Bounds for Static Black-Peg AB Mastermind

Christian Glazik¹, Gerold Jäger^{2(⊠)}, Jan Schiemann¹, and Anand Srivastav¹

¹ Department of Computer Science, Kiel University, Christian-Albrechts-Platz 4, 24118 Kiel, Germany

² Department of Mathematics and Mathematical Statistics, University of Umeå, 90187 Umeå, Sweden gerold.jaeger@math.umu.se

Abstract. *Mastermind* is a famous two-player game introduced by M. Meirowitz (1970). Its combinatorics has gained increased interest over the last years for different variants.

In this paper we consider a version known as the Black-Peg AB Game, where one player creates a secret code consisting of c colors on $p \leq c$ pegs, where each color is used at most once. The second player tries to guess the secret code with as few questions as possible. For each question he receives the number of correctly placed colors. In the *static* variant the second player doesn't receive the answers one at a time, but all at once after asking the last question. There are several results both for the AB and the static version, but the combination of both versions has not been considered so far. The most prominent case is n := p = c, where the secret code and all questions are permutations. The main result of this paper is an upper bound of $\mathcal{O}(n^{1.5\hat{2}5})$ questions for this setting. With a slight modification of the arguments of Doerr et al. (2016) we also give a lower bound of $\Omega(n \log n)$. Furthermore, we complement the upper bound for p = c by an optimal $(\lceil 4c/3 \rceil - 1)$ -strategy for the special case p=2 and arbitrary $c \geq 2$ and list optimal strategies for six additional pairs (p, c).

1 Introduction

Mastermind is a two players board game invented by Mordecai Meirowitz in 1970. In the original version the first player, called *codemaker*, chooses a secret code, consisting of four pegs, each with one of six possible colors. The second player, called *codebreaker*, tries to guess the secret code. Therefore, he asks *questions*, also in the form of four pegs in six colors. For each question, he receives an answer in the form of black and white pegs. The number of black pegs represents the number of colors occurring both in the question and the secret code, but on different positions. This game can be generalized for p pegs and c colors with $p, c \in \mathbb{N}$. In the *Black-Peg Game*, the codebreaker only receives the number of black pegs. AB Game is a restriction, where the secret code and each question contain every color at most once. This implies that $p \leq c$. For the case p = c,

© Springer International Publishing AG 2017

X. Gao et al. (Eds.): COCOA 2017, Part II, LNCS 10628, pp. 409–424, 2017.

https://doi.org/10.1007/978-3-319-71147-8_28

the secret codes and questions can be thought of as permutations. Here, every color is used exactly once, so the number of white pegs is the difference between p and the number of black pegs. Hence, this can be considered as a Black-Peg Game. In the *static* variant the second player doesn't receive the answers one at a time, but all at once after asking the last question. After that, he only has one more try to guess the code correctly.

1.1 Previous Work

Besides its popularity as board game, Mastermind has been of large interest also in theory (see the \mathcal{NP} -hardness proof in [14]) and practice (see an application in cryptography [6] and string and vector databases [1]).

Much research has been done in recent years in the area of Mastermind and its variants. The generalized version of Mastermind has been investigated in [9] and [4], in the latter a strategy with $\mathcal{O}(n \log \log n)$ questions is presented for the case p = c = n and is also adaptable to the Black-Peg variant. In [10], exact formulas are given for small p in Black-Peg Mastermind.

The best strategy known for AB Game with p = c = n needs $\mathcal{O}(n \log n)$ questions and is presented in [5]. They also give a lower bound of n questions for this setting. In [11] exact values and tight bounds for small p are presented.

Doerr et al. [4] also give an asymptotically tight bound of $\Omega(n \log n)$ questions for general Static Black-Peg Mastermind using probabilistic methods.

For p = 2 an optimal number of $\lceil (4c - 1)/3 \rceil$ questions is proven in [3] and an according strategy is presented in [8].

In this paper we consider a combination of these three variants, namely the Static Black-Peg AB Game, which, to the best of our knowledge, previously has not been considered in literature. For an overview of previous results about Classic and Static Mastermind for the case p = c = n see Table 1. Note that all known bounds equal the ones for the Black-Peg version.

Table	1.	Known	lower	and	upper	bounds	for	Classic	and	Static	Mastermind	for	p =
c = n.													

	Adaptive	Static				
(a) Lower bounds						
Classic	$\Omega(n)$	$\Omega(n\log n)$ [4]				
AB-Game	$\Omega(n)$	$\Omega(n \log n)$ (Ours)				
(b) Upper bounds						
Classic	$\mathcal{O}(n\log\log n)$ [4]	$\mathcal{O}(n\log n)$ [4]				
AB-Game	$\mathcal{O}(n\log n)$ [5]	$\mathcal{O}(n^{1.525})$ (Ours)				

1.2 Our Contribution

As the main result of this paper we show that for $n \in \mathbb{N}$ there is a feasible strategy for the Static Black-Peg AB Game on n pegs with n colors that uses $\mathcal{O}(n^{1.525})$ questions. A modification of the arguments of Doerr et al. [4] gives a lower bound of $\Omega(n \log n)$ questions. For the proof of the upper bound, we define the term "separation". Let S_n be the set of permutations of the set $\{1, \ldots, n\}$. We say that a question Q separates two possible secret codes (secrets) $X_1, X_2 \in S_n$ if Q yields different answers for them. A set of questions, called strategy, is feasible if every pair of possible secrets is separated by at least one question of the set.

First we show that there is a set of $\mathcal{O}(n^{1.525})$ questions such that every pair of possible secrets with Hamming distance of at most \sqrt{n} is separated by at least one question. For a prime n, we construct a set of $\mathcal{O}(n^{1.5})$ questions for that matter. If n is not a prime, the problem gets slightly more complicated. Here we start with a fairly simple feasible $(2n^2/3)$ -strategy. We then modify this strategy to get a set of $\mathcal{O}(n^{1.525})$ questions that for a secret gives us the placement of the last $n^{0.525}$ colors and the colors of the last $n^{0.525}$ pegs. A result of Baker et al. [2] for the difference between consecutive primes reveals that for sufficiently large nthere is a prime $p(n) \in [n - n^{0.525}, n]$, so we can use the mentioned $\mathcal{O}(p(n)^{1.5})$ questions to get the information of the first p(n) colors and pegs. Altogether for this part we use $\mathcal{O}(n^{1.525})$ questions.

For pairs of possible secrets with Hamming distance of at least \sqrt{n} there are considerably more separating questions, so we can use a different approach. We give a non-constructive proof that for every set of pairs of possible secrets with large Hamming distance, there is a question that separates at least a fraction of $\frac{1}{18\sqrt{n}}$ of it. So iteratively we can choose such a question and consider the set of pairs of possible secrets not yet separated. After $\mathcal{O}(n^{1.5} \log n)$ iteration steps every pair of possible secrets is separated by at least one question.

We complement the upper bound for p = c by an optimal $(\lceil 4c/3 \rceil - 1)$ strategy for the special case p = 2 and arbitrary $c \ge 2$. Furthermore, for small pand c, we have computed tighter upper bounds and exact values via randomized resp. brute force algorithms.

1.3 Organization of the Article

In Sect. 2, we introduce the basic definitions. A $(2n^2/3)$ -strategy is presented in Sect. 3. Section 4 contains a non-constructive proof of the upper bound of $\mathcal{O}(n^{1.525})$. We give a lower bound of $\Omega(n \log n)$ for p = c = n in Sect. 5. Section 6 is dedicated to our optimal strategies for 2 pegs. Finally, in Sect. 7, we present upper bounds and exact values for small p and c. We defer the proofs of some lemmata and examples to the full version of the paper.

2 Preliminaries

Let p denote the number of pegs and c the number of colors. In the AB Game, it holds that $p \leq c$. If p and c are fixed, we call the game (p, c)-Static Black-Peg AB

Game. The code chosen by the codemaker is called *secret*. The possible answers are written as 0B, 1B, 2B, ..., pB. For calculation purposes we often write *i* instead of *i*B when the context is clear. Each strategy for Static Black-Peg AB Game starts with r - 1 main questions which the codebreaker has to ask at the beginning of the game, and one *final question*, which has to be correct. We distinguish between a *feasible strategy*, where after the r - 1 main questions the secret is uniquely determined and an *infeasible strategy*, where after the r - 1main questions at least two secrets are possible. A strategy is called *optimal* if there is no feasible strategy with fewer questions. For fixed $p, c \in \mathbb{N}$, define sa(p, c) as the number of questions of an optimal strategy of (p, c)-Static Black-Peg AB Game. For a strategy we say that a peg *contains* $l \leq c$ colors if there are exactly *l* colors that occur in at least one main question on that peg. In the following let the pegs be numbered by $1, 2, \ldots, p$ and the colors by $1, 2, \ldots, c$.

In our context, secrets and questions are functions, i.e., mappings of the pegs $1, \ldots, p$ to the colors $1, \ldots, c$. Let Q be a question and X a possible secret. We write C(Q, X) = i if question Q would receive the answer iB for the secret X. Let X_1, X_2 be possible secrets. We say that a question Q separates X_1 and X_2 if $C(Q, X_1) \neq C(Q, X_2)$. For $n \in \mathbb{N}$ let $[n] := \{1, \ldots, n\}$. The Hamming distance $\Delta(X_1, X_2)$ of X_1 and X_2 is the number of pegs at which the corresponding colors are different. So $\Delta(X_1, X_2) = |\{i \in [p] \mid X_1(i) \neq X_2(i)\}|$. Note that in the following we consider the case p = c =: n. In this case the secrets and questions are permutations on [n]. We use the common one-line notation for permutations in the form (b_1, b_2, \ldots, b_p) , which means that 1 is mapped to b_1 , 2 is mapped to b_2 and so on. For $k \in \{0, \ldots, n\}$ the Rencontres number $D_{n,k}$ denotes the number of permutations in S_n with exactly k fixpoints and has the form

$$D_{n,k} = \frac{n!}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!}$$
 (see e.g. [12]).

3 A Feasible $\mathcal{O}(n^2)$ -Strategy for the Case n = p = c

We start presenting a feasible strategy with $\mathcal{O}(n^2)$ questions. The technique developed here will later be used to fill some gaps in the main strategy. We use the following questions.

Definition 1. For $n \in \mathbb{N}_{\geq 3}$ define the following questions by starting with the identity function and only changing the mapping of two or three elements: Let $i, j, k \in [n]$ be pairwise distinct.

$$P_{i,j}^{(2)}:[n] \longrightarrow [n], x \mapsto \begin{cases} j, & \text{if } x = i, \\ i, & \text{if } x = j, \\ x, & \text{otherwise.} \end{cases}$$

$$P_{i,j,k}^{(3)}:[n] \longrightarrow [n], x \mapsto \begin{cases} j, & \text{if } x = i, \\ k, & \text{if } x = j, \\ i, & \text{if } x = k, \\ x, & \text{otherwise.} \end{cases}$$

Let I denote the identity function on [n].

Keep in mind that $P_{i,j,k}^{(3)} = P_{j,k,i}^{(3)} = P_{k,i,j}^{(3)}$. When changing the mapping of two elements of a question, the difference of

the answers is at most 2, so there are five cases to consider. It is easy to see the conditions of each case. We denote the exclusive disjunction of events A and Bby $A \oplus B := (A \land \neg B) \lor (\neg A \land B)$.

Observation 1. Let $n \in \mathbb{N}_{\geq 3}$ and $i, j \in [n]$ be distinct. Let Q, X be permutations on [n].

$$(i) \ C\left(P_{Q(i),Q(j)}^{(2)} \circ Q, X\right) = C(Q, X) + 2 \iff Q(i) = X(j) \land Q(j) = X(i)$$

$$(ii) \ C\left(P_{Q(i),Q(j)}^{(2)} \circ Q, X\right) = C(Q, X) + 1 \iff Q(i) = X(j) \oplus Q(j) = X(i)$$

$$(iii) \ C\left(P_{Q(i),Q(j)}^{(2)} \circ Q, X\right) = C(Q, X) + 0 \iff Q(i), Q(j) \notin \{X(i), X(j)\}$$

$$(i) \ C\left(P_{Q(i),Q(j)}^{(2)} \circ Q, X\right) = C(Q, X) + 0 \iff Q(i), Q(j) \notin \{X(i), X(j)\}$$

$$(iv) C \left(P_{Q(i),Q(j)}^{(2)} \circ Q, X \right) = C(Q, X) - 1 \iff Q(i) = X(i) \oplus Q(j) = X(j)$$

$$(v) \ C\left(P_{Q(i),Q(j)}^{(2)} \circ Q, X\right) = C(Q,X) - 2 \Longleftrightarrow Q(i) = X(i) \land Q(j) = X(j)$$

We will show that with the questions of Definition 1 a feasible strategy can be constructed. The next lemmata provide some criteria to determine for given i, j whether peg *i* has color *j*.

Lemma 1. Let $n \in \mathbb{N}$ and $i \in [n]$. Let X be a possible secret on [n]. Then, X(i) = i if and only if $C(P_{i,j}^{(2)}, X) < C(I, X)$ for all $j \in [n] \setminus \{i\}$.

Lemma 2. Let $n \in \mathbb{N}_{\geq 3}$ and $i, j, k \in [n]$ be pairwise distinct. Let X be a possible secret on [n]. Then, X(i) = j if and only if one of the following conditions holds:

(i)
$$C(P_{i,j}^{(2)}, X) = C(I, X) + 2,$$

(

- $\begin{array}{l} (ii) \quad C(P_{i,j}^{(2)},X) = C(I,X) + 1 \ and \ C(P_{i,j,k}^{(3)},X) \geq C(P_{i,j}^{(2)},X), \\ (iii) \quad C(P_{i,j}^{(2)},X) = C(I,X) + 1, \ C(P_{i,j,k}^{(3)},X) = C(P_{i,j}^{(2)},X) 1 \ and \ C(P_{i,k}^{(2)},X) < C(I,X). \end{array}$

In the following lemma we show that the same result can be achieved with the question $P_{j,i,k}^{(3)}$ instead of the question $P_{i,j,k}^{(3)}$.

Lemma 3. Let $n \in \mathbb{N}_{\geq 3}$ and $i, j, k \in [n]$ be distinct. Let X be a possible secret on [n]. Then, X(i) = j if and only if one of the following conditions holds:

(i) $C(P_{i,i}^{(2)}, X) = C(I, X) + 2,$

 $\begin{array}{ll} (ii) \ C(P_{i,j}^{(2)},X) = C(I,X) + 1 \ and \ C(P_{j,i,k}^{(3)},X) = C(P_{i,j}^{(2)},X) - 2, \\ (iii) \ C(P_{i,j}^{(2)},X) = C(I,X) + 1, \ C(P_{j,i,k}^{(3)},X) = C(P_{i,j}^{(2)},X) - 1 \ and \ C(P_{i,k}^{(2)},X) \geq C(I,X). \end{array}$

Combining Lemmata 2 and 3 we construct a first feasible strategy.

Theorem 1. For $n \in \mathbb{N}_{\geq 5}$ there is a feasible strategy with at most $2n^2/3$ questions, so $\operatorname{sa}(n,n) \leq 2n^2/3$.

Proof. If for every distinct $i, j \in [n]$ the permutation $P_{i,j}^{(2)}$ is a question and there is a $k \in [n] \setminus \{i, j\}$ such that $P_{i,k}^{(2)}$ and $P_{i,j,k}^{(3)}$ (or $P_{j,i,k}^{(3)}$) are questions, this set of questions together with I forms a feasible strategy, because for every secret Xand every $i \in [n]$ we get X(i) by Lemmata 1, 2 and 3. Hedlund and Fort [7] showed that for every $n \in \mathbb{N}_{\geq 3}$ there is a set $\{T_1, \ldots, T_t\}$ with $t \leq n^2/6 + 1/3$ and $|T_{\tilde{t}}| = 3$ for all $1 \leq \tilde{t} \leq t$ such that for every pair $\{i, j\} \subset [n]$ there is a $\tilde{t} \in [t]$ with $\{i, j\} \subset T_{\tilde{t}}$. So, there is a set T of questions with $|T| \leq n^2/6 + 1/3$ such that for every $i, j \in [n]$ there is a $k \in [n] \setminus \{i, j\}$ with $P_{i,j,k}^{(3)} \in T$ or $P_{j,i,k}^{(3)} \in T$. Combined with the final question, the identity function I and the n(n-1)/2 questions $P_{i,j}^{(2)}$, we get a feasible *r*-strategy with $r \le 1 + 1 + n(n-1)/2 + n^2/6 + 1/3 \le 2n^2/3$, where $n \geq 5$. \square

Remark 1. Such a set T and thus the strategy can be easily drafted. For further details on how to construct the question set we refer to Theorem 1 of [7].

A Feasible $\mathcal{O}(n^{1.525})$ -Strategy for the Case n = p = c4

For improving the upper bound of $\mathcal{O}(n^2)$, we use this strategy just for a fraction of the number of colors and pegs. Lemmata 2 and 3 give clues about the colors of specific pegs. Since we need the concept of "separating pairs of possible secrets" in the following lemmata, we modify the result accordingly.

Remark 2. Let $n, r \in \mathbb{N}$ and $T = (Q_1, \ldots, Q_r) \in (S_n)^r$ be a strategy. There exist $X_1, X_2 \in S_n$ with $X_1 \neq X_2$ such that $C(Q_i, X_1) = C(Q_i, X_2)$ for all $i \in [r]$ if and only if the strategy is infeasible. Hence, we can prove T to be a feasible r-strategy by showing that for every $X_1, X_2 \in S_n$ with $X_1 \neq X_2$ there is an $i \in [r]$ such that Q_i separates X_1 and X_2 .

Lemma 4. Let $n \in \mathbb{N}_{\geq 3}$ and $t \in [n]$. Let $T^{(2)} := \left\{ P_{i,j}^{(2)} \mid i \in \{t+1,\ldots,n\}, j \in \{t, t\}\} \right\}$ $[n], i \neq j \Big\}, \ T^{(3)} := \Big\{ P^{(3)}_{i,j,1} \mid i \in \{t+1, \dots, n\}, j \in \{2, \dots, n\}, i \neq j \Big\} \ and \ T := \Big\{ P^{(3)}_{i,j,1} \mid i \in \{t+1, \dots, n\}, j \in \{2, \dots, n\}, i \neq j \Big\}$ $\{I\} \cup T^{(2)} \cup T^{(3)}$. Let $X_1, X_2 \in S_n$ with at least one of the following properties:

(i) There is an $i \in \{t+1,...,n\}$ with $X_1(i) \neq X_2(i)$. (ii) There is a $j \in \{t+1,...,n\}$ with $X_1^{-1}(j) \neq X_2^{-1}(j)$.

Then at least one question of T separates X_1 and X_2 .

4.1 Possible Secrets with Low Hamming Distance

In this subsection we depict how to separate pairs of possible secrets with low Hamming distance, i.e. a Hamming distance $\leq \sqrt{n}$. Let $t \in \{0, \ldots, n-1\}$ be a prime.

Definition 2. For $m, n \in \mathbb{N}$ with $m \ge n$ we denote by $\operatorname{rem}_n(m)$ (remainder) the unique integer $r \in \{0, \ldots, n-1\}$ such that there exists $q \in \mathbb{N}$ with $m = q \cdot n + r$. Let $n \in \mathbb{N}$, $t \le n$ be a prime, $k \in [t-1]$ and $l \in [t]$. Define the question $P(n, t, k, l) : [n] \longrightarrow [n]$ as

$$P(n,t,k,l)(b) = \begin{cases} \operatorname{rem}_t(k \cdot b + l) + 1, & \text{if } b \le t \\ b, & \text{if } b > t. \end{cases}$$

Lemma 5. For n, t, k, l as above, P(n, t, k, l) is a bijective function, i.e., a permutation.

The separation of a pair of possible secrets will be achieved as follows:

Lemma 6. Let $n \in \mathbb{N}$ and $X_1, X_2 \in S_n$ be possible secrets with $h := \Delta(X_1, X_2)$. Let a_1, \ldots, a_h be the pegs on which X_1 and X_2 have different colors. Let $Q \in S_n$ be a question with $Q(a_1) = X_1(a_1)$ and $Q(a_i) \neq X_2(a_i)$ for all $i \in [h]$. Then Q separates X_1 and X_2 .

If we have a pair of possible secrets X_1 , X_2 with low Hamming distance and a prime $t \leq n$ such that no condition of Lemma 4 is fulfilled, there is a suitable P(n, t, k, l) that separates X_1 and X_2 .

Lemma 7. Let $n \in \mathbb{N}$, $t \in \{\lceil \sqrt{n} \rceil, \ldots, n\}$ be a prime and $X_1, X_2 \in S_n$ be possible secrets with $2 \leq h := \Delta(X_1, X_2) \leq \sqrt{n}$. If $X_1^{-1}(b) = X_2^{-1}(b)$ for every $b \in \{t+1,\ldots,n\}$ and there is a peg $a \in [t]$ with $X_1(a) \neq X_2(a)$, then there exist $k \in [\lceil \sqrt{n} \rceil]$ and $l \in [t]$ such that P(n, t, k, l) separates X_1 and X_2 .

With the questions from Definitions 1 and 2 we can construct a strategy for separating pairs of possible secrets with low Hamming distance.

Strategy 1. (for $n \in \mathbb{N}$ sufficiently large)

- 1. Determine the largest prime p(n) in [n].
- 2. Take the identity function I as question.

3. For every $i \in \{p(n) + 1, ..., n\}, j \in [i-1]$ take question $P_{i,j}^{(2)}$.

- 4. For every $i \in \{p(n) + 1, ..., n\}, j \in \{2, ..., i 1\}$ take question $P_{i,i,1}^{(3)}$.
- 5. For every $k \in \lfloor \sqrt{n} \rfloor$ and $l \in [p(n)]$ take question P(n, p(n), k, l).

Lemma 8. Let $n \in \mathbb{N}$ be sufficiently large. Let $p(n) \in [n]$ be the largest prime in [n]. Then Strategy 1 has $\mathcal{O}(\max\{\sqrt{n} \cdot p(n), n \cdot (n-p(n))\})$ questions and every pair $X_1, X_2 \in S_n$ with $2 \leq \Delta(X_1, X_2) \leq \sqrt{n}$ is separated by at least one question.

Proof. In steps 2–4 of Strategy 1 we use less than $1 + (n - p(n)) \cdot 2n$ questions. In step 5 we add $\lceil \sqrt{n} \rceil \cdot p(n)$ questions. Overall the number of questions is in $\mathcal{O}(\max\{\sqrt{n} \cdot p(n), n \cdot (n - p(n))\})$. The claimed property of the strategy follows by Lemmata 4 and 7: Let X_1, X_2 be a pair of possible secrets with $2 \leq \Delta(X_1, X_2) \leq \sqrt{n}$. There are two cases: If there is an $i \in \{p(n) + 1, \ldots, n\}$ with $X_1(i) \neq X_2(i)$ or $X_1^{-1}(i) \neq X_2^{-1}(i)$, by Lemma 4 the questions of steps 2–4 are sufficient. Otherwise, there exists a color $a \in [p(n)]$ with $X_1(a) \neq X_2(a)$, because $X_1 \neq X_2$. So, X_1 and X_2 fulfill the properties of Lemma 7. Hence, at least one question of step 5 separates X_1 and X_2 .

For further specifying the bound mentioned in Lemma 8 we need an upper bound on the difference of consecutive primes. For the next theorem we use the following result of Baker et al.

Lemma 9 [2]. There exists an x_0 such that for all $x \ge x_0$ the interval $[x - x^{0.525}, x]$ contains at least one prime number.

Theorem 2. Let $n \in \mathbb{N}$ be sufficiently large.

- a) There exists a set T of $\mathcal{O}(n^{1.525})$ questions such that every pair $X_1, X_2 \in S_n$ with $2 \leq \Delta(X_1, X_2) \leq \sqrt{n}$ is separated by at least one question from T.
- b) If n is a prime, T has $\mathcal{O}(n^{1.5})$ questions.

Proof. Lemma 8 shows that Strategy 1 contains $\mathcal{O}\left(\max\left\{\sqrt{n}\cdot p(n), n(n-p(n))\right\}\right)$ questions. If n is a prime, we have n-p(n)=0, so there are $\mathcal{O}\left(n^{1.5}\right)$ questions. In general, $\sqrt{n}\cdot p(n) \leq n^{1.5}$ and $n \cdot (n-p(n)) \leq n^{1.525}$ because of Lemma 9. \Box

4.2 Possible Secrets with High Hamming Distance

We have yet to separate pairs of possible secrets with Hamming distance of $h \ge \sqrt{n}$. We depict this case as a problem on hypergraphs.

A hypergraph is a pair $\mathcal{H} = (V, \mathcal{E})$, where V is a finite set and \mathcal{E} is a set of subsets of V. We call elements of V vertices and the elements of \mathcal{E} edges. A vertex cover is a subset $U \subseteq V$ such that every edge $e \in \mathcal{E}$ contains at least one vertex of U. For a detailed description see e.g. [13].

We start by showing that for every pair X_1, X_2 there is a relatively large number of separating questions. Let $n \in \mathbb{N}$. We define the hypergraph $\mathcal{H} = (V, \mathcal{E})$ as follows:

- V is the set of questions, so $V := S_n$.
- For every pair of possible secrets X_1, X_2 with $\Delta(X_1, X_2) \ge \sqrt{n}$, we create an edge called H_{X_1, X_2} . So, $\mathcal{E} := \{H_{X_1, X_2} \mid X_1, X_2 \in S_n, \Delta(X_1, X_2) \ge \sqrt{n}\}.$
- An edge H_{X_1,X_2} consists of all questions Q separating X_1 and X_2 . So, for all $H_{X_1,X_2} \in \mathcal{E}$ we have $H_{X_1,X_2} = \{Q \in S_n \mid C(Q,X_1) \neq C(Q,X_2)\}.$

Lemma 10. Let $n \geq 6$. Let $X_1, X_2 \in S_n$ with $\Delta(X_1, X_2) \geq \sqrt{n}$ and $e := H_{X_1, X_2} \in \mathcal{E}$. Then $|e| \geq n! \cdot \frac{1}{18\sqrt{n}}$.

Similarly, there is always a question that separates a relatively large number of pairs of possible secrets with large Hamming distance.

Lemma 11. Let $n \ge 6$. For every subset $\emptyset \neq \mathcal{F} \subseteq \mathcal{E}$, there is a vertex $Q \in V$ with $|\{e \in \mathcal{F} \mid Q \in e\}| \ge \frac{1}{18\sqrt{n}} \cdot |\mathcal{F}|$.

We can iteratively pick such questions to separate as many pairs of possible secrets as possible. After $\mathcal{O}(n^{1.5} \log n)$ iteration steps every pair with high Hamming distance is separated.

Theorem 3. Let $n \ge 6$. There exists a set of $\mathcal{O}(n^{1.5} \log n)$ questions such that every pair $X_1, X_2 \in S_n$ with $\Delta(X_1, X_2) \ge \sqrt{n}$ is separated by at least one question.

Proof. We prove that a vertex cover T of \mathcal{H} with $\mathcal{O}(n^{1.5} \log n)$ vertices exists. This translates to a set \tilde{T} of questions with the needed property, because for every distinct pair $X_1, X_2 \in S_n$ there is a vertex $Q \in T$ that covers the edge $H_{X_1,X_2} \in \mathcal{E}$, so the corresponding question separates X_1 and X_2 . With Lemma 11, for every subset $\emptyset \neq \mathcal{F} \subseteq \mathcal{E}$ there is a vertex $Q \in V$ with $|\{e \in \mathcal{F} \mid Q \in e\}| \geq \frac{1}{18\sqrt{n}} \cdot |\mathcal{F}|$. Hence, for every subset $\emptyset \neq \mathcal{F} \subseteq \mathcal{E}$ we can pick a vertex $Q \in V$ that leaves at most $\left(1 - \frac{1}{18\sqrt{n}}\right) \cdot |\mathcal{F}|$ uncovered. Now we can start with the empty set $T = \emptyset$, and iteratively add vertices to T, which at the time are in at least a fraction of $\frac{1}{18\sqrt{n}}$ of the uncovered edges. After t steps the fraction of uncovered edges is at most $\left(1 - \frac{1}{18\sqrt{n}}\right)^t$. With $t := 36 \ln(n) \cdot n^{1.5}$, the fraction of uncovered edges after t steps is at most

$$\left(1 - \frac{1}{18\sqrt{n}}\right)^{36\ln(n)\cdot n^{1.5}} = \left(1 - \frac{1}{18\sqrt{n}}\right)^{18\sqrt{n}\cdot(2\ln(n)\cdot n)}$$
$$\leq \left(\frac{1}{e}\right)^{2\ln(n)\cdot n}$$
$$= \left(\frac{1}{n}\right)^{2n} \leq \frac{1}{(n!)^2}.$$

Since $|\mathcal{E}| < (n!)^2$, after about t iteration steps T is a vertex cover of \mathcal{H} . So, $\mathcal{O}(n^{1.5}\log(n))$ questions suffice for separating every pair $X_1, X_2 \in S_n$ with $\Delta(X_1, X_2) \ge \sqrt{n}$.

Finally, we can prove our main result.

Theorem 4. Let $n \in \mathbb{N}$ be sufficiently large. There exists a set T of $\mathcal{O}(n^{1.525})$ questions such that every pair $X_1, X_2 \in S_n$ is separated by at least one question from T. Therefore, T is a feasible strategy.

Proof. Remark 2 states that a strategy is feasible if every pair of possible secrets is separated by at least one question. Theorem 2 implies that there are $\mathcal{O}(n^{1.525})$

questions separating every pair of possible secrets with low Hamming distance. Because of Theorem 3, $\mathcal{O}(n^{1.5} \log n)$ questions are sufficient for separating pairs of possible secrets with high Hamming distance. Altogether, there exists a feasible strategy with $\mathcal{O}(n^{1.525})$ questions.

5 A Lower Bound of $\Omega(n \log N)$ for the Case n = p = c

In this section we present a lower bound of $\Omega(n \log n)$ questions for Static Black-Peg AB Game for n = p = c. The technique is based on information theory and adopted from [4]. Note that in the following we use the logarithm to the base 2 and denote it by "log".

We introduce a few notions and results from information theory.

Definition 3. Let X be a discrete random variable on a domain D. The entropy of X is defined as

$$H(X) := \sum_{x \in D} \Pr[X = x] \log \left(\frac{1}{\Pr[X = x]}\right).$$

Intuitively speaking, the entropy is a measure on how much information X will reveal in expectation.

We need the following well-known properties of entropy:

Lemma 12. Let X, Y be discrete random variables.

(i) $H((X,Y)) \le H(X) + H(Y)$. (ii) If X = f(Y) for some deterministic function f, then $H(X) \le H(Y)$.

Consider a possible secret $X \in S_n$ chosen uniformly at random (so $H(X) = \log(|S_n|) = \log(n!)$). Let s be the size of a feasible strategy. For $i \in [s]$ let Y_i be the answer to the *i*-th question. Since our strategy is feasible, the sequence $Y = (Y_1, \ldots, Y_s)$ determines X and hence we have $H(Y) \ge H(X)$ by Lemma 12 (ii). On the other hand, $H(Y) = H(Y_1, \ldots, Y_s) \le \sum_{i=1}^s H(Y_i)$ by Lemma 12 (i). Recalling the definition of the Rencontres number $D_{n,k}$, for any $i \in [s]$ we have $H(Y_i) = \sum_{k=0}^n \frac{D_{n,k}}{n!} \log\left(\frac{n!}{D_{n,k}}\right)$. Note that

$$D_{n,k} = \frac{n!}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} \le \frac{n!}{2k!} \text{ for any } k < n$$
(1)

and on the other hand

$$D_{n,k} \ge \frac{n!}{3k!} \text{ for any } k < n-1.$$
(2)

Moreover $D_{n,n} = 1$ and $D_{n,n-1} = 0$, so for any $i \in [s]$ and $n \ge 4$ we get

$$\begin{split} H(Y_i) &= \sum_{k=0}^{n} \frac{D_{n,k}}{n!} \log\left(\frac{n!}{D_{n,k}}\right) \\ &= \frac{\log(n!)}{n!} + \sum_{k=0}^{n-2} \frac{D_{n,k}}{n!} \log\left(\frac{n!}{D_{n,k}}\right) \\ &\leq \frac{\log(n!)}{n!} + \frac{1}{2} \sum_{k=0}^{n-2} \frac{\log(3k!)}{k!} \qquad (\text{Eq. (1),(2)}) \\ &= \frac{\log(n!)}{n!} + \frac{1}{2} \sum_{k=0}^{n-2} \frac{\log 3}{k!} + \frac{1}{2} \sum_{k=2}^{n-2} \frac{\log(k!)}{k!} \\ &\leq \frac{n \log n}{n!} + \frac{\log 3}{2} e + \frac{1}{2} \sum_{k=2}^{n-2} \frac{\log(k!)}{k!} \\ &\leq \frac{e}{5} + \frac{4e}{5} + \frac{1}{2} \sum_{k=2}^{n-2} \frac{\log(k!)}{k!} \qquad (n \ge 4) \\ &\leq e + \frac{1}{2} \sum_{k=2}^{n-2} \frac{1}{(k-2)!} \\ &\leq \frac{3}{2} e. \end{split}$$

So altogether we have $\log(n!) = H(X) \leq H(Y) \leq \frac{3se}{2}$ and hence $s \geq 2\log(n!)/3e = \Omega(n\log n)$.

6 An Optimal $(\lceil 4c/3 \rceil - 1)$ -Strategy for p = 2

In this section we present a $(\lceil 4c/3 \rceil - 1)$ -strategy for the case of p = 2 pegs and arbitrarily many colors $c \ge 2$. In the following let p = 2. Observe that a feasible strategy for Static Black-Peg Mastermind is automatically also a feasible strategy for the Static Black-Peg AB Game if in each question no color occurs twice. The idea of the following strategies for the Static Black-Peg AB Game and arbitrarily many colors $c \ge p = 2$ is to use the strategy for Static Black-Peg Mastermind for p = 2 pegs from [8] as basis, apply it to c - 2 colors and add two further main questions to this strategy, namely the questions (c - 2, c - 1) and (c - 1, c). In the following we introduce a $(\lceil 4c/3 \rceil - 1)$ -strategy for each $c \ge 3$ except c = 4, 5. We distinguish between the cases $c \equiv 0 \mod 3$, $c \equiv 1 \mod 3$, $c \equiv 2 \mod 3$. For the number $k := (\lceil 4c/3 \rceil - 2)$ of main questions it holds that

$$k = \begin{cases} \frac{4}{3} \cdot c - 2 = 4 \cdot \frac{c}{3} - 2 \equiv 2 \mod 4 \text{ for } c \equiv 0 \mod 3\\ \frac{4}{3} \cdot c - \frac{4}{3} = 4 \cdot \frac{c - 1}{3} \equiv 0 \mod 4 \text{ for } c \equiv 1 \mod 3\\ \frac{4}{3} \cdot c - \frac{5}{3} = 4 \cdot \frac{c - 2}{3} + 1 \equiv 1 \mod 4 \text{ for } c \equiv 2 \mod 3 \end{cases}$$
(3)

Strategy 2. $((\lceil 4c/3 \rceil - 1))$ -strategy for p = 2 and arbitrary $c \equiv 0 \mod 3$

- 1. Divide the k main questions into three blocks of questions, the first (k-2)/2 questions and the second (k-2)/2 questions, and two additional questions (c-2, c-1) and (c-1, c).
- The first peg contains the colors 1, 2,..., (k − 2)/2 in the first block and the colors (k − 2)/2 + 1, (k − 2)/2 + 1, (k − 2)/2 + 2, (k − 2)/2 + 2..., 3(k − 2)/4, 3(k − 2)/4(= c − 3) in the second block.
- 3. In the first two blocks, the second peg is received from the first peg by switching the role of the two blocks.
- 4. Finally, the secret has to be asked as final question Q_{k+1} .

Strategy 3. $((\lceil 4c/3 \rceil - 1)$ -strategy for p = 2 and arbitrary $c \equiv 1 \mod 3, c \neq 1, 4)$

- 1. Divide the k main questions into three blocks of questions, the first (k-2)/2 questions and the second (k-2)/2 questions, and two additional questions (c-2, c-1) and (c-1, c).
- 2. The first peg contains the colors 1, 2, ..., (k-2)/2 in the first block and the colors (k-2)/2 + 1, (k-2)/2 + 1, (k-2)/2 + 1, (k-2)/2 + 2, (k-2)/2 + 2, ..., 3(k-4)/4 + 1, 3(k-4)/4 + 1(= c 3) in the second block (note that the first number (k-2)/2 + 1 occurs three times, not only twice).
- 3. In the first two blocks, the second peg is received from the first peg by switching the role of the two blocks.
- 4. Finally, the secret has to be asked as final question Q_{k+1} .

Strategy 4. $((\lceil 4c/3 \rceil - 1)$ -strategy for p = 2 and arbitrary $c \equiv 2 \mod 3$, $c \neq 2, 5)$

- 1. Divide the k main questions into three blocks of questions, the first (k-3)/2 questions and the second (k-1)/2 questions, and two additional questions (c-2, c-1) and (c-1, c).
- 2. The first peg contains the colors 1, 2, ..., (k-3)/2 in the first block and the colors (k-1)/2, (k-1)/2, (k-1)/2 + 1, (k-1)/2 + 1, ..., 3(k-1)/4 1, 3(k-1)/4 1(=c-3) in the second block.
- 3. The second peg contains the colors (k-1)/2 + 1, (k-1)/2 + 1, (k-1)/2 + 1, (k-1)/2 + 2, (k-1)/2 + 2, \dots , 3(k-1)/4 1, 3(k-1)/4 1 in the first block and the colors 1, 2, ..., (k-1)/2 in the second block (note that the first number (k-1)/2 + 1 occurs three times, not only twice).
- 4. Finally, the secret has to be asked as final question Q_{k+1} .

Remark 3. Note that in all three strategies the colors c and c-2 are not used in the first peg and second peg, respectively, of the main questions at all.

As examples, the main questions of Strategy 2 for p = 2 and c = 12 with k = 14 questions (the first 12 questions can be found in Table 1(b) of [8]), Strategy 3 for p = 2 and c = 13 with k = 16 questions (the first 14 questions can be found in Table 1(c) of [8]), and Strategy 4 for p = 2 and c = 11 with k = 13

			Peg	1	2			
			Q_1	1	8			
Peg	1	2	Q_2	2	8			
Q_1	1	7	Q_3	3	8	Peg	1	2
Q_2	2	7	Q_4	4	9	Q_1	1	7
Q_3	3	8	Q_5	5	9	Q_2	2	7
Q_4	4	8	Q_6	6	10	Q_3	3	7
Q_5	5	9	Q_7	7	10	Q_4	4	8
Q_6	6	9	Q_8	8	1	Q_5	5	8
Q_7	7	1	Q_9	8	2	Q_6	6	1
Q_8	7	2	Q_{10}	8	3	Q_7	6	2
Q_9	8	3	Q_{11}	9	4	Q_8	7	3
Q_{10}	8	4	Q_{12}	9	5	Q_9	7	4
Q_{11}	9	5	Q_{13}	10	6	Q_{10}	8	5
Q_{12}	9	6	Q_{14}	10	7	Q_{11}	8	6
Q_{13}	10	11	Q_{15}	11	12	Q_{12}	9	10
Q_{14}	11	12	Q_{16}	12	13	Q_{13}	10	11
(a) $c =$	12, k	= 14.	(b) $c =$	13, k	= 16.	(c) $c =$	11, k	= 13.

Table 2. Examples for Strategies 2, 3 and 4 with p = 2.

questions (the first 11 questions can be found in Table 1(a) of [8]) are listed in Tables 2a, b and c, respectively.

This idea works for all $c \geq 2$ except for c = 2, 4, 5. The case c = 2 is trivial. Regarding the case c = 4, one optimal strategy for Static Black-Peg Mastermind for c - 2 = 2 contains the main question (1, 1) which is forbidden in the AB Game. Analogously, regarding the case c = 5, one optimal strategy for Static Black-Peg Mastermind for c - 2 = 3 contains the forbidden main question (2, 2). These cases are considered in the following observation.

Observation 2. (a) For c = 2, the strategy which consists only of the main question (1,2) is a feasible strategy which needs $(\lceil 4 \cdot c/3 \rceil - 1) = 2$ questions.

- (b) For c = 4, the strategy which consists of the four main questions (1,2), (1,3) (2,1), (3,1) is a feasible strategy for Static Black-Peg AB Game. It needs ([4 ⋅ c/3] − 1) = 5 questions.
- (c) For c = 5, the strategy which consists of the five main questions (1,2), (1,3), (2,1), (3,1), (4,5) is a feasible strategy for Static Black-Peg AB Game. It needs ([4 ⋅ c/3] 1) = 6 questions.

Theorem 5. The strategy of [8] for Static Black-Peg Mastermind for c-2 colors plus the additional main questions (c-2, c-1), (c-1, c) is a feasible and optimal $(\lceil 4c/3 \rceil - 1)$ -strategy for p = 2 and for the corresponding $c \ge 3$, $c \ne 4, 5$. I.e., $sa(2, c) = (\lceil 4c/3 \rceil - 1)$ for arbitrary $c \ge 2$.

We obtain the following interesting relations in comparison to the strategies of [8].

- Remark 4. (a) For $c \equiv 0 \mod 3$, the strategies of Theorem 5 for the Static Black-Peg AB Game need one question less than Strategy 1 from [8] for Static Black-Peg Mastermind.
- (b) For $c \equiv 1 \mod 3$, the strategies of Theorem 5 and Observation 2 for the Static Black-Peg AB Game need the same number of questions as Strategy 2 from [8] for Static Black-Peg Mastermind.
- (c) For $c \equiv 2 \mod 3$, the strategies of Theorem 5 and Observation 2 need one question less than Strategy 3 from [8] for Static Black-Peg Mastermind.

7 Optimal and Random Strategies and Future Work

In Sect. 6 we computed exact values for sa(2, c) for all $c \ge 2$ by giving optimal strategies. However, for pairs (p, c) with $3 \le p \le c$ it seems rather difficult to construct strategies that are optimal, or at least close to optimum with sensible computing effort. We tackled this problem using *random strategies*, i.e. strategies, where each main question is chosen randomly and uniformly distributed over all possible questions to compute tighter upper bounds on sa(p, c) for pairs (p, c) with larger p, c.

For this purpose we used a computer program which for small p, c finds optimal strategies by brute-force search and for larger p, c generates a random strategy of a given length and checks, whether the computed strategy is feasible. The program was implemented in the programming language C++, and all experiments were done on a standard desktop in a Unix-based system. Its source code is available online at [15].

The results for $1 \le p \le c \le 8$ can be found in Table 3.

In addition to the values for p = 2 we were able to compute exact values for sa(pc) for six additional pairs and upper bounds for several other cases.

Further, note that for *all* pairs (p, c), where we could (theoretically or by the program) validate exact values, at least one of 10,000 tested random strategies were optimal. The computed upper bounds turn out to be remarkably smaller than the strategy constructed in Sect. 3 (e.g. 18 questions vs. 41 questions for p = c = 8).

Regarding future work, this motivates both to theoretically investigate the behavior of random strategies and to improve the known upper bounds for the case p = c, but also for the case p < c.

c	p										
	1	2	3	4	5	6	7	8			
1	1	_	_	_	_	_	_	_			
2	2	2	_	-	_	—	-	_			
3	3	3	5	-	—	_	-	—			
4	4	5	5	5	—	—	-	—			
5	5	6	7	7	7	—	-	_			
6	6	7	≤ 8	≤ 10	≤ 11	≤ 9	-	_			
7	7	9	≤ 10	≤ 13	≤ 16	≤ 15	$\leq \! 13$	-			
8	8	10	≤ 12	≤ 15	$\leq \! 17$	≤ 19	≤ 20	$\leq \! 18$			

Table 3. Summary of results for values sa(p, c) for $p \le c \le 8$

Acknowledgments. The second author's research was supported by the Kempe Foundation Grant No. SMK-1354 (Sweden).

Furthermore, we would like to thank the anonymous referees for their valuable comments which significantly helped to improve the paper.

References

- Asuncion, A.U., Goodrich, M.T.: Nonadaptive mastermind algorithms for string and vector databases, with case studies. IEEE Trans. Knowl. Data Eng. 25(1), 131–144 (2013)
- Baker, R.C., Harman, G., Pintz, J.: The difference between consecutive primes II. Proc. Lond. Math. Soc. 83(3), 532–632 (2001)
- Cáceres, J., Hernando, C., Mora, M., Pelayo, I.M., Puertas, M.L., Seara, C., Wood, D.R.: On the metric dimension of cartesian products of graphs. SIAM J. Discret. Math. 21(2), 423–441 (2007)
- Doerr, B., Doerr, C., Spöhel, R., Thomas, H.: Playing Mastermind With Many Colors. J. ACM 63(5), 42:1–42:23 (2016). ACM
- El Ouali, M., Glazik, C., Sauerland, V., Srivastav, A.: On the query complexity of black-peg AB-mastermind. CoRR, abs/1611.05907 2016) http://arxiv.org/abs/ 1611.05907
- Focardi, R., Luccio, F.L.: Guessing bank PINs by winning a mastermind game. Theory Comput. Syst. 50(1), 52–71 (2012)
- Fort, M.K., Hedlund, G.A.: Minimal coverings of pairs by triples. Pac. J. Math. 8(4), 709–719 (1958)
- Jäger, G.: An optimal strategy for static black-peg mastermind with two pegs. In: Chan, T.-H.H., Li, M., Wang, L. (eds.) COCOA 2016. LNCS, vol. 10043, pp. 670–682. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-48749-6_48
- Jäger, G., Peczarski, M.: The number of pessimistic guesses in generalized mastermind. Inf. Process. Lett. 109(12), 635–641 (2009)
- Jäger, G., Peczarski, M.: The number of pessimistic guesses in generalized blackpeg mastermind. Inf. Process. Lett. **111**(19), 933–940 (2011)
- Jäger, G., Peczarski, M.: The worst case number of questions in generalized AB game with and without white-peg answers. Discret. Appl. Math. 184, 20–31 (2015)

- 12. Riordan, J.: Introduction to Combinatorial Analysis. Dover Books on Mathematics. Dover Publications, New York (2002)
- 13. Schrijver, A.: Combinatorial Optimization: Polyhedra and Efficiency. Algorithms and Combinatorics, 1st edn. Springer, Heidelberg (2003)
- Stuckman, J., Zhang, G.Q.: Mastermind is NP-complete. INFOCOMP J. Comput. Sci. 5(2), 25–28 (2006)
- 15. Source Code of the Computer Program of this Article. http://snovit.math.umu. se/~gerold/source_code_static_ab_game.tar.gz