_t taking the role of P_t, and letting

$$L(b) = (1+4\varepsilon)\frac{b}{n} + \varepsilon^2 m$$

gives

$$\begin{aligned}
\mathbb{P}[Y_i(H) > L(b_i(H))] &= \mathbb{P}\left[\sum_{t=1}^{b_i(H)} X_t > L(b_i(H))\right] \\
&\le \mathbb{P}[Y(H) \le \varepsilon mn] \cdot \mathbb{P}\left[\sum_{t=1}^{b_i(H)} X'_t > L(b_i(H))\right] + \mathbb{P}[Y(H) > \varepsilon mn] \\
&\le 1 \cdot \mathbb{P}[S_{b_i(H)} > L(b_i(H))] + 2e^{-mn} \\
&\le \mathbb{P}\left[\exists b \in [1, (1-2\varepsilon)mn],\ S_b > L(b)\right] + 2e^{-mn},
\end{aligned}$$

where the second to last step used Lemma 3.2 and the last step used that the value of $b_i(H)$ must lie in 1 and $(1-2\varepsilon)mn$.

To bound $\mathbb{P}\left[\exists b \in [1, (1-2\varepsilon)mn],\ S_b > L(b)\right]$, we partition $[1, (1-2\varepsilon)mn]$ into intervals $[b_{j-1}, b_j]$ (which we define below) and show that $\mathbb{P}\left[\exists b \in [b_{j-1}, b_j],\ S_b > L_j(b)\right]$ is small for all j, where $L_j(b)$ is some quantity upper bounded by $L(j)$. Taking a union bound will then give the desired result.

Let $b_0 := \frac{1}{2}\varepsilon^2(1-2\varepsilon)mn \ge 2$ for n sufficiently large. By taking $\lambda = 1$ in Lemma 3.11, we have

$$\mathbb{P}[\exists b \in [1, b_0],\ S_b > \varepsilon^2 m] = \mathbb{P}\left[S_{b_0} > \frac{2b_0 m}{(1-2\varepsilon)mn}\right] \le 3e^{-\frac{b_0 m}{(1-2\varepsilon)mn}} = 3e^{-\frac{1}{4}\varepsilon^2 m}. \tag{17}$$

Define $b_j = 2^j b_0$. Observe that for all $b \le b_j$ we have $\varepsilon^2 \ge 2^{1-j}\frac{b}{(1-2\varepsilon)mn}$. Thus for j such that $2^{1-j} \ge 4\varepsilon$ we find for n sufficiently large in terms of j,

$$\begin{aligned}
Pr[\exists b \in [b_{j-1}, b_j],\ S_b > \frac{b}{n} + \varepsilon^2 m] &\le \mathbb{P}\left[\exists b \in [b_{j-1}, b_j],\ S_b > (1-2\varepsilon+2^{1-j})\frac{bm}{(1-2\varepsilon)mn}\right] \\
&\le \mathbb{P}\left[\exists b \in [b_{j-1}, b_j],\ S_b > (1+2^{-j})\frac{bm}{(1-2\varepsilon)mn}\right] \\
&\le 48 \cdot 2^j e^{-\frac{2^j b_0}{2^{3j+9}(1-2\varepsilon)n}} = 48 \cdot 2^j e^{-2^{-2j-10}\varepsilon^2 m} \\
&\le 48\varepsilon^{-1} e^{-2^{-10}\varepsilon^4 m},
\end{aligned} \tag{18}$$

where this third inequality used Lemma 3.11.

Let $J = \left\lfloor \log_2(\varepsilon^{-1}) \right\rfloor - 1$, noting that we can apply the above bound up to $b_J = 2^{J-1}\varepsilon^2(1-2\varepsilon)mn \ge \frac{1}{8}\varepsilon(1-2\varepsilon)mn$. Observe that for $\varepsilon \le \frac{1}{8}$,

$$\frac{(1+4\varepsilon)b}{n} \ge \frac{(1+\varepsilon)b}{(1-2\varepsilon)n}.$$

Thus by Lemma 3.11 applied with $\lambda = \varepsilon$, we see that

$$\mathbb{P}\left[\exists b \in [b_J, (1-2\varepsilon)mn],\ S_b > \frac{(1+4\varepsilon)b}{n}\right] \le \frac{24(1-2\varepsilon)mn}{\varepsilon b_J} e^{-\frac{\varepsilon^3 b_J}{512(1-2\varepsilon)n}} \le 96\varepsilon^{-2} e^{\frac{-1+\varepsilon}{2048(1-2\varepsilon)}\varepsilon^4 m}.$$

Taking the union bound over this, (17), and (18) for the at most $-\log_2(\varepsilon) \le \varepsilon^{-1}$ values of $j \le J$ gives the result. □

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

3.6 Completing the proof

We need the following simple consequence of Lemma 3.2.

Lemma 3.13. *If n is sufficiently large and A is an event with $\mathbb{P}[A] = p$, then*

$$\mathbb{E}[Y(H)|A] < -1000 \log p + 1.$$

Proof. By reparametrising Lemma 3.2, we get that as long as $n > 6000$,

$$\mathbb{P}[Y(H) > 1000mx] \leq e^{-mx}$$

for any $x \geq 1$. In particular, even after conditioning on the event A,

$$\mathbb{P}[Y(H) > 1000mx|A] \leq p^{-1}e^{-mx}.$$

Choosing $x_0 = -\frac{1}{m} \log p$, we find that

$$\begin{aligned}\mathbb{E}[Y(H)|A] &\leq 1000mx_0 + \int_{x_0}^{\infty} \mathbb{P}[Y(H) > 1000mx|A]dx \\ &\leq 1000mx_0 + \int_{x_0}^{\infty} e^{-m(x-x_0)}dx = 1000mx_0 + \int_0^{\infty} e^{-mx}dx \leq -1000 \log p + 1,\end{aligned}$$

as $mx_0 = -\log p$ and the integral is absolutely bounded by 1. □

Finally we have all the tools to prove the main theorem.

Proof of Theorem 1.3. We will pick $\varepsilon = O((\log m/m)^{1/4})$, and show that for an appropriate such ε and n sufficiently large in terms of m and ε, $\mathbb{E}[Y(H)] \leq (1+\varepsilon)m$. To this end, we define the following three 'atypical' events: E_0, E_1, and E_∞.

- The event E_0 is the event that $Y_0(H) > (1+4\varepsilon)b_0(H)/n$, in other words that significantly more than the average number of subcritical guesses are correct.
- The event E_1 is the event that $Y_i(H) > (1+4\varepsilon)b_i(H)/n + \varepsilon^2 m$ for some $i \geq 1$, in other words that for some critical card i, significantly more than the average number of critical guesses of card i are correct.
- The event E_∞ is the event that there is at least one supercritical card. In this case, this single card is guessed at least $(1-\varepsilon)mn$ times.

Our goal will be to calculate the conditional expectation of $Y(H)$ depending on whether or not the exceptional events above occur. It will be convenient to group E_0 and E_1 together and define their union $A = E_0 \vee E_1$. Then,

$$\mathbb{E}[Y(H)] = \mathbb{P}[A]\mathbb{E}[Y(H)|A] + \mathbb{P}[\overline{A} \wedge \overline{E}_\infty] \cdot \mathbb{E}[Y(H)|\overline{A} \wedge \overline{E}_\infty] + \mathbb{P}[\overline{A} \wedge E_\infty]\mathbb{E}[Y(H)|\overline{A} \wedge E_\infty]. \tag{19}$$

We first observe that if none of the events E_0, E_1, and E_∞ occur, then the conditional expectation of $Y(H)$ is small. Indeed, we have

$$\mathbb{E}[Y(H)|\overline{A} \wedge \overline{E}_\infty] = \mathbb{E}[Y(H)|\overline{E}_0 \wedge \overline{E}_1 \wedge \overline{E}_\infty] \leq (1+5\varepsilon)m, \tag{20}$$

since all guesses must be subcritical or critical, and there are at most ε^{-1} distinct critical card types i.

Define $p_j = \mathbb{P}[E_j]$ for $j \in \{0,1\}$. We have by Lemmas 3.9 and 3.12 that for some absolute constants $c, c' > 0$,

$$p_0 \leq c'\varepsilon^{-2}e^{-c\varepsilon^4 m}$$

and

$$p_1 \leq c'\varepsilon^{-3}e^{-c\varepsilon^4 m},$$

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

where there is an extra multiplicative factor of ε^{-1} in the second inequality because there may be up to ε^{-1} critical cards. In particular, we have

$$\mathbb{P}[A] = \mathbb{P}[E_0 \vee E_1] \le p := c'\varepsilon^{-3}e^{-c\varepsilon^4 m}, \tag{21}$$

possibly with a different choice of c'. By Lemma 3.13 we find

$$\mathbb{E}[Y(H)|A] \le -1000 \log \mathbb{P}[A] + 1,$$

and so

$$\mathbb{P}[A]\mathbb{E}[Y(H)|A] \le -1000\mathbb{P}[A] \log \mathbb{P}[A] + \mathbb{P}[A] \le -1000p \log p + 1,$$

for p defined above. By picking an appropriate $\varepsilon = O((\log m/m)^{1/4})$, we find

$$\mathbb{P}[A]\mathbb{E}[Y(H)|A] \le m^{-\Omega(1)} + 1 < \varepsilon m \tag{22}$$

for m sufficiently large.

Finally, to control the third term of (19), note that if there is a supercritical card, at most m guesses are correct for that card (since there are a total of m copies of that card in the deck), and at most εmn guesses are made of any other card, so all other guesses are subcritical. In particular, including guesses of the unique supercritical card, there at most $b_0(H) \le 2\varepsilon mn$ subcritical guesses. Thus, by the definition of E_0, we get

$$\mathbb{E}[Y(H)|\overline{A} \wedge E_\infty] \le m + (1 + 4\varepsilon)(2\varepsilon m) \le (1 + 3\varepsilon)m.$$

In total, using (19), (20), (22), and the inequality above, we find that for m, n sufficiently large,

$$\mathbb{E}[Y(H)] \le \varepsilon m + \mathbb{P}[\overline{A}] \cdot (1 + 5\varepsilon)m \le (1 + 6\varepsilon)m = m + O(m^{3/4} \log^{1/4} m),$$

completing the proof. □

4. Concluding remarks

In this paper we proved results for two different feedback models. The first were bounds on $\mathcal{C}^{\pm}_{m,n}$, which is the most/least number of cards one can guess in the complete feedback model. For fixed m we determined $\mathcal{C}^{+}_{m,n}$ asymptotically, and for $\mathcal{C}^{-}_{m,n}$ we gave the correct order of magnitude.

Question 4.1. What is $\mathcal{C}^{-}_{m,n}$ asymptotically?

We note that a more careful analysis of the proof of Theorem 1.2 gives that this value is asymptotic to $n^{-1}\mathbb{E}[T_m]$ with T_m as in Lemma 2.1, so it suffices to compute this expectation.

Our main result was proving a tight asymptotic upper bound on $\mathcal{P}^{+}_{m,n}$, which is the most number of cards one can guess correctly in expectation in the partial feedback model. Specifically, we proved $\mathcal{P}^{+}_{m,n} = m + O(m^{3/4} \log m)$ provided n is sufficiently large. One consequence of this bound is the following. For $\pi \in \mathfrak{S}_{m,n}$, let $L(\pi)$ denote the largest integer p such that there exist $i_1 < \cdots < i_p$ with $\pi_{i_j} = j$ for all $1 \le j \le p$. Define $\mathcal{L}_{m,n} = \mathbb{E}[L(\boldsymbol{\pi})]$ where $\boldsymbol{\pi} \sim \mathfrak{S}_{m,n}$.

Corollary 4.2. *If n is sufficiently large in terms of m, we have*

$$\mathcal{L}_{m,n} \le m + O(m^{3/4} \log m).$$

Proof. Consider the strategy $\mathcal{G}$ in the partial feedback model, which guesses 1 until you guess one correctly, then 2 until you guess one correctly, and so on; and if you guess n correctly, play arbitrarily for the remaining trials. Then $P(\mathcal{G}, \boldsymbol{\pi}) \ge L(\boldsymbol{\pi})$, and the result follows from Theorem 1.3. □

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

The best lower bound we can prove is $\mathcal{L}_{m,n} = \Omega(m/\log m)$. This is because with high probability there is a 1 in the first roughly $n \log m$ positions, a 2 in the next $n \log m$ positions, and so on. We suspect that the upper bound of Corollary 4.2 is closer to the truth.

Conjecture 4.3. *If n is sufficiently large in terms of m, we have*

$$\mathcal{L}_{m,n} = m - o(m).$$

A trivial lower bound for $\mathcal{P}^+_{m,n}$ is m obtained by guessing a card type uniformly at random at each trial. A more complicated strategy gives $m + \Omega(m^{1/2})$ corrects guesses in expectation. We rigorously prove this in [9], and here we give a brief sketch of the strategy. Guess 1 a total of $mn/2$ times. If you guessed at least $\frac{1}{2}m + \sqrt{m}$ cards correctly, guess 2 for the rest of the game, otherwise keep guessing 1. In the latter scenario we always get exactly m correct guesses. One can show that the first scenario happens with some constant probability, and given this the expected number of 2's left in the second half of the deck is at least $m/2$, and in total this gives a lower bound of $m + \Omega(m^{1/2})$. We suspect that this lower bound is close to the truth.

Conjecture 4.4. *For all $\varepsilon > 0$ and n sufficiently large,*

$$\mathcal{P}^+_{m,n} = m + O(m^{1/2+\varepsilon}).$$

The current proof overshoots this bound at two points. The first is in Lemma 3.8 where we try and bound the probability that an 'adversarial' binomial distribution deviates significantly from its mean. Our proof of this lemma essentially only used a union bound, and it's plausible that more sophisticated techniques could decrease this error term.

The second point is in the bounds of Lemmas 3.9 and 3.12 where we bound the probability that the subcritical or critical guesses are much larger than average. We note that by adding in an error term of εm to the lower bound of Y_0 in Lemma 3.9, one can decrease the probability from roughly $e^{-\varepsilon^4 m}$ to $e^{-\varepsilon^3 m}$, so the central issue is the critical case, and it seems like new ideas are needed here.

Another problem of interest is bounding $\mathcal{P}^-_{m,n}$, the fewest number of cards one can guess correctly in expectation in the partial feedback model. In [9] we prove $\mathcal{P}^-_{m,n} \le m - \Omega(m^{1/2})$ using an analog of the strategy for $\mathcal{P}^+_{m,n}$. We also prove an asymptotic lower bound for $\mathcal{P}^-_{m,n}$ of $1 - e^{-m}$ by showing that one always has probability at least this of guessing at least one card correctly. Thus $\mathcal{P}^-_{m,n} = \Omega(1)$, which is again in sharp contrast to the complete feedback model where one can get arbitrarily few correct guesses in expectation. There is still a large gap between these bounds, and as in Theorem 1.3 we suspect that the partial feedback model does not allow one to guess significantly fewer guesses than in the no feedback model.

Conjecture 4.5. *If n is sufficiently large in terms of m, then*

$$\mathcal{P}^-_{m,n} \sim m.$$

The central difficulty in this setting is that there does not exist an analog of Lemma 3.1 which lower bounds $\mathbb{P}[\boldsymbol{\pi}_t = i]$ given that we have not guessed i many times and that we have guessed few cards correctly. For example, say we incorrectly guessed 1 a total of $(m-1)n$, so the remaining cards are m copies of 1. Then the probability that the next card is 2 is 0 despite the fact that we have not guessed 2 at all nor guessed any cards correctly.

Acknowledgements

The authors would like to thank Steve Butler, Sourav Chatterjee, Andrea Ottolini, Mackenzie Simper, and Yuval Wigderson for fruitful conversations.

https://doi.org/10.1017/S0963548321000134 Published online by Cambridge University Press

References

[1] Alon, N. and Spencer, J. H. (2004) *The Probabilistic Method*. John Wiley & Sons.
[2] Blackwell, D. and Hodges, J. L. (1957) Design for the control of selection bias. *Ann. Math. Stat.* **28** 449–460.
[3] Briggs, W. and Ruppert, D. (2005) Assessing the skill of yes/no predictions. *Biometrics* **61**(3) 799–807.
[4] Chung, F., Diaconis, P., Graham, R. and Mallows, C. L. (1981) On the permanents of complements of the direct sum of identity matrices. *Adv. Appl. Math.* **2**(2) 121–137.
[5] Ciucu, M. (1998) No-feedback card guessing for dovetail shuffles. *Ann. Appl. Prob.* **8**(4) 1251–1269.
[6] Diaconis, P. (1978) Statistical problems in ESP research. *Science* **201**(4351) 131–136.
[7] Diaconis, P. and Graham, R. (1981) The analysis of sequential experiments with feedback to subjects. *Ann. Stat.* **9**(1) 3–23.
[8] Diaconis, P., Graham, R. and Holmes, S. (2001) *Statistical Problems Involving Permutations with Restricted Positions*, *Lecture Notes-Monograph Series*, pp. 195–222.
[9] Diaconis, P., Graham, R. and Spiro, S. (2020). Guessing about Guessing: Practical Strategies for Card Guessing with Feedback. arXiv preprint arXiv:2012.04019.
[10] Efron, B. (1971) Forcing a sequential experiment to be balanced. *Biometrika* **58**(3) 403–417.
[11] Gural, A., Simper, M. and So, E. (2019) Information Theory in a Card-Guessing Game. https://theinformaticists.com/wp-content/uploads/2019/03/Info_Theory_Project_Card_Guessing_Game.pdf
[12] Liu, P. (2019) Asymptotic analysis of card guessing with feedback. arXiv preprint arXiv:1908.07718.
[13] Pehlivan, L. (2009) On top to random shuffles, no feedback card guessing, and fixed points of permutations. PhD dissertation, University of Southern California.
[14] Proschan, M. (1991) A note on Blackwell and Hodges (1957) and Diaconis and Graham (1981). *Ann. Stat.* **19**(2) 1106–1108.
[15] Samaniego, F. and Utts, J. (1983) Evaluating performance in continuous experiments with feedback to subjects. *Psychometrika* **48**(2) 195–209.

Cite this article: Diaconis P, Graham R, He X and Spiro S (2022). Card guessing with partial feedback. *Combinatorics, Probability and Computing* **31**, 1–20. https://doi.org/10.1017/S0963548321000134