

Upper bounds for the necklace folding problems

Endre Csóka* Zoltán L. Blázsik† Zoltán Király‡ Dániel Lenger§

May 27, 2020

Abstract

A necklace can be considered as a cyclic list of n red and n blue beads in an arbitrary order, and the goal is to fold it into two and find a large cross-free matching of pairs of beads of different colors. We give a counterexample for a conjecture about the necklace folding problem, also known as the separated matching problem. The conjecture (given independently by three sets of authors) states that $\mu = \frac{2}{3}$, where μ is the ratio of the ‘covered’ beads to the total number of beads.

We refute this conjecture by giving a construction which proves that $\mu \leq 2 - \sqrt{2} < 0.5858$. Our construction also applies to the homogeneous model: when we are matching beads of the same color. Moreover, we also consider the problem where the two color classes not necessarily have the same size.

1 Introduction

In the last decades, essentially the same problem known as the necklace folding or the separated matching problem appeared in many areas of mathematics. The problem has two variants which we call the heterogeneous and the homogeneous model. Consider a necklace which consists of $N = 2n$ beads, n red and n blue ones. In both models, the aim is to find a specific folding of the necklace defined as follows. A matching M consists of $|M|$ mutually disjoint pairs of beads.

In the heterogeneous model each pair consists of one red and one blue bead while in the homogeneous model each pair consists of two beads of the same color. The matched pairs will be also called matching edges.

*Alfréd Rényi Institute, Budapest, Hungary. The research was supported by ERC Synergy grant No. 810115.

†MTA–ELTE Geometric and Algebraic Combinatorics Research Group, Budapest, Hungary. The research was supported by the Hungarian National Research, Development and Innovation Office, OTKA grant no. SNN 132625.

‡Department of Computer Science, Eötvös Loránd University, Budapest, Hungary. This research was partially supported by a grant (no. FK 132524) from the National Development Agency of Hungary, based on a source from the Research and Technology Innovation Fund, and in part by the Project no. ED-18-1-2019-0030 (Application-specific highly reliable IT solutions) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme funding scheme.

§Department of Computer Science, Eötvös Loránd University, Budapest, Hungary.

A matching is *cross-free* if no two matching edges cross each other. That is, if the two matching edges are ab and cd , then one arc between a and b is disjoint from the set $\{c, d\}$ while the other arc contains this set entirely.

A *secant* partitions the necklace into two continuous arcs, A_1 and A_2 . A matching is *secant-respecting* if for each matching edge, one end is in A_1 while the other end is in A_2 . We call a matching *proper* if it is cross-free and secant-respecting.

Let us remark here that if we drop the secant-respecting condition, then one can easily prove that there is always a cross-free matching consisting of n edges in the heterogeneous (and $n - 1$ edges in the homogeneous) model.

Proper matchings were called *separated* matchings in [2, 3, 7] where the same problem was considered in a geometric setup. We have n red and n blue points on a circle, the matching edges are considered as segments. A matching is non-crossing if the corresponding segments are pairwise disjoint and a non-crossing matching is separated if there is a straight-line that intersects the interior of each of its segments.

Let M be a proper matching. The size $|M|$ of the matching is the number of its edges. A bead is *covered* if it is contained in a matched pair, the number of the covered beads is clearly $2|M|$. Remember that a necklace consists of $N = 2n$ beads, half of them is red, the other half is blue (i.e., it is balanced). For an even integer N , let $\mathcal{N}(N)$ denote the set of possible balanced necklaces with N beads, and let $\mathcal{M}(L)$ denote the set of proper matchings for a given necklace L in the heterogeneous model, and $\mu(N, L) = \max_{M \in \mathcal{M}(L)} 2|M|$, i.e., the maximum number of covered beads in a proper matching. Moreover, let $\mu(N) = \min_{L \in \mathcal{N}(N)} \mu(N, L)$. Thus $\mu(N)$ is the maximum number of coverable beads in the ‘worst’ necklace. We are interested in $\frac{\mu(N)}{N}$, the ratio of the covered beads to the total number of beads. Remark that it is the same as $|M|/n$. Finally, let $\mu = \limsup_{N \rightarrow \infty} \mu(N)$. For the homogeneous model, we similarly define $\mu^{\text{hom}}(N)$ and μ^{hom} .

It is trivial that there is a proper matching of size $n/2$ in any given necklace for both models. In the heterogeneous model, one can take an arbitrary secant which cuts the necklace into two arcs each of which containing n beads. Since the number of blue and red beads are the same, therefore in one of the arcs there are at least as many blue beads as red ones, and in the other arc the opposite is true. Thus we can create a proper matching using the beads of the majority color from each arc. In the homogeneous model, one can take an appropriate secant for which the two arcs have the same number of blue beads. Then there is a proper matching of size $\lfloor n/2 \rfloor$. That is, $\mu \geq \frac{1}{2}$ and $\mu^{\text{hom}} \geq \frac{1}{2}$.

It was very exciting that for 20 years there were no significant improvements about this lower bound, only about the additional $o(n)$ term. However, very recently Mulzer and Valtr [11] managed to improve the lower bound of μ to $(1/2 + \varepsilon)$ for some absolute constant $\varepsilon > 0$.

The story regarding the upper bound is more diversified. Originally only the heterogeneous model was studied. Lyngsø and Pedersen [1] in 1999 proved that $\mu \leq 2/3$, and they conjectured that $\mu = 2/3$. Later independently Kynčl, Pach and Tóth [2, 3] and Brevier, Preissmann and Sebő [5] proved the same upper bound and formulated the same conjecture.

Conjecture 1. [1, 2, 3, 5] *In the heterogeneous model, there is always a proper matching of size at least $2n/3 - o(n)$, i.e., $\mu = 2/3$.*

Actually, in [2, 3] a more refined conjecture can be found. For a necklace $L \in \mathcal{N}(N)$, let $\text{mono}(L)$ denote the number of maximal monochromatic arcs, i.e., there is $\text{mono}(L)$ color changes in the necklace, or in other words, the necklace L consist of $\text{mono}(L)/2$ red arcs and $\text{mono}(L)/2$ blue arcs.

Conjecture 2 ([2, 3]). *If we restrict ourselves to necklaces L where $\text{mono}(L)/2 = k$, then for every constant k there is always a proper matching of size at least $\frac{2k-1}{3k-2}n - o(n)$ in the heterogeneous model.*

However, for the strict connection to Erdős problem about non-crossing alternating path, it is enough to assume that $k = o(n)$ (see below). In this case Conjecture 2 can be read as follows.

Conjecture 3. *In the heterogeneous model, there is always a proper matching of size at least $2n/3 - o(n)$, i.e., $\mu = 2/3$, if restrict ourselves for necklaces $L \in \mathcal{N}(N)$ where $\text{mono}(L) = o(n)$.*

Surprisingly, there are several connections between our problem and some interesting questions from different topics in mathematics. In the sequel, we are going to mention some of these examples as a motivation to our study. The following problem is due to Erdős from the late 80's.

Problem 1. *Determine or estimate the largest number $\ell = \ell(N)$ such that, for every set of $N/2$ red and $N/2$ blue points on a circle, there exists a non-crossing alternating path consisting of ℓ vertices.*

Kynčl, Pach and Tóth [2, 3] disproved the original conjecture of Erdős (stating that $\ell(N) = \frac{3}{4}N + o(N)$), and showed the following:

Theorem 1 ([2, 3]). *There exist constants $c, c' > 0$ such that $\frac{1}{2}N + c\sqrt{\frac{N}{\log N}} < \ell(N) < \frac{2}{3}N + c'\sqrt{N}$.*

Moreover, they conjectured that the upper bound is asymptotically tight.

Conjecture 4 ([2, 3]). $|\ell(N) - \frac{2}{3}N| = o(N)$.

Given a necklace $L \in \mathcal{N}(N)$, let $\ell(L)$ denote the maximum length of a non-crossing alternating path. They also proved the following.

Theorem 2 ([2, 3]). $\ell(N) - 2 \cdot \text{mono}(L) - 1 \leq \mu(N) \leq \ell(N)$.

About the same time Abellanas et al. [4] showed a very similar construction for the same upper bound. In 2010, Hajnal and Mészáros [6, 7] improved the lower bound on $\ell(N)$ to $N/2 + \Omega(\sqrt{n})$, and also gave a class of configurations reaching the upper bound.

Mészáros [8, 7] investigated separated matchings and found new families of constructions containing at most $\frac{4}{3}n + O(\sqrt{n})$ points in any separated matching. Furthermore, she showed that if the discrepancy is at most three, then there are at least $\frac{4}{3}n$ points in the maximum separated matching.

Our main theorem (Theorem 3) disproves Conjecture 4 as well by using Theorem 2.

Interestingly, the above mentioned problems are closely related to some applied questions about the structure of proteins and some very natural questions about drawing some geometric graphs with non-crossing straight-line edges, too. In 1999, Lyngsø and Pedersen [1] studied folding algorithms in the two dimensional Hydrophobic-Hydrophilic model (2D HP) for protein structure formation. They provided some approximation algorithms so that the approximation ratio depends exactly on the size of the largest proper matching in our terminology, and conjectured that there always exist a proper matching of size at least $2n/3$.

Moreover, there are some connections between these problems and the investigation of subsequences in circular words over the binary alphabet. One can rephrase Conjecture 1 with this terminology as it states that every binary circular word of length N with equal number of zeros and ones has an antipalindromic linear subsequence of length at least $2N/3 - o(N)$. Recently, independently from our work, Müllner and Ryzhikov [9, 10] gave a construction (which is essentially the same as our simple construction) that yields an upper bound of $2N/3 + o(N)$ for both the heterogeneous and the homogeneous models (in this latter model we are looking for a palindromic linear subsequence). It seems that they were the first who studied the homogeneous model, and they made the following conjecture.

Conjecture 5 ([9, 10]). $\mu^{\text{hom}} = 2/3$.

We disprove all Conjectures above. Furthermore, we improve the best known upper bound significantly by proving the following theorem.

Theorem 3. *For Construction 2, the size of the maximum proper matching is at most $(2 - \sqrt{2})n + o(n)$ in both the heterogeneous and the homogeneous models (i.e., $\mu \leq 2 - \sqrt{2}$ and $\mu^{\text{hom}} \leq 2 - \sqrt{2}$). Moreover, Construction 2 gives an infinite series of necklaces where $\text{mono}(L) = o(n)$.*

Remark 4. *It is not obvious how this theorem disproves Conjecture 2. Without the details we sketch the transition. By Theorem 3, there exists a specific necklace L_1 with N_1 beads where $\mu(N_1, L_1) < 0.6$. Let $k = \text{mono}(L_1)/2$. We are giving a counterexample to Conjecture 2 for this k , i.e., an infinite series of necklaces L_i with N_i beads where $\text{mono}(L_i) = 2k$ and $\mu(N_i, L_i) < 0.6$.*

We get L_i from L_1 by replacing every bead by i beads of the same color. So $N_i = iN_1$, and obviously $\text{mono}(L_i) = \text{mono}(L_1) = 2k$. Using the fact that in bipartite graphs the weight of the maximum fractional matching is the same as the weight of the maximum matching, it is not hard to prove that $\mu(N_i, L_i) = \mu(N_1, L_1)$.

Remark 5. *The problem itself, and also our construction can be defined in a measurable sense, i.e., necklace is a circle with a measurable two-coloring on its points, and for a proper matching we also require it to be measure-preserving. This is a natural generalization of the discrete problem. Although this language was very useful for finding our counterexample, we present our result in the more classical language of discrete objects. If we used the measurable definition, we may omit the terms $o(N)$ everywhere.*

The paper is organized as follows. In Section 2, we are going to provide our simple construction (Construction 1) and prove that any proper matching in it has got at most

$2n/3$ edges, furthermore it works for both models. We will modify the previous construction in order to further improve the upper bound and prove our main result, Theorem 3, in Section 3. Matter of fact, the improved construction (Construction 2) also works for both models.

Section 4 is devoted to the study of the unbalanced case where one color occurs more than the other. We give new upper bounds for the case where the number of red beads is between $N/3$ and $2N/3$ by slightly changing the previous construction. However, here we need to distinguish between the two models. Quite surprisingly, the upper bounds which we are able to acquire for the two models are the same accidentally, but the constructions are not exactly the same.

2 The simple construction

First, we give an upper bound of $2n/3$ for both μ and μ^{hom} , i.e., for both the heterogeneous and the homogeneous models. After writing down this section we learned that in a recent work by Müllner and Ryzhikov [9, 10], the same construction has been already given. However, we decided to keep this section as a gentle introduction to our main result.

Construction 1. *Let $s \geq 2$ be a integer parameter, and let $n = s^{s+1}$. The necklace consists of s large arcs, each having s^s blue and s^s red beads. Let L_1, \dots, L_s denote the large arcs.*

L_i is divided into s^{s-i} red and s^{s-i} blue monochromatic arcs, the colors alternates. A monochromatic arc inside L_i always consists of s^i beads.

For analyzing this construction, we fix an optimal proper secant-matching pair in either the homogeneous or the heterogeneous model, and denote this optimal matching by M . The secant may split at most two large arcs, call them L_p and L_r . If, e.g., one end of the secant is between the large arcs L_j and L_{j+1} , then let $p = j$. We may assume that $p < r$ (if $p = r$, then every matching edge has one end in L_p , so $|M| \leq n/s$).

Let x be the first M -matched bead in the sequence L_1, L_2, \dots , and $y \in L_q$ the bead matched to x ; if $q < p$, then set $q = p$. Thus $p \leq q \leq r$, and we have the following property: any bead in $L_{p+1} \cup \dots \cup L_{q-1}$ is matched into some L_k for $1 \leq k \leq p$ and any bead in $L_{q+1} \cup \dots \cup L_{r-1}$ is matched into some L_k for $r \leq k \leq s$. We consider large arcs L_p, L_q and L_r as exceptional.

Let $M' \subseteq M$ consist of those matching edges for which no end-vertex is in the set $L_p \cup L_q \cup L_r$. Our first goal is to show that $|M'|$ cannot be too large.

Let $\ell_{i,j}$ denote the j th monochromatic arc of L_i , its length is $|\ell_{i,j}| = s^i$. If x and y are the first and last M' -matched bead in $\ell_{i,j}$, and $M'(x)$ and $M'(y)$ are their matched partners, then we assign the arc spanned by $M'(x)$ and $M'(y)$ to $\ell_{i,j}$, denote this arc by $M'(\ell_{i,j})$.

We partition a subset of the beads. Let $P(i, j) = \ell_{i,j} \cup M'(\ell_{i,j})$ for every i whenever $p < i < q$ or $r < i \leq s$, and for every $j \leq 2s^{s-i}$. Clearly these $P(i, j)$ sets are pairwise disjoint. The idea is the following. For every edge of M' , exactly one of its end-vertex is in some L_i for either $p < i < q$ or $r < i \leq s$. The other end-vertex is in some L_k where $k < i$. Note that the monochromatic arcs in L_k are shorter than the monochromatic arcs in L_i , so each monochromatic arc in $M'(\ell_{i,j})$ has length at most s^{i-1} .

Lemma 6. *In the set $P(i, j)$ at most $(2/3 + 1/s)$ fraction of the beads are M' -covered. This is true in both the heterogeneous and the homogeneous models.*

Proof. Suppose that there are λs^i M' -covered beads in $\ell_{i,j}$. A monochromatic arc intersecting $M'(\ell_{i,j})$ is called covered if it contains at least one M' -covered vertex, and uncovered otherwise. If we list the monochromatic arcs intersecting $M'(\ell_{i,j})$ in the appropriate direction, the sizes are (not strictly) monotone increasing. The first one and the last one are covered, and for every covered monochromatic arc (except the last one) the next monochromatic arc is uncovered. The last monochromatic arc has length at most s^{i-1} , thus we get that inside $M'(\ell_{i,j})$ the number of M' -covered beads is at most the number of M' -uncovered beads plus s^{i-1} , that is the number of M' -uncovered beads is at least $\lambda s^i - s^{i-1}$. Thus in $P(i, j)$ the number of M' -covered beads is exactly $2\lambda s^i$ and the number of M' -uncovered beads is at least $(1 - \lambda)s^i + \lambda s^i - s^{i-1} = s^i - s^{i-1}$. So it is enough to prove that

$$2\lambda s^i \leq (2/3 + 1/s)((1 + 2\lambda)s^i - s^{i-1}),$$

or equivalently $6\lambda s^2 \leq (2s + 3)((1 + 2\lambda)s - 1) = (2 + 4\lambda)s^2 + (1 + 6\lambda)s - 3$, i.e., $(2 - 2\lambda)s^2 + (1 + 6\lambda)s \geq 3$ which is evident because $s \geq 1$ and $0 \leq \lambda \leq 1$. \square

Theorem 7. *For Construction 1, the size of the maximum proper matching is at most $2n/3 + o(n)$ in both the heterogeneous and the homogeneous models.*

Proof. The subpartition defined above covers some $n' \leq 2n$ beads, among them $(2/3 + 1/s)n' \leq (4/3 + 2/s)n$ are M' -covered. The number of beads covered by $M \setminus M'$ can be bounded by $6s^s$. So the total number of covered beads is at most $(4/3 + 2/s)n + 6s^s = 4n/3 + 8s^s = 4n/3 + o(n)$, as if n tends to ∞ , then s also tends to ∞ , so $8s^s/n = 8/s$ tends to zero. \square

3 Proof of Theorem 3

We present here our main construction showing that the size of the maximum proper matching is at most αn where α can be arbitrary close to $2 - \sqrt{2} = 0.5857\dots < 0.5858$.

Construction 2. *Let $s \geq 2$ be a integer parameter, and let $n = s^{5s+1}$. The necklace consists of s large arcs, each having s^{5s} blue and s^{5s} red beads. Let L_1, \dots, L_s denote the large arcs.*

L_i is divided into s^{2s-i} red and s^{2s-i} blue arcs, the colors alternates. Let $\ell_{i,j}$ denote the j^{th} arc of L_i , where $1 \leq j \leq 2s^{2s-i}$. The arc $\ell_{i,j}$ always consists of s^{3s+i} beads. In the next step, we will change the color of some beads in each $\ell_{i,j}$ in the following way. Let $\lambda \leq \frac{1}{2}$ be a positive parameter and for a fixed i , let's divide each $\ell_{i,j}$ into s^{s+2i} intervals of size s^{2s-i} and in each of these tiny intervals, change the color of $\lfloor \lambda s^{2s-i} \rfloor$ beads backwards from the clockwise end of the tiny interval. We will refer to those beads whose color were changed as dust in $\ell_{i,j}$. (See Figure 1).

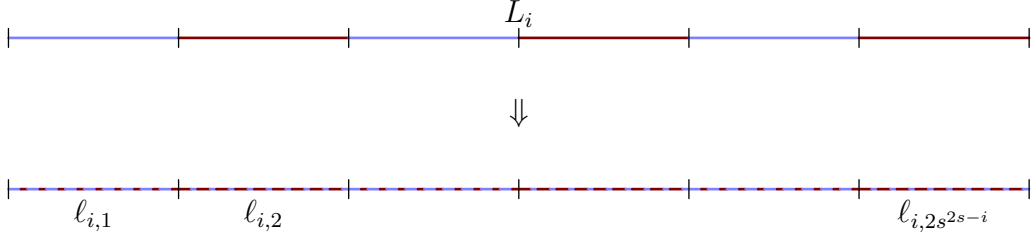


Figure 1: The intervals L_i and $\ell_{i,j}$

First we bound $\text{mono}(L)$ for a necklace L given by this construction in order to prove the last statement of Theorem 3.

$$\text{mono}(L) = \sum_{i=1}^s \sum_{j=1}^{2s^{2s-i}} 2s^{s+2i} = 4s^{3s} \sum_{i=1}^s s^i \leq 4 \frac{s^{s+1} - 1}{s - 1} s^{3s} \leq 8s^{4s},$$

which is $O(n^{4/5}) = o(N)$.

We will see that for $\lambda = 1 - \frac{1}{\sqrt{2}}$, as s tends to ∞ , we will get the desired bound, i.e., the upper bound on the size of the proper matching tends to $2 - \sqrt{2}$. We will use the little-o notation, e.g., $n/s = o(n)$.

For analyzing this construction we fix an optimal proper secant-matching pair in either the homogeneous or the heterogeneous model, and denote this optimal matching by M . The secant may split at most two large arcs, call them L_p and L_r . If, e.g., one end of the secant is between the large arcs L_j and L_{j+1} , then let $p = j$. We may assume that $p < r$ (if $p = r$, then every matching edge has one end in L_p , so $|M| \leq n/s = o(n)$).

Let $M' \subseteq M$ consist of those matching edges for which no end-vertex is inside the set $L_p \cup L_r$. Obviously, $|M| \leq |M'| + 2n/s = |M'| + o(n)$. We call a pair of indices (g, h) *bonded*, if there exists at least one edge of M' connecting L_g and L_h .

Lemma 8. *The number of the bonded pairs is at most $s - 3$.*

Proof. Consider the auxiliary graph with vertices $\{1, 2, \dots, s\} \setminus \{p, r\}$, and connect two vertices if the corresponding pair is bonded. There is no cycle in this graph, otherwise a cycle yields a crossing in M . Thus the auxiliary graph may have at most $(s - 2) - 1$ edges. \square

Let I be an interval. If x and y are the first and last M' -matched bead in I , and $M'(x)$ and $M'(y)$ are their matched partners, then we assign the arc spanned by $M'(x)$ and $M'(y)$ to I , denote this arc by $M'(I)$.

Let (g, h) be a bonded pair, where $g < h$. For $i \in \{1, \dots, s^{2s-g}\}$ let's call a pair of intervals $(\ell_{g,2i-1}, \ell_{g,2i})$ (g, h) -*regular*, if there exists $j \in \{1, \dots, 2s^{2s-h}\}$ such that $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \subseteq \ell_{h,j}$. Let's call an edge of M' *regular* if one of the end-vertices are in a (g, h) -regular pair for some $g < h$. Denote the set of regular edges by M'' . An edge of M' is called a (g, h) -edge if its end-vertices are in L_g and in L_h , respectively; and irregular if it is not regular.

Lemma 9. $|M'| \leq |M''| + 6n/s = |M''| + o(n)$.

Proof. Consider a bonded pair (g, h) for some $g < h$. First we are going to bound the number of irregular (g, h) -edges.

Take an $i \in \{1, \dots, s^{2s-g}\}$ for which $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \cap L_h \neq \emptyset$ but $(\ell_{g,2i-1}, \ell_{g,2i})$ is not (g, h) -regular. It means that either there exist a ‘bad index’ $j \in \{1, \dots, 2s^{2s-h} - 1\}$ such that both $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \cap \ell_{h,j}$ and $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \cap \ell_{h,j+1}$ are non-empty. We also call $j = 0$ bad if $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \cap L_{h-1}$ is non-empty (where $L_0 = L_s$), and $j = 2s^{2s-h}$ bad if $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \cap L_{h+1}$ is non-empty (where $L_{s+1} = L_1$).

Moreover, any j can be a bad index at most once therefore the number of such i ’s is at most $2s^{2s-h} + 1$. Even if all the beads in these non- (g, h) -regular pairs are M' -covered, we got rid of at most $(2s^{2s-h} + 1) \cdot 2|\ell_{g,2i}| = (2s^{2s-h} + 1) \cdot (2s^{3s+g}) < 6s^{5s-(h-g)} \leq 6s^{5s-1}$, since $h > g$. This is true for any bonded pairs, hence altogether we lost at most $(s-3)6s^{5s-1} < 6s^{5s} = 6n/s = o(n)$ M' -edges. \square

From now on, we estimate the number of regular edges. For the sake of simplicity, we will omit the floor and ceiling functions, because the difference in the result is again $o(n)$.

For $g < h$, let’s fix a bonded pair (g, h) . Consider a (g, h) -regular pair of intervals $(\ell_{g,2i-1}, \ell_{g,2i})$ such that $M'(\ell_{g,2i-1} \cup \ell_{g,2i}) \subseteq \ell_{h,j}$.

Until this point, there was no difference between the homogeneous and the heterogeneous case. In the sequel, there still won’t be any significant difference, the calculations work the same way in both cases. We now present the calculation for the homogeneous case, and we will assume that the ‘main’ color of $\ell_{h,j}$ is blue (i.e. the dust is red), the ‘main’ color of $\ell_{g,2i-1}$ is red, thus the ‘main’ color of $\ell_{g,2i}$ is blue.

Let’s denote the *efficiency* of the matching M'' on an interval I with

$$\text{eff}(I) = \frac{\# \text{ of } M'' - \text{covered beads in } I \cup M'(I)}{|I \cup M'(I)|}.$$

In the following lemma, we will show that the efficiency cannot exceed $2 - \sqrt{2} + o(1)$ for a suitable λ .

Lemma 10. $\text{eff}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq 2 - \sqrt{2} + o(1)$ if $\lambda = 1 - \frac{1}{\sqrt{2}}$.

Proof. Recall that $\ell_{g,2i-1}$ (and also $\ell_{g,2i}$) is divided into s^{s+2g} red and s^{s+2g} blue monochromatic intervals. We will call them $\ell_{g,2i-1}^{(\text{red},1)}, \dots, \ell_{g,2i-1}^{(\text{red},s^{s+2g})}, \ell_{g,2i-1}^{(\text{blue},1)}, \dots, \ell_{g,2i-1}^{(\text{blue},s^{s+2g})}$, where the *red* and *blue* indicates the color of the interval.

Also recall that, $\ell_{h,j}$ is divided into blue and red monochromatic intervals (the color alternates) whose sizes are $(1 - \lambda)s^{2s-h}$ and λs^{2s-h} , respectively. We define numbers a_k, b_k, c_k and d_k for $1 \leq k \leq s^{s+2g}$ in the following way.

Assume that the number of M'' -covered beads from $\ell_{g,2i-1}^{(\text{red},k)}$ is x . Then $a_k = \frac{x}{\lambda s^{2s-h}}$, i.e., the necessary number of small red intervals (dust) from $\ell_{h,j}$ to cover that many beads.

Similarly, assume that the number of M'' -covered beads from $\ell_{g,2i-1}^{(\text{blue},k)}$ is x . Then $b_k = \frac{x}{(1-\lambda)s^{2s-h}}$, i.e., the necessary number of blue intervals from $\ell_{h,j}$ to cover that many beads.

We define c_k and d_k in the same way for $\ell_{g,2i}^{(\text{red},k)}$ and $\ell_{g,2i}^{(\text{blue},k)}$, respectively. (See Figure 2.) It is easy to see that $a_k \leq \frac{1-\lambda}{\lambda} s^{h-g}$, $b_k \leq \frac{\lambda}{1-\lambda} s^{h-g}$, and $c_k, d_k \leq s^{h-g}$.

The number of M'' -covered beads in $(\ell_{g,2i-1} \cup \ell_{g,2i}) \cup M'(\ell_{g,2i-1} \cup \ell_{g,2i})$ is

$$2 \left(\sum a_k \lambda s^{2s-h} + \sum b_k (1 - \lambda) s^{2s-h} + \sum c_k \lambda s^{2s-h} + \sum d_k (1 - \lambda) s^{2s-h} \right).$$

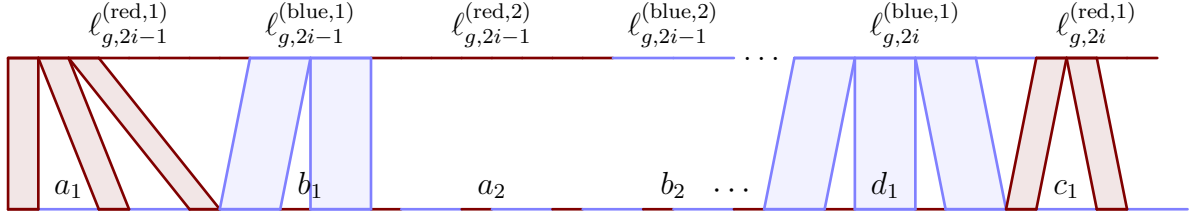


Figure 2: The definition of a_k, b_k, c_k and d_k

In all of the sums, k runs from 1 up to s^{s+2g} , so we use the following shorthands. Let

$$A = \sum_{k=1}^{s^{s+2g}} a_k, \quad B = \sum_{k=1}^{s^{s+2g}} b_k, \quad C = \sum_{k=1}^{s^{s+2g}} c_k, \quad D = \sum_{k=1}^{s^{s+2g}} d_k.$$

Obviously, $|\ell_{g,2i-1} \cup \ell_{g,2i}| = 2s^{3s+g}$. We will give a lower bound on $|M'(\ell_{g,2i-1} \cup \ell_{g,2i})|$. To cover the M'' -matched beads in $\ell_{g,2i-1}^{(\text{red},k)}$, we need at least a_k monochromatic red interval from $\ell_{h,j}$, thus at least $a_k - 1$ monochromatic blue interval remained unused, so $M'(\ell_{g,2i-1}^{(\text{red},k)}) \geq (a_k - 1)s^{2s-h}$. Similarly $M'(\ell_{g,2i-1}^{(\text{blue},k)}) \geq (b_k - 1)s^{2s-h}$, $M'(\ell_{g,2i}^{(\text{red},k)}) \geq (c_k - 1)s^{2s-h}$ and $M'(\ell_{g,2i}^{(\text{blue},k)}) \geq (d_k - 1)s^{2s-h}$.

Altogether, we get that $\text{eff}(\ell_{g,2i-1} \cup \ell_{g,2i})$ is at most

$$\begin{aligned} & 2 \frac{\lambda A s^{2s-h} + (1-\lambda) B s^{2s-h} + \lambda C s^{2s-h} + (1-\lambda) D s^{2s-h}}{[\sum (a_k - 1) s^{2s-h} + \sum (b_k - 1) s^{2s-h} + \sum (c_k - 1) s^{2s-h} + \sum (d_k - 1) s^{2s-h}] + 2s^{3s+g}} = \\ & = 2 \frac{\lambda A + (1-\lambda) B + \lambda C + (1-\lambda) D}{[A + B + C + D - 4s^{s+2g}] + 2s^{3s+g}/s^{2s-h}} = \\ & = 2 \frac{\lambda A + (1-\lambda) B + \lambda C + (1-\lambda) D}{A + B + C + D - 4s^{s+2g} + 2s^{s+g+h}}. \end{aligned}$$

Let's denote this last expression by $\text{eff}(A, B, C, D)$. First, we will show that this expression is monotone increasing in B and D .

$$\text{eff}(A, B, C, D) = 2(1-\lambda) + 2 \frac{(2\lambda - 1)(A + C) - (1-\lambda)(2s^{s+g+h} - 4s^{s+2g})}{A + B + C + D + (2s^{s+g+h} - 4s^{s+2g})}.$$

As $\lambda \leq \frac{1}{2}$ and $s^{s+g+h} \geq s^{s+2g+1} \geq 2s^{s+2g}$, we have that $2\lambda - 1 \leq 0$ and $2s^{s+g+h} - 4s^{s+2g} > 0$, so the numerator of the second term is negative. Thus we can increase the value of this expression by choosing B and D as large as possible which yields:

$$\text{eff}(A, B, C, D) \leq \text{eff}\left(A, \frac{\lambda}{1-\lambda} s^{s+g+h}, C, s^{s+g+h}\right).$$

We will do the same trick for A and C .

$$\text{eff}\left(A, \frac{\lambda}{1-\lambda} s^{s+g+h}, C, s^{s+g+h}\right) =$$

$$\begin{aligned}
&= 2 \frac{\lambda A + (1 - \lambda) \frac{\lambda}{1 - \lambda} s^{s+g+h} + \lambda C + (1 - \lambda) s^{s+g+h}}{A + \frac{\lambda}{1 - \lambda} s^{s+g+h} + C + s^{s+g+h} + 2s^{s+g+h} - 4s^{s+2g}} = \\
&= 2 \frac{\lambda A + \lambda C + s^{s+g+h}}{A + C + \left(\frac{\lambda}{1 - \lambda} + 3\right) s^{s+g+h} - 4s^{s+2g}} = \\
&= 2\lambda + 2 \frac{\left[1 - \lambda\left(\frac{\lambda}{1 - \lambda} + 3\right)\right] s^{s+g+h} + 4\lambda s^{s+2g}}{A + C + \left(\frac{\lambda}{1 - \lambda} + 3\right) s^{s+g+h} - 4s^{s+2g}} = \\
&= 2\lambda + 2 \frac{\frac{2\lambda^2 - 4\lambda + 1}{1 - \lambda} s^{s+g+h} + 4\lambda s^{s+2g}}{A + C + \left(\frac{\lambda}{1 - \lambda} + 3\right) s^{s+g+h} - 4s^{s+2g}}.
\end{aligned}$$

If $\lambda = 1 - \frac{1}{\sqrt{2}}$, then $2\lambda^2 - 4\lambda + 1 = 0$, so

$$\begin{aligned}
&\text{eff} \left(A, \frac{\lambda}{1 - \lambda} s^{s+g+h}, C, s^{s+g+h} \right) = \\
&= 2\lambda + 2 \frac{4\lambda s^{s+2g}}{A + C + \left(\frac{\lambda}{1 - \lambda} + 3\right) s^{s+g+h} - 4s^{s+2g}}.
\end{aligned}$$

The numerator of the second term is $\Theta(s^{s+2g})$ while the denominator is $\Theta(s^{s+g+h})$ because $A \leq \frac{1 - \lambda}{\lambda} s^{s+g+h}$ and $C \leq s^{s+g+h}$. Thus

$$\text{eff} \left(A, \frac{\lambda}{1 - \lambda} s^{s+g+h}, C, s^{s+g+h} \right) = 2\lambda + O(s^{g-h}) = 2 - \sqrt{2} + o(1). \quad \square$$

We have already proved that the number of those edges in the matching M which are not in M'' is negligible. We can partition the rest of the edges (the regular ones) into disjoint subsets according to the (a, b) -bonded pairs determined by their beads. In every such (a, b) -bonded pair (for some $a < b$), we can repeat the argument of Lemma 10. Hence, we can conclude that in this construction the size of a proper matching is at most $(2 - \sqrt{2} + o(1)) n \approx 0.5858n$. \square

Remark 11. One can similarly deal with the end of the proof of Lemma 10 in the heterogeneous case and then conclude that $\mu \leq 2 - \sqrt{2}$, too.

4 Unbalanced necklaces

In this section, we consider the case when the number of red and blue beads are different. There are two possible measurements of unbalancedness we can use. On the one hand, we can define that the number of red beads is φn , while the number of blue beads is n – thus the total number of beads is $N = (1 + \varphi)n$. On the other hand, we can first fix N , the total number of beads, and we define the number of red beads as pN , where $0 < p < 1$. (As before, we will omit integer parts). Of course any of these parameters defines the other one:

$$p = \frac{\varphi}{1 + \varphi},$$

$$\varphi = \frac{p}{1-p}.$$

In the calculations we will use φ , however, we will show our final results both as a function of φ and as a function of p . We consider only the case $1/2 \leq \varphi \leq 2$, i.e., $1/3 \leq p \leq 2/3$.

Now let $\mathcal{N}_\varphi(N)$ denote the set of possible necklaces with $\varphi n + n$ beads where φn ones are red and n ones are blue. Let $\mathcal{M}(L)$ denote the set of proper matchings for a given necklace L in the heterogeneous model. Moreover, let $\mu_\varphi(N) = \min_{L \in \mathcal{N}_\varphi(N)} \max_{M \in \mathcal{M}(L)} 2|M|$. We are interested in $\frac{\mu_\varphi(N)}{N}$, the ratio of the covered beads to the total number of beads. Finally let $\mu_\varphi = \limsup_{N \rightarrow \infty} \mu_\varphi(N)$. For the homogeneous model we similarly define $\mu_\varphi^{\text{hom}}(N)$ and μ_φ^{hom} .

In this section we give a modified version of Construction 2. Let t, u, v, w be real numbers such that $t + w = \varphi$, $u + v = 1$, $t \geq u$ and $v \geq w$. For the sake of simplicity, we will omit the floor and ceiling functions again even in the description of the construction.

Construction 3. Let $s \geq 2$ be a integer parameter, and let $n = s^{5s+1}$. The necklace consists of s large arcs, each having φs^{5s} red and s^{5s} blue beads. Let L_1, \dots, L_s denote the large arcs.

L_i is divided into s^{2s-i} red and s^{2s-i} blue arcs, the colors alternates. Let $\ell_{i,j}$ denote the j^{th} arc of L_i , where $1 \leq j \leq 2s^{2s-i}$. If j is odd, then the arc $\ell_{i,j}$ always consists of $(t+u)s^{3s+i}$ beads. If j is even, then the arc $\ell_{i,j}$ always consists of $(v+w)s^{3s+i}$ beads.

In the next step, we will change the color of some beads in each $\ell_{i,j}$ in the following way. If j is odd, then let's divide each $\ell_{i,j}$ into s^{s+2i} intervals of size $(t+u)s^{2s-i}$, and in each of this tiny intervals, change the color of us^{2s-i} beads backwards from the clockwise end of the tiny interval from red to blue. If j is even, then let's divide each $\ell_{i,j}$ into s^{s+2i} intervals of size $(v+w)s^{2s-i}$, and in each of this tiny intervals, change the color of ws^{2s-i} beads backwards from the clockwise end of the tiny interval from blue to red. We will refer to those beads whose color were changed as dust in $\ell_{i,j}$.

Notice that we cannot use the same formula for a_k, b_k, c_k, d_k since there is no λ in the current setup, but the definition of these parameters will be the same. Namely, these parameters will again denote the necessary number of "small" intervals (of size $O(s^{2s-h})$) to cover all the beads in the matching of the corresponding "big" intervals (of size $O(s^{2s-g})$) from the opposite side. Intentionally, we do not write down these definitions precisely because there would be too many very similar definitions depending on the color of the corresponding intervals. Otherwise, we will use the same notations. All estimations before the introduction of the efficiency notion in the analysis of Construction 2 work exactly the same way.

In order to properly calculate the efficiency, we have to distinguish the homogeneous and heterogeneous case.

4.1 Homogeneous case

The calculation will be different, if $\ell_{h,j}$ is blue, and if it is red. First, consider the blue case. We will use the same argument as in Lemma 10.

Therefore, we can get that $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i})$ is at most

$$\text{eff}_{\text{blue}}(A, B, C, D) = 2 \frac{wA + vB + wC + vD}{(v+w)(A+B+C+D) + (\varphi+1)s^{s+g+h} + \Theta(s^{s+2g})},$$

where $A \leq \frac{t}{w}s^{s+g+h}$, $B \leq \frac{u}{v}s^{s+g+h}$, $C \leq s^{s+g+h}$, $D \leq s^{s+g+h}$.

Again, this expression is monotone increasing in B and D , since $v \geq w$, so

$$\begin{aligned} \text{eff}_{\text{blue}}(A, B, C, D) &\leq \frac{2}{v+w} \frac{w(A+C) + v(\frac{u}{v}+1)s^{s+g+h}}{A+C + (\frac{u}{v}+1)s^{s+g+h} + \frac{\varphi+1}{v+w}s^{s+g+h} + \Theta(s^{s+2g})} = \\ &= \frac{2w}{v+w} + \frac{2}{v+w} \frac{((v-w)(\frac{u}{v}+1) - \frac{\varphi+1}{v+w}w)s^{s+g+h} + \Theta(s^{s+2g})}{A+C + (\frac{u}{v}+1)s^{s+g+h} + \frac{\varphi+1}{v+w}s^{s+g+h} + \Theta(s^{s+2g})}. \end{aligned}$$

Thus if we choose the parameters such that $(v-w)(\frac{u}{v}+1) - \frac{\varphi+1}{v+w}w = 0$, then we can conclude that $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2w}{v+w} + o(1)$.

Using the fact $u+v=1$, we get the quadratic equation $(\frac{v}{w})^2 - (\varphi+1)(\frac{v}{w}) - 1 = 0$ with one positive root

$$\psi_1 = \frac{\varphi+1 + \sqrt{(\varphi+1)^2 + 4}}{2},$$

thus $v = \psi_1 w$, and $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2}{\psi_1+1} + o(1)$.

In the second case where $\ell_{h,j}$ is red, one can go through the similar arguments as above and can get to the point in which we choose the parameters such that the main term of the corresponding numerator is 0. Then by using the fact, $t+w=\varphi$, we get that $(\frac{t}{u})^2 - \frac{\varphi+1}{\varphi}(\frac{t}{u}) - 1 = 0$, which has one positive root

$$\psi_2 = \frac{\frac{\varphi+1}{\varphi} + \sqrt{\left(\frac{\varphi+1}{\varphi}\right)^2 + 4}}{2},$$

thus $t = \psi_2 u$, and $\text{eff}_{\text{red}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2}{\psi_2+1} + o(1)$.

It is easy to check that the following system of equations has a unique solution. Also, if $\frac{1}{2} \leq \varphi \leq 2$, then $u, v, w, t \geq 0$.

$$t+w=\varphi$$

$$u+v=1$$

$$v=\psi_1 w$$

$$t=\psi_2 u$$

Altogether we get that at most $\max\left(\frac{2}{\psi_1+1}, \frac{2}{\psi_2+1}\right) + o(1)$ fraction of the beads are matched. By symmetry, we can assume that $\varphi \geq 1$. Then $\psi_1 \geq \psi_2$, thus $\frac{2}{\psi_1+1} \leq \frac{2}{\psi_2+1}$, and

$$\frac{2}{\psi_2+1} = \frac{4}{\varphi+3+\sqrt{\varphi^2+2\varphi+5}}.$$

One can get $\frac{\varphi}{1+\varphi}$ as a trivial lower bound (dashed green) for this case if we choose a secant which cut the red beads into two equal parts and consider a matching which covers all the red beads (and do not care about blue beads at all). In particular, for $\varphi=2$ we have equality. (See Figures 3 and 4.)

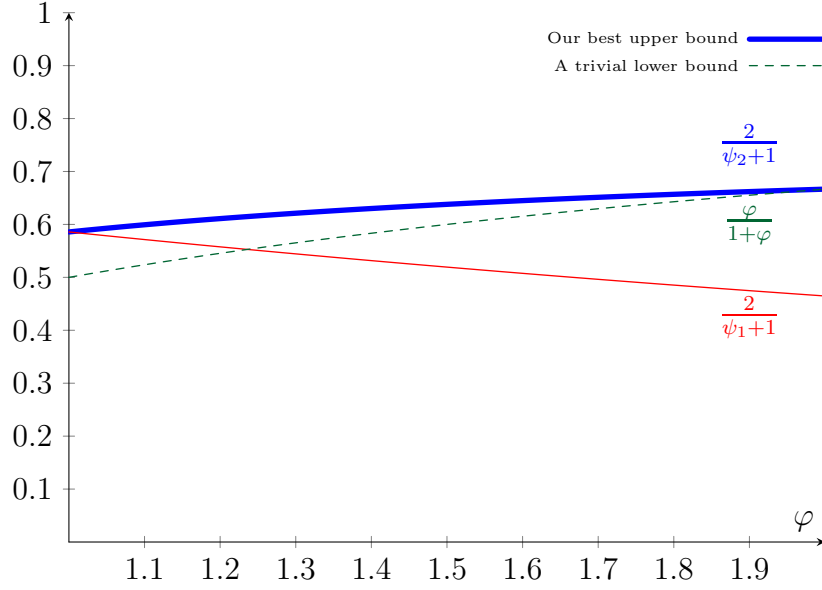


Figure 3: Lower and upper bounds in the homogeneous case as a function of φ , in the case $1 \leq \varphi \leq 2$.

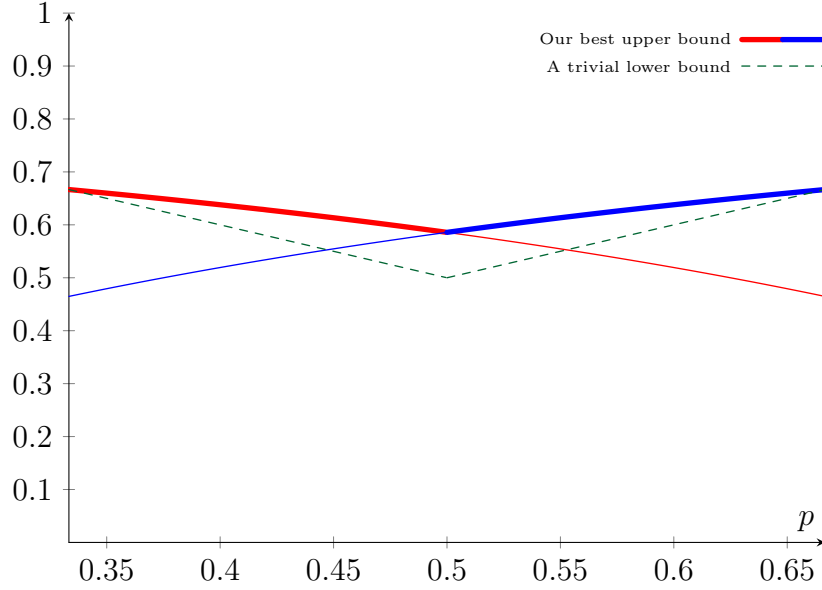


Figure 4: Lower and upper bounds in the homogeneous case as a function of p , in the case $1/3 \leq p \leq 2/3$.

4.2 Heterogeneous case

The calculation will be different again depending on the color of $\ell_{h,j}$. First, consider the case when $\ell_{h,j}$ is blue. We will use the same argument as in Lemma 10.

Therefore, we can get that $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i})$ is at most

$$\text{eff}_{\text{blue}}(A, B, C, D) = 2 \frac{vA + wB + vC + wD}{(v + w)(A + B + C + D) + (\varphi + 1)s^{s+g+h} + \Theta(s^{s+2g})},$$

where $A \leq \frac{t}{v}s^{s+g+h}$, $B \leq \frac{u}{w}s^{s+g+h}$, $C \leq \frac{w}{v}s^{s+g+h}$, $D \leq \frac{v}{w}s^{s+g+h}$.

Again, this expression is monotone increasing in A and C , since $v \geq w$, so

$$\begin{aligned} \text{eff}_{\text{blue}}(A, B, C, D) &\leq \frac{2}{v+w} \frac{w(B+D) + v\left(\frac{t}{v} + \frac{w}{v}\right)s^{s+g+h}}{B+D + \left(\frac{t}{v} + \frac{w}{v}\right)s^{s+g+h} + \frac{\varphi+1}{v+w}s^{s+g+h} + \Theta(s^{s+2g})} = \\ &= \frac{2w}{v+w} + \frac{2}{v+w} \frac{((v-w)\left(\frac{t}{v} + \frac{w}{v}\right) - \frac{\varphi+1}{v+w}w)s^{s+g+h} + \Theta(s^{s+2g})}{B+D + \left(\frac{u}{v} + 1\right)s^{s+g+h} + \frac{\varphi+1}{v+w}s^{s+g+h} + \Theta(s^{s+2g})}. \end{aligned}$$

Thus if we choose the parameters such that $(v-w)\left(\frac{t}{v} + \frac{w}{v}\right) - \frac{\varphi+1}{v+w}w = 0$, then we can conclude that $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2w}{v+w} + o(1)$.

Using the fact, $t+w = \varphi$, we get the quadratic equation $\left(\frac{v}{w}\right)^2 - \frac{\varphi+1}{\varphi}\left(\frac{v}{w}\right) - 1 = 0$ with one positive root which we called ψ_2 , and $\text{eff}_{\text{blue}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2}{\psi_2+1} + o(1)$.

If $\ell_{h,j}$ is red, then we get that $\left(\frac{t}{u}\right)^2 - (\varphi+1)\left(\frac{t}{u}\right) - 1 = 0$, which has one positive root ψ_1 , and $\text{eff}_{\text{red}}(\ell_{g,2i-1} \cup \ell_{g,2i}) \leq \frac{2}{\psi_1+1} + o(1)$.

It is easy to check that the following system of equations has a unique solution, and u, v, w, t are always positive.

$$\begin{aligned} t + w &= \varphi \\ u + v &= 1 \\ v &= \psi_2 w \\ t &= \psi_1 u \end{aligned}$$

Altogether we get that at most $\max\left(\frac{2}{\psi_1+1}, \frac{2}{\psi_2+1}\right) + o(1)$ fraction of the beads are matched, which is the same as in the homogeneous case. By symmetry, we can assume that $\varphi \geq 1$. Then $\psi_1 \geq \psi_2$, thus $\frac{2}{\psi_1+1} \leq \frac{2}{\psi_2+1}$, and

$$\frac{2}{\psi_2+1} = \frac{4}{\varphi+3+\sqrt{\varphi^2+2\varphi+5}}.$$

One can get $\frac{1}{1+\varphi}$ as a trivial lower bound (dashed green) for this case if we choose a secant which cut the blue beads into two equal parts. Since $\varphi \geq 1$, one can find a matching which covers all the blue beads from the side which contains at most as many red beads as the other side. On the other hand, one can get $\frac{2}{1+\varphi}$ as a trivial upper bound (dotted yellow) because in every edge of the matching exactly one of the endpoints have color blue. In particular, for $\varphi = 2$ we have equality with the trivial bound, thus our construction gives a non-trivial bound only for $\varphi < 2$. (See Figures 5 and 6.)

Acknowledgement

This research was started during the 9th Emléktábla Workshop, 2019. The authors are thankful for the organizers for inviting them. We also thank D. Pálvölgyi, G. Damásdi, T. Fleiner and Zs. Jankó for valuable questions and observations.

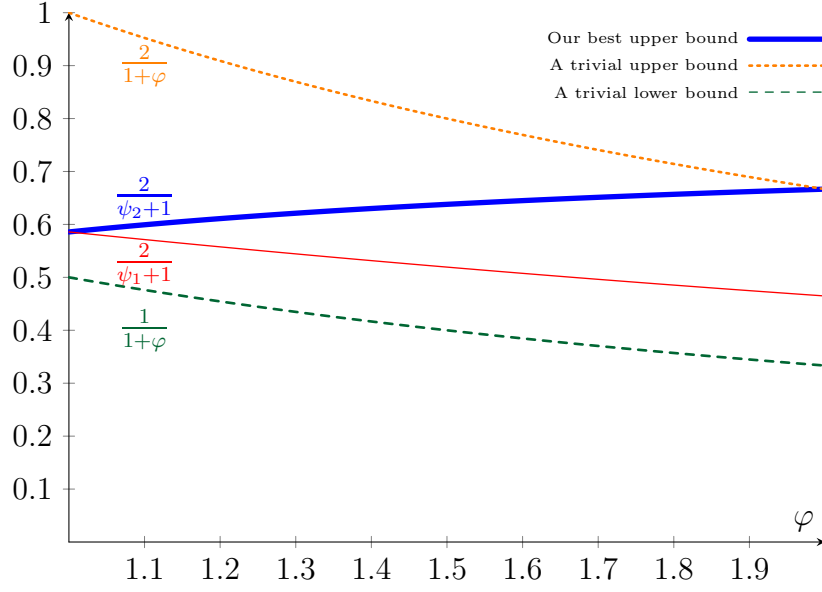


Figure 5: Lower and upper bounds in the heterogeneous case as a function of φ , in the case $1 \leq \varphi \leq 2$.

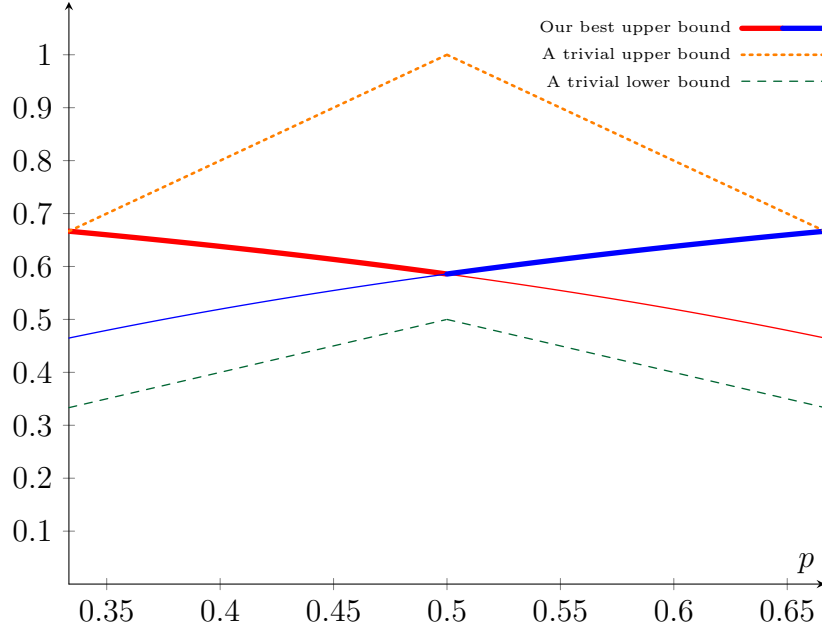


Figure 6: Lower and upper bounds in the heterogeneous case as a function of p , in the case $1/3 \leq p \leq 2/3$.

References

- [1] R. B. Lyngsø and C. N. S. Pedersen, Protein Folding in the 2D HP Model, *BRICS Report Series*, **RS-99-16**, (1999).
- [2] J. Kynčl, J. Pach and G. Tóth, Long alternating paths in bicolored point sets, in Graph Drawing (J. Pach, ed.), Lecture Notes in Computer Science **3383**, Springer-Verlag,

- Berlin, (2004), pp. 340–348.
- [3] J. Kynčl, J. Pach and G. Tóth, Long alternating paths in bicolored point sets, *Discrete Mathematics*, **308**, (2008), pp. 4315–4322.
 - [4] M. Abellanas, A. Garcia, F. Hurtado and J. Tejel, Caminos alternantes, in: *X Encuentros de Geometria Computacional (in Spanish)*, Sevilla, (2003), pp. 7–12.
 - [5] G. Brevier, M. Preissmann and A. Sebő, personal communication (2004).
 - [6] P. Hajnal and V. Mészáros, A note on noncrossing path in colored convex sets, manuscript, (2010).
 - [7] V. Mészáros, Extremal problems on planar point sets, *Ph.D. thesis*, doktori.bibl.u-szeged.hu/688/1/mvdoktori.pdf, (2011).
 - [8] V. Mészáros, Separated matchings and small discrepancy colorings. *Computational Geometry, Lecture Notes in Comput. Sci.*, **7579**, Springer, Cham, (2011), pp. 236–248.
 - [9] C. Müllner and A. Ryzhikov, Palindromic Subsequences in Finite Words. *arXiv*, **1901.07502**, (2019).
 - [10] C. Müllner and A. Ryzhikov, Palindromic subsequences in finite words. In Proc 13th Int. Conf. Language and Automata Theory and Applications (LATA), (2019), pp. 460–468.
 - [11] W. Mulzer and P. Valtr, Long alternating paths exist. *arXiv*, **2003.13291**, (2020).